Molecular Analysis of Tuberculosis Transmission in Household Contacts Using a Culturomics Approach

Astika Sewcharran (1822059)

Supervisor 1: Prof. Bavesh Kana

Supervisor 2: Dr. Bhavna Gordhan

Centre of Excellence for Biomedical TB Research (CBTBR)

A dissertation submitted to the Faculty of Health Science, University of the Witwatersrand, Johannesburg, in fulfilment of the requirements for the degree of Master of Science.

Johannesburg, 2025

Declaration

I, Astika Sewcharran, declare that this dissertation is my own, unaided work. It is being submitted for the degree of Master of Science at the University of the Witwatersrand, Johannesburg. It has not been submitted before for any other degree or examination at any other university.

Astika Sewcharran

Signed on the 26th day of March 2025 in Johannesburg.

Dedication

In memory of my father,

Ashwin Sewcharran

1972 - 2010

Conference Presentations from this Study

1. Household TB Transmission Using Culturomics Assays presented at the FSPI Whole Genome Sequencing as a Tool for Surveillance of *Mycobacterium bovis* and Rifampicin-Resistant *Mycobacterium tuberculosis* in Madagascar and South Africa in Cape Town, 2022
2. Assessing Household Tuberculosis Transmission using detection of Differentially Culturable Tubercle Bacteria (HoTT-DCTB) presented at the FSPI final hybrid in-person/ online meeting held at the National Institute of Communicable Disease (NICD) in Johannesburg, 2022
3. Differential culturability of *M. tuberculosis* uncovers transmission links within households which are missed by standard diagnostic tests – Eposter with oral presentation at The Union World Conference on Lung Health 2024 in Bali, Indonesia.

Publications During the Study

1. Ealand C, Peters J, Jacobs O, **Sewcharran A**, Ghoor A, Golub J, Brahmbhatt H, Martinson N, Dangor Z, Lala SG, Kana B. Detection of Mycobacterium tuberculosis Complex Bacilli and Nucleic Acids from Tongue Swabs in Young, Hospitalized Children. Front Cell Infect Microbiol. 2021 Jun 14;11:696379. doi: 10.3389/fcimb.2021.696379. PMID: 34195103; PMCID: PMC8238041.
2. Gordhan BG, **Sewcharran A**, Letsoalo M, Chinappa T, Yende-Zuma N, Padayatchi N, Naidoo K, Kana BD. Detection of differentially culturable tubercle bacteria in sputum from drug-resistant tuberculosis patients. Front Cell Infect Microbiol. 2022 Sep 9;12:949370. doi: 10.3389/fcimb.2022.949370. PMID: 36159642; PMCID: PMC9500503.
3. Ealand CS, **Sewcharran A**, Peters JS, Gordhan BG, Kamariza M, Bertozzi CR, Waja Z, Martinson NA, Kana BD. The performance of tongue swabs for detection of pulmonary tuberculosis. Front Cell Infect Microbiol. 2023 Sep 6;13:1186191. doi: 10.3389/fcimb.2023.1186191. PMID: 37743867; PMCID: PMC10512057.
4. Gordhan BG, Padarath K, **Sewcharran A**, McIvor A, VanNieuwenhze MS, Waja Z, Martinson N, Kana BD. Clinical Strains of *Mycobacterium tuberculosis* Representing Different Genotype Families Exhibit Distinct Propensities to Adopt the Differentially Culturable State. Pathogens. 2024 Apr 12;13(4):318. doi: 10.3390/pathogens13040318. PMID: 38668273; PMCID: PMC11054447.
5. Bhavna G. Gordhan[§], **Astika Sewcharran[§],** Ziyaad Waja, Neil Martinson, Mariana Araujo-Pereira, Alessandra de Souza Saboia, Maria Cristina S. Lourenço, Adriano Gomes-Silva, Valeria C. Rolla[8*], Bavesh D. Kana[1*] and the Associative BRICS Research in Covid-19 and TB (ABRICOT) Research Group. Oral rinse as an alternative diagnostic specimen for detection of tuberculosis. Submitted to American Journal of Respiratory and Critical Care Medicine (24th March 2025). **Equal first author.**
6. **Astika Sewcharran[†],** Bhavna G. Gordhan[†], Serene Keenan, Dhanishta Patel, Linrui, Tang, Ziyaad Waja, Barun Mathema, Neil Martinson and Bavesh D. Kana. The differential culturability of *Mycobacterium tuberculosis* has the propensity to produce higher positivity rates in infected individuals and their respective household contacts. *In preparation.*

Abstract

Tuberculosis (TB), caused by *Mycobacterium tuberculosis* (*Mtb)*, continues to be a major global public health concern. While molecular diagnostics have advanced, culture is still the gold standard for TB diagnosis. However, certain sub-populations of *Mtb* in clinical samples, termed differentially culturable tubercule bacteria (DCTB), cannot grow under standard culture conditions, resulting in them being missed during diagnosis, particularly in paucibacillary disease. This also impacts transmission studies, which are based on the premise that all bacteria in a clinical sample have equivalent abilities to grow therefore, complicating transmission network mapping. This study employed a culturomics approach, integrating various culture techniques, quantitative polymerase chain reaction (qPCR), and whole genome sequencing (WGS) to investigate TB transmission dynamics within households of TB patients. In 247 household contacts (HHCs) of 100 index participants, 112 (45%) tested positive for *Mtb* in at least one assay. The standard Mycobacterial Growth Indicator Tube (MGIT) assay identified the majority of *Mtb*-positive cases, however, the most probable number (MPN) assay used to detect DCTB, and colony-forming unit (CFU) assays detected 55 additional cases missed by MGIT culture across the 16-month timeline. Combining MGIT culture and MPN assays increased the yield of TB detection in HHCs from 14.2% to 22.7%. MPN assays also identified an additional 15 (26%) transmission links within households, which were missed by routine culture. An in-house molecular diagnostic method, qPCR identified 59 *Mtb*-positive HHCs that were culture negative. These findings highlight the limitations of the current culture techniques in identifying all populations of *Mtb*. Integrating routine culture with MPN/CFU assays provides a more comprehensive understanding of *Mtb*-positive HHCs and the communities they reside in thereby, contributing to improved TB control strategies.

Acknowledgements

I would like to express my heartfelt gratitude to:

Prof. Bavesh Kana, my supervisor, for providing invaluable mentorship and the opportunities that made the completion of this degree possible. Your expert guidance and unwavering support have been instrumental throughout this journey.

Dr. Bhavna Gordhan, my supervisor, for stepping in with dedication and care. Your insightful guidance and commitment played a crucial role in shaping the final stages of this dissertation, I would not have reached this stage without your help, and I am truly appreciative of your contributions.

The National Institute of Health (NIH) for the financial support that made this research possible.

The CBTBR team for your encouragement and the positive environment that made this process feel less intimidating.

My family - thank you for your unwavering belief in me and for being a constant source of motivation.

Ryan – thank you for your patience, encouragement, and the many sacrifices you made along the way. I am deeply grateful for you.

Table of Contents

List of Abbreviations and Symbols

Abbreviation	Full Term
%	Percentage
µl	Microliter
µm	Micrometer
µM	Micromolar
^{o}C	Degrees Celsius
Ag85	Antigen 85
AIDS	Acquired Immunodeficiency Syndrome
ART	Antiretroviral Therapy
ARV	Antiretroviral Drugs
BCF	Binary Call Format
BCG	Bacillus Calmette–Guérin
BD	Becton Dickinson
BMI	Body Mass Index
BWA	Burrows-Wheeler Aligner
CBTBR	Centre of Excellence for Biomedical TB Research
CDC	Centers for Disease Control and Prevention
CF	Culture Filtrate
CFU	Colony Forming Unit
CTAB	Cetyltrimethylammonium Bromide
DCS	Differentially Culturable State
DCTB	Differentially Culturable Tubercle Bacteria
DMN-Tre	4-N,N-Dimethylamino-1,8-naphthalimide conjugate of Trehalose
DNA	Deoxyribonucleic Acid
DOT	Directly Observed Therapy
E	Efficiency
E. coli	*Escherichia coli*
EDQ	Eclipse Dark Quencher
EMB	Ethambutol
g	Grams
G+C	Guanine and Cytosine

gDNA	Genomic DNA
HHC(s)	Household Contact(s)
HIV	Human Immunodeficiency Virus
HoTT DCTB	Assessing Household Tuberculosis Transmission using detection of Differentially Culturable Tubercle Bacteria
IGRA	Interferon Gamma Release Assay
INH	Isoniazid
IPT	Isoniazid Preventive Therapy
IQR	Interquartile Range
L	Litre
LAMP	Loop-mediated Isothermal Amplification
LF-LAM	Lateral Flow Urine Lipoaribinomannan
LJ	Löwenstein-Jensen
LOG	Logarithm
LPA	Line Probe Assay
LTBI	Latent Tuberculosis Infection
kg	Kilogram
km	kilometer
M	Molar
M. bovis	*Mycobacterium bovis*
MDR	Multidrug Resistant
MGB	Minor Groove Binding Probe
MGIT	Mycobacterium Growth Indicator Tube
MIRU-VNTR	Mycobacterial Interspersed Repetitive Unit Variable Number Tandem Repeats
mL	Milliliter
mm	Millimeter
mM	Millimolar
MPN	Most Probable Number Assay
M. smeg	*Mycobacterium smegmatis*
Mtb	*Mycobacterium tuberculosis*
MTBC	Mycobacterium Tuberculosis Complex
n	Number
NAAT	Nucleic-Acid Amplification Tests
ng	Nanogram

NTC	No Template Control
NTM	Non-tuberculosis Mycobacteria
OADC	Oleic Acid-Albumin-Dextrose-catalase
OD	Optical Density
PANTA	Polymyxin, Amphotericin B, Nalidixic Acid, Trimethoprim and Azlocillin
PBS	Phosphate Buffered Saline
PCR	Polymerase Chain Reaction
pH	Potential of Hydrogen
p-value	Probability
PZA	Pyrazinamide
qPCR	Quantitative Polymerase Chain Reaction
R^2	Coefficient of Determination
RFLP	Restriction Fragment Length Polymorphism
RI	Resuscitation Index
RIF	Rifampin
RIFAFOUR	Rifampin, Isoniazid, Ethambutol and Pyrazinamide
Rpf	Resuscitation Promoting Factor
rpm	Revolutions Per Minute
RR	Rifampin-Resistance
RRDR	Rifampin Resistance-Determining Region
SARS-CoV-2	Severe Acute Respiratory Syndrome Coronavirus 2
SNP	Single Nucleotide Polymorphism
Spoligotyping	Spacer Oligonucleotide Typing
TB	Tuberculosis
TE	Tris and EDTA Buffer
TTP	Time to Positivity
TST	Tuberculin Skin Test
UN	United Nations
VCF	Variant Call Format
VNTR	Variable Number of Tandem Repeats
WGS	Whole Genome Sequencing
WHO	World Health Organization
W/V	Weight/Volume

XDR	Extensively Drug Resistant
Z-N stain	Ziehl-Neelsen Stain

List of Figures

List of Tables

1. Introduction

1.1 Tuberculosis Disease

Tuberculosis (TB), caused by the pathogen *Mycobacterium tuberculosis* (*Mtb*), results in significant mortality and morbidity, worldwide (1). The World Health Organisation (WHO) Global TB Report 2024 reports an estimated 10.8 million individuals infected with TB, with approximately 1.25 million TB related deaths. The reported incident rate was 150-400 per 100 000 population in most of the 30 high TB burden countries, including South Africa (Figure 1) (2) which contributed to 3% (324 000) of TB cases and 4% (50 000) of TB deaths worldwide (3). Whereas, low burden countries such as Canada, Australia and New Zealand have yearly incidence rates of less than 10 cases per 100 000 population (4, 5).

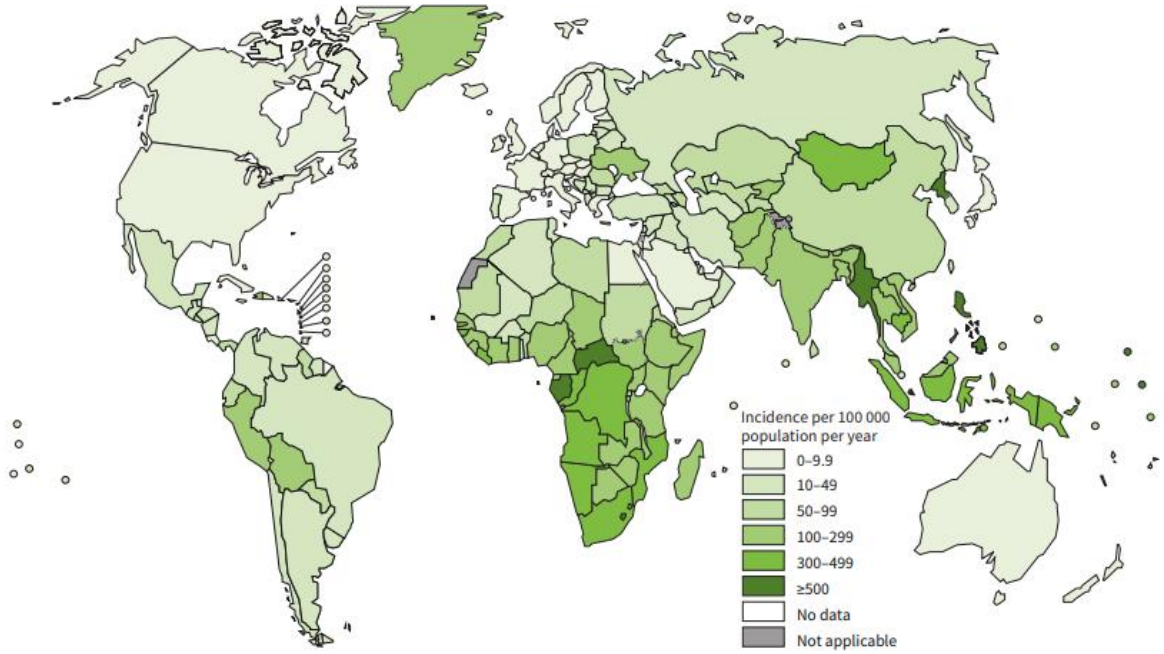


Figure 1: Global Trends in TB Incidence in 2023. The highest TB incidence rates in 2023 were reported in parts of Africa and Asia, ranging from 10 and > 500 cases per 100 000 population, followed by South America and North America with incidence rates ranging from 0 to 299 cases per 100 000 population.
Image taken from WHO's Global Tuberculosis Report 2024.

Despite the African and European regions making the most progress in reducing TB incidence rates, the burden of TB is still disproportionately high within Africa. South Africa, for example, has an incidence rate of 300-499 cases per 100 000 population, placing it among the highest-burden countries (2). Therefore, TB research in South Africa needs to be a national priority to reduce the TB disease burden, especially in populations at a higher risk of TB infection.

Social stratification based on wealth, ethnicity and education are high risk factors for TB that affect the most vulnerable communities. Individuals at the lower levels of the stratification system are more likely to contract TB as they experience (I) overcrowding and poor ventilation in both their working and living environments, (II) malnutrition and weaker immune systems with an increased risk of developing active TB disease and (III) lack of basic education and understanding of health care and the associated preventative measures for risk factors such as alcohol abuse, smoking and human immunodeficiency virus (HIV). These individuals also experience a lack of access to adequate health care facilities (6, 7). Acknowledging the inequalities of these communities by health sectors and enforcing policies to alleviate poverty will aid in reducing the risk of progression from infection to active TB disease, while improving the health care of individuals with active TB **(7)**.

Tuberculosis is an ancient disease, as research in human skeletons showed infections dating back thousands of years (8). *Mtb* bacteria are transmitted when an individual with pulmonary TB creates an aerosolized droplet encapsulating the bacteria which are expelled through sneezing or coughing. TB transmission is dependent on several factors, including (I) the extent of lung damage in an individual with pulmonary TB, (II) the probability of a droplet being inhaled by the exposed individual based on the ventilation of the air, the number of bacteria present and the duration of exposure, and (III) whether the exposed individual is immunosuppressed. Inhaled bacterial particles, ranging from 1-5 μm in size, can travel to the alveoli where the virulence of the bacteria, coupled with the bactericidal abilities of the host alveolar macrophage determine whether a successful infection will occur or not. Furthermore, when the alveolar macrophage ingests the bacteria, the host, unaware of the invasion, will allow replication of the bacteria without triggering any inhibiting factors (9). During the initial phase of mycobacterial replication within the host, the bacteria can circulate throughout the body via the lymphatic system and ultimately infect other organs apart from the lungs. Individuals with normal functioning cellular immune response form granulomas in the lungs where the replication of *Mtb* is restricted thereby, creating a period of latent infection. Progression from latent TB infection (LTBI) to active TB disease can range from a few weeks to several years, or it may not progress at all (9). The inability to predict which individuals will progress to active TB significantly hinders targeted preventative strategies thereby, sustaining the global TB burden. There is no definitive method to predict LTBI progression, and the limited feasibility of preventative therapy increases the risk of active disease and transmission. This uncertainty complicates public health strategies, since it hinders the effective implementation of

preventative interventions. Addressing this knowledge gap is essential to prevent further increases in the global TB burden and achieve significant reductions in incidence rates. However, progress is further constrained by limitations of existing preventative measures, including the variable effectiveness of the bacillus Calmette-Guérin (BCG) vaccine, which remains the only approved vaccine for TB, regardless of being developed nearly a century ago (10).

Despite the availability of the BCG vaccine, *Mtb* continues to evolve to evade host immunity (11). The BCG vaccine was developed through multiple passages of *Mycobacterium bovis (M. bovis),* leading to several genetic changes, including the deletion of the RD-1 locus. This locus encodes nine genes, some of which are part of the ESX-1 secretion system responsible for producing virulence proteins such as ESAT-6 and CFP-10. The deletion of the RD-1 locus contributes to the attenuation of the strain, as it is directly associated with the virulence of *Mtb* (12). However, the effectiveness of the BCG vaccine is variable due to differences in protocols and the number of passages of the strain in laboratories across the globe (11). While the BCG vaccine marks a significant milestone in TB prevention, its limitations further complicate the challenges of TB control.

Although the BCG vaccine provided partial protection, TB remained a public health challenge until the mid-20th century, when effective TB drugs were developed. Drugs such as isoniazid (INH), rifampin (RIF), and ethambutol (EMB) transformed the treatment of TB, forming the basis of modern TB therapy. However, the emergence of drug-resistant (DR) TB strains has complicated treatment and continues to threaten global TB control efforts (13).

1.2 Challenges of TB Control

Progress in TB control has been hindered by several factors, more specifically, the development of drug-resistant TB strains, social and economic factors and comorbid diseases such as HIV - acquired immunodeficiency syndrome (AIDS). Drug-resistant TB strains emerge due to several factors, including (I) inadequate treatment when patients who do not adhere to or complete the protracted treatment as prescribed by healthcare professionals, (II) in underdeveloped communities, the availability of medication is often irregular and untimely, leading to intermittent stock shortages, (III) overpopulated and under resourced government health care centers have a longer turnaround time of laboratory tests resulting in a delay in the diagnosis and treatment of TB, and (IV) misdiagnosis or unsuitable administration of treatment.

Furthermore, DR-TB spreads from individuals harboring the resistant bacteria, facilitating the spread of resistant strains within households and communities (14). DR-TB strains remain a challenge as they require a tailored regimen of drugs which have adverse side effects.

1.2.1 TB-HIV-AIDS Co-infection

Mtb and HIV fuel each other when co-infecting the human host, resulting in a weakened host immune response with ultimate premature mortality in the absence of adequate treatment. HIV-1, the most common strain of HIV, targets the CD4+ T cells and macrophages thus weaking the immune system and inhibiting the body's ability to fight off illnesses; while AIDS is the condition that occurs when the individual with HIV has a severely weakened immune system increasing the risk of opportunistic infections and diseases.

The alveolar macrophage, the principal target for *Mtb* infection, can also be a site for HIV infection. A compromised immune system exacerbates the proliferation of *Mtb* in the lungs of infected individuals due to poor sensitization or priming of the immune system to affect a sustained immune response against the pathogen. The development of granulomas, which contain the infection, requires macrophages, CD4+ and CD8+ T cells. These cells are depleted in HIV infected individuals, leading to a higher risk of extra-pulmonary TB and a poor prognosis (15). Patients with HIV-1 on antiretroviral therapy (ART), have these cellular populations restored but still remain at risk to develop TB infection (16).

The relationship between HIV and TB is particularly problematic in Southern Africa, with an estimated 270 000 HIV-positive individuals developing TB disease, annually (17). Individuals with HIV-TB co-infection are at a higher risk of poor outcomes which are attributed to several factors, including a weakened immune system, an increased risk of disseminated TB and the challenges arising from drug-drug interactions. Globally, approximately 671 000 TB cases have been reported across all age groups living with HIV infection, with 461 000 occurring within the African region (5). In 2022, TB-HIV co-infection attributed to 167 000 reported deaths. These cases represents a reduction from the 190 000 reported deaths in 2021 (8). The decrease in mortality from TB is likely attributed to the increased rollout of ARV drugs . Studies have shown that the accessibility to ART has significantly reduced TB-HIV coinfection related deaths by enhancing the immune system, which leads to a reduction in viral load and thereby, improves tolerance to TB treatment (5). Although significant progress has been made in reducing the mortality caused by the TB-HIV co-infection, the emergence of the severe acute

respiratory syndrome coronavirus 2 (SARS-CoV-2) pandemic has posed new challenges, reversing the advancements in TB control by disrupting healthcare services and reallocating resources to address the global health crisis (18).

1.2.2 The Impact of SARS-CoV-2 on the TB Pandemic

The SARS-CoV-2 pandemic has hindered the progress towards the United Nations (UN) sustainable development goals and the WHO End TB strategy, creating a false impression of a reduction in TB disease worldwide (18). During the pandemic, many were afraid to visit clinics and hospitals due to the fear of contracting Covid-19, compounded by the redirection of important TB services to address the Covid-19 pandemic. Consequently, the number of newly diagnosed and relapsed TB cases decreased from 7.1 million in 2019 to 5.8 million in 2020. In 2021, there was a resurgence of the TB disease, with a reported 6.4 million new and relapsed cases, mainly owing to the undiagnosed cases during the pandemic (19).

Since SARS-CoV-2 is a respiratory disease, the interventions undertaken to prevent the spread of infection within and outside of households were expected to consequently reduce the spread of TB disease. To determine the impact of Covid-19 on the TB burden, a mathematical model using data from high TB burden countries such as China, India and South Africa were used to estimate the effect of social distancing interventions undertaken during the Covid-19 pandemic on TB disease. Social interactions in high-contact areas such as schools, public transport, entertainment, and workplace settings were investigated. The study indicated that social distancing reduced TB disease incidence, rather than TB deaths (20, 21). The global disruptions caused by the SARS-CoV-2 pandemic highlighted the need for a deeper understanding of the *Mycobacterium tuberculosis* Complex (MTBC), allowing for better preparation of future challenges.

1.3 The *Mycobacterium tuberculosis* Complex

The MTBC consists of a closely related group of bacteria that can cause TB in various hosts, including humans and bovines. The species within the MTBC include *Mtb, M. bovis, Mycobacterium africanum* (*M. africanum), Mycobacterium microti (M. microti), Mycobacterium caprae (M. caprae), Mycobacterium pinnipedii (M. pinnipedii) and Mycobacterium canettii (M. canettii)* (22). One of the characteristics of the pathogens within the MTBC is the presence of the IS6110 region which is a specific insertion sequence exclusive to species within the MTBC allowing for *Mtb* detection and classification. *Mtb* grows relatively

slowly with a replication turn-around time of approximately 16-22 hours (23). *Mtb* displays low levels of genetic diversity, approximately 1 single nucleotide polymorphism (SNP) for every two kilobase pairs (16). The *Mtb* genome is 4.4 million base pairs long, with an approximate G+C (Guanine and Cytosine) content of 65% and a high rate of repetitive insertion deoxyribonucleic acid (DNA) sequences, with 16 copies of the direct repeat insertion sequence, IS6110.

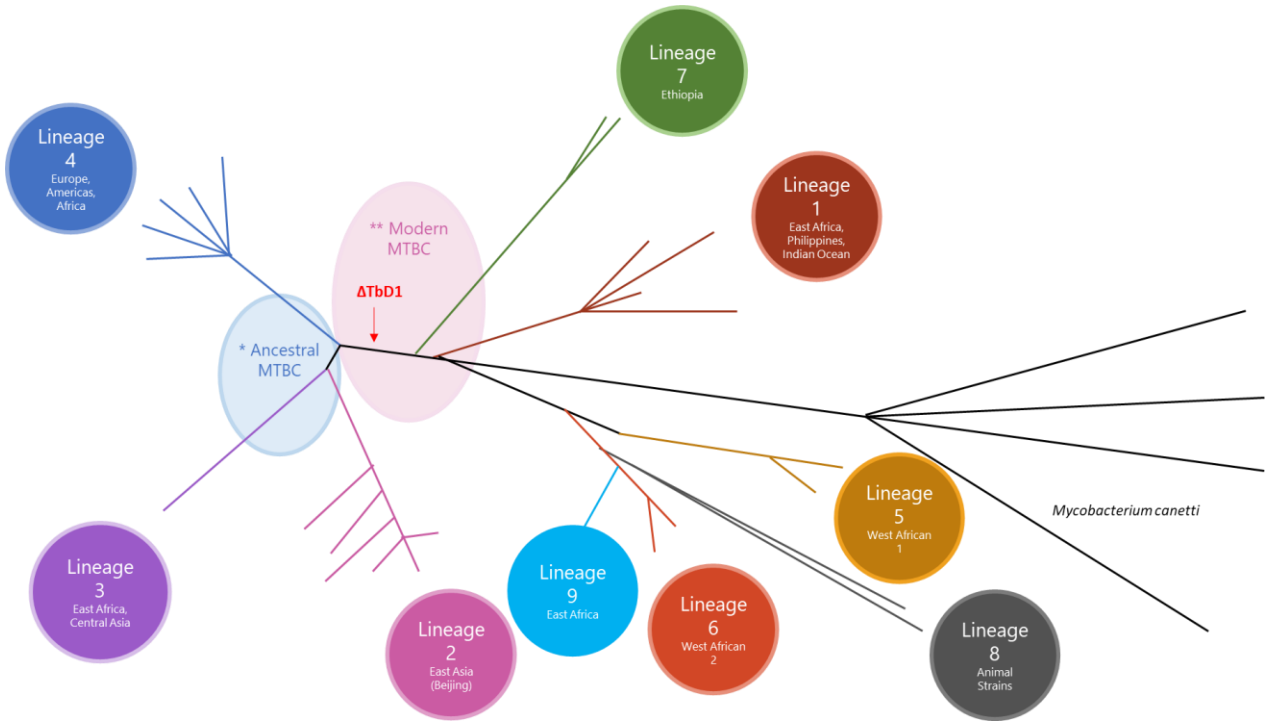


Figure 2: Phylogenetic Tree of the Evolution of the MTBC Lineages. The modern MTBC lineages (2,3 and 4) have a large section of deleted TbD1 segments, which are responsible for the widespread global tuberculosis epidemic evolved from the ancestral lineages (1, 5, 6 and 7) of *Mtb*. Lineages L8 and L9 from East Africa branch between M. africanum (L5–L6) and the modern lineages. The branch lengths are an approximate depiction of relatedness between the strains.
Image adapted from Brites & Gagneux, 2015.

Within the MTBC, nine key lineages of *Mtb* (lineages 1, 2, 3, 4 and 7) and *M. africanum* (lineages 5 and 6) have been recognized across the globe as TB disease causing agents (24). Lineages 1 and 7 are part of the ancestral strains (light blue section in Figure 2), from which the modern clade (light pink section in Figure 2) has evolved. Lineages 2 (East Asian), 3 (East-African Indian) and 4 (Euro-American) are part of the modern clade of strains causing TB and are often associated with epidemics, which have evolved over more than 600 years since the first *Mtb* lineage was identified. The modern clade of strains arose after the deletion of a 2153

bp segment in the genome, called the *Mtb*-specific deletion 1 region TbD1 (Figure 2) (25). *M. africanum* lineage 6 shares a common ancestor with animal-infecting *Mtb* strains (26) while lineages 8 and 9, identified in East Africa, branch between *M. africanum* and the modern lineages (27, 28).

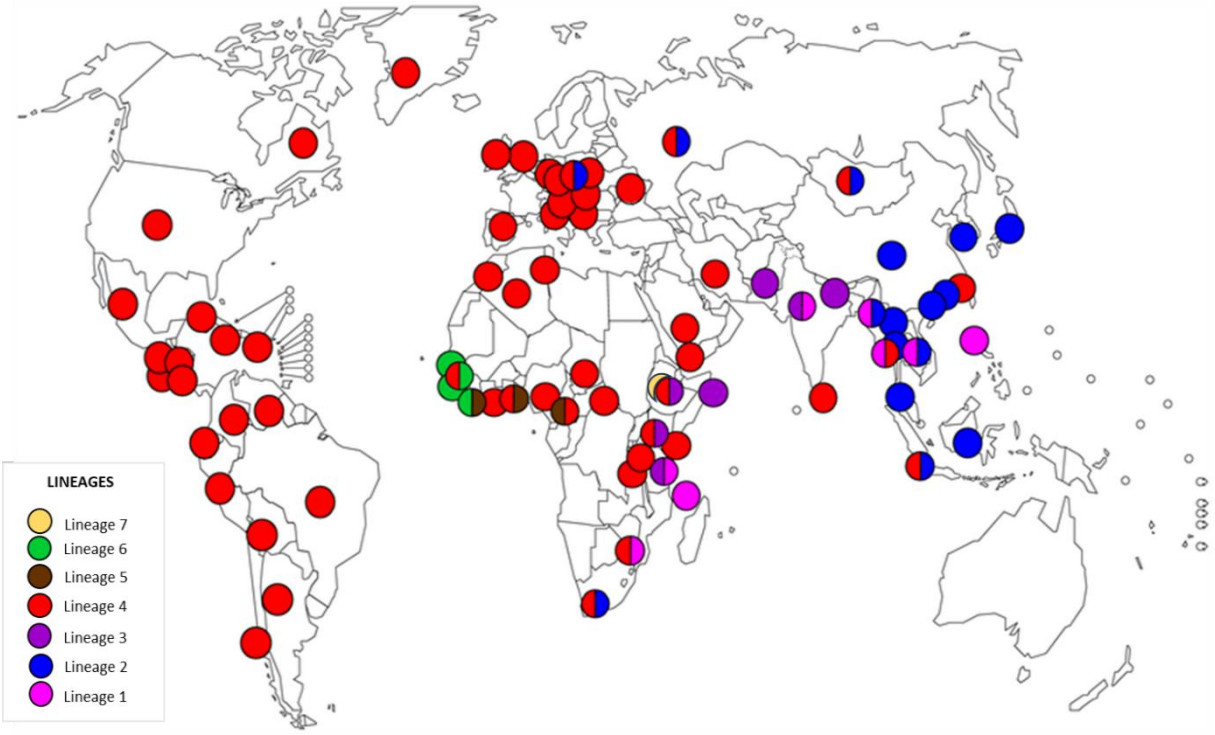


Figure 3: Global Phylogenetic Distribution of the Seven Major *Mtb* Lineages. Lineage 4 from the modern clade is the most dominant strain across Africa, Northern and Southern America and Europe, while lineage 2 dominates across parts of Asia.
Image taken from Comas & Gagneux, 2009 with minor edits.

A distinguishing characteristic of TB in South Africa is the presence of several major lineages of *M. tuberculosis*, including lineages 1-4. Among these, lineage 4 is the most prevalent, accounting for approximately 87.4% of cases. Specific genotypes within lineage 4, such as the KZN strain, have been associated with extensively drug-resistant TB (XDR-TB) outbreaks in the KwaZulu-Natal region (29). Lineages 5 and 6, traditionally associated with *M. africanum*, are rarely found in South Africa and are typically linked to individuals of West African origin (30). Beijing strains (lineage 2) are less common nationally but have been reported in specific areas, such as the Eastern and Western Cape provinces, raising concerns about potential higher transmissibility (31).

.

Each of the 7 lineages has been shown to differ in pathogenicity and virulence, including the extent of disease, transmissibility, drug resistance, protective efficacy against the BCG vaccine, and the propensity to adopt LTBI. Further differences have also been noted in the sub-lineages of the 7 major lineages, such as the severity of disease and the ability of these strains to cause disease (18). Lineages 3 (purple) and 4 (red) are the most widespread globally as indicated in Figure 3 (32) and are likely more virulent than the other MTBC lineages. Ancestral lineage 1 (pink) continues to exhibit prevalence in parts of Africa and Asia, indicating its persistent virulence and transmissibility. However, ancestral lineage 7 (green) is restricted to Ethiopia as shown in Figure 3. Given the genetic diversity and varying pathogenicity of the MTBC lineages, accurate and advanced TB diagnostic tools are important to identify and characterize the strains responsible for TB disease.

1.4 TB Diagnosis

An accurate diagnosis of TB is essential for the prevention and treatment of the disease. TB diagnostic methods are aimed at either detecting the bacteria or the host's immune response to the infection (33). Common methods include microscopy, culture, radiological methods, immune-based assays and molecular diagnostics. Emerging advanced diagnostic methods such as Urine-based Lipoarabinomannan (LAM) tests which detect TB antigens in urine, are being developed to detect *Mtb* antigens using non-invasive sampling techniques. The 2024 WHO consolidated guidelines on TB recommend molecular assays specifically, GeneXpert *MTB*/RIF or Xpert *MTB*/RIF ultra, as the initial diagnostic tool for TB and RIF-resistance detection (2). For diagnostic testing without suspected drug resistance, loop-mediated isothermal amplification (LAMP) and lateral flow urine lipoarabinomannan (LF-LAM) assays are recommended. Follow up testing for additional drug resistance employs low complexity nucleic-acid amplification tests (NAATs), line probe assays (LPA), targeted next generation sequencing and high complexity NAATs for comprehensive resistance profiling (33).

1.4.1 Culture Diagnostic Assays

Despite advancements in new diagnostic methods such as NAATs and the Xpert *MTB*/RIF/Ultra/XDR assay (34), culture remains the gold standard for TB diagnosis. Culture has a high sensitivity and the ability to detect viable *Mtb*, providing insight into the extent of disease and response to treatment. Furthermore, culture facilitates drug susceptibility testing, making it an important factor in TB management. However, culture is resource intensive, has a

long turnaround time and requires biosafety facilities which necessitates development of quicker diagnostic technologies that do not require highly skilled personnel.

1.4.1.1 Bacterial Detection Using Solid Media

Solid growth media such as Middlebrook 7H10/7H11 agar or Löwenstein-Jensen (LJ) medium is used to culture *Mtb* bacteria with a turn-around time for growth of up to 42 days. Solid media allows for the growth of *Mtb* colonies which can be identified by morphological or biochemical analysis. Despite their high specificity in bacterial detection, solid growth assays have a low recovery rate especially in samples with a low bacterial load. Given the prolonged time of 2-6 weeks for *Mtb* to grow on solid media, there is an additional risk of contamination as the sputum sample carries with it the lung and oral microbiome which may be resistant to the decontamination process. Hence, solid culture methods are being phased out of routine diagnostic laboratories although they continue to play a significant role in resource-limited settings where liquid culture systems are not available (35, 36).

1.4.1.2 Detection of Mycobacterial using Mycobacterium Growth Indicator Tube Culture

The Mycobacterium growth indicator tube (MGIT) system manufactured by Becton-Dickinson (BD), was developed to improve the turnaround time for *Mtb* detection. MGIT culture is a liquid-based assay containing 7H9 Middlebrook broth and a ruthenium pentahydrate fluorescent oxygen sensor located at the bottom of the tube which allows for the detection of growth. Growth is detected as a measure of oxygen depletion in the media, which is indirectly proportional to the number of bacteria present and directly proportional to the fluorescent signal emitted (37). The time to positivity (TTP) for growth is reduced by at least 2 weeks in comparison to solid media. Although the MGIT system supports the growth of *Mtb*, it is not specific for *Mtb.* The sputum matrix contains several other respiratory and oral bacteria that may resist decontamination and as a result can grow in the MGIT media. In cases where the MGIT result is questionable, the TBc (tuberculosis complex) identification kit can be used to detect the MPT64 protein, only produced by members of the MTBC to validate and confirm the presence of *Mtb* (38). Additionally, the liquid MGIT culture system has the potential to allow growth of differentially culturable populations of *Mtb* that may display limited growth on solid cultures but are able to grow in liquid media (39).

1.4.1.3 Differentially Culturable Tubercle Bacteria

Differentially culturable tubercle bacteria (DCTB) are organisms that are incapable of growing in routine diagnostic media (liquid or solid media) but can be resuscitated in liquid limiting dilution assays, supplemented with or without culture filtrate (CF) containing growth stimulatory molecules such as resuscitation promoting factors (Rpf). Rpfs are a group of proteins secreted by mycobacteria that play an important role in chronic infection due to their involvement in reactivation of dormant bacteria (40). Rpfs are proteins that stimulate the growth of non-replicating mycobacterial cells (41) that were shown to be drug tolerant thus, have important implications in the treatment of TB (42). In a study carried out by Chengalroyen and colleagues (2016), 3 distinct groups of DCTB populations in the sputum of drug susceptible treatment naïve TB patients were described. These included (I) CF-dependent DCTB, (II) Rpf-independent DCTB and (III) CF-independent DCTB. The study also showed an association of DCTB with HIV infection status where HIV-TB co-infected patients had lower DCTB populations (43). Since currently no single diagnostic culture method identifies all prevailing tubercle bacilli populations in sputum, the presence of these complex DCTB populations underscores the need for improved diagnostic tools.

To better understand the response of the DCTB population during treatment, a longitudinal study was conducted on treatment naïve TB patients with drug-susceptible TB. These individuals were monitored over the 6-month treatment period for any changes in DCTB populations (44). Using liquid limiting dilutions with and without CF containing Rpfs to detect the DCTB population, in addition to solid culture methods, demonstrated that the DCTB population cleared at a slower rate in comparison to routine culture over the 6 months. Treatment responses in this study were categorized into 4 patterns based on their DCTB loads: (I) Classic bi-phasic bacterial clearance – defined by a rapid decline in bacterial load within the first week with a slower rate of bacterial clearance thereafter; (II) early non-responders with slower clearance – this group of participants showed no significant decline in their bacterial load within the first week and then showed a gradual decline in bacterial load; (III) paradoxical worsening with an increase in bacterial load upon treatment initiation – the participants in this group showed an increase in bacterial load by CFU and MPN assays within the first week of treatment and (IV) non-responders that showed no change in bacterial load. Although the MGIT and CFU assays showed a decline in bacterial load over the treatment period, bacterial detection by the MPN assay remained positive across the treatment period. Surprisingly, DCTB populations could be detected at the end of treatment in two-thirds of individuals completing

treatment. Individuals who showed the presence of DCTB at this point of treatment were associated with a higher positivity rate in MGIT culture, GeneXpert, and smear at 2 months post-treatment. These findings indicate that specific populations of drug-tolerant bacteria can be identified using MPN/CFU assays which would otherwise be missed using routine diagnostic culture. These findings also suggest that the 6-month TB treatment may be insufficient to completely eradicate all bacterial populations (44).

In another study, the response of DCTB populations in sputum from drug resistant (DR)-TB participants over 24 months showed that patients with multidrug resistant (MDR) TB yielded a greater quantity of DCTB compared to patients with RIF resistant (RR) TB (45) . Whilst the CFUs declined to undetectable levels, the DCTB population was still detectable up to 16 weeks post-treatment. This study showed that MPN/CFU assays could be valuable in diagnosing and monitoring treatment response in DR-TB patients (45). While culture-based methods provide insight into bacterial load and cell viability, microscopy-based methods offer a complementary approach which enables direct visualization of bacterial cells in clinical samples.

1.4.2 Microscopy Based Diagnostic Testing

Microscopy-based diagnostic testing are tools often used in conjunction with culture-based assays. These methods are able to visualize the bacterial cells in clinical samples, typically through techniques that stain components of the bacterial cell wall, most commonly, the mycolic acid layer (46). Microscopy based methods are relatively cost effective and have a faster turnaround time, however, there are limitations that directly affect the outcome of this test. The presence of other respiratory pathogens can lead to false positive results, while lower bacterial loads reduce detection sensitivity. Additionally, the quantity and quality of sputum samples collected influence diagnostic accuracy.

The WHO recommends smear microscopy as a key diagnostic technique and South Africa has adhered to these recommendations and included it under its guidelines for TB diagnosis (47). More recently, the use of fluorogenic dyes have emerged as an alternative, with the promise to enhance the identification of viable cells only, providing a more accurate understanding of the bacterial populations harboured by the patient (46).

1.4.2.1 Smear Microscopy

Diagnostic smear microscopy detects the presence of whole mycobacterial cells and is reliant on dyes such as the Ziehl-Neelsen (Z-N) and the auramine O stains that can detect the mycolic acid layer in the cell wall that is specific to mycobacterial species (46). TB diagnosis in low to middle-income countries is generally carried out using Z-N sputum smear microscopy due to its relatively fast output and inexpensive cost (48). Studies have shown that in individuals with extra-pulmonary TB and HIV, smear-based diagnosis of TB is not useful as the bacterial load in the sputum is often low, and the test is not sensitive enough, thus missing cases of active TB (49, 50). These undiagnosed cases still have the capacity to transmit the disease thereby, adding to the overall disease burden and increased mortality rates as without a diagnosis, these individuals do not receive treatment (51). To address some of the limitations of smear microscopy, advancements in the field of microscopy for TB diagnosis have been made, including the development of fluorogenic dyes to improve diagnostic sensitivity and gain further understanding of cell viability.

1.4.2.2 Fluorogenic Dye Based Microscopy

The fluorogenic dye, 4-N,N-Dimethylamino-1,8-naphthalimide conjugate of Trehalose (DMN-Tre), is being investigated in research settings as an alternative TB diagnostic tool. DMN-Tre exhibits a fluorescence signal once the dye is metabolized into the mycobacterial cell wall, enabling the selective identification of live cells only (52). The outer cell wall of mycobacteria is essential for its survival and resistance to host response and environmental stresses. DMN-Tre specifically binds to the outer cell wall, allowing the dye to enter and label bacterial cells thereby, distinguishing between metabolically active (viable) and metabolically inactive (non-viable) cells. This is achieved by the three isoforms of antigen 85 (Ag85) which are secreted to assist in the synthesis of the outer cell wall components (a) arabinogalactan, (b) trehalose monomycolate and (c) trehalose dimycolate. These isoforms are almost identical at the active site, which is responsible for their enzymatic activity in mycolic acid transfer, giving them a similar molecular structure (53). While microscopy can differentiate bacterial cell viability, it is limited in sensitivity and specificity. Molecular testing complements this approach by detecting *Mtb* DNA and genetic resistance markers, providing a more sensitive and specific means of diagnosing TB.

1.4.3 Molecular Testing

Molecular diagnostics have transformed TB and drug-resistance detection by providing highly sensitive and specific identification of *Mtb* DNA and genetic resistance markers. Compared to culture-based methods, molecular assays offer a faster turnaround time with increased accuracy. Molecular tests include amplification of the targeted DNA sequences thereby, enabling detection of *Mtb* in samples with low bacterial loads. Additionally, molecular testing provides insight into drug resistance profiles of *Mtb*, enabling a targeted treatment regimen for improved patient outcome (33).

1.4.3.1 Tuberculin Skin Test and Interferon Gamma Release Assay

Present-day procedure for LTBI detection and diagnosis is the tuberculin skin test (TST) or Interferon Gamma Release Assay (IGRA) which detects immune system markers for TB. TSTs are cost effective compared to IGRA however, their specificity is limited by false positives due to BCG vaccination and non-tuberculous mycobacterial (NTM) infections. Conversely, IGRAs are expensive, yet highly specific as they detect *Mtb* specific antigens: ESAT-6, CFP-10 and TB 7.7 (10). LTBI diagnosis is more common in developed countries due to the lower disease burden and the ability of health systems to diagnose LTBI. In developing countries however, the sheer burden of TB disease together with the cost of testing and providing preventative therapy makes it impossible to incorporate LTBI screening into routine surveillance (54). Currently, LTBI tests are unable to predict whether an individual will progress to active TB. Therefore, the value of these tests can be only be increased by being coupled with tests that have the ability to predict disease progression (55). While TSTs and IGRA focus on detecting immune markers for LTBI infection, advancements in molecular diagnostics to provide a rapid and automated approach were developed to bridge the gap between screening and confirmatory testing (56).

1.4.3.2 Cepheid's *MTB*/RIF/Ultra/XDR GeneXpert

The GeneXpert *MTB*/RIF/Ultra/XDR assay developed by Cepheid is a clinical diagnostic method that detects the presence of *Mtb* DNA directly in a sputum sample. The process begins with combining the proprietary buffer with the sputum sample, which lyses the bacterial cell to release the DNA. The sample is then loaded into a single use cartridge, and the remainder of the process is automated. Bacterial DNA is filtered and concentrated before being combined with polymerase chain reaction (PCR) reagents and real time PCR is performed.

In addition to detecting TB DNA, the GeneXpert can simultaneously distinguish between DR and drug susceptible TB. Currently, the GeneXpert assay is the mainstay of TB diagnosis worldwide due to its modular nature, fast turnaround time and the ability to simultaneously detect drug resistance. The most common drug resistance is the resistance to RIF. The GeneXpert assay specifically targets the *rpoB* gene, with mutations in its 81 base pair RIF resistance-determining region (RRDR) indicating RIF resistance (57-59). Although GeneXpert does not directly quantify the number of organisms, the PCR cycle threshold can be used as an indicator to infer the bacterial burden in a sample. Despite its high sensitivity in *Mtb* DNA detection, the GeneXpert test cannot determine the viability of *Mtb* cells.

1.4.4 The Hain Assay

The Hain assay is a second-generation LPA, which is a DNA-DNA hybridization assay designed for the rapid detection of drug resistance in *Mtb* performed on MGIT-positive samples using multiple probes (60). While it is not a direct diagnostic assay, it is a commonly used PCR and reverse hybridization-based method for rapid identification of resistance mutations to RIF and INH (37). The mutations and correlating resistance profiles for the first- and second-line drugs are shown in Table 1. The Hain assay detects mutations in specific regions within the genome that confers resistance to a particular drug. For example, a mutation in the *rpoB* gene is linked to RIF resistance, while mutations in the *inhA* and *KatG* genes are linked to INH resistance (61). The level of resistance is variable, depending on the specific mutation within the resistance-associated region. This variability determines the degree of resistance that the bacteria will exhibit towards the drug, ranging from low to high levels of resistance. The Hain assay is still one of the most common assays for the detection of resistance due to its high sensitivity (98.9%) and specificity (100%), short turnaround time (~ 5 hours), and relatively low cost. However, it is not a strain-typing or lineage identification tool; it does not categorize bacterial strains for epidemiological or transmission analysis. Despite its high specificity and sensitivity, it cannot be used as a point-of-care test in diagnostic laboratories due to its high complexity, cost, and dependence on culture based assays (61, 62).

Table 1: Genomic Mutations in *Mtb* and their Correlation with Drug Resistance Profiles.

Mutation Region	Resistance to drug	Resistance Level
rpoB	RIF	Low - High

inhA	INH	Low
KatG	INH	High
gyrA/gyrB	Fluoroquinolone	Low – High
rrs/eis	Aminoglycosides	Low – High
embB	EMB	Low - High

**Information referenced from Pei et al., 2024*

1.4.4.1 Quantitative Polymerase Chain Reaction

The high number of direct repeat insertion sequences in the IS6110 region, specific to species within the MTBC, allows for more accurate detection of bacteria using PCR. The PCR method amplifies the IS6110 target region using enzymes and are highly sensitive allowing for trace amounts of DNA to be amplified. The primers are complementary sequences to the target region and the 4 nucleotides–dATP, dCTP, dGTP, and dTTP are the building blocks of the resultant PCR product which are linked together by DNA polymerase, as depicted in Figure 4A.

Whilst the in-house quantitative PCR (qPCR) uses primers and a minor groove binding probe (MGB) specifically designed to detect and amplify the IS6110 insertion element, making the detection of *Mtb* more specific. The MGB probe is labelled with a 5' fluorescent reporter dye and a 3' Eclipse Dark Quencher (EDQ). The probe binds to the double helix at the target site, and during amplification the quencher is separated from the reporter dye, allowing fluorescence to be emitted each time a strand of DNA is amplified. Fluorescence is measured in real time, enabling detection and quantification of *Mtb* DNA, as shown in Figure 4B.

Using qPCR for detecting *Mtb* is comparable to Cepheid's GeneXpert assay in sensitivity. A study carried out by Kim et al. (63) showed that in smear positive cases, GeneXpert performed better, while in smear negative cases with lower bacterial loads, qPCR was more effective. Overall, qPCR was found to be more sensitive than GeneXpert. Some of the drawbacks of qPCR are that it uses a limited number of probes in a single reaction therefore, it may not be ideal for a pathogen that harbors multiple mutations. Secondly, qPCR requires expensive equipment and reagents, as well as skilled technicians.

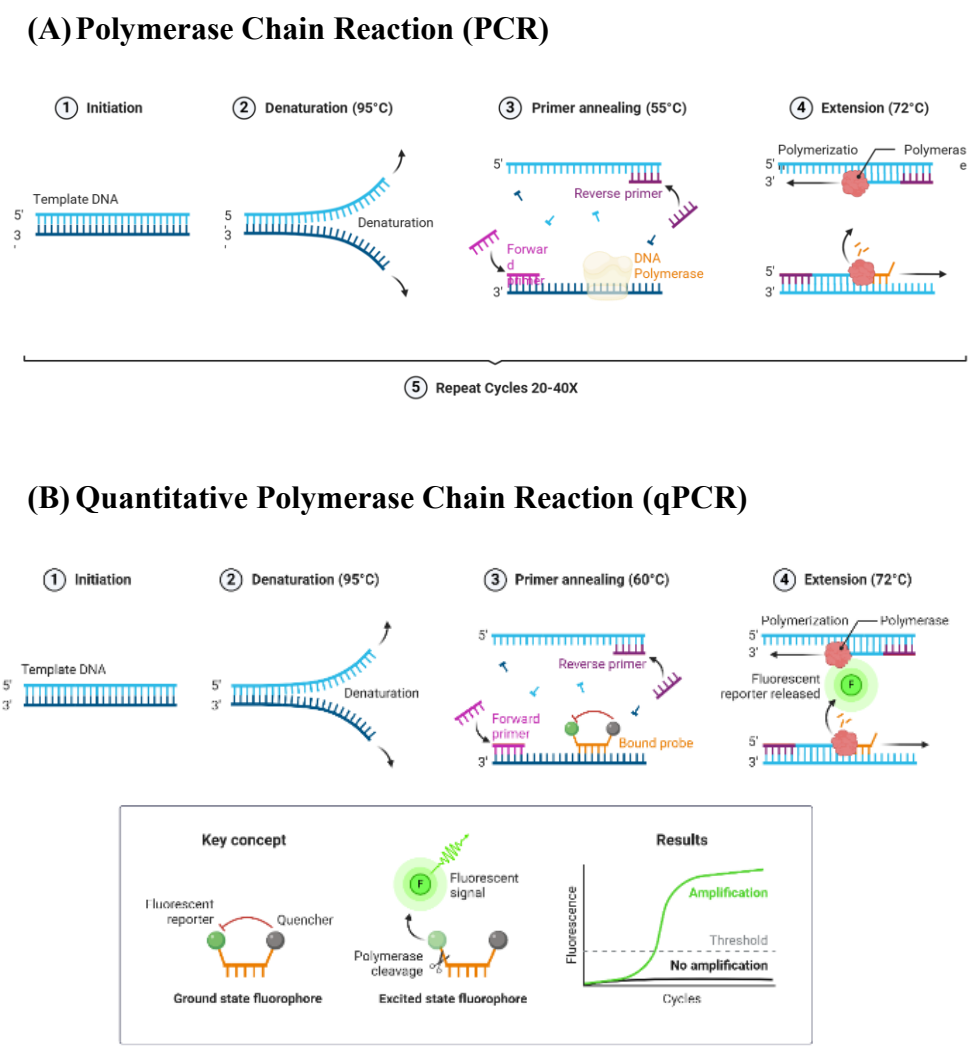


Figure 4: Schematic representation of DNA Amplification by Polymerase Chain Reaction and Quantitative Polymerase Chain Reaction. Both techniques use a thermal cycling process consisting of 3 main steps: denaturation, annealing of primers and extension of the DNA strand. (A) PCR: provides qualitative data visualized by gel electrophoresis. (B) qPCR monitors amplification in real time using fluorescent probes, which enables both qualitative and quantitative data. The fluorescent signal is proportional to the concentration of DNA present in a sample.
Created in BioRender, https://BioRender.com

1.4.5 Chest X-rays

Chest X-rays are rapid imaging tools used for easy identification of chest abnormalities. Chest radiographs can be used to detect *Mtb* infection by visualizing cavities in the lung tissues formed by the granulomas during TB pathogenesis. In active pulmonary TB cases, infiltrates and/or cavities are generally visible on the apex of the lungs together with the detection of lymphadenopathy or abnormal size/shape of the lymph nodes (64, 65). While chest X-rays are

valuable for diagnosing TB, they have limitations. Identification of a cavity on an X-ray confirms the presence of a lesion but cannot definitively diagnose TB. This is because the X-ray image only reveals the anatomical abnormality and not the specific causative agent, which is essential for a conclusive diagnosis as lesions can be caused by other pathogens or infections (66).

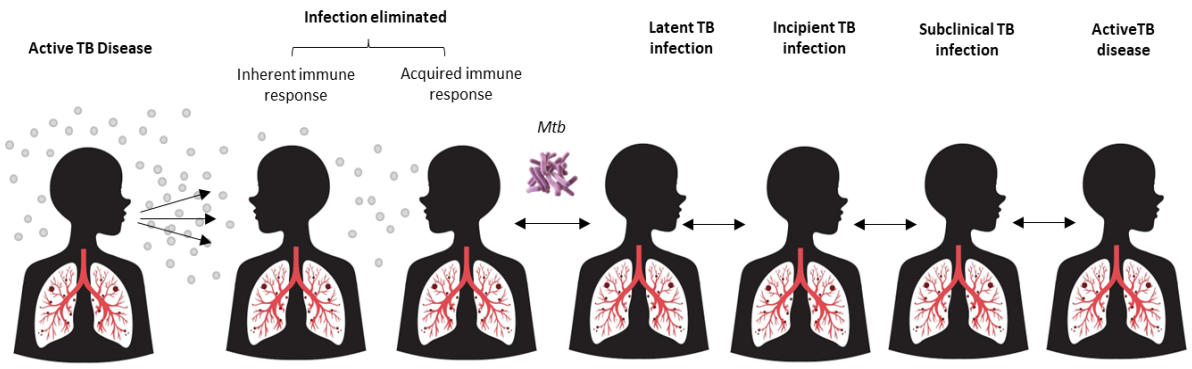
1.5 Tuberculosis Disease States

The TB infection cycle begins when aerosols containing M. tuberculosis are expelled by an infected individual and inhaled by a susceptible person. These droplet nuclei travel through the respiratory tract to the alveoli, where infection is established (Figure 5). Infection can progress along a continuum, from latent TB infection (LTBI) to incipient TB, subclinical TB, and ultimately active disease.

Each stage differs in infectiousness, detectability, and recommended treatment. LTBI and incipient TB are TST/IGRA positive but culture and smear negative, asymptomatic, and non-infectious; preventative therapy is recommended. Subclinical TB shows intermittent culture positivity, mild symptoms, and occasional infectiousness; multidrug therapy is advised. Active TB is culture positive, may be smear positive, symptomatic, and highly infectious, requiring full multidrug treatment.

The World Health Organization describes TB infection as a continuum, reflecting gradual progression toward active disease. Globally, about one-third of the population is infected, yet only 5–10% develop active TB without treatment. Immunocompromised individuals face a 16-21-fold higher risk, particularly within the first two years post-exposure. The majority either eradicate the bacteria or maintain immune control without clinical disease (67-69).

Despite this framework, the mechanisms determining why only a subset of infected individuals progress to active TB remain incompletely understood. The different disease states are discussed in detail below.



TST/IGRA	**Negative**	**Positive**	**Positive**	**Positive**	**Positive**	**Usually positive**
Culture	Negative	Negative	Negative	Negative	Intermittently positive	**Positive**
Smear	Negative	Negative	Negative	Negative	Usually negative	**Positive or Negative**
Infectious	No	No	No	No	Sporadically	**Yes**
Symptoms	None	None	None	None	Mild or none	**Mild to severe**
Recommended treatment	**None**	**None**	**Preventative therapy**	**Preventative therapy**	**Multidrug therapy**	**Multidrug therapy**

Figure 5: Spectrum of TB Disease. TB infection ranges from early clearance to active disease. LTBI and incipient TB are asymptomatic and microbiologically negative, subclinical TB is microbiologically positive with mild symptoms, and active TB is symptomatic, transmissible, and requires multidrug therapy.
Image adapted from Pai et al., 2016.

1.5.1 Latent Tuberculosis Infection

LTBI is one of the early stages in the TB life-cycle of infection and is defined by the WHO as a state of incessant immune response to *Mtb* antigens but without clinical or diagnostic confirmation of active TB disease (70). LTBI can be diagnosed indirectly by detecting a cellular immune response to *Mtb* antigens using either TST or IGRA (70). Whether this latent infection progresses to active TB is dependent on a host of variables, including immune response and comorbidities. However, in most healthy individuals, LTBI can be harbored for a lifetime without any symptoms. A study by Kunkel et al. showed that the risk of latent TB reactivation can be significantly reduced by placing individuals on a 9-month INH (INH9) preventative therapy (IPT). This strategy will only be effective if the prevalence of TB remains in control as per the WHO's reduced TB strategy, as IPT alone may not prevent TB reactivation in high prevalence settings due to ongoing transmission and higher exposure risks (71). Despite compelling evidence supporting the benefits of IPT, concerns regarding its lengthy duration,

the associated risk of liver disease, and the potential for developing drug-resistant TB, has been met with hesitation among individuals in high TB-burden countries. In response, several clinical trials have explored improved and shortened treatment regimens.

The first trial involved a weekly dose of INH and rifapentine (3HP) over 3 months and was shown to be as effective as INH in preventing reactivation. The regimen has been recommended by the WHO and CDC as it has shown improved adherence and fewer adverse events (72). The second trial involved a daily dose of RIF (4R) over 4 months, with a similar efficacy to INH9 and 3HP in preventing reactivation and with fewer cases of hepatoxicity (67). The latest trial, which involved a 28-day daily treatment course of INH and rifapentine (known as 1HP), has shown promise in addressing the concerns around treatment duration, preventing reactivation and adverse events (68, 69, 73). However, the drug regimens may not yet be recognized by U.S. (United States) guidelines due to several reasons, including (I) the trials are being conducted outside of the U.S., (II) only PLWH are being enrolled in the study, (III) inclusion of participants without confirmed TB infection and (IV) 1HP being more expensive than other regimens (74).

1.5.2 Incipient TB

Incipient TB is a disease phase between LTBI and subclinical TB. It represents the early stage, when the host's immune response is challenged to control the replication of *Mtb*, leading to an increase in viable bacilli within granulomas. Our understanding of incipient TB largely stems from experimental research in non-human primates, and we currently lack validated methods to diagnose this condition in humans. During this phase, tissue damage is minimal, and the likelihood of the patient being infectious is low (75). In this TB disease state, the infected individual will most likely progress to active disease in the absence of intervention, but will not display clinical symptoms, abnormal x-ray images, or bacterial evidence of active disease (76). Recent scientific advances have focused on the potential of blood-based biomarkers to identify incipient TB or to predict its progression to active disease. Host immune biomarkers that are specifically and differentially expressed during *Mtb* exposure or infection have emerged as promising tools for TB detection. Numerous studies have been conducted to identify and validate RNA (ribonucleic acid)-based biomarkers with high specificity for TB, offering new prospects for the early identification and management of individuals at risk of developing active TB (49, 69).

1.5.3 Subclinical TB

Subclinical TB is caused by viable bacteria that also do not display clinical symptoms in the infected individual, but the presence of the infection can be detected by radiographic imaging or microbiological assays. The viable and metabolically active nature of *Mtb* during the subclinical TB stage indicates the imminent advancement to full-blown disease (77).

1.5.4 Active TB

The last stage in the spectrum of TB disease, where viable bacteria result in the presentation of clinical symptoms, is called the active TB stage. Active disease is detected using the gold standard culture method and GeneXpert MTB/RIF/Ultra/XDR (76, 77). Individuals with clinically confirmed TB typically exhibit a spectrum of symptoms, including a persistent cough, chest pains, weakness, fatigue, fever, night sweats, and weight loss. However, these symptoms are not exclusive to TB, necessitating the development of the WHO-recommended four-symptom screen (W4SS) as an initial, non-invasive screening approach. The W4SS assesses for the presence of cough, fever, weight loss, and night sweats however, it is important to note that it relies on an individual's subjective understanding and willingness to report their symptoms, and as such, it does not provide a definitive diagnosis (78). Understanding the progression through the TB disease spectrum is important for guiding the appropriate treatment strategies, as the stage of disease influences the type of treatment required to prevent further progression, transmission and development of drug resistance.

1.6 Tuberculosis Treatment

Patients infected with susceptible TB are placed on a standard 6-month treatment therapy consisting of daily doses of 4 drugs, collectively called RIFAFOUR: (I) RIF, (II) INH, (III) EMB and (IV) pyrazinamide (PZA), as recommended by the WHO (78). Patients are first placed on an intensive phase of treatment (2 months), where the regimen is aimed at rapid reduction of bacterial load to contain infectiousness and reduce symptoms. The WHO further recommends that individuals be placed on continuation therapy which is aimed at completely eradicating the bacteria by targeting slow growing or dormant bacteria that may have survived the intensive phase to prevent relapse or the development of drug resistance. This continuation phase involves less frequent doses of treatment for a duration of 4 months, where patients are prescribed a reduced regimen, typically consisting of RIF and INH. (78). To ensure that patients take their medications as prescribed and complete the full course of treatment, some healthcare

facilities use a strategy called Directly Observed Therapy (DOT), which entails a family member or healthcare worker observing the patient take their daily doses of treatment (79).

PLWH who are not already on ART, are advised to carry out the 6-month RIF treatment and concurrently with TB treatment, begin ART (80). The WHO has released a 4-month treatment plan for individuals 12 years and over, consisting of INH, rifapentine, moxifloxacin, and PZA. At the end of the 6-month treatment period, if the individual still presents with a positive smear result, drug resistance against any of the 4 treatment drugs would need to be ruled out before the second line therapy is initiated. This may not always be possible in areas where drug susceptibility tests are not easily available, however as per the new WHO guidelines, an informed decision needs to be made before treatment is extended, based on drug susceptibility test results, the patients' clinical condition, their adherence to treatment, and/or microbiological evidence (78).

DR *Mtb* is classified into 4 groups: (I) RR is the resistance to RIF and includes resistance to RIF found in any of the other categories; (II) mono-resistant (MR) strains are resistant to a single first-line TB drug, excluding RIF; (III) MDR strains are resistant to both INH and RIF with resistance to other first line drugs, and (IV) Extensively drug-resistant TB (XDR-TB) is defined as MDR-TB with additional resistance to any fluoroquinolone and at least one Group A drug, such as bedaquiline or linezolid (81).

For the treatment of MDR or RR TB and pre-XDR TB, the duration may vary based on disease severity and resistance profiles. Pre-XDR-TB is defined as MDR or RR-TB with additional resistance to any fluoroquinolone. A longer, individualized regimen is required, typically lasting 18-20 months for patients with XDR or pre-XDR TB. A 6-9 month treatment regimen is used for MDR or RR-TB patients and the 18-month or longer regimen is reserved for patients with fluoroquinolone resistant MDR or RR-TB patients or those with complex resistance profiles (82). Drugs used to treat DR-TB are categorized into three groups: Group A - levofloxacin or moxifloxacin, bedaquiline and linezolid. Group B - clofazimine, and cycloserine or terizidone; and Group C - EMB, delamanid, PZA, imipenem–cilastatin or meropenem, amikacin (or streptomycin), ethionamide or prothionamide, and p-aminosalicylic acid. The 6-month treatment regimen, known as the BPaL regimen, consists of bedaquiline, pretomanid, linezolid, and moxifloxacin. In instances where resistance to fluoroquinolones has been excluded and MDR/RR-TB has been confirmed, a 9 month regimen is recommended instead of an 18-month regimen due to the efficacy and safety profile of the shorter regimen

(83). In the 18 months or longer regimens for MDR/RR-TB patients, a 4-drug treatment regimen is advised, consisting of levofloxacin or moxifloxacin, bedaquiline and linezolid and at least one drug from group B, in comparison to the combination of second line drugs in the 9-month regimen such as moxifloxacin, clofazimine, EMB, and PZA (82).

Kanamycin and capreomycin are not recommended in the treatment of MDR/RR-TB patients on longer regimens due to the potentially irreversible adverse effects such as ototoxicity (damage to the ear) and nephrotoxicity (damage to the kidney). There are less harmful drugs available such as bedaquiline and linezolid which are effective against MDR/RR-TB thereby, also increasing adherence to treatment therapy. It is important that throughout the treatment process, patients adhere to regular follow-up appointments, to monitor side effects and treatment progress so that treatment can be adjusted if necessary. Several cases of relapsed TB disease occur as patients fail to complete their prescribed medication. This non-adherence, results in several adverse effects such as the development of drug resistance to second line drugs with fatality in severe cases and continuous transmission of these resistant strains (82).

1.7 Tuberculosis Strain Typing

Mtb strain typing has been widely used since the early 1990's for various reasons, including molecular epidemiology, regional distribution of strains, monitoring of transmission dynamics, identification of resistant strains, detection of strains not detectable by conventional contact tracing thereby, aiding in transmission inferences and differentiation between reinfection and reactivation (84). Several molecular methods have been developed for improved strain typing and more recently the advent of whole genome sequencing (WGS) has revolutionized our understanding of TB disease dynamics.

1.7.1 IS6110-Restriction Fragment Length Polymorphism

IS6110-RFLP had been long considered the gold standard for MTBC detection and has played a significant role in detecting outbreaks, however, this is no longer the case. The RFLP technique, developed by Botstein et al., utilizes restriction enzymes to cleave DNA at specific sites, generating multiple fragments of varying lengths. These fragments are then separated by gel electrophoresis, followed by transfer to nitrocellulose filters via Southern blotting. Subsequently, the fragments are labeled with radioactive DNA probes and visualized using photographic film (85). The number of copies of IS6110 within the *Mtb* genome varies depending on the strain, with some having up to 25 copies per genome, while a small minority

have a single copy (86). Lok et al., however, discovered a population of strains predominantly in Southeast Asia that do not have any copies of IS6110 insertion sequences, using the RFLP method (87). The disadvantages of IS6110 typing is that it is a labor-intensive process, time-consuming, its unable to detect strains with less than 5 copies of IS6110, and inter-laboratory comparisons of the complex RFLP banding patterns is complicated (88).

1.7.2 Spoligotyping

Spoligotyping is a crude method of strain typing that is based on PCR amplification of the highly polymorphic direct repeat region of the *Mtb* genome (89). The direct repeat region is PCR amplified using specific primers designed to detect this region. The presence or absence of the 43 unique spacers in this region is detected by hybridization to a membrane with the pre-printed spacer probes. Each *Mtb* strain has a unique combination of the spacer sequences and the pattern of these spacers are used to determine the strain type (90). Online spoligotyping repositories such as SpotClust (https://tbinsight.cs.rpi.edu/run_tb_lineage.html) contains the binary and octal codes of *Mtb* isolates which are used to determine strain type from the spoligotype pattern (91).

1.7.3 Mycobacterial Interspersed Repetitive Unit - Variable Number Tandem Repeats

Variable number of tandem repeats (VNTRs) are regions in the *Mtb* genome where short nucleotide sequences are arranged as a tandem repeat. Each of these VNTRs varies in length among strains, creating a unique genetic fingerprint. These VNTRs are inherited alleles, which can be used in *Mtb* detection and identification. In clinical settings, VNTRs involve extracting bacterial DNA, amplifying the target loci using PCR and analyzing the size of the resulting DNA fragments to confirm the presence of *Mtb*. The unique genetic fingerprints can be used to distinguish between new and relapse TB cases. However, VNTRs are complex and expensive, making it difficult to implement in resource limited settings (92).

The Mycobacterial Interspersed Repetitive Unit (MIRU) - VNTR is a PCR based genotyping method which can also be used to determine genetic relatedness of strains in the MTBC. MIRU-VNTR targets the repetitive interspersed mycobacterial units, by amplifying the internationally standardized 24-loci to determine the size and number of repeated units (93). The sizing of the PCR sequences can be done either using standard agarose gel electrophoresis or automated sequence readers. One of the limitations of MIRU-VNTR is that it can be difficult to distinguish

between different strains of *Mtb*. However, this can be overcome if used in conjunction with spoligotyping or the inclusion of hypervariable loci specific to different strains of *Mtb*, particularly in regions with similar genotypes (94).

1.7.4 Whole Genome Sequencing

WGS has emerged as the superior tool for TB research for a variety of reasons. WGS has high specificity and sensitivity to detect mutations in regions associated with drug resistance, thus allowing for a more targeted approach to treatment therapies in cases of MDR/RR-TB. In addition to resistance prediction, WGS is used for lineage determination and cluster analysis, enabling categorization of strains into lineages and identification of closely related strains for epidemiological studies. In studies of transmission dynamics, WGS enables the inference of transmission within households and communities by determining genetic relatedness of strains with a high degree of accuracy. This particularly aids in outbreak investigations to identify the source of infection and track the spread of strains. Furthermore, WGS is also useful for evolutionary studies by providing information on the emergence of drug resistance in certain strains. It can also be used in clinical management to predict the possibility of drug resistance based on genetic markers. Overall, WGS has transformed our understanding of TB epidemiology, drug resistance and pathogenesis thus providing knowledge for improved interventions and strategies for better management and control of TB disease (95, 96).

It is essential to sequence as many MTBC strains as possible to build libraries of SNPs to determine relatedness and find correlates of clinical consequences, genetic relatedness and the origin of certain strains of *Mtb* as well as aid in identifying transmission dynamics (97). Although WGS technology has been a breakthrough for medical science, analyses of the data has proved to be challenging for many scientists, due to the fast moving development of software platforms (98). The ability to determine lineages and clusters using WGS helps integrate transmission mapping with drug resistance prediction, providing comprehensive insights into TB epidemiology. WGS has filled the gap for transmission inferences enabling precise identification of the relatedness between strains amongst infected individuals in households and TB endemic communities. Constant evolution of bioinformatics tools is improving our understanding on transmission directionality which is much need for clinical management and control of TB.

1.8 Transmission Dynamics

Molecular studies have shown that less than 20% of TB transmission in the adult population occurs within a household (99). This probability increases when social contact studies are included (100). Household contact (HHC) TB transmission studies involve following HHCs for the duration of the index participant's TB treatment. Tracing the origin of LTBI in a HHC is challenging, as the individual could have acquired the infection either from the index patient or from infected individuals outside of the household, which complicates the understanding of transmission dynamics. However, it is assumed that HHCs who test positive after the index participant begins treatment, most likely acquired the infection within the household (101). A crude method for confirming transmission links is by strain typing to determine whether both the index participant and the HHC are infected with the same *Mtb* strain.

Determining the source of infection through molecular techniques such as WGS can provide valuable insight into the mechanics of transmission. This allows for more targeted intervention strategies, including the administration of more specific preventative treatment therapies to HHCs to reduce the risk of progression to active TB disease (102). However, as previously mentioned, the limitations of standard culture techniques make it difficult to detect early infection in HHCs, highlighting the need for more sensitive diagnostic approaches, particularly for the detection of subclinical TB. This asymptomatic group, which is also partially responsible for the global TB epidemic, is particularly difficult to detect as they harbor a low bacterial load that prevails with this form of infection. Individuals with subclinical TB remain untreated, and continue to transmit the bacteria thereby, perpetuating the cycle of transmission and complicating disease control. Therefore, it is crucial to identify individuals with subclinical TB and treat them appropriately to assist in reducing the global TB burden (103-105).

In this study, we aimed to investigate whether the application of different culture techniques, including those that detect DCTB, together with molecular techniques, allows for greater detection of mycobacteria in HHCs with low bacterial loads. We hypothesize that detection of DCTB using the MPN assays will detect the bacterial population in HHCs that will otherwise be missed by routine culture methods thereby, allowing more transmission inferences Moreover, qPCR will be used for early detection of *Mtb* in HHCs as its high sensitivity enables detection of low bacterial loads.

2. Study Aims and Objectives

2.1 Aims

a. To assess *Mtb* transmission within households of active index participants by using culture and molecular based assays to understand *Mtb* prevalence within a household.
b. To determine transmissibility of *Mtb* within households with HIV-positive versus HIV-negative index participants

2.2 Study objectives

a. To assess *Mtb* transmission to HHCs in 100 households with an active TB index participant, using MPN/CFU assays.
b. To determine the *Mtb* genome equivalents in infected index participant and their respective HHCs in 100 households by qPCR. The sensitivity of the qPCR will allow for detection and quantification of the bacterial load in these individuals which often is undeterminable by culture methods.
c. To analyze the genotype of the *Mtb* isolated from each positive individual in the respective household (index participants and HHCs) using WGS. The sequence data will be used to determine transmission links within the individual households.
d. Using the clinical demographics data obtained from the primary HoTT DCTB (Assessing Household Tuberculosis Transmission using detection of Differentially Culturable Tubercle Bacteria) study, *Mtb* transmission between HHCs in households with either a HIV-positive or a HIV-negative index participant will be compared.

3. Materials and Methods

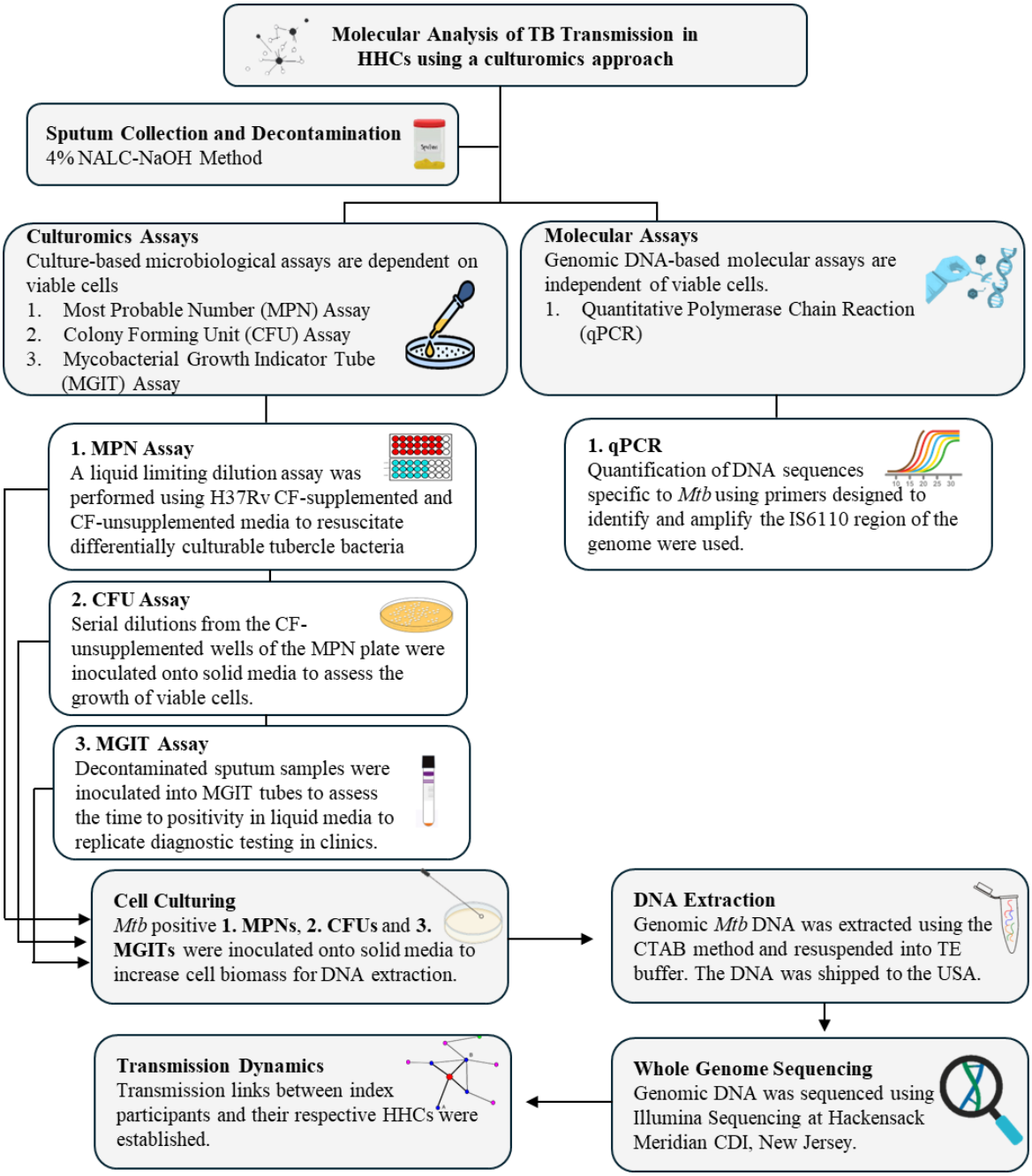


Figure 6: Overview of Study Methodology and Schema. Infected participants with drug sensitive and drug-resistant TB were recruited at the clinics based on the eligibility criteria before their respective HHCs were enrolled. Each participant provided two sputum samples (spot and overnight) that was combined before testing for the presence of *Mtb* using 3 different culturomics assays: (1) MPN assay, (2) CFU assay, (3) MGIT assay and a molecular assay qPCR. *Mtb* cells from the culturomics assays were cultured to increase the biomass for gDNA extraction using the CTAB method. The DNA was outsourced for whole genome sequencing (WGS). Data was used to establish transmission links within the household based on SNP distances.

3.1 Study Population

This study was nested in a larger clinical research study conducted at the CBTBR focusing on household TB transmission, the HoTT DCTB study. A cohort of 100 index participants, aged 12 years and older together with their HHCs of older than 5 years of age, were recruited from 2 sites in South Africa, Klerksdorp town within the Matlosana Municipality (North West Province) and Botshabelo (Free State Province). Both sites reported high annual TB incidence rates of 600 per 100 000, as shown in Figure 7 (106).

Figure 7: Location of Study Sites. Map of South Africa illustrating the two provinces (black pins) where the clinics were set up for participant recruitment and enrolment in the study. Clinics in both the Free State and Northwest reported high annual TB incidences.
Map image taken from https://www.dreamstime.com/illustration/south-africa-map.html

3.2 Ethics

The primary HoTT-DCTB study was approved by the Human Research Ethics Committee of the University of the Witwatersrand (HREC number:190910). The ethics certificate is supplied in appendix A. The work described in this study is covered under this ethics approval; additionally, the researcher has been listed as a laboratory investigator on the protocol.

3.3 Study Design

For the primary study, recruiters at both the approved clinics identified and enrolled eligible and consenting index participants with either DR or drug susceptible TB together with their respective HHCs. However, in this component of the work, only drug susceptible index participants were included. Both index participants and HHCs were followed up for a 16-month period across 6 timepoints (baseline, month 2, month 4, month 8, month 12 and month 16) and at each time point the participants provided an overnight and spot sputum sample. The sputum specimens from the index participants and their respective HHCs were analysed using the in-house molecular based test, qPCR and various culturomics assays (MPN assay, CFU assay and MGIT assay) to determine the presence and yield of *Mtb*. Any specimen with confirmed *Mtb* was then cultured to obtain the desired biomass for DNA extraction and WGS followed by bioinformatics analysis to identify possible transmission inferences within households, as shown in Figure 6.

3.4 Inclusion and Exclusion Criteria

The inclusion criteria for participants were as follows: (I) Index participants had to be in possession of a hard copy of a public sector-requested laboratory result confirming *Mtb* detection on a sputum specimen. Individuals with trace GeneXpert results were not included. (II) Participants had to have a history of TB symptoms dating at least 2 weeks prior to enrolment. (III) Participants had to be over the age of 12 years. For those below the age of 18 years, assent from the minor as well as parental consent was required. (IV) All participants 18 years and older had to provide written informed consent. (V) Participants also had to provide consent for the research team to visit their household and access their routine clinical records (including but not limited to mycobacterial and HIV-related results). (VI) Participants had to provide a HIV test result not older than 3 months or be willing to be tested for HIV. (VII) All index and HHCs had to produce two baseline sputum samples (a spot and overnight) of at least 5 mL each. Participants were excluded from the study if (I) they were hospitalized; (II) they were planning to relocate away from the clinic sites; (III) they were on TB treatment for more than 2 days in the absence of a smear result or 3-4 days if the smear result was negative or scanty; (IV) the sputum sample was received more than 48 hours after expectoration by the lab; (V) there was a lack of eligible HHCs or (VI) both baseline sputum samples were culture negative. In the primary study, participants were recruited into 2 arms: (1) drug susceptible TB and (2) DR-TB. For this study, only drug susceptible index participants were included.

3.5 Sputum Decontamination

The raw sputum specimens were decontaminated using the NALC-NaOH method. A 1:1 volume of sterile 4% NALC-NaOH solution was added to the raw sputum specimen and incubated at room temperature for 15 minutes, vortexed at 5-minute intervals and subsequently washed by making up the volume to 45 mL with 0.01 M phosphate buffered saline (PBS), pH 7.4. The sputum was then centrifuged at 3000 rpm for 15 minutes at 4^{o}C to pellet the organisms. The supernatant was discarded, and the remaining pellet resuspended in 2 mL of Middlebrook 7H9 (4.7 g/L) media, supplemented with 10% oleic acid-albumin-dextrose-catalase (OADC), 0.2% glycerol and 0.2% Tween80 (7H9 media). The decontaminated sputum was assessed for the presence of mycobacteria by the MPN assay, CFU assay, MGIT culture and qPCR. Any residual sputum sample was frozen at -80^{o}C. A detailed list of media components and methodology is supplied in Appendix B.

3.6 Culturomics Assays

Culturomics assays are a set of culture-based microbiological assays that are used to comprehensively characterize a population of *Mtb* from clinical samples and involve the culture of bacteria under an extensive range of culture conditions to promote growth of unculturable or differentially culturable organisms that are typically difficult to culture using standard methods. In this study, 3 culturomics techniques were applied to sputum samples collected from index patients and their HHCs: (I) The MPN assay, a liquid limiting dilution series that uses a Poisson distribution to quantify the bacterial growth in liquid media. A ten-fold dilution of decontaminated sputum samples with and without CF (depicted as ‘A’ in Figure 8) was grown in a 48-well plate, (II) the CFU assay quantifies the number of bacteria that can grow on solid media. This was carried out by inoculation of the 10-fold dilutions from the media wells of the MPN assay onto 7H11 Middlebrook plates (depicted as ‘B’ in Figure 8), and (III) decontaminated sputum samples were inoculated into MGIT tubes (depicted as ‘C’ in Figure 8), supplemented with OADC enrichment and 4% (w/v) polymyxin, amphotericin B, nalidixic acid, trimethoprim and azlocillin (PANTA). Mycobacterial growth in the MGIT assay was confirmed using the BD Bactec TBc ID test, which detects the antigen MPT64 which is specific to *Mtb*.

3.6.1 Bacterial Culturing and Culture Filtrate Preparation

The clinical strain *Mtb* H37Rv was cultured by inoculating 1 mL of stock culture (OD_{600nm} of 0.6-0.9) into 8 mL of 7H9 and allowed to grow for 3 days to an OD_{600nm} of ~ 1. The inoculum was added to 42 mL of 7H9 media and incubated at 37 ºC for 3 days to grow to an OD_{600nm} of ~ 0.8. Consecutive cultures of *Mtb* were set up so that CF could be prepared fresh each day as previously described by Chengalroyen (2016) (42). Briefly, the mycobacterial cultures were pelleted by centrifuging at 4200 rpms for 8 minutes at 22 ºC and the supernatant filtered through a 0.22 µM filter attached to a sterile 50 mL syringe.

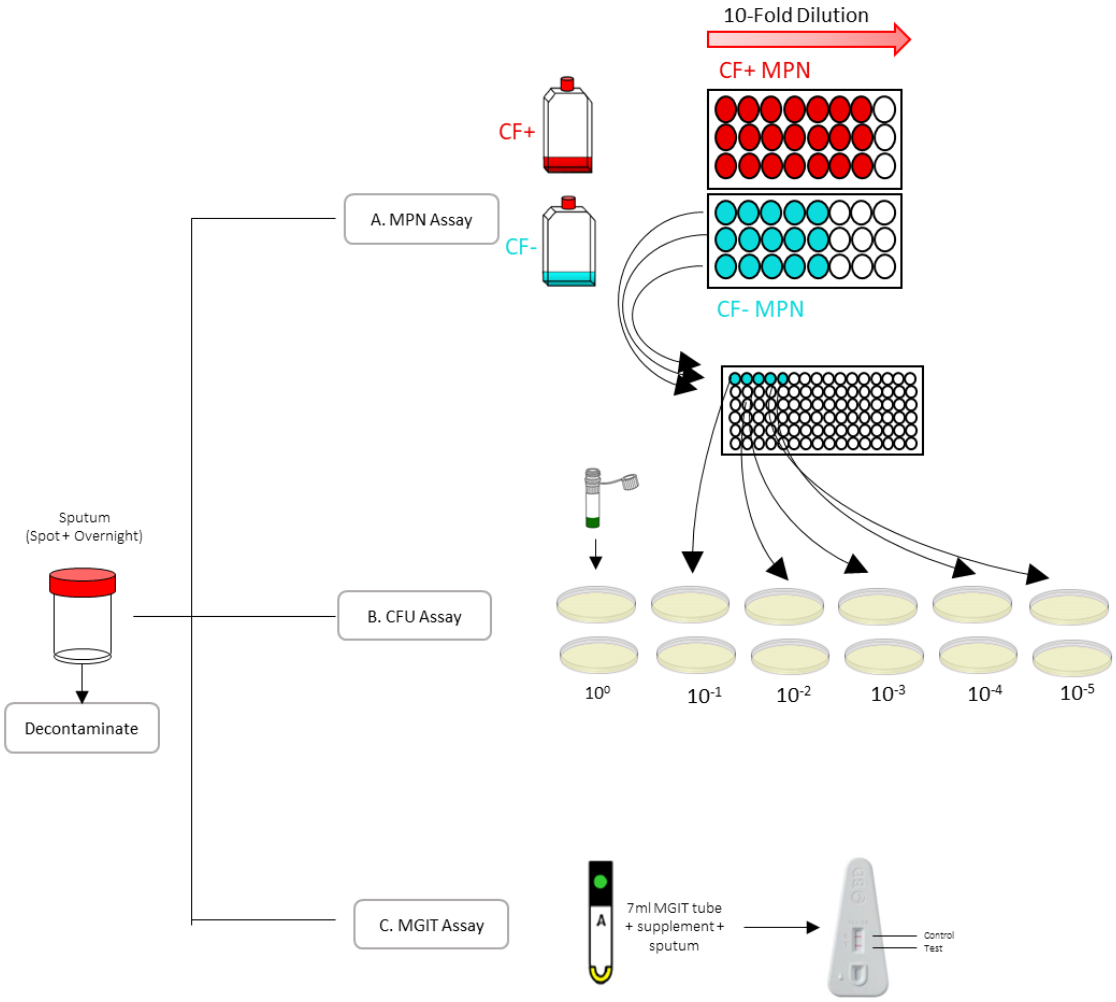


Figure 8: Illustration of the Culturomics Assays. The 2 sputum samples were pooled and decontaminated for testing in the MPN assay, CFU assay and the MGIT assay. The MPN assay was conducted with and without CF supplementation. Aliquots of the first five dilutions from the triplicate media wells (blue) in the MPN assay were spread onto 7H11 media and the plates incubated for 4 weeks. The MGIT tubes were supplemented with PANTA-OADC and inoculated with the decontaminated sputum sample. All tubes that flagged positive for growth on the MGIT machine were confirmed for the presence of *Mtb* on the BD Bactec MGIT TBc ID test.

The filtered CF was diluted 50% with 7H9 media and supplemented with PANTA for the management of possible contamination in the decontaminated sputum sample. To confirm the absence of contamination of the CF, a 1 mL aliquot was incubated at 37 ºC for 6 weeks with an additional 100 µL aliquot spread on 7H11 media and the plate incubated at 37 ºC for 4 weeks.

3.6.2 The Most Probable Number Assay

The MPN assay was carried out in 48-well (Nunc) microtitre plates as described by Chengalroyen et al., with some modifications (42). Briefly, CF from wild type H37Rv was prepared as described above (47). Four hundred and fifty microliters of the Middlebrook 7H9 media supplemented in equal proportions with CF and 4% PANTA (CF+MPN) was dispensed across the top 3 columns in all 8 rows of the 48-well plate (Figure 8), with Middlebrook 7H9 media supplemented with 4% PANTA (CF-MPN) dispensed into the bottom 3 columns of all 8 rows of the plate. Fifty microliters of the decontaminated sputum specimen were inoculated into the first row of the 48-well plate and subsequently diluted 10-fold across the plate with change of tips after each dilution. The 48-well plates were sealed using biohazard tape and incubated at 37 ºC for 6 weeks and scored visually for growth using an inverted mirror. One hundred microliters of cells from wells that showed sparse growth were spread on Middlebrook 7H11 (20.5 g/L) media, supplemented with 10% OADC and 0.2% glycerol (7H11 media) and the plates were incubated for 4 weeks at 37 ºC to confirm the presence of *Mtb*. The total number of bacteria in the MPN assays were estimated using the MPN calculator software available at: http://www.wiwiss.fu.berlin.de/fachbereich/vwl/iso/ehemalige/wilrich/index.html.

The resuscitation index (RI) was derived by determining the difference in growth between the CF-supplemented and CF-unsupplemented liquid media (MPN value) and solid media (CFU/mL) using the formula below. The quantum of DCTB was determined by calculating the logarithm (log) RI.

$$RI = \frac{\text{MPN Value}}{CFU/mL}$$

$$\boldsymbol{DCTB = \log RI}$$

3.6.3 Colony Forming Units

The CFU assay was performed by taking aliquots from media wells in the MPN assay plate (CF-MPN) as described by Saito et al (107). Briefly, for index participants at baseline, 70 µL from each replicate of the first 5 dilutions were pooled into a 96 well plate and mixed by

pipetting up and down (Figure 8). One hundred microliters from each dilution ($10^0 – 10^{-5}$) was spread in duplicate onto 7H11 media. The plates were incubated at 37 ºC for 4 weeks before colonies were counted and the CFU/mL calculated using the formula:

$$\frac{CFU}{mL} = \frac{Average\ no.of\ colonies\ *\ Total\ dilution\ factor}{Volume\ of\ innoculum\ plated\ in\ mL}$$

3.6.4 Mycobacterial Growth Indicator Tube Assay

MGIT assays were carried out according to the BD MGIT manufacturer's instructions. Briefly, each tube was inoculated with 800 µL of MGIT OADC-PANTA and 300 µL of decontaminated sputum. The tubes were placed into the Bactec MGIT machine until positive growth was flagged by the machine. Growth is determined by the depletion of oxygen in the tube and is automatically detected by the machine, for up to 41 days, after which, the sample is deemed negative. All positively flagged cultures were confirmed for *Mtb* growth using the BD Bactec MGIT TBc ID Test. A 100 µL aliquot was inoculated into the well of the strip and incubated at room temperature for 15 minutes. The sample was confirmed positive for *Mtb* growth when 2 pink bands were observed: one control band and one test band. If only the control band was visualized, the sample was considered negative for *Mtb*.

3.6.5 Cell Culture to Increase Biomass for DNA Extraction

Aliquots from MPN wells and MGIT tubes that flagged positive for growth were plated on Middlebrook 7H11 (20.5 g/L) media, supplemented with 10% OADC and 0.5% glycerol incubated at 37 ºC for 4 weeks. After incubation, culture plates that showed confluent *Mtb* growth were scraped, and the cells resuspended in 500 µL of 1x TE buffer in preparation of genomic DNA (gDNA) extraction. Low yield CFU colonies were cultured in 2 mL of Middlebrook 7H9 and incubated at 37 ºC until the required biomass was obtained (~OD_{600} 0.9) after which the gDNA was extracted.

3.7 Genomic DNA Extraction: Cetyltrimethylammonium Bromide Method

The cetyltrimethylammonium bromide (CTAB) method of DNA extraction was carried out as described by Bjorn-Mortensen et al. (108). Briefly, bacterial cells resuspended in TE buffer were heat killed at 80 °C for 30 minutes. Lysozyme (10 mg/mL) was added to the killed cells and incubated for 1 hour at 37 °C. Next, 10% SDS and 20 mg/mL proteinase K was added and the mixture incubated at 65 °C for 1.5 hours. Thereafter, 100 µL each of 4.1% NaCl and 10%

CTAB was added, and the cells were incubated at 65 °C for 15 minutes. The DNA was then purified by adding 700 µL of chloroform:isoamyl alcohol (24:1) to the cell suspension, mixed thoroughly and centrifuged at 13000 rpm for 10 minutes. The upper aqueous phase containing the gDNA was transferred to a 1.5 mL O-ring Eppendorf tube containing 700 µL of isopropanol. The DNA was harvested by centrifugation for 20 minutes at 13000 rpm. After discarding the supernatant, the pellet was washed with 70% ethanol and centrifuged for 10 minutes at 13000 rpm. The supernatant was discarded and the pellet dried in a SpeedVac for 15-20 minutes. The DNA was resuspended in 120 µL TE buffer and the concentration measured using the Nanodrop. Agarose gel electrophoresis which allows for the separation of DNA fragments based on their size, was used to view the quality of the DNA. A 1% agarose gel was prepared in TAE buffer with 3 µL of the DNA stain and 10 000x SYBR safe concentrate. A 5 µL aliquot of DNA was mixed with 2 µL of 6x loading dye and the sample was loaded into the gel well. The DNA was separated in TAE buffer at 80-100 V. Molecular weight markers were not used as the gDNA was being analysed for quality and not for size determination. Gels were viewed under UV light using the Vacutec G-Box SYNGENE system and images were captured using GeneSnap software.

3.8 Genomic DNA Sequencing

gDNA sequencing was carried out by the Hackensack Meridian Health Centre for Discovery and Innovation in New Jersey, USA. Briefly, the process involved shearing the DNA obtained from the CTAB DNA extraction method into ~250 bp fragments using a Covaris E220 focused-ultrasonicator. The TruSeq WGS DNA preparation kit was used to prepare the samples before sequencing using the Illumina HiSeq platform, generating 125 bp pair ended reads that were analysed by collaborators of the HoTT DCTB study at the University of Columbia, using a custom bioinformatics pipeline (109).

3.9 Quantitative Analysis of Samples Using qPCR

Fifty microliters of decontaminated sputum sample were heat killed at 95 °C for 5 minutes. The sample was centrifuged at 13000 rpm for 5 minutes, resulting in a pellet and the DNA aqueous phase (lysate). The lysate was harvested and 8.5 µL was used for qPCR. Primers designed to amplify the IS6110 region were used, together with the MGB probe to detect the presence of *Mtb* gDNA (Table 2). To validate the PCR assay, positive controls of DNA isolated from H37Rv and *M. bovis* and as negative controls DNA from *Mycobacterium smegmatis (M. smeg)*

and *Escherichia coli (E.coli)* were used. PCR cycling conditions were as follows: 95 °C for 30 seconds, 95 °C for 15 seconds and 60 °C for 60 seconds which were repeated for a further 39 cycles.

Table 2: Primers and Probe used for *Mtb* Detection by qPCR.

Primer/Probe	Oligonucleotide Sequence (5'-3')	Annealing Temperature of Primer	Storage Temperature
IS6110_F2	GGGTAGCAGACCTCACCTATG	60 °C	
IS6110_R2	AGCGTAGGCGTCGGTGA		-20 °C
IS6110_MGB	FAM-TCGCCTACGTGGCCTTT-MGBQ	N/A	

3.10 Data Analysis

Microsoft Excel was used for data organisation and graph construction. GraphPad Prism 10 was used for constructing of graphs and statistical data analysis. Statistical significance between non-parametric groups was determined using t-tests and Mann-Whitney U tests, while Chi-squared tests and Fisher's exact tests were used to assess statistical significance between proportions. Regression analysis was conducted to determine the relationship between bacterial populations. A 95% confidence interval was used, with significance thresholds set at * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$ and **** $p<0.0001$.

Additionally, a custom bioinformatics pipeline, designed by Maha Farhat (86) was adapted by Prof. B. Mathema (Columbia University) to analyse the raw sequence files. Reads were trimmed using PRINSEQ, with an average Phred score threshold of 20. The reads were decontaminated to remove all non-*Mtb* isolates using Kraken. Reads were aligned to the reference genome, H37Rv (GenBank NC000962.3), using Burrows-Wheeler Aligner (BWA) - MEM. Duplicate reads were removed using PICARD. Variants were called using Pilon tools. Strains with less than 10x coverage were excluded from the analysis. A detailed list of the software and bioinformatics tools used in this study is appended in Appendix C.

4. Results

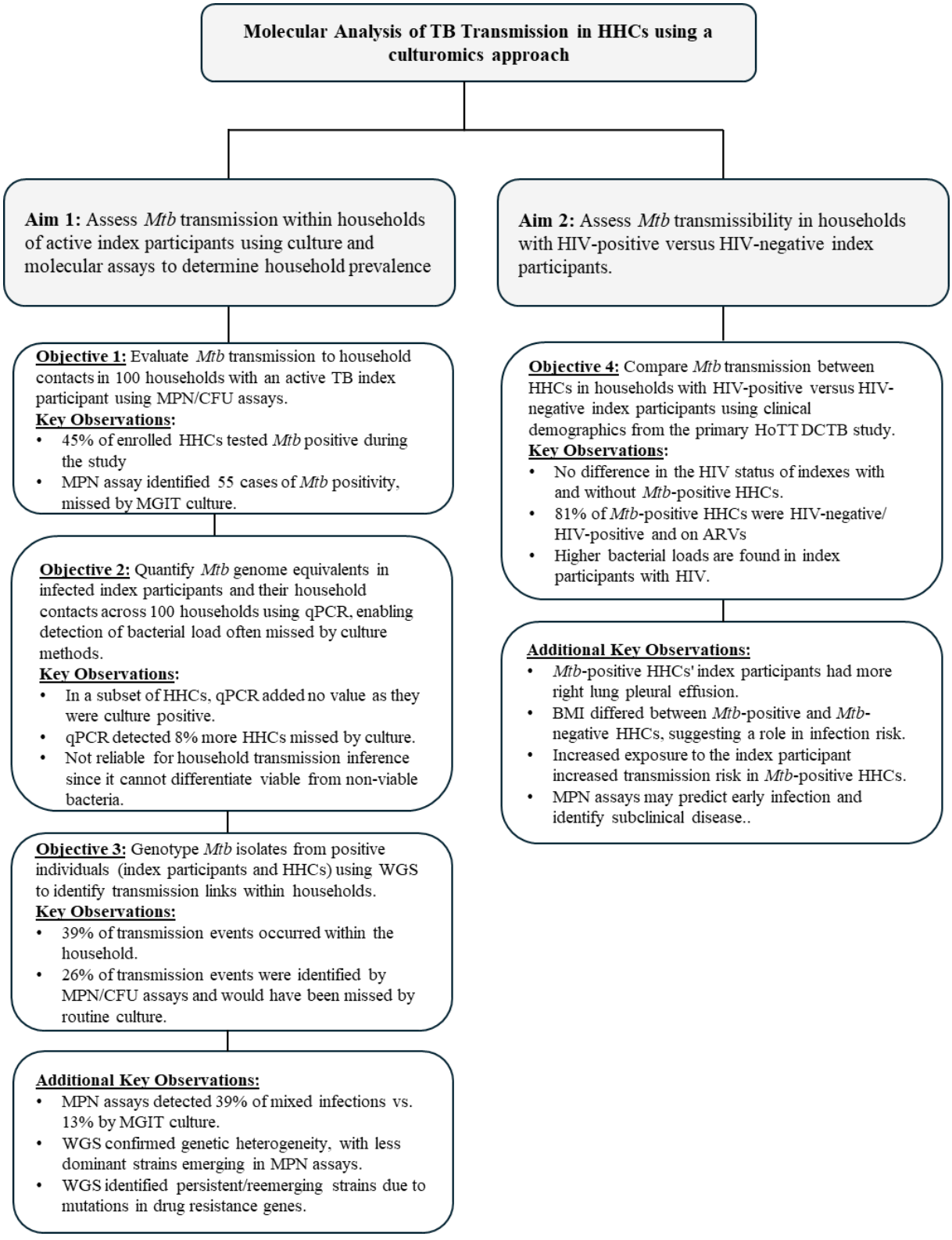


Figure 9: Overview of Results from this Study. The four objectives across two aims highlight the value of MPN assays, used alongside routine culture assays, in detecting undiagnosed *Mtb* cases within the households of TB-positive individuals. Additionally, MPN assays contribute to identifying greater genetic heterogeneity of *Mtb* within individuals, emphasizing their utility in uncovering genetic diversity that may be missed with standard culture methods.

4.1 Selection Criteria

The HoTT DCTB parental study enrolled 395 index participants (138 DR and 257 drug-susceptible). From the drug-susceptible group, 100 participants (50 per site) and their 247 household contacts (HHCs) were followed over 16 months at six timepoints (Figure 10). During follow-up, 27 index participants and 33 HHCs were terminated due to death, loss to follow-up, relocation, withdrawal of consent, failure to produce baseline sputum, or other study-specific reasons. Terminated participants were retained in the analysis to preserve potential transmission data within households. The 100 index participants comprised of 60% male whilst for the 247 HHCs, 34% were male. A greater proportion of index participants were HIV-positive (52%) compared to the 23.9% HIV positivity in HHCs. The median viral load in HIV-positive index participants was 8371 copies/mL, while the median CD4 count was 173 cells/μL and 490 cells/μL in HHCs. Viral load data in HHCs were not collected in this study. The median age was 38 years for index participants and 26 years in HHCs. The median body mass index (BMI) was 18.5 kg/m^2 in index participants and 22 kg/m^2 in HHCs. Baseline laboratory data showed a median MGIT TTP of 149.5 hours for index participants and negative at 984 hours in HHCs. The CF-supplemented and CF-unsupplemented MPN showed a log median of 5.4 and 3.6 respectively in index participants and 0.0 in HHCs. The median log CFU was 3.9 in index participants and 0.0 in HHCs. CF-dependent and CF-independent DCTB yielded median log values of 0.7 and 0.0 in index participants and 0.0 in HHCs respectively (see Table 3). The DNA from bacteria identified in any assay was subsequently extracted and WGS was conducted to assess transmission links.

Table 3: Demographics and Laboratory Data of Study Participants

	Index Participants (n=100)	**HHCs (n=247)**
Variables (Index: n=100; HHC: n=247)		
Demographics		
Sex:		
Women (%)	40 (40)	164 (66)
Men (%)	60 (60)	83 (34)
Age, year, median (IQR)	38 (27 – 48)	26 (14 – 46)
BMI, kg/m^2, median (IQR)	18.5 (16.7 – 21.3)	22.0 (18.2 – 28.0)
HIV Status at baseline		
Positive, n (%)	52 (52.0)	59 (23.9)

Negative, n (%)	48 (48.0)	188 (76.1)
CD4 Count*, cells/μL, median (IQR)	173 (50 – 259)	490 (304 – 755)
Viral Load*, copies/mL, median (IQR)	8371 (421 – 353000)	Not Done
Laboratory Data at Baseline		
MGIT TTP, hours, median (IQR)	149.5 (106 – 233.5)	984 (584 – 984)
Log Bacterial load by MPN assay		
CF-dependent MPN, log median (IQR)	5.4 (3.5 – 6.8)	0.0 (0.0 – 0.0)[α]
CF-independent MPN, log median (IQR)	3.6 (1.7 – 6.0)	0.0 (0.0 – 0.0) [α]
CFU/mL, log median (IQR)	3.9 (0.0 – 5.9)	0.0 (0.0 – 0.0) [α]
Amount of DCTB (MPN/CFU)		
CF-dependent DCTB, log median (IQR)	0.7 (0.2 – 1.8)	0.0 (0.0 – 0.0) [α]
CF-independent DCTB, log media (IQR)	0.0 (0.0 – 0.9)	0.0 (0.0 – 0.0) [α]

*Only in HIV-positive participants
[α] Does not indicate complete negativity at baseline.

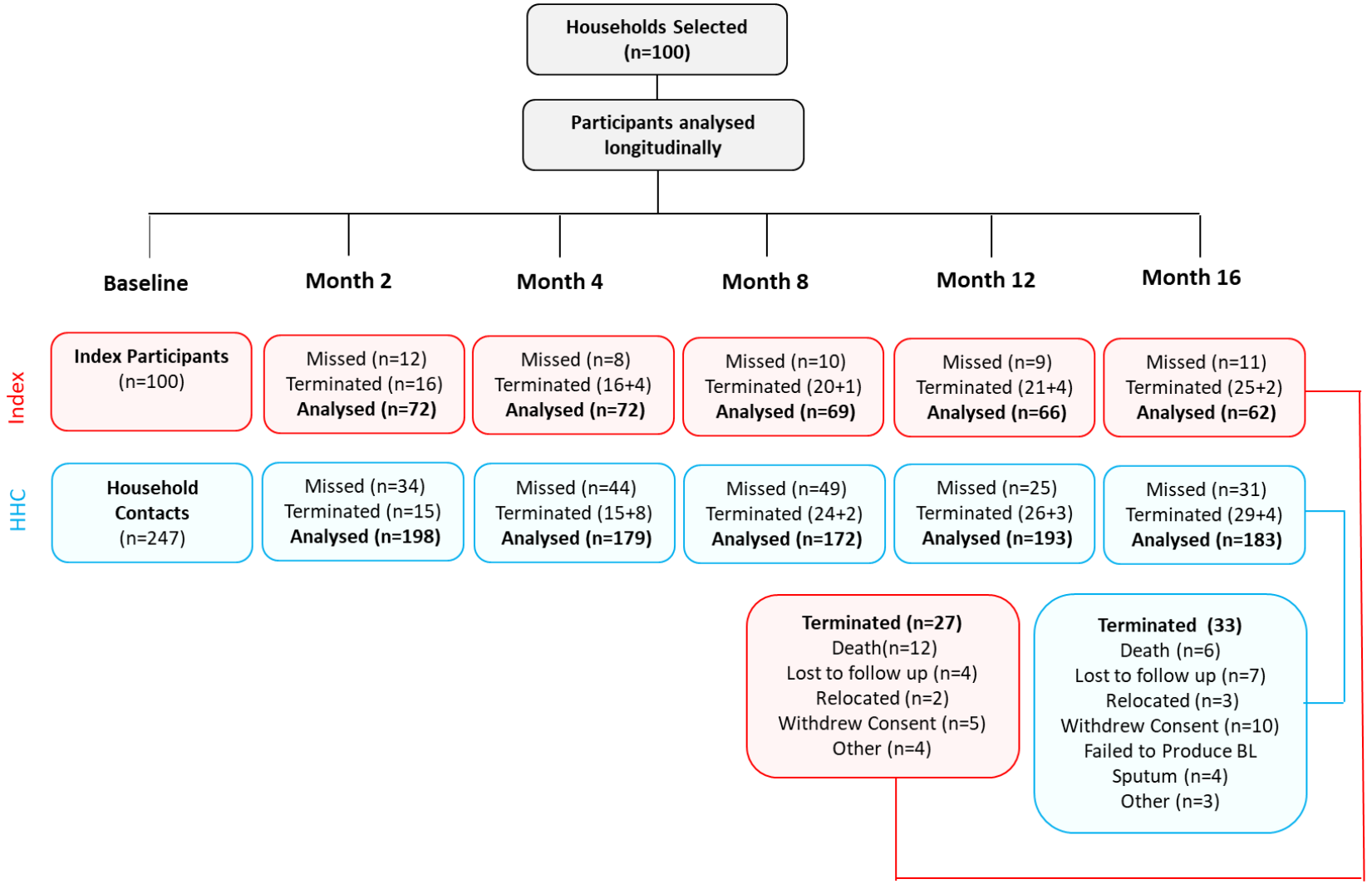


Figure 10: Participant Disposition Chart of Study Participants. Illustration of the number of index participants and HHCs that missed each visit and those that were terminated across the 16-month period. Participants were terminated due to: (1) death, (2) lost to follow up, (3) relocation, (4) withdrew consent, and (5) other. All participants that completed any study visit were considered for data analysis.

4.2 Identifying *Mtb* Prevalence Within a Household by Culturomics Assays

Mtb transmission within households is typically assessed using routine diagnostic methods which have known limitations. For instance, the GeneXpert assay has a limited ability to distinguish between viable and non-viable cells, while smear microscopy stains both viable and non-viable cells. As a result, there are potential gaps in the identification of all bacterial populations present within the host which limits the ability to map all possible transmission links within households. In an attempt to identify the missing link in transmission, this study explored the various culture and molecular methodologies to increase the probability of positive detection of *Mtb* from sputum samples. A spot and overnight sputum specimen was collected from index participants and HHCs at each of the 6 timepoints. These samples were combined and various culturomics and molecular assays were applied to detect even the most minimal quantities of *Mtb*. The MGIT assay is a diagnostic technique and was used as a comparator for MPN/CFU assays, distinguishing bacterial populations that would be detected or remain undetected under routine clinical conditions.

Any MPN wells with growth were plated onto 7H11 media to confirm *Mtb* positivity and at the same time obtain bacterial biomass for DNA extraction. After confirmation of *Mtb* growth in the MGIT tubes, the culture was also plated onto Middlebrook 7H11 media to obtain the required biomass for DNA extraction. The DNA from these samples as well as from the CFU plates was extracted and subjected to WGS to identify the mycobacterial lineage distribution within households. This information will allow for enhanced mapping of transmission links and provides a better sense of *Mtb* transmission within households and communities.

4.2.1 Optimization of Culturomics and Molecular Assays

Culturomic experiments and molecular assays were optimized to ensure streamlined protocols that would yield accurate results. The following approaches were considered during the optimization process: (1) sample type, (2) assay sensitivity and specificity, (3) quality control measures and (4) optimal data analysis.

4.2.1.1 Comparison of Spot and Overnight Sputum Samples

The HoTT DCTB parent study initially assessed spot and overnight sputum samples separately. Due to the large sample size of the study, analysing 2 samples from a single participant became unfeasible as the study progressed. To address this, a decision was made to pool the two samples and assess it as a single sample. However, before implementing this step, it was important to understand the culturability and the quantum of *Mtb* obtained from each of the individual

samples. The baseline spot and overnight sputum samples from the index participants were tested separately for viable bacteria (MGIT assay) and for DCTB (MPN and CFU assay). The results showed no significant difference in the bacterial yield between the spot (Figure 11A, teal column) and overnight samples (Figure 11A, blue column) in both the supplemented (CF+) and unsupplemented (CF-) MPN assay (Figure 11A). Although higher CFU bacterial loads (p=0.0458) were obtained in the overnight sputum sample compared to the spot sputum sample (Figure 11A).

Overnight sputum samples also had a lower TTP (p=0.0169) in the MGIT assay (Figure 11B) in comparison to the spot sputum sample. Overall, the overnight sputum sample has higher bacterial load and appears as a superior quality sample. However, since the spot sputum sample also contain a significant amount of bacteria, all downstream processing of the sputum was done by combining both sputum samples as this would increase the bacterial load and enhance detection of viable bacteria and DCTB.

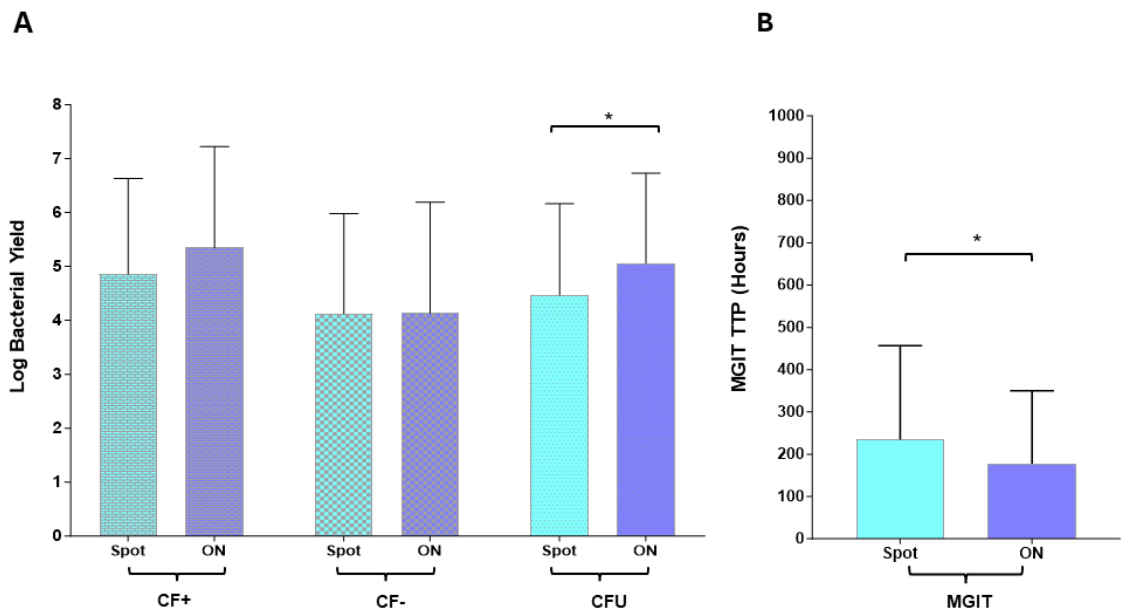


Figure 11: Comparison of Spot and Overnight Sputum Samples by different Culture Assays. (A) Column graph showing the difference in bacterial yield between spot (teal) and overnight (ON - blue) sputum samples assessed by CF-supplemented MPN, CF-unsupplemented MPN and CFU. (B) Column graph depicting the TTP in hours for the spot and ON sputum samples in the MGIT assay. Statistical significance was determined using the Mann-Whitney U test, with a 95% confidence interval. Differences were observed in the CFU assay (p=0.0458) and MGIT assay (p=0.0169). No significant differences were observed between spot and overnight samples in the CF-supplemented and unsupplemented MPN assay.

4.2.1.2 Comparison of Sputum Samples from Study Sites

Participants were recruited from 2 study sites located in different provinces within South Africa: the North West Province and the Free State Province. Sputum samples collected from

participants were transported to the laboratory in Johannesburg, Gauteng, under strict temperature-controlled conditions for laboratory analysis. The Matlosana site is 174 km (2.5 hours) away from the laboratory whilst the Botshabelo site is 445 km (5 hours) away. Samples from Matlosana were transported to the laboratory on the day of collection due to the shorter distance and travel time. Whereas samples from Botshabelo were batched and stored at 4 ºC and transported within 72 hours after collection. All samples were processed within 72 hours of being received at the laboratory. Based on this, we wanted to ensure that no artifacts were introduced into the data due to working with sites across large distances.

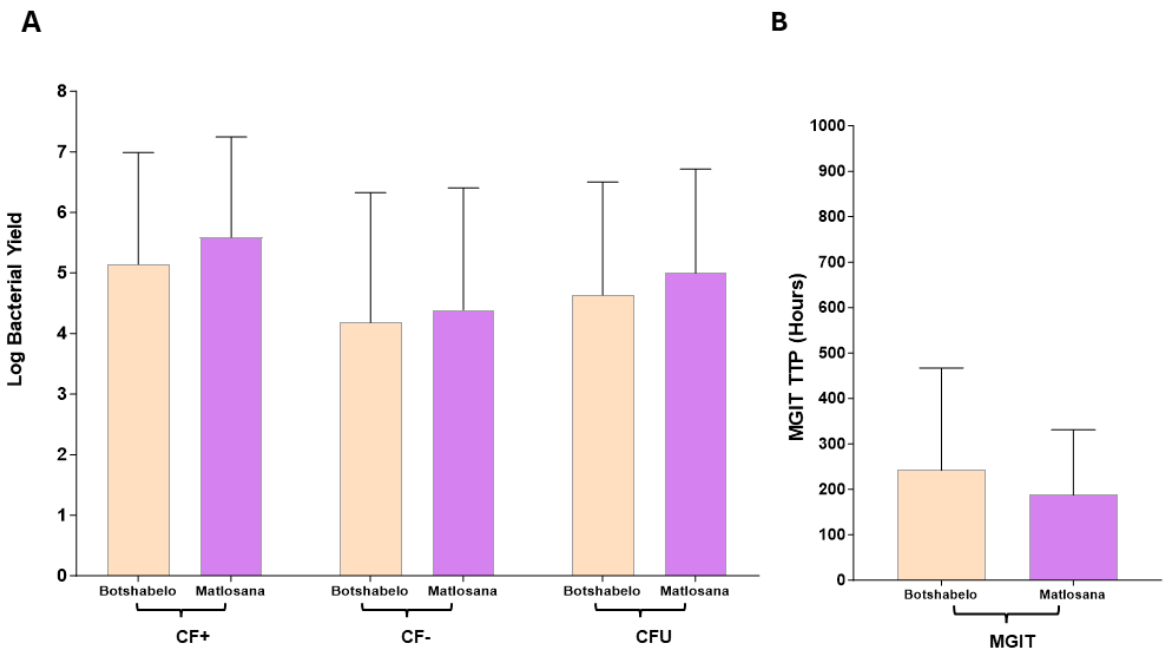


Figure 12: Comparison of *Mtb* Detection in Samples from the Two Study Sites by Different Culture Assays. (A) Column graph showing bacterial yield in sputum samples between the Botshabelo (peach) and Matlosana (purple) study sites by CF-supplemented MPN, CF-unsupplemented MPN and CFU assays. (B) Column graph depicting the TTP in hours for sputum samples collected at each site in the MGIT assay. Statistical significance was determined using the Mann-Whitney U test, with a 95% confidence interval. No significant differences in growth were observed between sputum samples collected at either site in both the assays.

Baseline sputum samples from 50 index participants from each site were compared and indicated no significant differences in bacterial yield with the MPN and CFU assays (Figure 12A) as well as TTP in the MGIT assay (Figure 12B). Hence, the data obtained for samples from both sites were combined and analysed as a single dataset.

4.2.1.3 Comparison of DNA Yield from Cells Cultured in Liquid Media and Solid Media.

Sputum samples that showed *Mtb* growth in the MPN and MGIT assays were cultured on 7H11 media to increase biomass for optimal DNA yield for WGS. Colonies from CFU plates were directly resuspended in TE buffer for DNA extraction. A comparison of DNA yield from bacteria cultured in liquid and solid media was performed on 2 participants. For liquid media culturing, inoculum from MGIT tubes and from CF+ and CF- in the MPN was grown in 7H9 Middlebrook medium for 6 weeks followed by DNA extraction. Additionally, inoculum from the MGIT tube and the MPN wells was plated onto 7H11 Middlebrook agar plates and incubated for 4 weeks after which gDNA was extracted. DNA yield and quality was assessed using a 1% gel electrophoresis. For DNA extracted from *Mtb* cells grown in liquid media, no bands were observed on the agarose gel (Figure 13; lanes 2, 4 and 6), suggesting a poor yield of DNA. Whilst the same samples cultured on solid media showed distinct bands of DNA (Figure 13; lanes 1, 3 and 5).

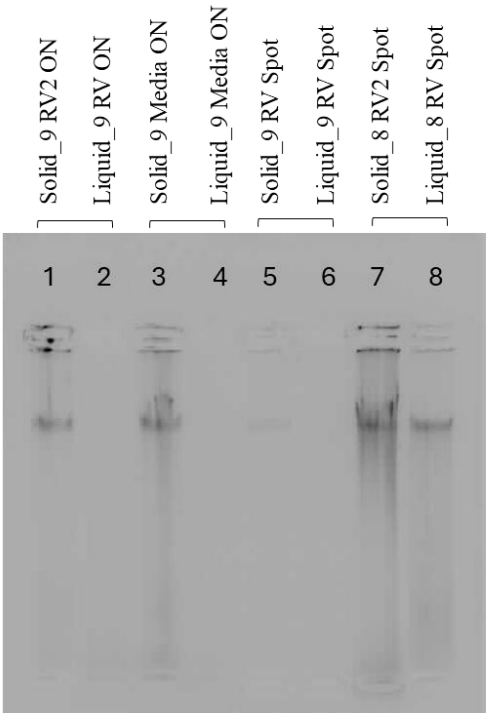


Figure 13: Gel Electrophoresis of DNA Extracted from Solid and Liquid Media. 1% Agarose Gel of DNA extracted from sputum samples from 2 participants cultured in solid and liquid media to determine DNA quality.

Table 4: Comparison of Concentration of DNA Extracted from Liquid and Solid Media. 2 µL of DNA from a Total Volume of 120 µL was Measured Using the Nanodrop.

Sample type	Concentration (ng/µL)	
	Liquid Culture	Solid Media
HH09 Overnight - CF+	14.8	1304.5
HH09 Overnight - CF-	29	1062.2
HH09 Spot - CF+	8.1	840.3
HH08 Spot - CF+	499.5	2268.2

DNA concentration measurements using the nanodrop showed that the DNA yield is several folds greater from *Mtb* grown on solid media compared to bacteria grown in liquid media (Table 4). Given these results, solid media culturing was selected as the preferred method of cell culturing for all samples analysed in this study, after emerging from different liquid culture assays to ensure optimal WGS.

4.2.2 Bacterial Load in Households at Baseline

Baseline bacterial load of an index participant is valuable as it provides insight into disease extent and potential for transmission to exposed individuals within the household. Quantification of the bacterial load at baseline also provides meaningful interpretation of the effectiveness, or lack thereof, of the TB treatment regimen over the 16-month timeline. Whilst assessment of the possible *Mtb* bacterial load at baseline in exposed HHCs of the index participant can highlight the transmission dynamics already at play within the household or within the community.

4.2.2.1 Assessment of Bacterial Load in Index Participants and HHCs at Baseline by Culture Assays

One of the exclusion criteria for recruitment of index participants in this study was that participants should not take more than 3 doses of TB treatment. Therefore, assessment of the bacterial load in the baseline sputum sample in index participants is a reflection of the extent of disease prior to treatment. Individuals with high bacterial loads pose a greater potential to transmit bacteria to their respective HHCs and other individuals within the community. Assessment of the DCTB population in the presence and absence of CF supplementation in the baseline sample of the index participants showed that CF supplementation (CF+ MPNs) yielded a higher ($p=0.0005$) bacterial load compared to CF independent (CF- MPNs) culture conditions and CFUs ($p=0.0004$). The unsupplemented MPNs (CF-MPNs) and the CFUs showed no significant difference in bacterial yield (Figure 14B). Culture conditions are illustrated in Figure 14A.

DCTB populations were evaluated as the log of the quotient of the MPN and CFU values. This confirmed that CF supplementation (CF-dependent DCTB) enhanced the detection of DCTB ($p<0.0001$) compared to CF-independent DCTB populations. HHCs that are exposed to *Mtb* infection may be asymptomatic therefore, do not get tested or treated, but unknowingly continue to transmit the bacteria – these individuals are classified as carrying subclinical disease and significantly contribute to the global TB burden (110).

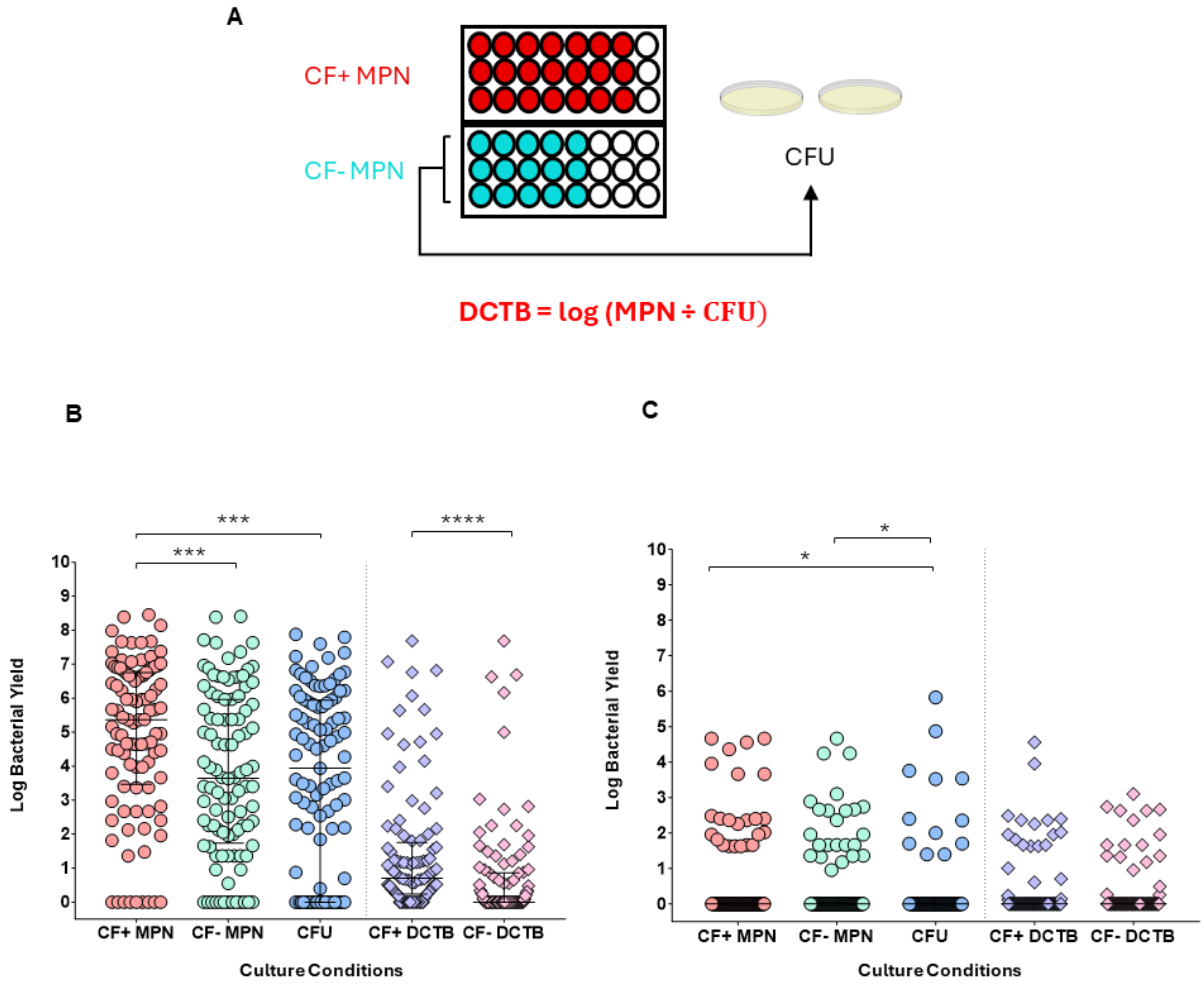


Figure 14: Bacterial Load in Index Participants and HHCs by MPN/CFU Assays at Baseline. Scatter plot showing median bacterial yield by MPN and CFU assays, as well as CF-supplemented and unsupplemented DCTB populations. (A) Overview of culture conditions, (B) bacterial load in index participants, and (C) bacterial load in HHCs at baseline. CF+ (pastel red) and CF− (mint green) represent MPN assays, CFU assays are shown in blue, and CF+ (purple) and CF− (pink) represent DCTB populations. In index participants, CF-supplemented media yielded significantly higher bacterial loads than CF-unsupplemented media ($p = 0.0004$), with no further differences observed. In HHCs, CF+ MPNs yielded higher bacterial loads than CFUs ($p=0.0386$), and CF− MPNs also differed significantly from CFUs ($p=0.0439$). No additional differences were detected. Statistical significance was determined using the Mann-Whitney U test (95% confidence interval).

Assessment of DCTB populations in HHCs at baseline could identify individuals that display no clinical symptoms but have been infected by the index participant. The WHO has established the W4SS criteria (cough, night sweats, fatigue and weight loss) to classify individuals as symptomatic. Since these individuals do not meet the diagnostic requirements for routine TB testing, they would typically remain undiagnosed. However, early identification would allow for prompt intervention with treatment therapies to prevent disease progression.

Figure 14C shows the bacterial load in HHCs at baseline, identified by the CF-supplemented and unsupplemented MPN and CFU assays. The bacterial load at baseline shows no significant differences between CF-supplemented and unsupplemented MPN assay conditions. However, the bacterial yield with CF-supplemented MPNs show a higher bacterial load when compared to CFUs ($p=0.0386$). Bacteria identified by the CF-unsupplemented MPN assay also yielded a higher bacterial load compared to CFUs ($p=0.0439$). The supplemented and unsupplemented DCTB populations show no differences under both culture conditions (Figure 14C). These data suggest that the high number of *Mtb* positivity identified in HHCs by the MPN assay (DCTB) and not by solid media (CFU) is indicative of a major gap in contact tracing and therefore, better testing methodologies need to be developed with appropriate policy implementation for individuals with subclinical TB disease.

4.2.2.2 Bacterial Load in Households at Baseline by MGIT Assay

The MGIT assay is one of the main diagnostic tests used in clinical settings for the detection of *Mtb* in sputum samples from individuals displaying classical TB symptoms. The MGIT assay is highly sensitive and is able to detect the presence of *Mtb* with a short turnaround time, making it an easy and efficient test in the clinical setting. MGIT results are interpreted alongside GeneXpert results, smear microscopy, and clinical evaluation before a confirmed TB diagnosis is established. The baseline bacterial load in index participants and HHCs was evaluated using the MGIT assay (Figure 15). The results showed that sputum from index participants has a significantly shorter TTP compared to the HHCs ($p<0.0001$), as expected since the bacterial load is higher in individuals with active TB. Interestingly, the MGIT assay was able to detect *Mtb* in 28 (11.3%) of the HHCs.

56

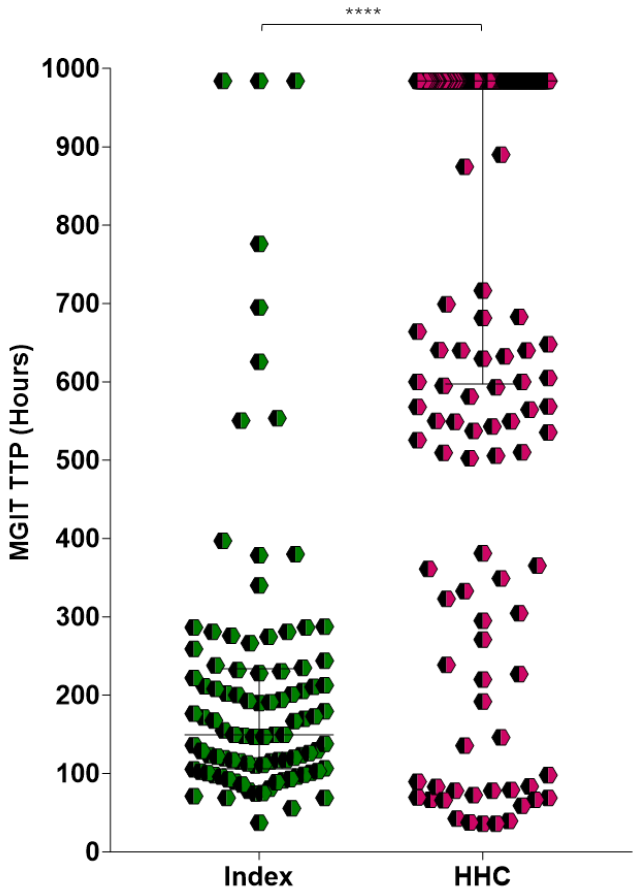


Figure 15: MGIT TTP of Baseline Sputum from Infected Index Participants and HHCs. Baseline sputum samples from TB infected index participants and their respective HHCs was assessed for bacterial growth in the MGIT Bactec assay. The scatter plot shows a significant difference ($p<0.0001$) in the median TTP in hours for index participants and HHCs.

4.3 Assessment of Longitudinal Bacterial Load in Households

Participants in this study were followed up for 16 months which allowed monitoring of the bacterial load at various time points over this period in both the index participants and HHCs. This assessment allowed for a deeper understanding of TB transmission dynamics within households during and after treatment of the infected index participant and anticipate this knowledge will help develop more effective strategies for TB control and prevention.

4.3.1 Longitudinal Assessment of Bacterial Load by MPN and CFU Assays in Sputum from Index Participants and HHCs

Once an individual has been clinically diagnosed with active pulmonary TB, a treatment regimen is formulated based on the drug susceptibility testing indicated by the GeneXpert test. Drug susceptible participants are placed on a 4-drug treatment regimen for a period of 6 months. Within the first 2-4 weeks, most of the bacterial load is significantly reduced however,

prolonged presence of bacteria post treatment may be an indication of drug resistance. In this study, sputum samples were obtained from index participants at baseline, 2-months, 4-months, 8-months, 12-months and 16-months.

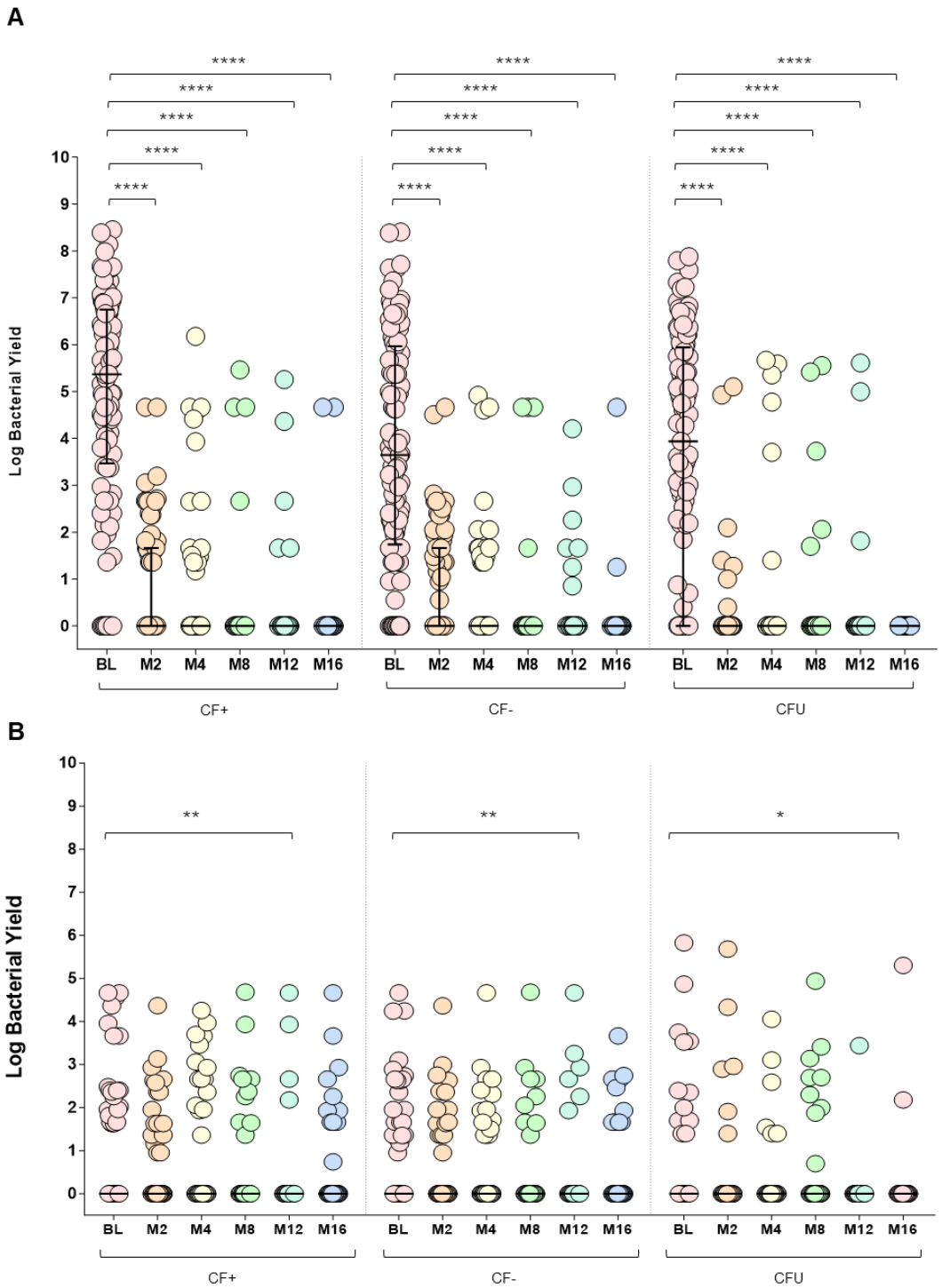


Figure 16: Longitudinal Bacterial Load in Index Participants by MPNs and CFUs. Median bacterial yield was compared in (A) index participants and (B) HHCs from CF-supplemented and CF-unsupplemented MPN and CFU assays. Significant differences in bacterial load were observed in index participants across all timepoints ($p<0.0001$). In HHCs, bacterial recovery varied significantly between baseline and later timepoints under specific assay conditions. Statistical significance was determined using the Mann-Whitney U test (95% CI).

The sputum sample at each time point was subjected to culture and molecular testing to quantify the presence of *Mtb*. As expected, due to the bactericidal effect of the TB treatment, the bacterial load and number of individuals testing positive post enrolment was significantly reduced. Both supplemented and unsupplemented MPN assays together with CFU assessment of sputum samples showed the highest bacterial load at the baseline visit, with 94 of the 100 enrolled index participants showing *Mtb* positivity (Figure 16A).

After the baseline visit, the CF-supplemented MPN group showed a significant reduction in the bacterial load at months 2 and 4 ($p<0.0001$), consistent with the participants being on treatment. Assessment for *Mtb* growth in sputum from HHCs showed bacterial presence in the MPN and CFU assay at baseline up to month 16 (Figure 16B). Significant differences are observed between the baseline and month 12 visit in the CF-supplemented ($p=0.0014$) and CF-unsupplemented ($p=0.0084$) MPNs and between the baseline and month 16 visit in the CFU assay ($p=0.0330$). This was expected as these individuals were not on TB treatment. These data suggest that although HHCs when exposed to TB infected individuals do not display significant symptoms, they continue to harbour low levels of bacteria which is a potential reservoir for progression to active TB disease under conducive conditions.

More importantly, the MPN and CFU assays were able to identify DCTB populations in both the index participants and the HHCs, which typically remain undetected by standard culture conditions. Prior studies have shown that DCTB populations in diseased individuals persist beyond the 6-month treatment period (44), a similar trend observed in the index participants in this cohort (Figure 17A). The presence of CF did not enhance the recovery of DCTB since at each time point equivalent amounts of DCTB were detected with or without CF at each timepoint. There was a significant decline in the DCTB population at month 2 compared to baseline under both the CF+ ($p=0.0017$) and CF- ($p=0.0043$) culture conditions. The presence of persistent low-level of bacteria beyond the 6-month treatment period under DCTB culture conditions and CFU assay, suggests that the standard culture methods used at clinics to detect *Mtb* clearly are insensitive to detect the DCTB population. These data also highlight that first line drugs used to treat drug susceptible TB is unable to kill all populations of *Mtb.* Undetected DCTB populations during routine testing can have profound effects on the treatment response and outcome.

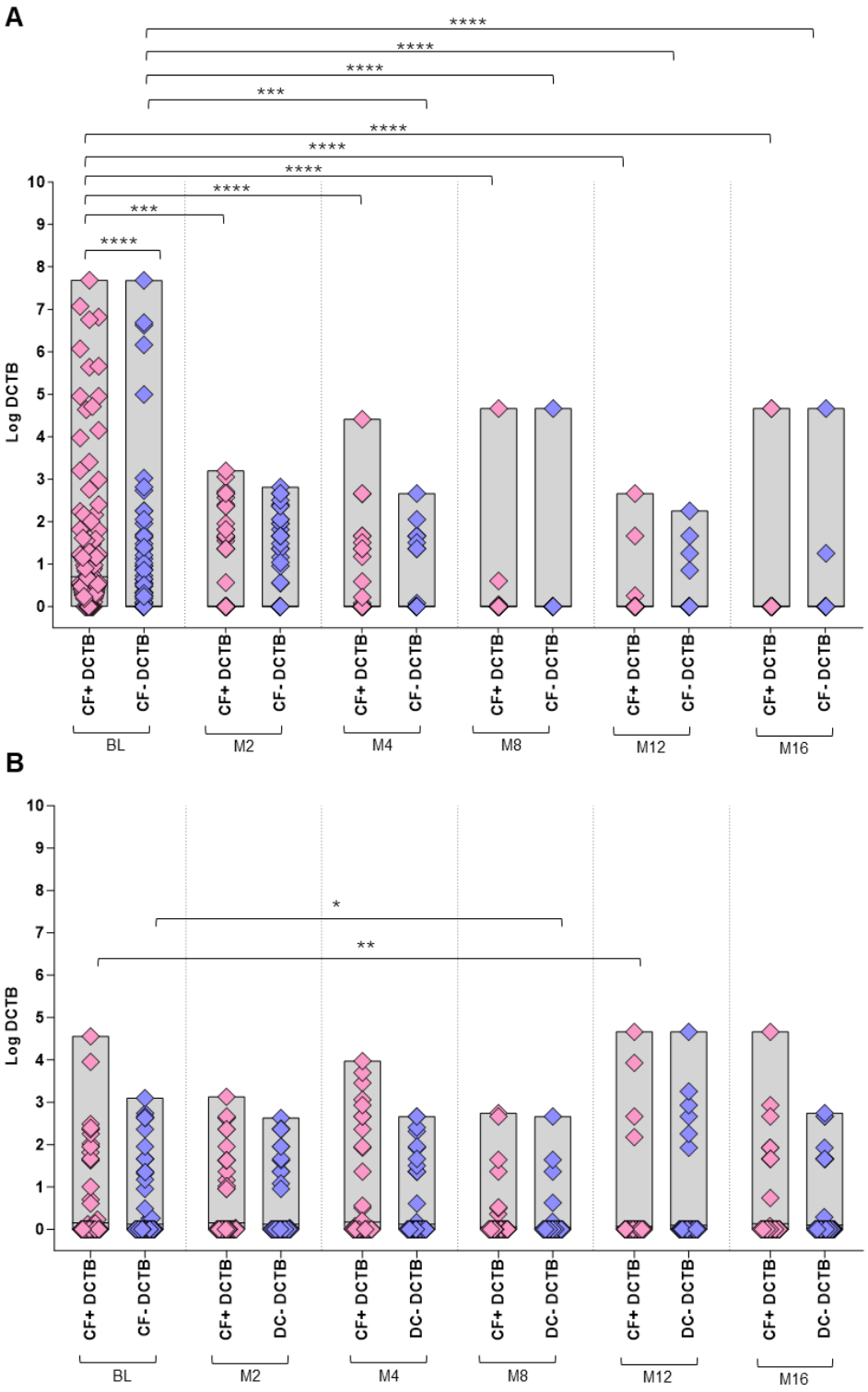


Figure 17: Longitudinal DCTB Populations in Index Participants and HHCs. Floating bar graphs of CF-supplemented and CF-unsupplemented DCTB populations in (A) index participants and (B) HHCs from baseline to month 16. In index participants, significant differences between CF+ and CF- DCTB populations were observed at baseline ($p<0.0001$), with significant differences persisting across timepoints up to month 16. In HHCs significant differences were observed for CF+ DCTB populations between baseline and month 12 ($p=0.0045$) and in CF- DCTB between baseline and month 8 ($p=0.0476$). Statistical significance was determined using the Mann-Whitney U Test with a 95% confidence interval.

HHCs already have a low bacterial load and the presence of DCTB populations can mask the detection of *Mtb* using routine diagnostic tests in these individuals. Evaluation of CF-supplemented and CF-unsupplemented DCTB populations in HHCs from baseline to month 16 showed no differences under each of the culture conditions tested (Figure 17B). However, there was a significant difference in the quantum of CF- DCTB at baseline and month 8 (p=0.0476) and CF+ DCTB at baseline and month 12 (p=0.0045). This decrease in the amount of DCTB could be attributed to an enhanced immune response to control the infection.

Given the above observations, it was important to understand the mycobacterial population dynamics in index participants before and after treatment completion, as this will be an indicator of whether an individual is truly cured or if they have converted to any other disease state such as latent or subclinical in the spectrum of TB disease. Incorrect declaration of an individual as cured can have serious implications for TB control, as these individuals may have converted to subclinical disease state and may still be transmitting the disease to those around them.

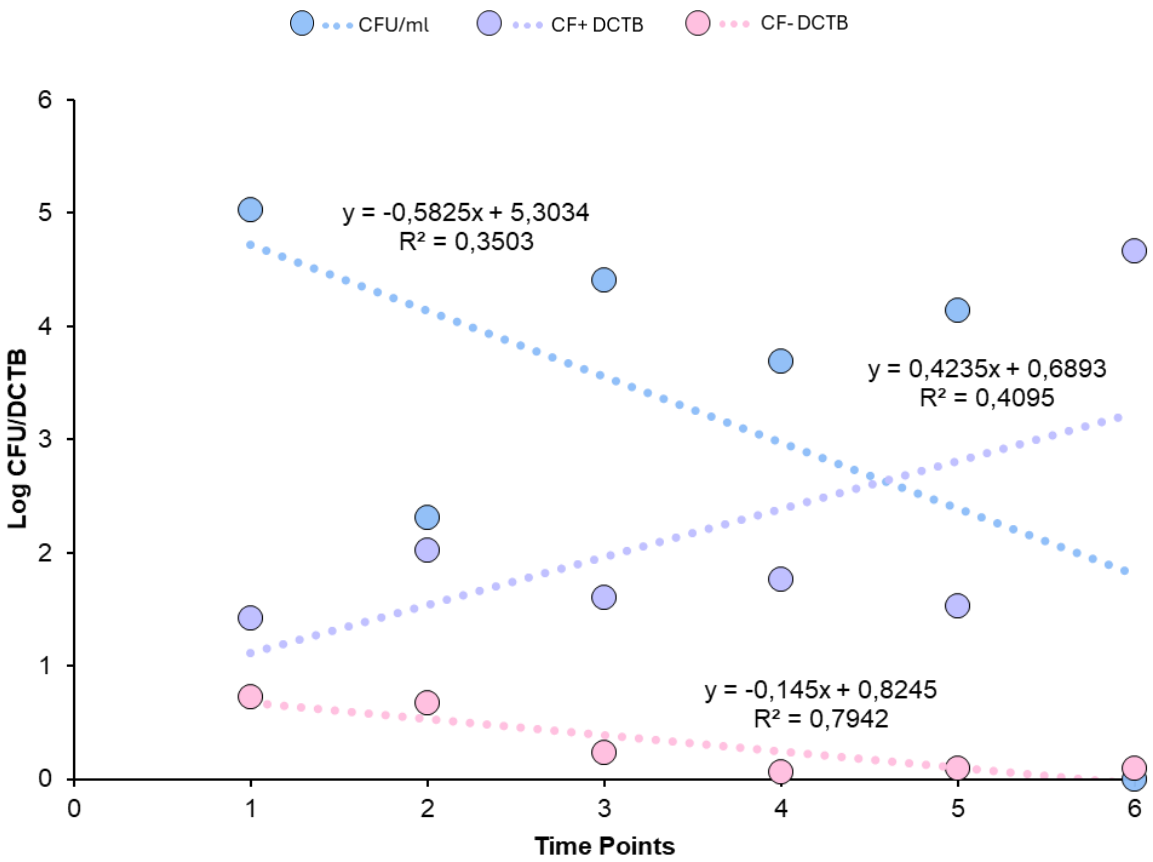


Figure 18: Longitudinal Regression Analysis of the Mean Culturable and Non-Culturable Bacteria in Index Participants. Scatter plot with trendlines showing the average CFU/ml (culturable bacteria; blue) and CF-supplemented DCTB (purple) and CF-unsupplemented DCTB (pink) (non-culturable bacteria) in index participants from baseline to month 16. CFU/ml and CF- DCTB both have a linear decline trendline, while CF+ DCTB has a linear incline trendline.

To assess the mycobacterial population dynamics in this cohort, a regression analysis of the average culturable bacteria (CFU/mL) in comparison to the non-culturable bacteria (CF+/CF-DCTB) from baseline to month 16 for the index participants was carried out (Figure 18). The culturable bacteria (blue) showed a linear decline trendline with a slope of -0.5825, indicating that culturable bacteria decreases over time. A similar trend was observed for the CF-unsupplemented non-culturable bacteria (pink), with a slope of -0.145. Conversely, CF-supplemented non-culturable bacteria (purple) shows an increase in the average quanta of bacteria detected, with a slope of 0.4235. These data once again suggest that CF supplementation plays an important role in identifying differentially culturable populations of bacteria.

4.3.2 Longitudinal Assessment of Bacterial Load in Households by the MGIT Assay

Since the MGIT assay is used in the clinical setting as a proxy for bacterial clearance after treatment, sputum samples collected longitudinally over the 16-month period were tested by MGIT culture to determine the *Mtb* burden during and after treatment. The longitudinal TTP in index participants by MGIT assay showed a significant increase in the TTP from month 2 to month 16 compared to the baseline samples ($p<0.0001$) (Figure 19A). This was as expected since the bactericidal effects of the TB treatment reduces the number of viable organisms over time hence, the increased TTP. There were no significant differences observed between the last timepoint before end of treatment (month 4) and any timepoints after treatment (months 8, 12 and 16).

However, MGIT positive index participants post the 6-month treatment regimen were still identified at months 8, 12 and 16, suggesting that some bacterial populations were not responsive to treatment, or the treatment was ineffective due to the development of drug resistance over the treatment period. Despite an increase in the TTP, the presence of *Mtb* in index participants post treatment is indicative of either treatment failure, reinfection or reemergence. These factors may contribute to transmission within the household and within the community. Assessment of sputum from HHCs in the MGIT assay showed a significant difference in TTP over the different timepoints compared to the baseline sample (Figure 19B).

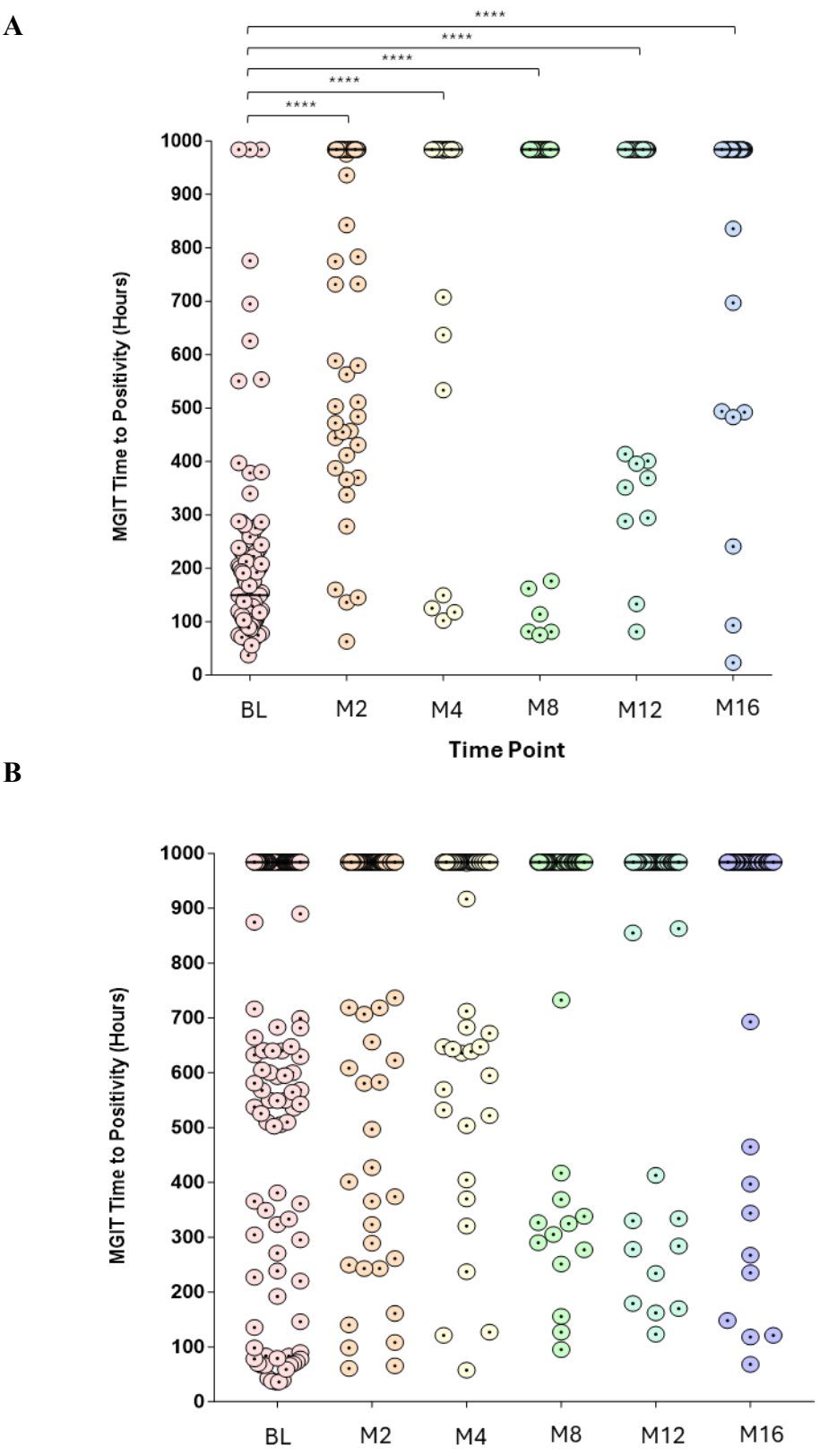


Figure 19: Longitudinal MGIT TTP in Index Participants and HHCs. Scatter plot showing the median MGIT TTP in (A) index participants and (B) HHCs from baseline to month 16. Significant differences in the TTP were observed between baseline and all subsequent timepoints for both index participants and HHCs. The Mann-Whitney U test was used to determine statistical significance, with a 95% confidence interval.

Given the high proportion of index participants (26, 26%) with persistent *Mtb* bacteria post-treatment, it was important to investigate their treatment outcomes. These data are presented in Table 5, where pink blocks indicate culture positivity, grey blocks indicate culture negativity, and white blocks indicate a missed visit.

Table 5: Treatment Outcomes of Index Participants with Persistent *Mtb* After Treatment Completion

			Month 8			Month 12			Month 16		
Group	**Participant**	**Treatment Outcome**	**MGIT**	**MPN**	**CFU**	**MGIT**	**MPN**	**CFU**	**MGIT**	**MPN**	**CFU**
1	HH08B	Cured[α]	missed visit			-	-	-	+	-	-
	HH11B	Cured	-	-	-	+	-	-	-	-	-
	HH24B	Cured	-	-	+	missed visit			-	-	-
	HH35B	Cured	+	-	-	-	-	-	-	-	-
	HH36B	Cured	-	-	-	-	-	-	+	-	-
	HH44B	Cured	-	-	-	+	-	-	-	-	-
	HH45B	Cured	-	-	-	-	-	-	+	-	-
	HH58B	Cured	-	-	-	-	+	-	-	-	-
	HH12	Cured	-	-	-	-	+	-	-	-	-
	HH25	Cured	-	-	-	-	+	+	missed visit		
	HH30	Cured	-	-	-	+	-	-	-	-	-
	HH33	Cured	-	-	-	+	-	-	-	-	-
	HH37	Cured	missed visit			-	+	-	-	-	-
	HH44	Cured	-	-	-	-	-	-	+	-	-
	HH48	Cured	-	-	-	-	-	-	+	-	-
	HH54	Cured	-	-	-	+	-	-	-	-	-
	HH58	Cured	-	-	-	+	-	-	-	-	-
	HH63	Cured	+	-	-	-	-	-	-	-	-
2	HH21B	Treatment Extended	+	+	+	+	-	-	+	+	-
	HH65	Treatment Extended	-	-	+	-	-	-	-	-	-
	HH55B	Treatment Extended[α]	+	+	+	-	-	-	-	-	-
	HH06	Treatment Extended[¥α]	+	+	+	+	+	+	-	-	-
3	HH16	TB Recurrence	-	-	-	-	-	-	+	+	-
	HH59	TB Recurrence	-	-	-	-	+	-	-	-	-
	HH41B	TB Recurrence	+	+	-	+	+	+	+	-	-
4	HH11	Treatment Failure	-	-	-	-	+	-	-	-	-

¥ - Placed on Rif-resistance treatment regimen.
α – Mutation in resistance gene

Of the 26 participants, 18 (69%) were categorized as cured (Group 1) at the 6-month treatment completion timepoint. However, each of these participants were culture positive by at least 1 assay after treatment completion, with one participant (HH08B) developing a mutation in the

rpoB gene encoding for Rif resistance. These participants were likely misdiagnosed as per the South African guidelines that outline that the same test conducted at baseline to diagnose a participant should be used upon treatment completion, to confirm treatment success or failure. Therefore, the current clinical diagnostic tests used to detect bacterial presence was unable to identify the persisting bacteria within these individuals. Group 2 consists of 4 (15%) participants who required treatment extension. Two participants developed drug resistance mutations, however, only 1 was correctly identified and placed on an appropriate 2nd line regimen. The third group consisted of 3 (11.5%) participants who had TB recurrence and were retreated. The 4th group consisted of a single participant who defaulted on treatment and was classified as a treatment failure.

4.3.3 Comparison of Assay-Specific *Mtb* Prevalence in Households

Employing MPN/CFU assays in conjunction with routine diagnostic methods can assist with substantial reduction in undiagnosed TB cases thereby, mitigating the overall burden of the disease. Herein, index participants and HHCs were analysed at each timepoint, with participants divided into 4 groups: (1) MGIT positive but MPN/CFU negative; (2) MGIT negative but MPN positive; (3) MGIT and MPN negative but CFU positive and (4) *Mtb*-negative on all assays.

As expected, 94% of index cases at baseline were MGIT positive, with positivity declining over time but remaining detectable throughout the study period (Figure 20A). At month 2, MGIT culture identified 39% of index cases, 11% at month 4, 9% at month 8, 14% at month 12, and 13% at month 16. The MPN assay was able to identify *Mtb* in index cases that MGIT culture was unable to, 3% at baseline, 21% at month 2, 13% at month 4, and 9% at month 12. Moreover, the CFU assay detected 3% of index cases at month 8 that both the MPN and MGIT assays were unable to detect. These results are shown in Figure 20A. Figure 20B shows that from the 247 HHCs enrolled, 112 (45%) had a positive result in at least one assay over the 16-month study period. The MGIT assay yielded the highest number of *Mtb*-positive cases at each timepoint: 11% at baseline, 12% at month 2, 11% at month 4, 6% at months 8 and 12, and 5% at month 16 (Figure 20B).

The MPN assay detected *Mtb* in samples that the MGIT assay was unable to detect: 8% of all positive cases at baseline, 7% at month 2, 5% at month 4, 2% at month 8, 3% at month 12, and 7% at month 16. This totalled to 55 individuals across all timepoints in whom *Mtb* was detected by the MPN assay but not by standard culture methods. Additionally, the CFU assay detected *Mtb* in samples that the MGIT and MPN assays were unable to: 2% at baseline and 1% at months 4, 8 and 12.

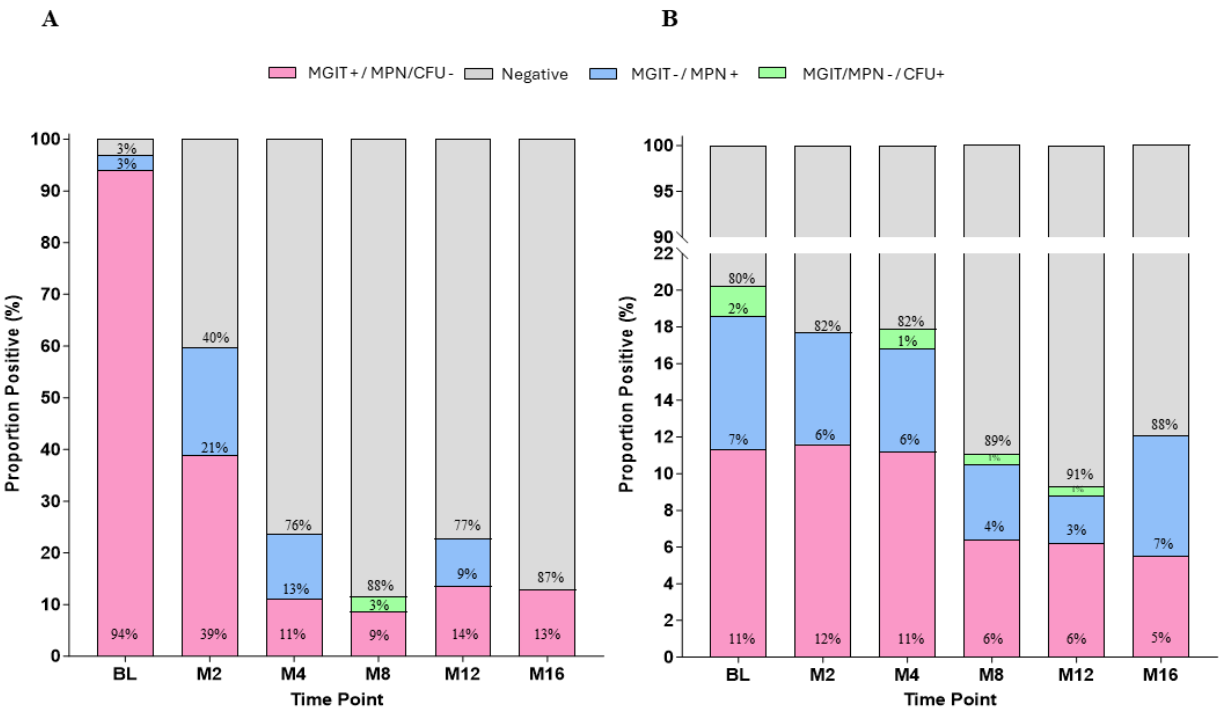


Figure 20: Longitudinal Proportion of Index Participants and HHCs with *Mtb* detected by the MGIT, CFU and MPN assay. Stacked column graph illustrating the longitudinal detection of *Mtb* in (A) index participants and (B) HHCs, by MGIT assay (pink), MPN assay(blue) and CFU assay (green). Grey bars indicate participants with no detectable *Mtb* growth in any of the assays.

Most HHC tracing studies are typically based on a single visit to the contacts of infected index participants. This may provide information on HHCs that are infected at that current time, however, it does not account for infected HHCs that that are in the incipient or subclinical stage of the TB cycle, or those that are yet to be infected suggesting that disease dynamics within a household of an infected individual is in a constant flux. In our analysis of the HHCs, 50 (45%) of all positive cases were identified at baseline by at least 1 of the assays, whilst over the longitudinal follow up period an additional 62 (55%) *Mtb*-positive cases were identified as follows: 26 (23%) at month 2, 11 (10%) at month 4, 7 (6%) at month 8, 11 (10%) at month 12, and 7 (6%) at month 16. Furthermore, 45 (40%) of the positive individuals were either identified by MPN or CFU assays at the first timepoint. Subsequently their bacterial load increased,

allowing them to become identifiable by MGIT assays at a subsequent timepoint, or these individuals were only identifiable by MPN/CFU assays at a later time point in the study timeline.

One limitation of culture-based assays is the long turnaround time, which may require as much as 6 weeks for confirmation of results. During this period, the bacteria continue to replicate, allowing for disease progression in the individual and further transmission to individuals in the household and the community. Therefore, a faster method to detect *Mtb* within individuals will be beneficial in reducing the TB burden. Molecular testing, which is highly specific and sensitive could address this need.

4.4 Identifying *Mtb* Prevalence Within a Household by Molecular Assays

HHCs present with low bacterial load often without clinical symptoms in the early stages of infection and therefore, current culture methods are unable to identify these bacteria. The limitation with culturomics assays is the long turnaround time of obtaining results. The ability to identify the presence of *Mtb* in HHCs early would allow for individuals to begin preventative treatment before the bacterial load increases and the extent of disease worsens. qPCR is a fast and sensitive method that can detect and quantify bacterial load in these individuals within 3-4 hours as opposed to 4-6 weeks of culture time.

4.4.1 Optimization of the qPCR Technique

The IS6110 qPCR technique was optimized by performing multiple reactions with *Mtb* DNA to generate a standard curve with an efficiency of approximately 100%, together with a coefficient of determination (R^2) value close to 1 indicating a positive linear relationship between the dilutions. The red lines in the amplification graphs on the left shown in Figure 21 indicate the 10-fold dilution of the standard curve, which correlates to the efficiency (E) of the qPCR reactions shown as the standard curve graphs on the right.

The first attempt at generating the standard curve with the controls resulted in an efficiency of 80.5% with an R^2 value of 0.992 (Figure 21A). With improved pipetting, the efficiency increased to 104.5% and the R^2 value increased to 0.995 in the second attempt (Figure 21B). The no template control (NTC) and negative control (DNA from *M. smeg*) showed amplification only after cycle 35 due to non-specific binding. This amplification is, however, insignificant.

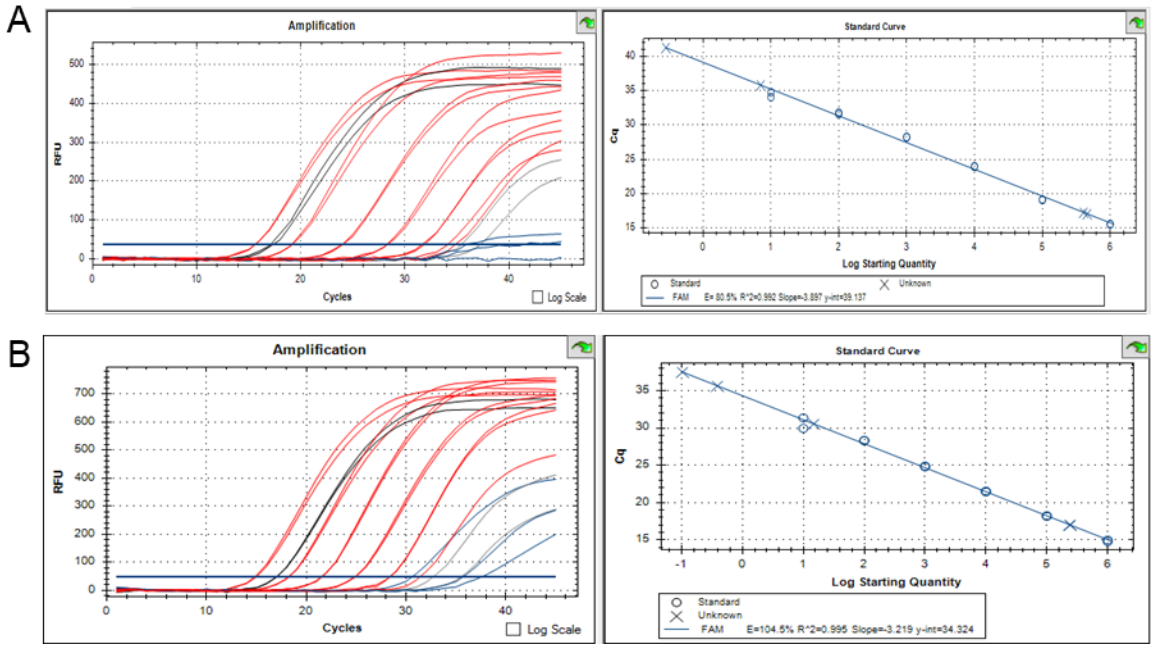


Figure 21: Optimization of the technique used in qPCR analysis. The red lines show the 10-fold dilution of the DNA from *Mtb* used to generate the standard curve in (A) attempt 1 and (B) attempt 2. The dark blue line depicts the positive control, DNA amplified from *M. bovis*. The light blue lines depict the negative control (DNA from *M. smeg*) and the no template control.

4.4.2 Detection of *Mtb*-Positive HHCs by qPCR

Once optimal results were achieved with the optimization, qPCR was performed on clinical sputum samples as a comparative molecular assay measure to the culture assays. The aim was to determine whether qPCR could be used as a rapid detection assay for *Mtb* infection in HHCs.

Sputum samples were decontaminated, after which a 50 µL aliquot was removed for qPCR. The sample was heat killed, centrifuged and the lysate was used. In a considerable proportion of HHCs, qPCR did not have any additive value, as sputum samples from these HHCs were also *Mtb*-positive on culture assays (MPN, CFU and MGIT; Figure 22). At baseline, 20.2% of HHCs were positive by culture and qPCR, while 72.9% were *Mtb*-negative. qPCR identified *Mtb* in an additional 6.9% of HHCs that would otherwise have been missed. Subsequent timepoints showed similar findings with qPCR and culture positivity in HHCs at month 2 to month 16 as follows: 17.7%, 17.8%, 11.0%, 9.3% and 12.1% respectively. Similarly, there were large proportions of HHCs that were *Mtb*-negative at each timepoint: 79.3% at month 2, 73.7% at month 4, 84.9% at month 8, 86.5% at month 12, and 86.3% at month 16. qPCR was able to detect *Mtb* in an additional 3.0% of HHCs at month 2, 8.4% at month 4, 4.1% at months 8 and 12, and 1.6% at month 16. Overall, 56 (8%) cases of individuals that were culture negative, but

qPCR positive were identified. These data suggest that qPCR is a valuable molecular tool for early detection of *Mtb* in HHCs which otherwise would have been missed under routine or specialist culture conditions.

4.4.3 Identification of Early Infection in HHCs by qPCR

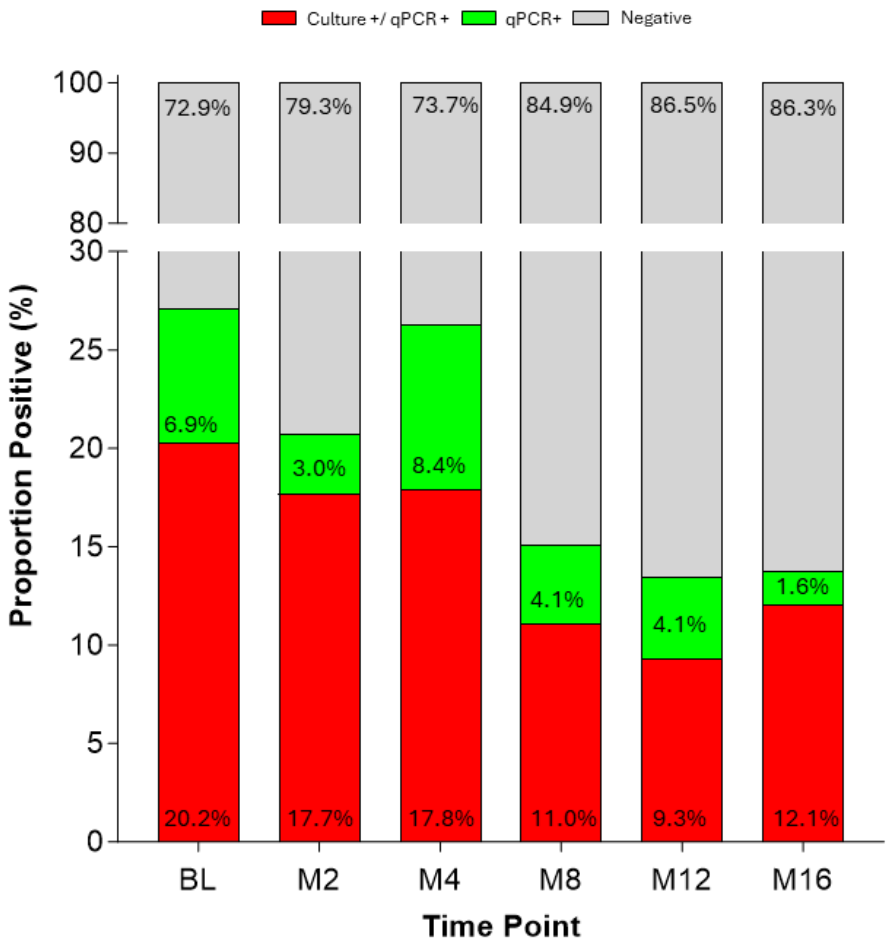


Figure 22: Longitudinal Proportion of *Mtb*-Positive HHCs by Culture and qPCR. Stacked bar graph showing the proportion of HHCs positive by culture (MGIT, MPN and CFU assays) and qPCR (red bars) and HHCs exclusively positive by qPCR (green bars). HHCs in which *Mtb* was not detected is represented by the grey bars.

Upon TB exposure most individuals are able to overcome the infection due to their immune response and therefore, do not require any invention. However, in cases of sustained *Mtb* infection, early detection of the infection followed up with preventative therapies is crucial in reducing the burden of TB. To understand this complexity of infection we analysed the qPCR data to assess whether these data can be used for not only early detection of *Mtb* infection but also to assess how the individual progresses to active disease in the absence of treatment.

Ten HHCs that were qPCR-positive but culture-negative at baseline were analysed to assess when they became culture positive over the 16-month timeline. The duration of transition from

qPCR positive to culture positive varied between the 10 HHCs. Four individuals tested qPCR positive at baseline and showed culture positivity only 12 months later, while 3 individuals were qPCR positive at baseline and became culture positive 2 months later. HH27.2B had a 10-month gap between qPCR positivity and culture positivity, whilst HHCs HH59.1 and HH55.2B had an 8-month gap and a 4-month gap respectively, before presenting as culture positive.

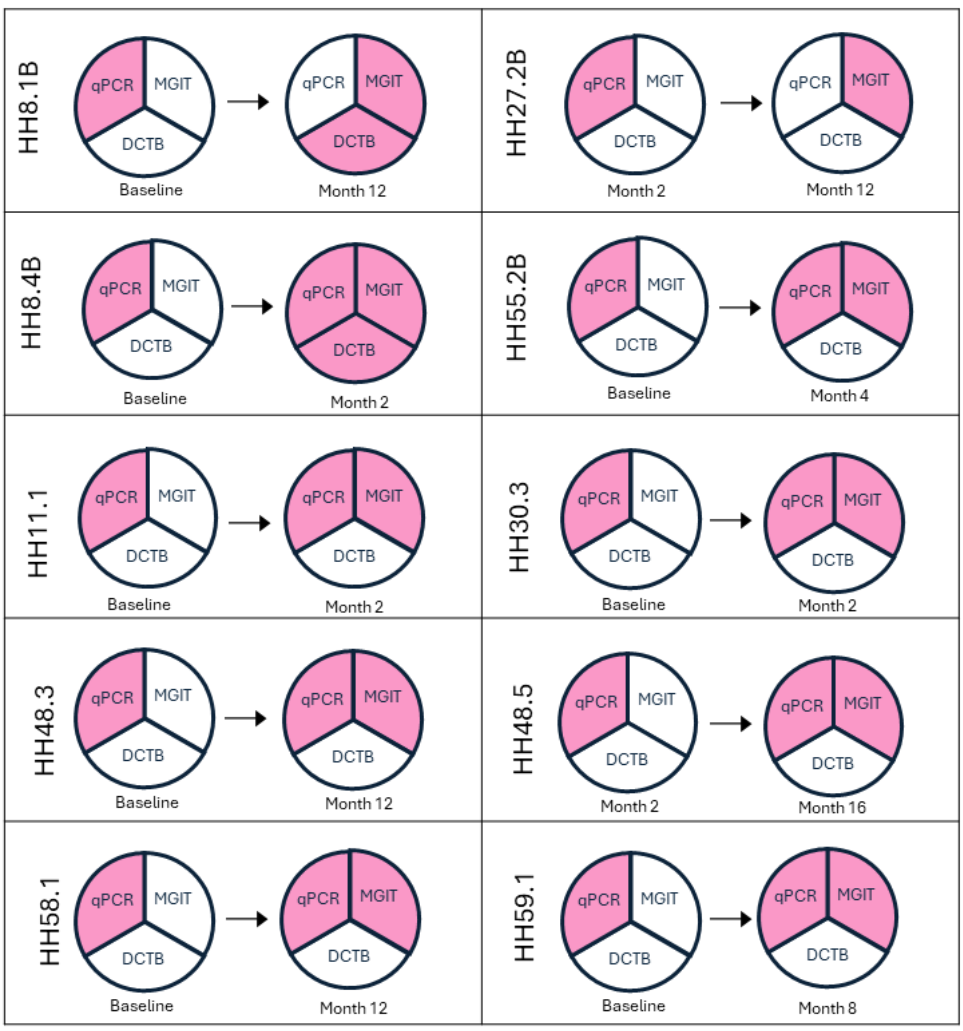


Figure 23: Early Detection of Progression to *Mtb* Culture Positivity by qPCR. Graphic illustration of HHCs that tested positive by qPCR at an earlier timepoint and subsequently progressed to culture-positive at a later timepoint. Pink represents assay positivity while white denotes negative results.

qPCR, although useful for early detection of *Mtb* in HHCs, has its limitations. Firstly molecular-based assays cannot distinguish between viable and non-viable bacteria and secondly as shown in Figure 23, a positive qPCR result does not equate to a positive culture result and can take several months before the infecting bacteria can be measured. Therefore, the presence of *Mtb* DNA in an individual may not necessarily indicate recent infection and is not a reliable tool for inference of household transmission.

4.5 Transmission Dynamics of *Mtb* within Households

To investigate whether HHCs that exhibited growth on molecular or culturomic tests acquired the infection from the index participant, samples from culture-positive HHCs and their respective index participants were collected and cultured at each available time point before gDNA extraction. A subset of these DNA samples was first assessed using spoligotyping to determine the genotype families present in the index participant and assess whether the same genotype family was circulating within their respective HHCs. Additionally, all available DNA samples underwent WGS for further extensive analysis.

4.5.1 Transmission in the Household by Spoligotyping

Spoligotyping is a cost-effective method for classifying *Mtb* strains into broad genotype families. However, it does not provide detailed genomic resolution or detect specific mutations. WGS offers a higher resolution by identifying SNPs and genetic variants, making it the preferred method for precise transmission inference. Therefore, in this project WGS of DNA *Mtb* isolated from infected index participants and their respective HHCs was the end point for transmission inferences. Since WGS was carried out with collaborators in the USA, permits for the export of the DNA was required.

Whilst this process was in progress, spoligotyping was used in the interim to determine the genotypic family of *Mtb* isolated from *Mtb*-positive HHCs and their respective index participants. Unlike WGS, which assigns specific lineages (e.g., lineage 1–7), spoligotyping generates binary codes that classify *Mtb* strains into broad genotype families (e.g., Beijing, LAM, S-family, T1-family) rather than specific lineages.

DNA samples from 9 households were spoligotyped and the binary codes were inputted into the SpotClust program (https://tbinsight.cs.rpi.edu/about_spotclust.html) to determine the genotype family of each sample. The 43-spacer sequence pattern is represented by black dots, where the presence of a dot denotes the presence of a spacer, while its absence denotes the absence of a spacer. In Figure 24, index participants in 5 (A, B, C, E, and F) of the 9 households showed different genotype families (55.6%) from their HHCs, while in the other 4 households (D, F, G and I), the index participant and at least 1 of their respective HHCs shared the same genotype family (44.4%).

Comparison of genotype families obtained by spoligotyping and WGS in households D, F, G and I show that in 3 of the 4 households, the SNP differences between the sequences of each sample were below 13, confirming a possible transmission event. However, in household I, the number of SNPs was greater than the 13-SNP threshold, indicating that a transmission event did not occur within this household. The data, therefore, indicates that inferring transmission based solely on genotype family may lead to an overestimation of transmission within households and the community.

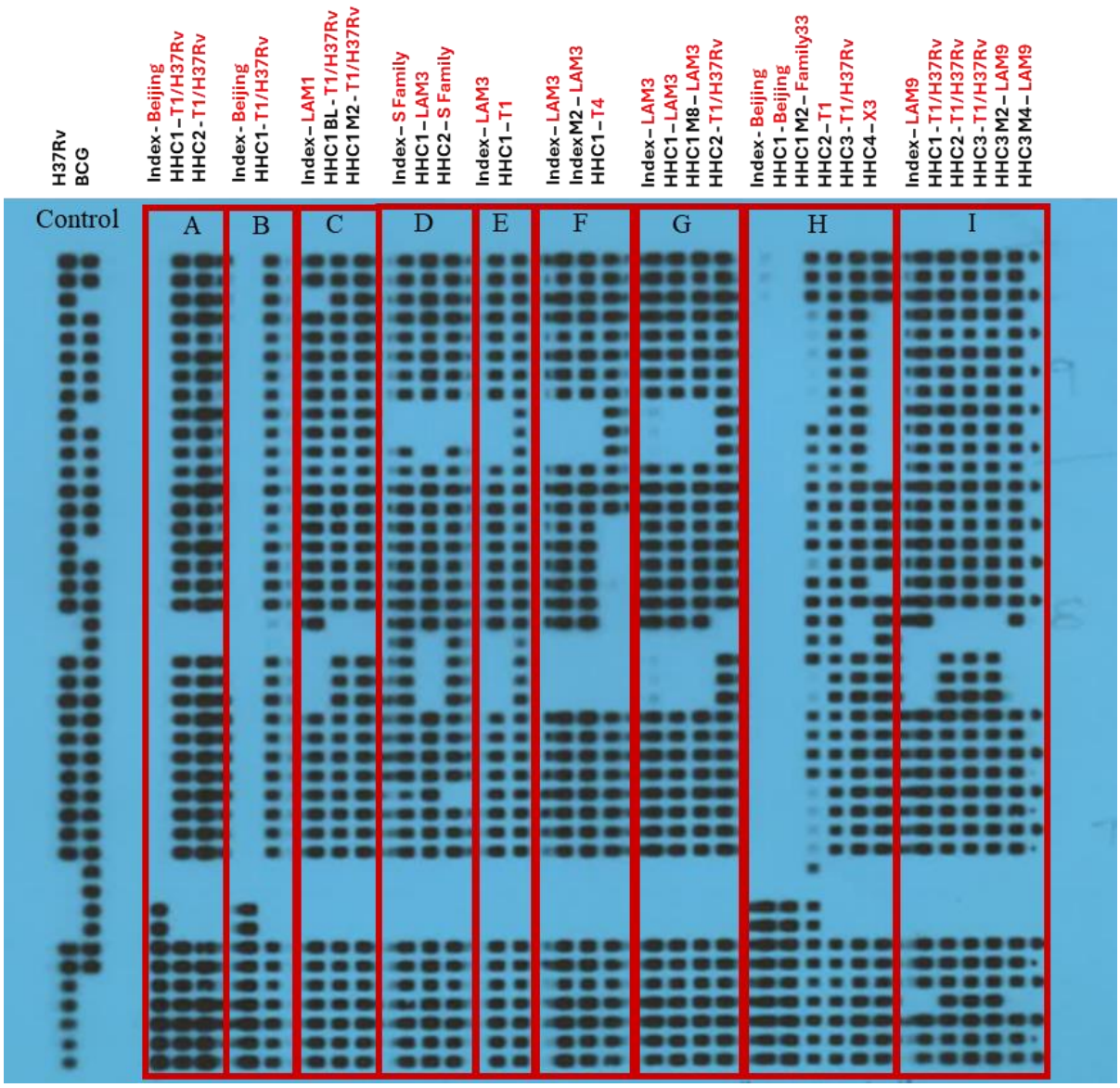


Figure 24: Spoligotyping result showing genotypic lineages of *Mtb* isolates from nine households. The 43 spacer oligonucleotides form a binary pattern representing genotype families from samples of 9 households (A-I), each with at least one *Mtb*-positive HHC. Index participants are in the first lane of each group (red), followed by HHCs in chronological order. DNA from *Mtb* H37Rv and BCG were used as positive controls. BL – Baseline, M2 = Month 2, M4 = Month 4, M8 = Month 8, M12 = Month 12.

A key limitation of spoligotyping DNA extracted from culture assays is the time required for bacterial growth. To address this, and while awaiting DNA export permits, we attempted to spoligotype decontaminated sputum samples directly, without culturing, to determine whether genotype families could be identified. Sample selection was guided by qPCR to prioritize those positive for *Mtb*.

4.5.1.1 Spoligotyping of Decontaminated Sputum Samples to FastTtrack Transmission Linking

We hypothesized that spoligotyping directly from decontaminated sputum would exhibit similar sensitivity to qPCR, as both assays utilize PCR amplification to increase detectable DNA. However, since spoligotyping targets the *Mtb* direct repeat region rather than a highly conserved gene, its sensitivity may be influenced by DNA integrity and abundance in sputum samples.

If a correlation between qPCR and spoligotyping can be made, this would assist in lowering the turnaround time in identifying transmission links within the household. To investigate this, a subset of 32 decontaminated sputum samples from index participants and HHCs were assessed for the number of genomes detectable by qPCR as well as spoligotyping to determine genotype families.

The qPCR data were divided into 4 categories based on the number of genomes detected for each sample: (1) Trace - 0-10, (2) Low - 11-100, (3) Medium – 101-1000 and (4) High - >1001 genomes. Correlation of the qPCR genome equivalents and the spoligotyping data showed that all samples from the ‘high’ category were detectable by spoligotyping, whilst only a subset of samples in the ‘medium’ category and a single sample in the ‘low’ category were detectable (Figure 25). No samples in the ‘trace’ category were detected, indicating that spoligotyping would not be useful in identifying *Mtb* genotype families in individuals with low levels of bacteria as spoligotyping is a hybridization method which requires more DNA. To further enhance our understanding of transmission in households, the DNA from all subsequent samples cultured by MPN/CFU assays and diagnostic culture methods was assessed by WGS.

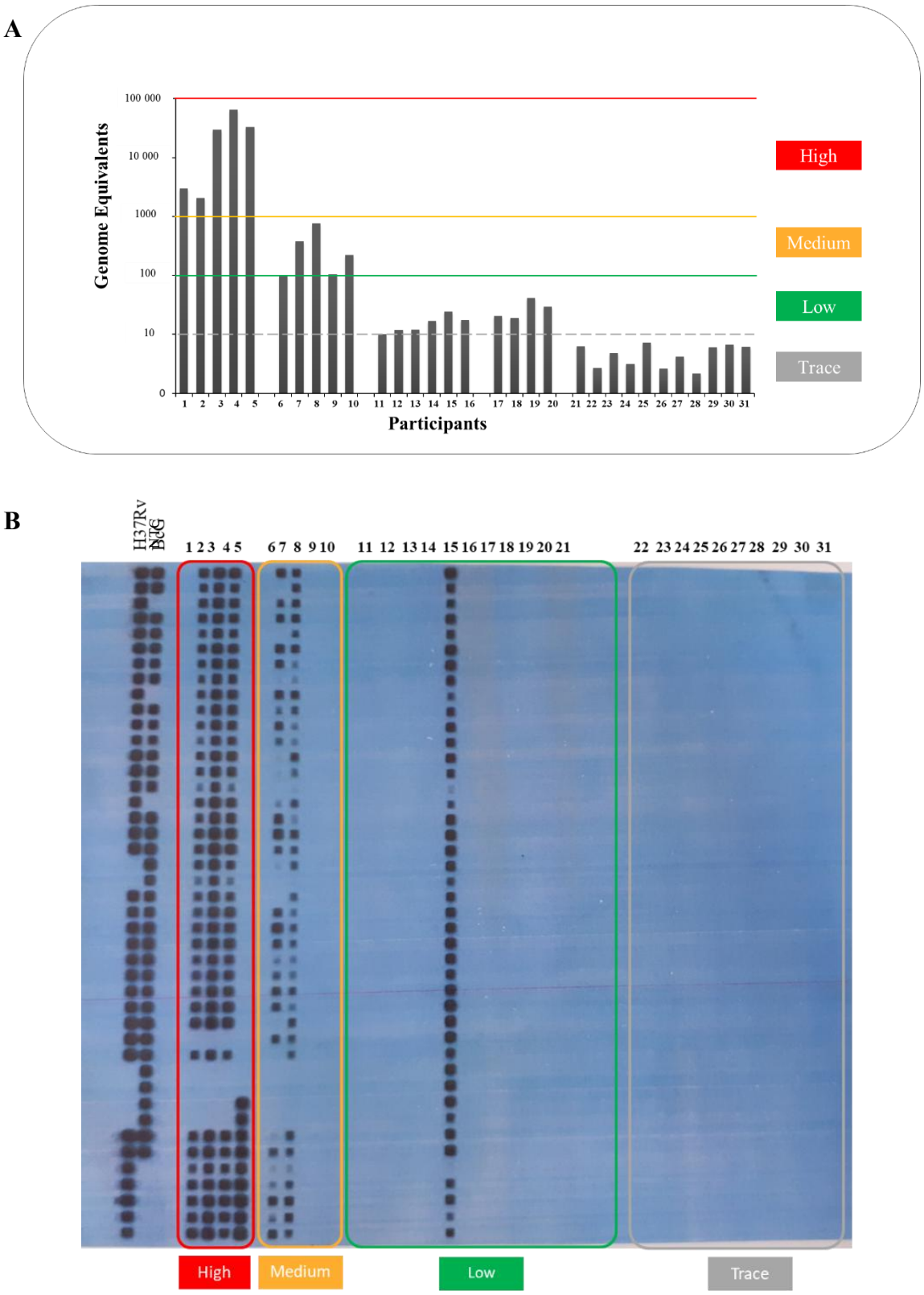


Figure 25: Assessment of qPCR positive sputum samples by spoligotyping. (A) Column graph showing the genome equivalents by qPCR of participants. The number of genome equivalents were divided into 4 categories: High, Medium, Low and Trace. (B) Spoligotyping membrane showing the binary code for lineages detected in each of qPCR genome equivalent categories. An NTC served as the negative control, while H37Rv and BCG were used as positive controls.

4.5.1.2 Whole genome sequencing to establish TB Transmission in Households.

Spoligotyping only analyses the direct repeat region within the *Mtb* genome whilst, WGS provides a more precise distinction between different genotype families strains. This power of discrimination allows for more informed decisions for possible transmission inferences within HHs and within communities. WGS also allows for identification of genetic variations between genotype families and differences within subfamilies. WGS allows for the identification of genomic variation between different genotype families as well as mutations, which have implications for the behavioural variations of the mycobacteria during infection. Hence, the DNA isolated from *Mtb* samples were analysed by WGS to assess if the infecting strain is the same or different between individuals within the same household, to establish a plausible transmission link between these individuals.

4.5.1.3 Bioinformatic Analysis of Whole Genome Sequencing Data

The WGS data were analysed by lineage calling and SNP matrix generation using a customized bioinformatics pipeline. Each genome sequence was aligned to the reference genome, H37Rv using BWA-MEM for sequence alignment. To correct potential sequencing errors and improve sequence accuracy, Pilon was used for realignment. Variant Call Format (VCF) tools were then used to identify variant sites found within the genome, including insertions, deletions and SNPs. VCF tools were used to extract SNPs from the VCF file and to determine the SNP differences between strains of the same genotype lineage. A SNP matrix was then generated using VCF tools to visualize the nucleotide present at each SNP site across samples. To prevent overestimation of transmission within households and to ensure the most accurate representation of transmission events, a SNP cutoff of ≤13 between isolates was applied. This threshold was chosen based on its alignment with studies in high-burden settings, where higher SNP thresholds are often required to account for greater genetic diversity. Literature reports a range of SNP thresholds used to define transmission events, with values ranging from ≤5 SNPs (111) to <50 SNPs (112). For example, a study in a low-incidence country applied a 12-SNP cutoff to identify transmission clusters (113), while another study used ≤5 SNPs for definite transmission and ≤15 SNPs for probable transmission (114). This study's choice of ≤13 SNPs represents a balance between sensitivity and specificity, consistent with thresholds used in other high-burden settings.

Of the 112 *Mtb*-positive HHCs, WGS data were successfully obtained for 74. From these 74 HHCs, a total of 515 *Mtb* DNA samples were isolated from six different culture conditions, where available. Thirty-eight HHCs were excluded from the analysis as *Mtb* isolates from various culture assays could not be sufficiently grown for DNA extraction. Additionally, seven

HHCs from five households were excluded due to low sequencing coverage. As a result, data analysis was conducted on 67 HHCs from 47 households, while the remaining 48 households did not have any *Mtb*-positive HHCs.

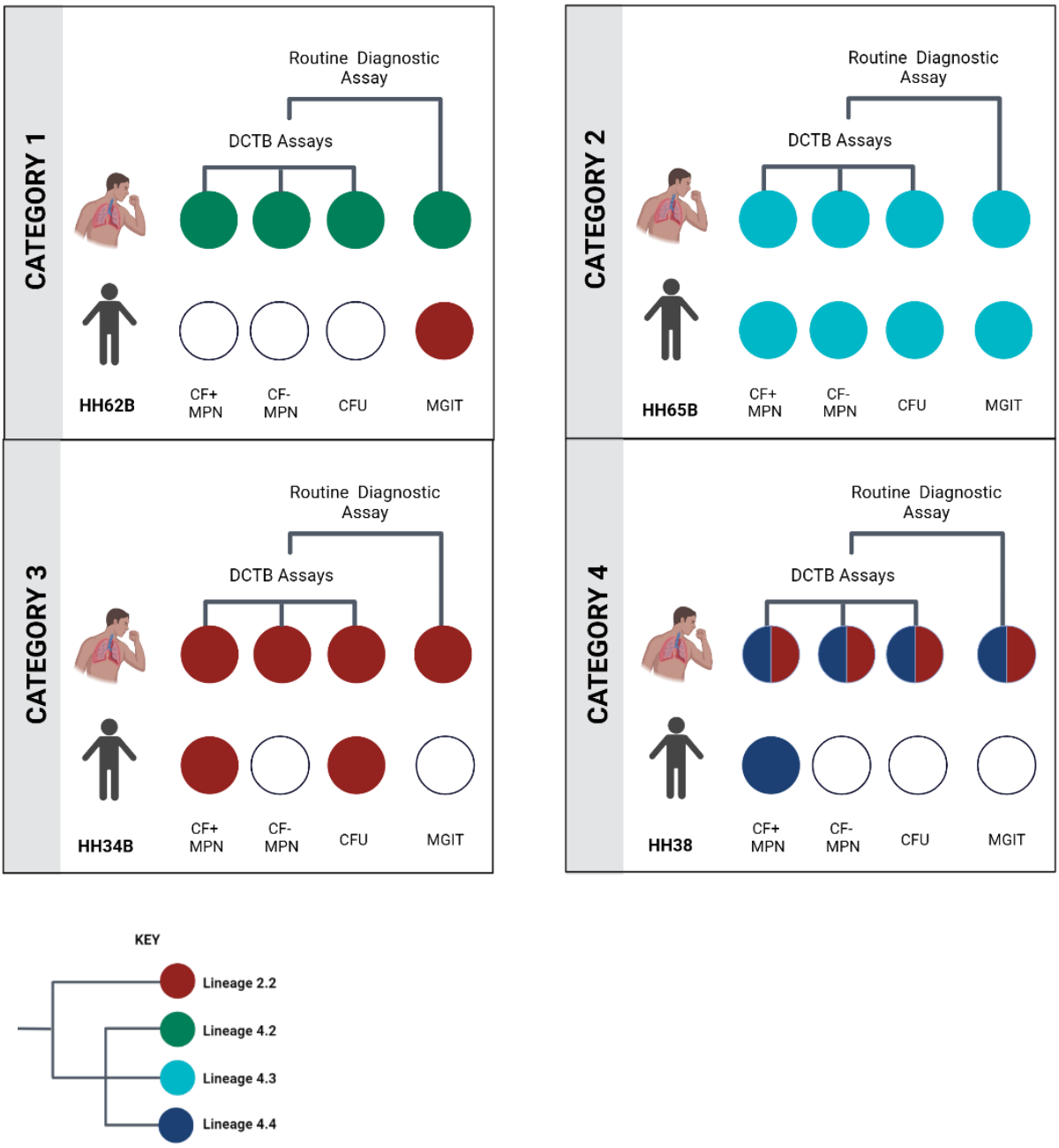


Figure 26: Illustrative Representation of Four Possible of Transmission Inferences. The circles represent the different genotype families and lineages identified in each assay; clear circles indicate no growth. Category 1 comprises households with positive HHCs, however no transmission can be inferred within the household. Category 2 encompasses households with potential transmission events inferred by both diagnostic and MPN/CFU assays, while category 3 comprises households with potential transmission events inferred solely by MPN/CFU assays, within the SNP cutoff of ≤13. Category 4 comprises of households where transmission events cannot be inferred as the number of SNPs are >13, and *Mtb* in the HHC can only be detected by MPN/CFU assays.

Based on the contribution of each culture assay in assigning possible transmission links, the households were grouped into 4 categories: (1) No transmission inference can be made within the household, MPN/CFU assays do not contribute to the identification of positive HHCs. (2) Transmission inference can be made within the household based on identification of positive HHCs using both routine culture techniques and MPN/CFU assays. (3) Transmission inference can be made within the household, by identification of positive HHCs with MPN/CFU assays only. (4) No transmission inference can be made within the household, as routine cultures would likely be negative, suggesting no infection. However, MPN/CFU assays identified *Mtb* that would have been missed by routine culture.

Figure 26 illustrates a representative HH from each category, showing the genotype families identified by MPN/CFU assays and diagnostic assays. Coloured circles indicate the different genotype families detected in the index participant and their respective HHC, white circles represent no growth. The number of households and transmission events is presented in Table 6. Fourteen (23%) of the households were classified into category 1 and 22 (37%) households were placed in category 4 as the strains identified in the HHCs and index participants were either different or exceeded the SNP cutoff. Although transmission inference could not be made between the index and the respective HHCs in category 4, the MPN/CFU assay does provide an additive value as all HHCs in this category were identified by the MPN/CFU assay only. These data imply that the HHCs placed in categories 1 and 4 most likely acquired TB outside of the household, within the community.

In category 2, HHCs in 8 households (13%) were identified by both routine diagnostic and MPN/CFU assays. The same genotype family and lineage was identified for bacteria isolated from the index and HHCs, suggesting clear transmission links between these individuals within the household. In category 3, transmission inferences can be made in 16 households (27%), based on detection of *Mtb* in HHCs by MPN/CFU assays only and therefore, would have been missed with routine culture testing. Cumulatively 69 transmission events were identified amongst 112 (67%) *Mtb*-positive HHCs, and out of these, 43 (57%) transmission events were identified by MPN/CFU assays that would have been missed with routine diagnostic tests only. This confirms the value MPN/CFU assays for diagnostic testing in identifying the missing links in transmission. An illustrative depiction of all households within the 4 categories are supplied in Figure S1 of Appendix D.

Table 6: Categorization of Transmission Inference in *Mtb*-Positive Households

	Households, n (%)	Transmission Events, n (%)
Categories		
(1) No transmission inferred within the household, n (%)	14 (23.0)	15 (22.0)
(2) Possible HH transmission inferred by both MGIT and MPN/CFU assays, n (%)	8 (13.0)	9 (13.0)
(3) Possible HH transmission inferred by MPN/CFU assays only, n (%)	16 (27.0)	18 (26.0)
(4) No transmission inferred within the household, but HHC can only be identified as *Mtb*-positive by MPN/CFU assays, n (%)	22 (37.0)	27 (39.0)

Forty-two (61%) transmission events within this cohort occurred outside the household, confirming that community transmission contributes more to the spread of TB compared to household transmission. Assessment of the prevalence of specific genotype families within a region may help uncover the patterns of community transmission and allow identification of the potential source of infection within these 42 individuals. To address this, we next looked at the prevalence of dominant genotype families at each site.

4.5.1.4 Comparison of the Prevalence of Different Genotype Families at the Two Study Sites

Mtb transmission within communities generally occur from dominant circulating strains within the area. The proportion of the four genotype families identified from sequence analysis of samples from participants at each study site was compared to ascertain which strains were circulating in the two area.

In Matlosana, in the North West Province (Figure 27A), LAM (lineage 4) is predominately responsible for TB disease in the region (77%), followed by Beijing (lineage 2) at 23%. Botshabelo, in the Free State Province (Figure 27B), shows a similar prevalence of the LAM genotype family (69%) and Beijing genotype family, (23%). However, in this region a small proportion of individuals were infected with the Indo-Oceanic (lineage 1; 35%), and East African Indian (lineage 3; 5%), lineages. These findings indicate that the dominant genotype families at both study sites are largely similar, suggesting that relying solely on genotype family prevalence may not be sufficient to infer transmission links.

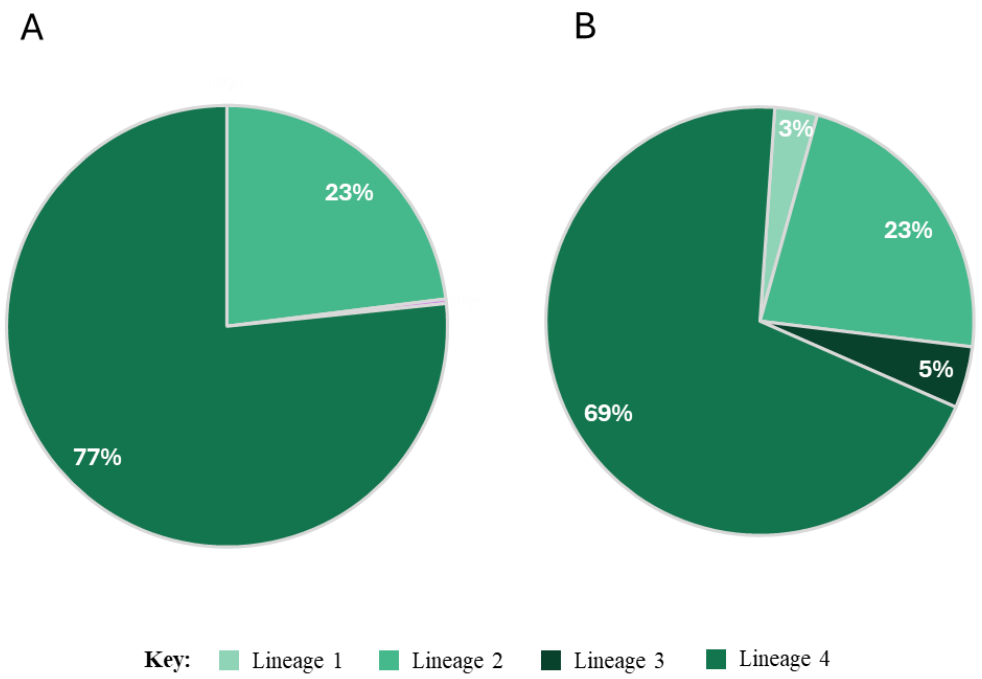


Figure 27: Mycobacterial genotype family and lineage prevalence at the 2 study sites. Pie chart showing the proportion of each lineage prevalent in (A) Matlosana and (B) Botshabelo.

4.6 Detection of Intra-Host Genetic Heterogeneity

Hosts often harbour mixed infections of different *Mtb* genotype families or even different strains within the same family. These strains may exhibit distinct behaviours and may require different conditions to emerge when cultured in vitro. Employing a diverse array of culturing techniques can increase the probability of isolating less-dominant strains thereby, facilitating targeted treatment. The consequence of these strains going undetected may result in treatment failure or relapse disease. WGS provides an abundance of information on phylogenetics, drug resistance and virulence factors that gives insight into the biology and evolution of the bacterium.

4.6.1 Detection of Mixed Bacterial Genetic Heterogeneity in a Single Host

Mtb infection in humans is not necessarily limited to a single genotype family and may involve multiple variants of the bacterium within a single host. This genetic heterogeneity can have implications for disease progression, development of drug resistance and overall treatment outcome. Identification of all subpopulations within a single host can significantly enhance our understanding of disease pathogenesis that can advance efforts in the global fight against the TB epidemic. Tailoring of TB treatment therapies to target specific populations of the bacteria can potentially improve treatment efficacy and restrict the emergence of drug resistance.

Analysis of samples with mixed infections under the different culture conditions showed that CF-supplemented MPN assays were able to identify the highest proportion of all mixed infections (9/23, 39%), followed by CF-unsupplemented MPNs (7/23, 30%) and CFUs (4/23, 17%), while MGIT assays identified the least number of mixed infections (3/23, 13%) (Figure 28). CF uncovered greater genetic heterogeneity in sputum from infected individuals, further emphasizing the value of MPN assays to detect all populations of *Mtb* that are missed under standard culture conditions.

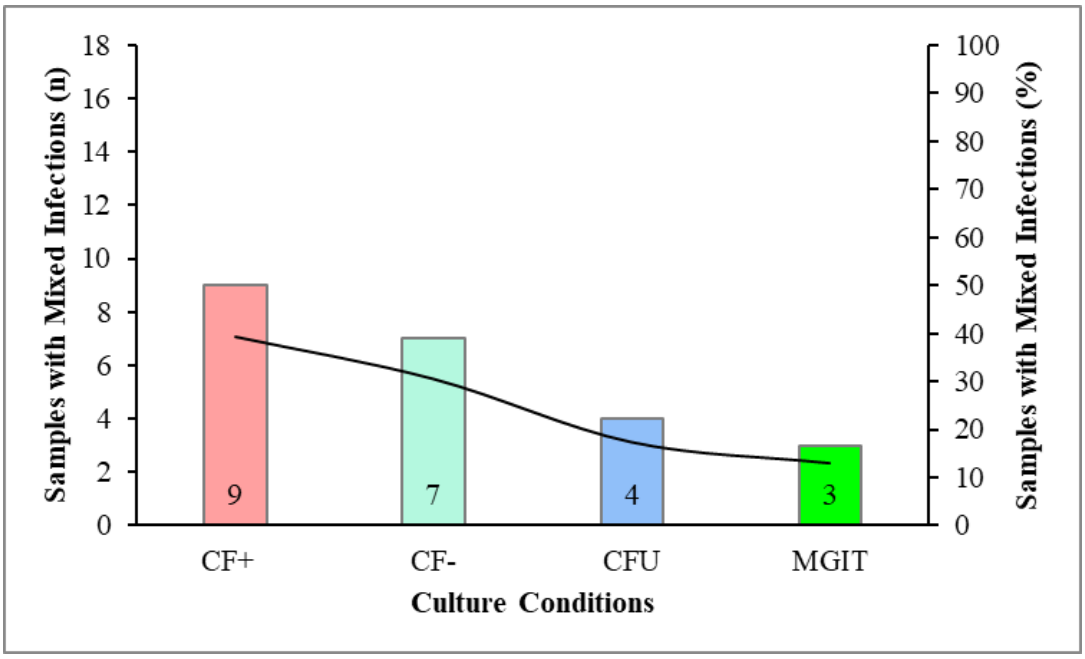


Figure 28: Detection of mixed infections in single sputum specimens. Combined column and line graph showing the number and proportion (%) of specimens with mixed infections isolated under different culture conditions. A total of 515 bacterial specimens sequenced from 47 households showed the presence of mixed infections in 23 of the samples. Five samples had more than 2 strains.

4.6.2 The Impact of Culture Conditions on Uncovering Genetic Heterogeneity

Mtb cultured under different growth conditions can lead to the emergence of different genotype families or subpopulations belonging to different lineages within the same genotype family. These subpopulations likely arise as the bacteria adapt to changing environmental conditions during infection (in vivo) and during growth in-vitro. *Mtb* from sputum (in vivo) when cultured using different culture methods, can face challenges such as differences in nutrient availability, presence of adequate growth supplements, or varying oxygen concentrations which impacts on the growth of specific subpopulations that are better suited to survive and replicate under particular conditions. These subpopulations can exhibit differences in their genetic makeup, including mutations or gene expression patterns, compared to the original infecting strain.

Mycobacterial genetic heterogeneity can have severe implications for TB diagnosis and treatment as treatment may not be effective against all the strains. It was hypothesized that different culture methods create conditions that selectively promote the growth of pre-existing subpopulations within the host to emerge.

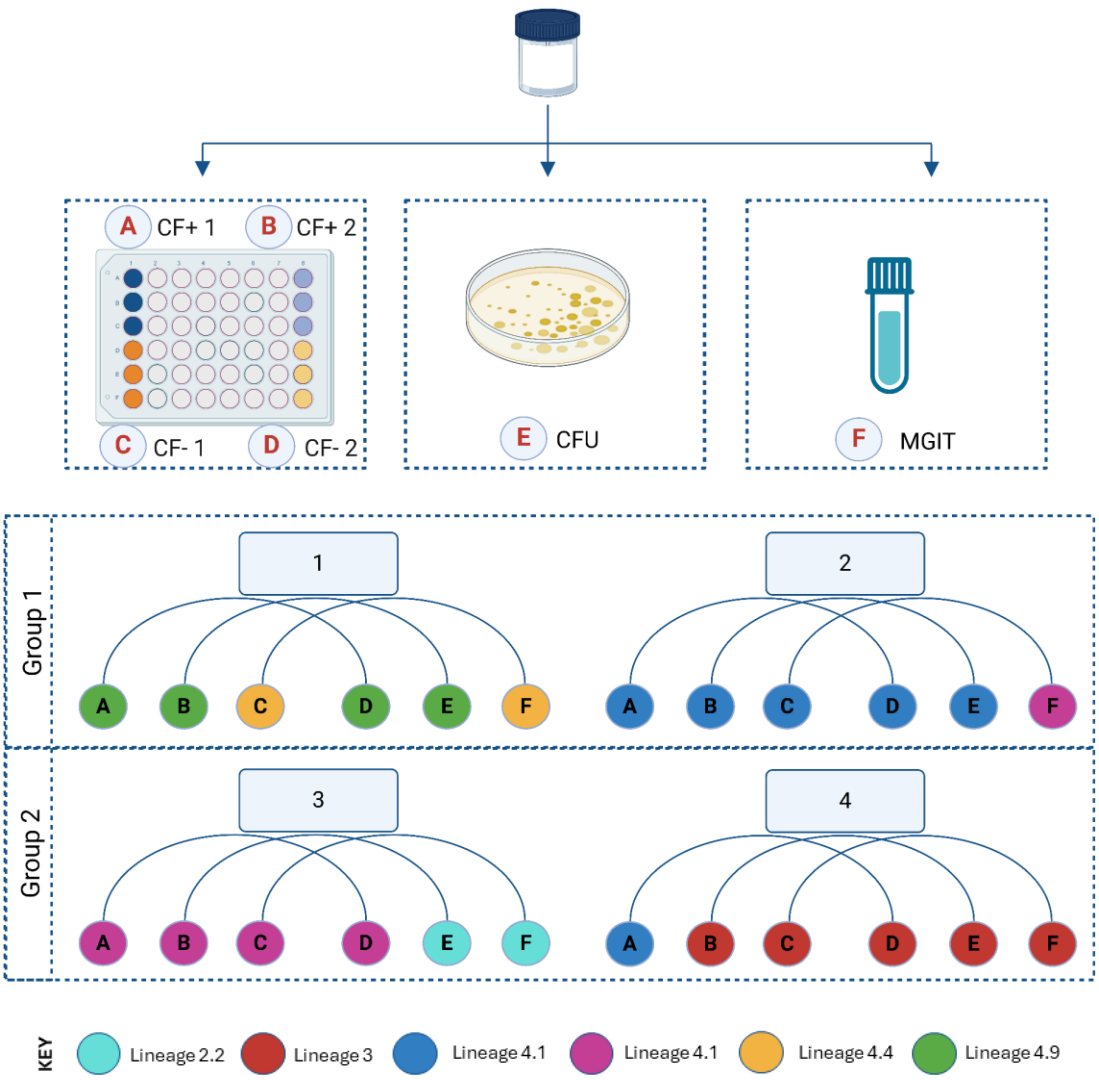


Figure 29: Graphical Illustration of Mycobacterial Genetic Heterogeneity in Infected Individuals. Illustrative depiction of the emergence of different *Mtb* strains within a single individual at a single time point cultured under different growth conditions. Growth conditions: (A) CF+ low dilution of cells, (B) CF+ highest dilution of cells, (C) CF- low dilution of cells, (D) CF- highest dilution of cells, (E) CFU plates and (F) MGIT culture. Each colour represents a different genotype lineage.

The different genotypic lineages identified within a single sample were categorized into 2 groups (Group 1 – detection of different strains in the same genotype family; Group 2 – detection of different genotype families). These categorizations were based on strains detected

under the 6 culture conditions. Conditions A-D were derived from the MPN assay, where cells were taken from different dilutions and inoculated onto solid media: (A) CF+ low dilution of cells, (B) CF+ highest dilution of cells, (C) CF- low dilution of cells and (D) CF- highest dilution of cells. Condition (E) was derived from CFU plates and culture condition (F) was inoculated onto solid media from MGIT culture. This is shown in Figure 29. In the first group, the MPN/CFU assays detected a different strain within the same genotypic family compared to the MGIT assay. Participant 1 showed that the MPN/CFU assay identified lineage 4.9, which would remain undetected as the MGIT assay detected lineage 4.4.

A similar pattern can be seen in participant 2, where the MGIT culture identified lineage 4.3, while the MPN and CFU assays detected 4.1. In participant 3, the MGIT and CFU assay identified Lineage 2.2 (Beijing), while the MPN assay identified lineage 4.3 (LAM). Similarly, in participant 4, the MGIT assay together with the CFU, unsupplemented MPN and diluted CF+ all identified lineage 3 (East African Indian), while the least diluted CF-supplemented MPN detected lineage 4.1 (LAM). All participants within the 2 groups can be found in Figure S2 of Appendix D. Thus, WGS of mycobacteria isolated from infected index participants and HHCs grown under different culture conditions showed that different culture conditions have the power to detect different genotype families as well as genotype lineages within these genotype families, already present in the host during infection. This further confirms the importance of the MPN assay in uncovering genetic heterogeneity within hosts.

These patterns of genetic heterogeneity within individuals further demonstrate that as the dominant strain is diluted in the MPN assay by 10-fold dilutions, the less-dominant strain emerges. This effect is particularly evident in participant HH32, where a mixed infection with Beijing (blue) and LAM (red) is observed in the CF-supplemented and unsupplemented MPN, CFU and MGIT culture (Figure 30). Using different culture methods on a single sputum sample clearly shows the power of culturomics assays in uncovering the mycobacterial genetic heterogeneity pre-existent within hosts. This becomes particularly useful in understanding treatment failures, especially in cases of heteroresistance, where the treatment regimen may not effectively target all subpopulations of *Mtb.*

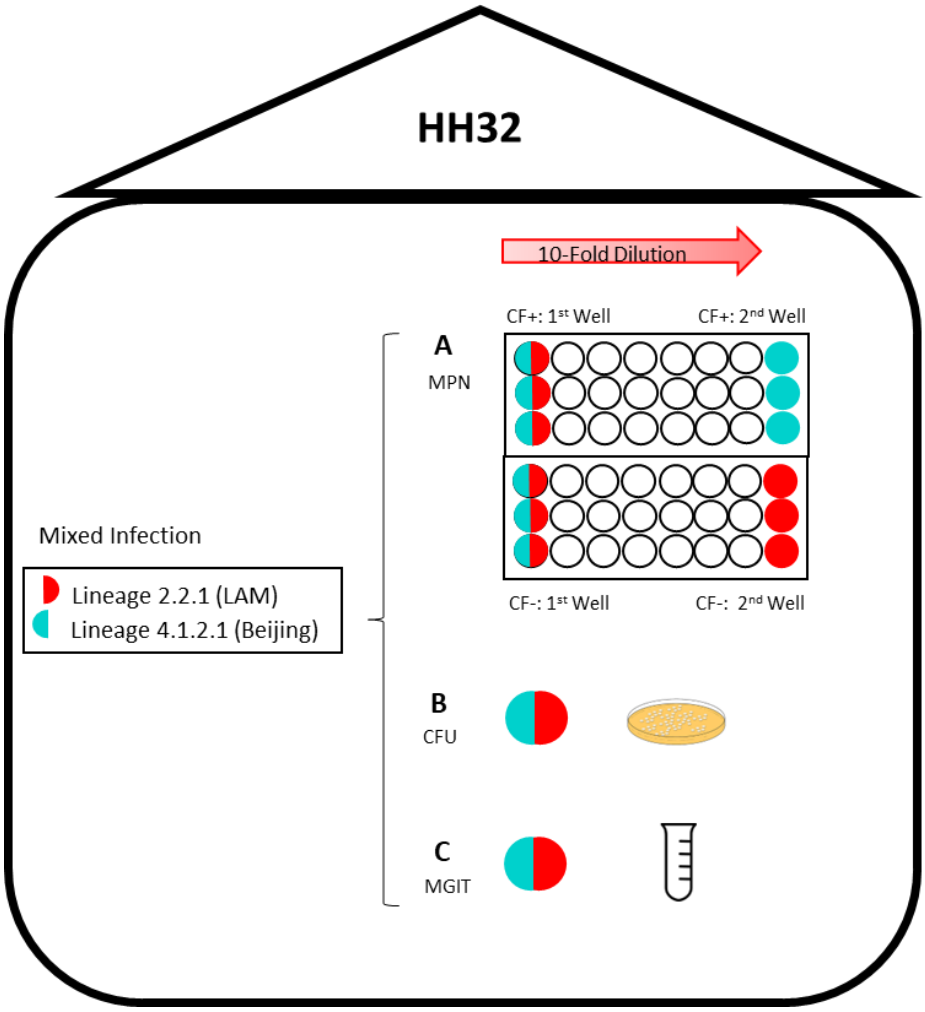


Figure 30: Separation and Identification of Mixed Strain Lineages Using the MPN Assay. The sputum sample from the index participant of household 32 was tested for *Mtb* growth using various culture assays. All assays detected a mixed infection for this participant. Genomic sequence analysis of bacteria detected in the MPN assay allowed for identification of two genotype families.

4.6.3 Longitudinal Change in Genotype Family Lineages

Mycobacteria experience stresses when exposed to TB treatment. To survive, the mycobacteria will mutate its gDNA specifically in the target region of the drug. In order to see how TB disease progresses within individuals during treatment, the WGS from 32 individuals including index participants and HHCs at each available timepoint was analysed (Figure 31). The participants were grouped into 3 categories based on the patterns observed over the 16-month timepoint. Group A consists of individuals harbouring the strain identified at baseline, which persists longitudinally throughout the study period. Additionally, over time, some individuals acquired additional *Mtb* strains, potentially from infected individuals within the household or from community exposure. In some individuals, mixed strains were identified at baseline; however, as treatment progressed, one strain – possibly less dominant or more responsive to treatment -

was lost. The persistence of a strain at the month 8 timepoint is suggestive of ineffective treatment. For example, participants 6 and 10 consistently showed the presence of lineage 4.3 from baseline to month 16, whilst participant 13 is presenting with lineage 3, suggesting that these strains were unresponsive to treatment and most likely developed drug resistance or tolerance over the 6-month treatment period.

Group B consists of individuals that consistently harbour the same strains at each timepoint. Participants 16-23 show the same strain at baseline until month 4, with eradication of these lineages post month 4, indicating successful treatment. Participants 24 and 25 are examples of late infection over the study timeline. Interestingly, participant 15 shows infection with lineage 4.1 at baseline and month 2, after which no growth was detected for 12 months suggesting treatment success. However, lineage 4.1 reemerges at month 16, which most likely is due to relapse disease or reinfection. The HHC of participant 15 presented with lineages 4.1 and 4.3 at month 12 therefore, we can infer that the index participant was reinfected by the HHC between months 12 and 16.

Group C consists of 6 individuals exhibiting different strains at each timepoint. Participants 27, 28 and 32 demonstrate diverse lineages throughout the study timeline, potentially indicating multiple infection events or the suppression of less dominant strains by more dominant ones. For instance, participant 31 has a mixed infection with lineages 3.1 and 4.4 at baseline and month 2, which subsequently clears. At month 4, a mixed infection with lineages 2.2 and 4.9 appears, which clears by month 8. However, lineage 4.9 reemerges at month 12, possibly due to multiple infection events. Identifying and understanding these dynamic infection patterns with mixed strains, will allow for more effective and targeted TB treatments, ultimately improving patient outcomes and reducing the likelihood of treatment failure and disease relapse. Observing the complexity of mixed infections within a single participant and the varying patterns of *Mtb* positivity within participants over time prompted an investigation into the factors that may influence why some index patients are more likely to transmit *Mtb* to their respective HHCs and the factors that make HHCs more susceptible to infection. Therefore, we next examined the health status of both indexes and HHCs to identify whether a single factor or combination of factors contributes to (1) increased transmission from index participants and (2) increased risk of infection in HHCs.

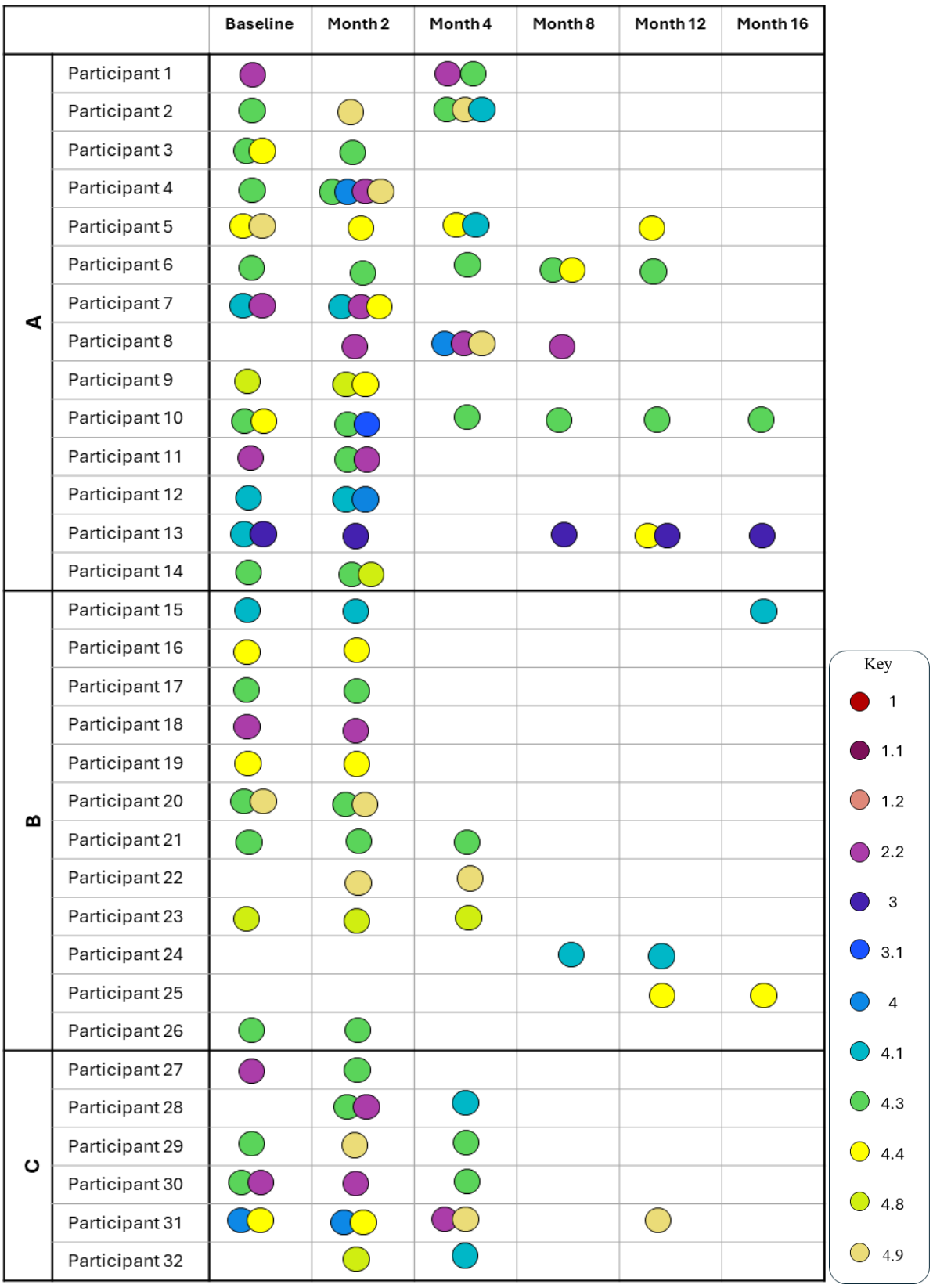


Figure 31: Longitudinal Genetic Heterogeneity within Individuals. Illustrative depiction of the different strains within a single individual over time. Thirty-two individuals were divided into 3 groups based on the observed patterns of heterogeneity. Each color represents a different genotypic lineage.

4.6.4 Factors that Influence the Transmission Potential of Index Participants

The burden of disease, which refers to the severity and extent of TB disease within an individual, plays a crucial role in determining how likely they will transmit the infection to others in contact with them. Individuals with a higher burden of disease often have a greater quantum of bacteria linked with clinical symptoms resulting in an increased chance of spreading the disease.

4.6.4.1 Effect of the Extent of Disease on the Transmission Potential of Index Participants

The extent of TB disease in index participants refers to the degree of disease progression and the impact the disease has on the body. This includes the severity of clinical symptoms, and dissemination of the disease within the host. TB disease primarily affects the lungs, but it can also spread to other parts of the body. The body's immune response to infection results in the formation of granulomas which indicate successful containment of the infection. Over time, a weakened immune system may either allow for the bacteria to reactivate or the centre of the granuloma may liquify due to necrotic material, leaving an open cavity in the lung, known as cavitation.

The left and right lungs differ functionally and anatomically and may influence the process of infection and progression to active TB disease. The right lung is shorter and wider, with 3 lobes while the left lung is longer and narrower with 2 lobes. The bronchus of the right lung is wider, making it more susceptible to aspirated materials such as in the case of TB infections. The anatomical differences between the 2 lungs also influence *Mtb* infection and manifestation. The upper lobe in the right lung is more conducive to *Mtb* growth due to the higher oxygen tension and also facilitates the spread of the bacteria, leading to higher incidences of cavitary lesions (115).

Chest x-rays of index participants at baseline, month 2 and month 12 were examined to monitor disease extent during and after treatment. The participants' lung health was assessed to establish whether it increased the likelihood of transmission in households. A comparison of the radiographical results at baseline, separated into index participants with at least one positive HHC and those without positive HHCs showed that index participants that transmitted to HHCs had cavities and infiltrates in both lungs, and adenopathy and pleural disease/effusion in their right lungs (Figure 32). While index participants that did not transmit the disease to their HHCs have adenopathy and pleural disease/effusion in their left lungs, although these findings did not

have any statistical significance. However, a significant difference was observed in the occurrence of pleural effusion in the right lung of index participants with *Mtb*-positive HHCs (p=0.0014), who had a higher prevalence of this condition compared to those without *Mtb*-positive HHCs. No other significant differences were observed between the 2 groups. Therefore, lung pathology alone does not appear to be a reliable predictor of transmissibility in index participants.

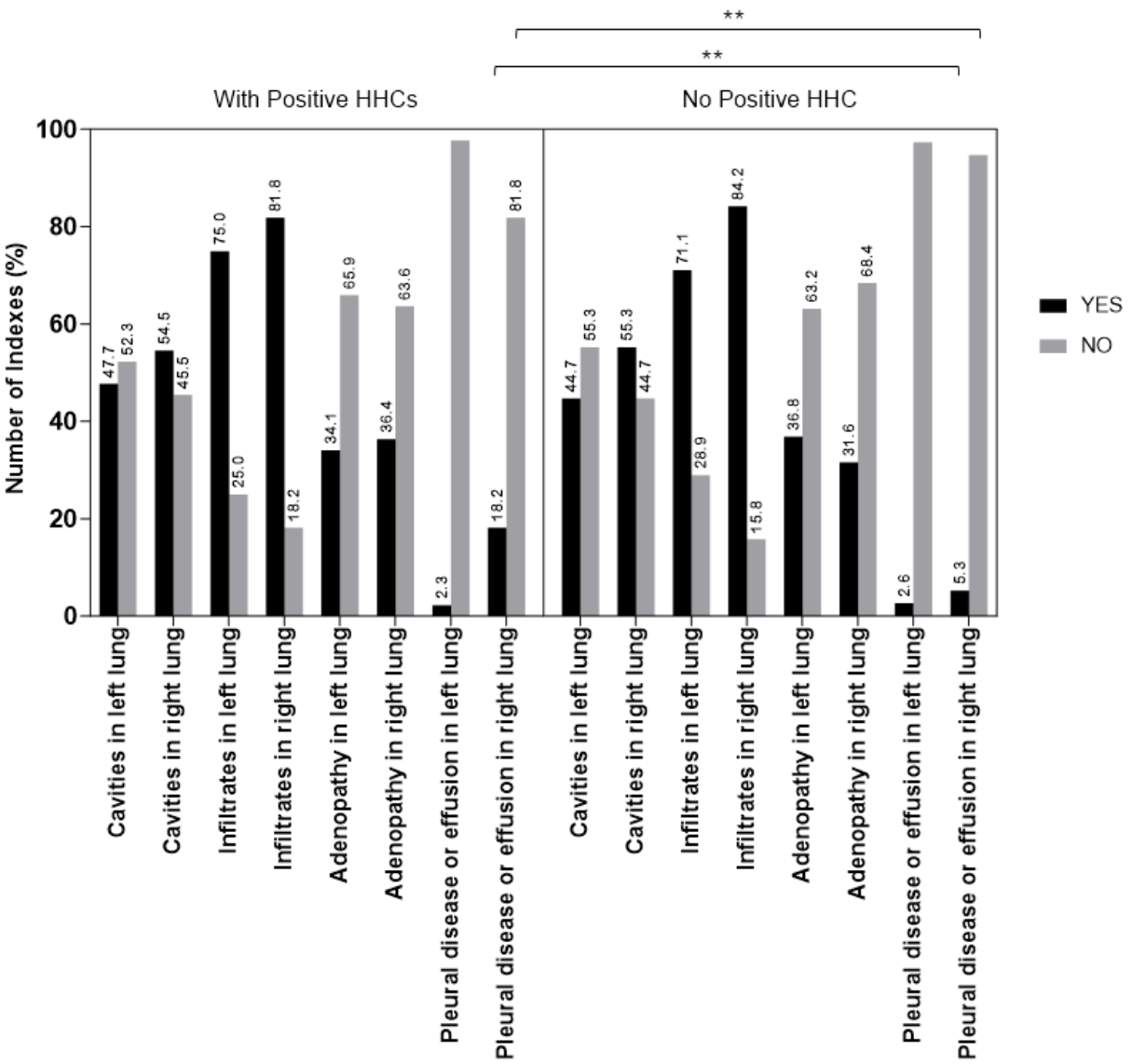


Figure 32: Chest X-ray Results of Index Participants at Baseline. Comparison of lung pathology based on chest x rays taken from index participants at baseline, with a positive HHC and those without positive HHCs. A significant difference in pleural effusion of the right lung was observed between the two groups (χ^2 test, p = 0.0014, 95% CI).

At month 2, 8 of the available 68 X-rays from index participants indicated a worsened lung pathology from baseline whilst 4 participants showed normal pathology in both lungs and the remaining 56 indicated no change. At month 12, only 1 index participant had a worsened lung pathology, 4 were normal in both lungs, 6 had improved lung pathology and the remaining 26 indicated no change. Improved lung pathology indicates a reduction in the overall extent of disease, with fewer cavitations and infiltrates compared to previous report findings (data not shown).

4.6.4.2 Assessment of Bacterial Load in HIV-positive and HIV-negative Index Participants

To assess if HIV infection impacts the transmission from TB infected index participants to their HHCs. The bacterial load in the baseline sputum from both HIV-positive and HIV-negative index participants was assessed by the CFU assay and the MPN assay with and without supplementation. HIV status influenced differences in bacterial load by CF+ supplemented MPN and CFU assays. A significant difference was observed in the CF-supplemented MPNs ($p=0.0260$) and CF-unsupplemented MPNs ($p=0.0015$), with HIV-positive index participants yielding a lower quantum of bacteria in comparison to HIV-negative indexes (Figure 33A). Baseline sputum from HIV-positive individuals on ARV therapy showed a significant difference in the CF-supplemented MPNs ($p=0.0258$), those on ARV therapy yielded higher bacterial loads than those without (Figure 33B). No effect on the bacterial load was observed between participants on or not on ARVs when assessed by unsupplemented MPNs and CFU assay, as shown in Figure 33B.

HIV-positive index participants stratified by CD4 count (<200 and > 200 cells/µL) showed no differences between the 2 groups (Figure 33C). The MGIT TTP of HIV-positive and HIV-negative index participants showed no significant differences (Figure 33D). Overall, the data indicate that while HIV status influences bacterial load as detected by CF-supplemented and unsupplemented MPN assays, it does not significantly impact bacterial load measured by CFU assay or MGIT TTP. Given the complex interplay between HIV and TB, it is essential to further investigate whether these bacterial load differences translate into variations in TB transmission dynamics between HIV-positive and HIV-negative index participants, particularly in relation to their HHCs.

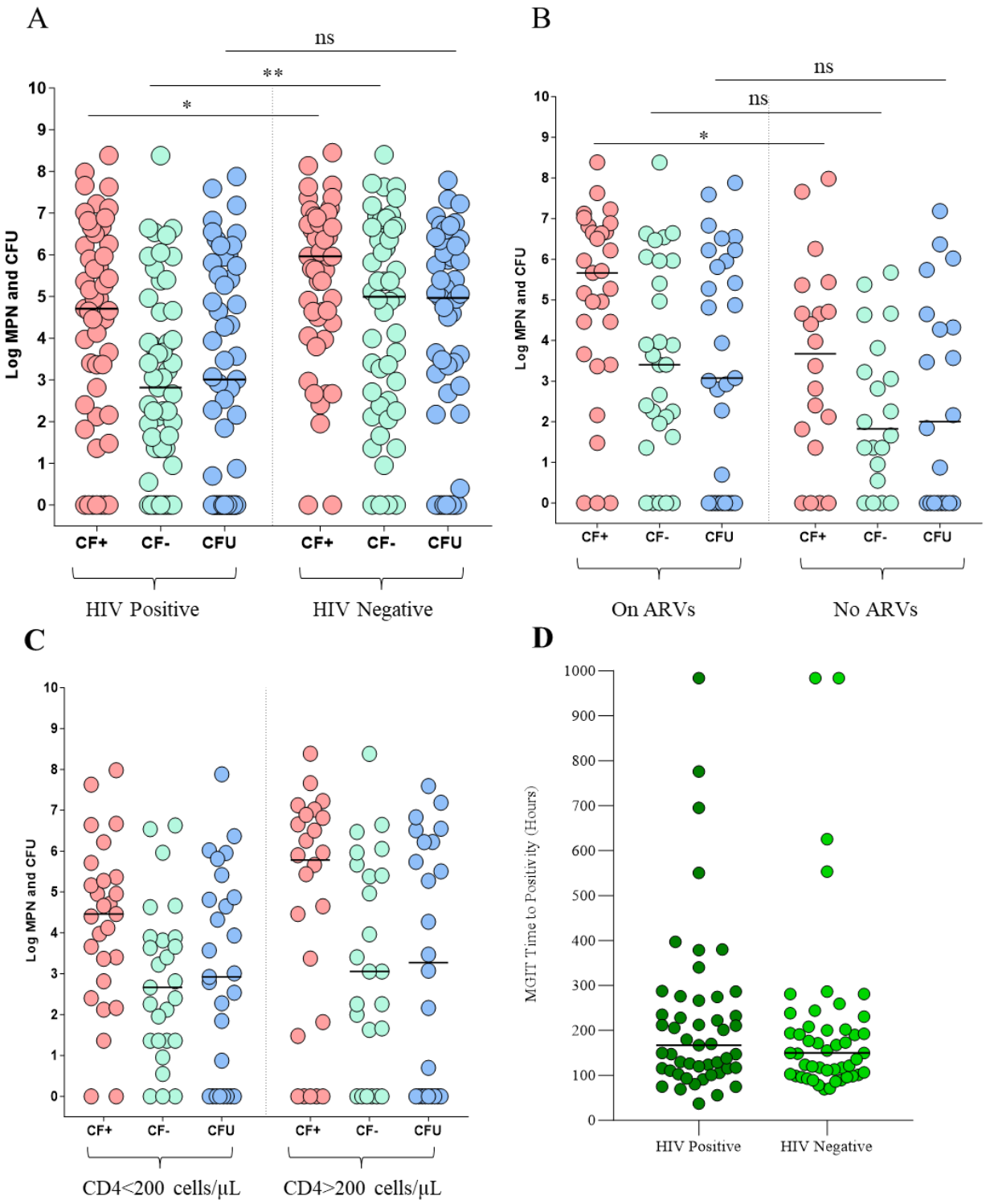


Figure 33: Comparison of Bacterial load in HIV-Positive and HIV-Negative Index Participants. Scatter plot showing the median bacterial load in by CF-supplemented MPNs (pink), CF-unsupplemented MPNs (green) and CFUs (blue). (A) HIV positive versus HIV negative indexes. Significant differences are observed in the CF-supplemented MPNs (p=0.0260) and unsupplemented MPNs (p=0.0015). (B) Indexes receiving and not receiving ARVs. A significant difference can be observed in the CF-supplemented MPNs (p= 0.0258). (C) HIV positive indexes stratified by CD4 count (< 200 versus > 200 cells/µL). No significant differences were observed between the 2 groups. (D) Comparison of HIV positive and HIV negative index participants by MGIT culture showed no significant differences.

4.6.4.3 Assessment of Differences in TB Transmission Between HIV-Positive and HIV-Negative Index Participants with and without an *Mtb*-Positive HHC

The interaction between HIV and TB is complex, and individuals with HIV-TB coinfection may not inherently transmit TB more easily compared to HIV-negative TB patients. However, factors such as higher bacillary loads in HIV-positive index cases can increase transmissibility. Also, HIV-positive individuals on ARVs have a stronger immune system than those that remain untreated for HIV. Therefore, investigating the differences between HIV-positive and HIV-negative individuals and their propensity to transmit TB is important in understanding transmission dynamics.

In our study cohort, an equal number of index participants (n=26) with a positive HHC were identified in both the HIV-positive and HIV-negative group (Figure 34). In the HIV-positive group, 16 out of 26 (61.5%) participants were on ARVs. In comparison, 52% of index participants without a positive HHC, were HIV-positive, with 13 out of 25 (52%) on ARVs.

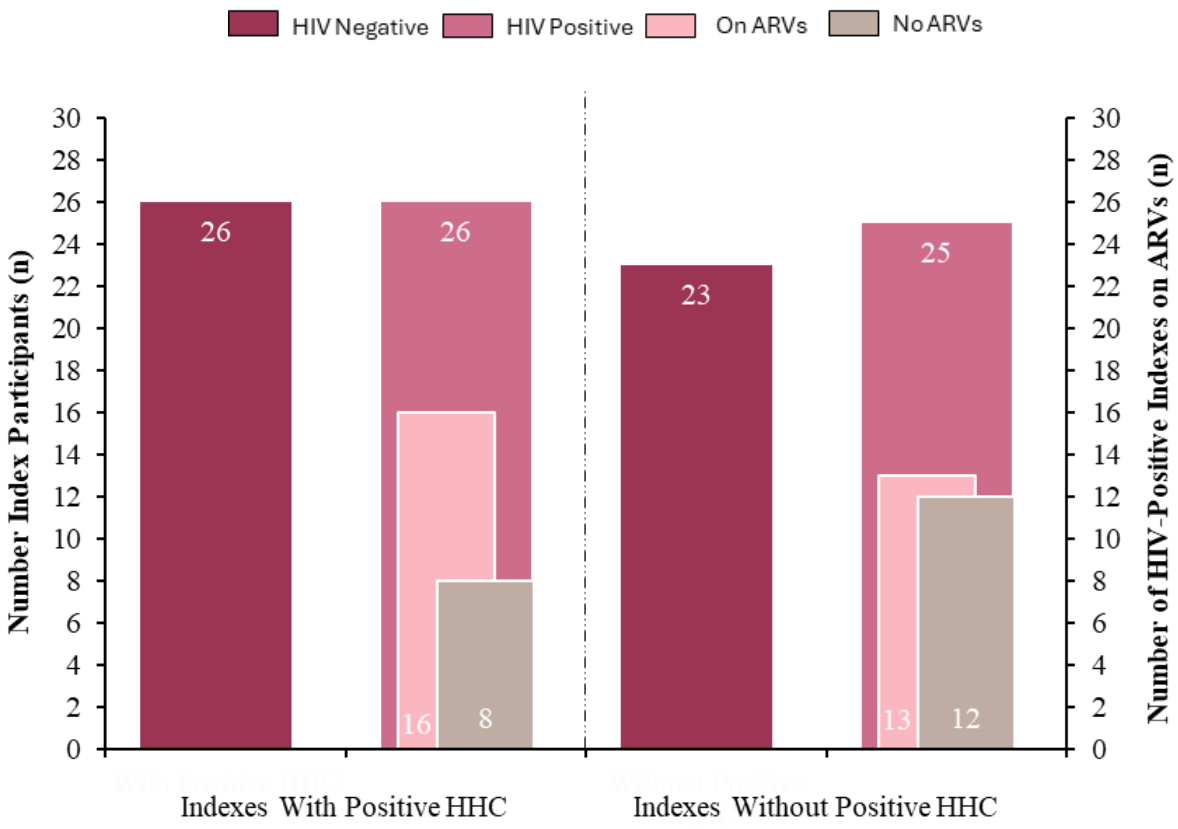


Figure 34: HIV Status of Index Participants in Households with Possible Transmission Events. Column graph showing the number of HIV negative and HIV positive index participants in households with possible transmission events. HIV Positive index participants on ARVs are shown nested in the HIV positive column. The ARV status of 2 HIV-positive participants with *Mtb*-positive HHCs were unknown.

Studies have shown that HIV-positive individuals on effective ARVs exhibit similar immune responses as HIV-negative individuals (80, 116). Based on this, there are more HIV-negative individuals and HIV-positive individuals on ARVs in the group with an *Mtb*-positive HHC (81%) compared to the group of indexes without an *Mtb*-positive HHC (75%). However, the difference between the 2 groups is not significant enough to conclude that transmissibility is increased in individuals without HIV infection and in individuals with HIV infection but on ARV treatment.

4.6.4.4 Assessment of whether Immunosuppression of HHCs Increases the Risk of Acquiring TB Infection

Upon infection, the individual's immune system plays a significant role in determining whether TB infection will progress from the latent stage to active TB disease. Several factors influence an individual's immune response, including HIV infection and comorbidities such as diabetes and hypertension. To fully understand transmission dynamics within a household, it is important to assess the health status of HHCs to identify key contributing factors to an increased risk of infection. Herein, the health status - including HIV status, BMI and the presence of comorbid conditions, together with lifestyle activities of HHCs with positive *Mtb* detection were analysed as tabulated in Table 7.

HIV weakens the immune system, targeting the CD4+ T cells, thereby, predisposing individuals to increased risk of TB infection. HIV-positive individuals often present atypically, with no evidence of cavitation and less pronounced respiratory symptoms, which may complicate diagnosis and treatment of TB. A comparison of the demographics and exposure to the index participants of the 112 HHCs from which a positive *Mtb* result was obtained and the 135 HHCs that were *Mtb*-negative are also shown in Table 7. Majority of the HHCs were female in both the *Mtb*-positive (63.4%) and *Mtb*-negative (68.9%) groups, and 23.2% and 24.4% of individuals in the 2 groups were HIV-positive, respectively. This suggests that although HIV infection is a known risk factor for TB, majority of HHCs that acquired infection were HIV-negative and therefore, did not play a significant role in the increased risk of TB infection within this cohort. There are also no significant differences between the HIV status of index participants with *Mtb*-positive HHCs and those without *Mtb*-positive HHCs.

BMI is another key indicator of immune health, with malnutrition and obesity both linked to a weakened immune response. The median BMI of HHCs was 22.1 kg/m^2 in the *Mtb*-positive group and 21.9 kg/m^2 in the *Mtb*-negative group. In the *Mtb*-positive group, 35.1% of participants were within the healthy BMI range of 18.5-24.9, while 64.8% fell into the

unhealthy BMI range – 30.6% were underweight and 34.2% were overweight. In the *Mtb*-negative group, 24.6% of participants were within the healthy BMI range, while 60.4% were in the unhealthy BMI range – 24.6% were underweight and 35.8% were overweight. Statistical analysis using chi-squared testing showed a significant difference in the BMI distribution between the 2 groups (p=0.0041), suggesting that BMI may be a contributing factor to *Mtb* infection risk in this cohort. However, BMI alone is not a definitive determinant of infection risk.

Table 7: Comparison of Demographics and Exposure to Infected Index Participants Between *Mtb*-Positive and *Mtb*-Negative HHCs

	HHCs Positive *Mtb* result (n=112)	HHCs Negative *Mtb* result (n=135)	Statistical Significance
Variables, n=247			
Demographics			
Sex			
Women, n (%)	71 (63.4)	93 (68.9)	ns
Men, n (%)	41 (36.6)	42 (31.1)	ns
Age, years, median (IQR)	26 (13 – 46)	25 (14 – 46)	ns
BMI, kg/m^2, median (IQR)	22.1 (17.8 – 27.3)	21.9 (18.5 – 29.8)	**
Under Weight (BMI < 18.5kg/m^2), n (%)	34 (30.6)	33 (24.6)	ns
Normal (BMI 18.5 - 24.9kg/m^2), n (%)	39 (35.1)	53 (39.6)	ns
Overweight (BMI > 24.9kg/m^2), n (%)	38 (34.2)	48 (35.8)	ns
HIV Status at Baseline, n (%)			
Positive	26 (23.2)	33 (24.4)	ns
Negative	86 (76.8)	102 (75.6)	ns
Comorbid Conditions at Any Timepoint			
Diabetes, n (%)	2 (1.8)	1 (0.7)	ns
Hypertension, n (%)	15 (13.4)	26 (19.3)	ns
Other medical conditions, n (%)$^{\beta}$	6 (5.4)	9 (6.7)	ns
Smoker			
Current, n (%)	22 (19.6)	34 (25.2)	ns
Former, n (%)	3 (2.7)	2 (1.5)	ns
Exposure to index case			
Sleep in the same room as index, n (%)	29 (25.9)	34 (25.2)	ns
Time spent in the same room as index case per night, hours, median (IQR)	8 (8 – 10)	8 (8 – 9)	ns

Share a bed with index, n (%)	16 (14.3)	22 (16.3)	ns
Duration spent together per week, hours			
Eat meals together, hours, median (IQR)	6 (4 – 9)	5 (3 – 7)	ns
Watch TV, or play video games, hours, median (IQR)	15 (7 – 21)	10 (5 – 20)	**
Drink alcohol together, hours, median (IQR)	7 (4.8 – 7.8)	7.5 (2.5 – 9.3)	ns
Prepare meals together, hours, median (IQR)	6 (2 – 8)	5 (1 – 7)	ns
Playing board or card games, hours, median (IQR)	9 (5.5 – 14)	5 (4.3 – 5.8)	****

[β] Epilepsy, high cholesterol, asthma, renal failure
** $p<0.01$, **** $p<0.0001$

Less than 2% of both groups were diabetic, while 15 (13.4%) of the *Mtb*-positive group and 26 (19.3%) of the *Mtb*-negative group had hypertension. Six (5.4%) of the *Mtb*-positive group and 9 (6.7%) of the *Mtb*-negative group had other medical conditions. Twenty-two (19.6%) of the *Mtb*-positive group were current smokers, while 34 (25.2%) of the *Mtb*-negative group were current smokers. Overall, 81% of participants had no comorbid diabetes or hypertension (*Mtb*-positive group - 95 (84.8%); *Mtb*-negative group - 105 (77.8%)), indicating that comorbid conditions do not significantly impact the risk of infection in HHCs.

Exposure time of HHCs to the infected index participant was also assessed as a risk factor for infection. In the *Mtb*-positive group, 25.9% of HHCs sleep in the same room as the index, while 25.2% in the *Mtb*-negative group do the same. Additionally, 14.3% and 16.3% of HHCs share a bed with the index participant, respectively. The median time spent together in the same room as the index participant was 8 hours in both groups. Further analysis of time spent together participating in various household activities showed variable durations between the two groups (Table 7). The median time spent eating meals together was 6 hours per week in the *Mtb*-positive group, compared to 5 hours in the *Mtb*-negative group. Watching TV or playing video games together had the highest median time, with 15 hours per week in the *Mtb*-positive group and 10 hours in the *Mtb*-negative group, with a statistically significant difference of $p=0.0075$. Similarly, the time spent playing board or card games per week was significantly higher in the *Mtb*-positive group (9 versus 5 hours, $p<0.0001$). The median time spent preparing meals together per week was marginally higher in the *Mtb*-positive group (6 versus 5 hours), as was the median time spent drinking alcohol together (7 versus 7.5 hours). These findings suggest

that increased exposure to the index participant may contribute to a higher risk of transmission in the *Mtb*-positive group.

4.6.5 Subclinical TB disease Within Households

Next, we assessed the role of WGS in understanding subclinical TB disease, characterized by the presence of *Mtb* in an individual who does not exhibit the 4 typical symptoms of active TB: fever, persistent cough, weight loss or night sweats. Despite the absence of these overt symptoms, individuals with subclinical TB can still transmit the bacteria to others, contributing to the spread of the disease. This makes subclinical TB a significant public health concern, as it often goes undiagnosed and untreated, potentially leading to a substantial number of new undetected TB cases. The subjectivity of symptoms is a significant downside in TB diagnostics screening. Each individual's perception of symptom severity and their decision to report these symptoms can vary greatly. In this study, 69 of 112 (61.6%) HHCs with a positive *Mtb* result, reported an absence of symptoms across the 16-month timeline.

Figure 35 is an illustrative depiction of 16 HHCs that either (1) progressed from *Mtb*-negative (black circles) to *Mtb*-positive and symptomatic (purple circles), or (2) regressed from *Mtb*-positive and symptomatic (teal) to *Mtb*-positive and asymptomatic (pink) over the course of 16 months. Symptomatic was determined by having reported at least 1 of the following symptoms: cough, fatigue, weight loss, night sweats, pleuritic chest pains.

At baseline, 7 of the 16 (43.8%) HHCs tested positive for *Mtb*, despite being asymptomatic. These HHCs later developed symptoms at subsequent timepoints. Four additional HHCs tested negative for *Mtb* at baseline but became positive at later timepoints, initially without symptoms, and later developed symptoms. Two HHCs presented with symptoms at baseline but tested negative for *Mtb*, later becoming *Mtb*-positive. Additionally, three HHCs tested positive for *Mtb* and exhibited symptoms at the month 2 visit. These observations highlight the importance of early and accurate detection of subclinical TB disease to allow for timely intervention which can prevent the progression of subclinical disease to active disease and reduce the risk of transmission to others.

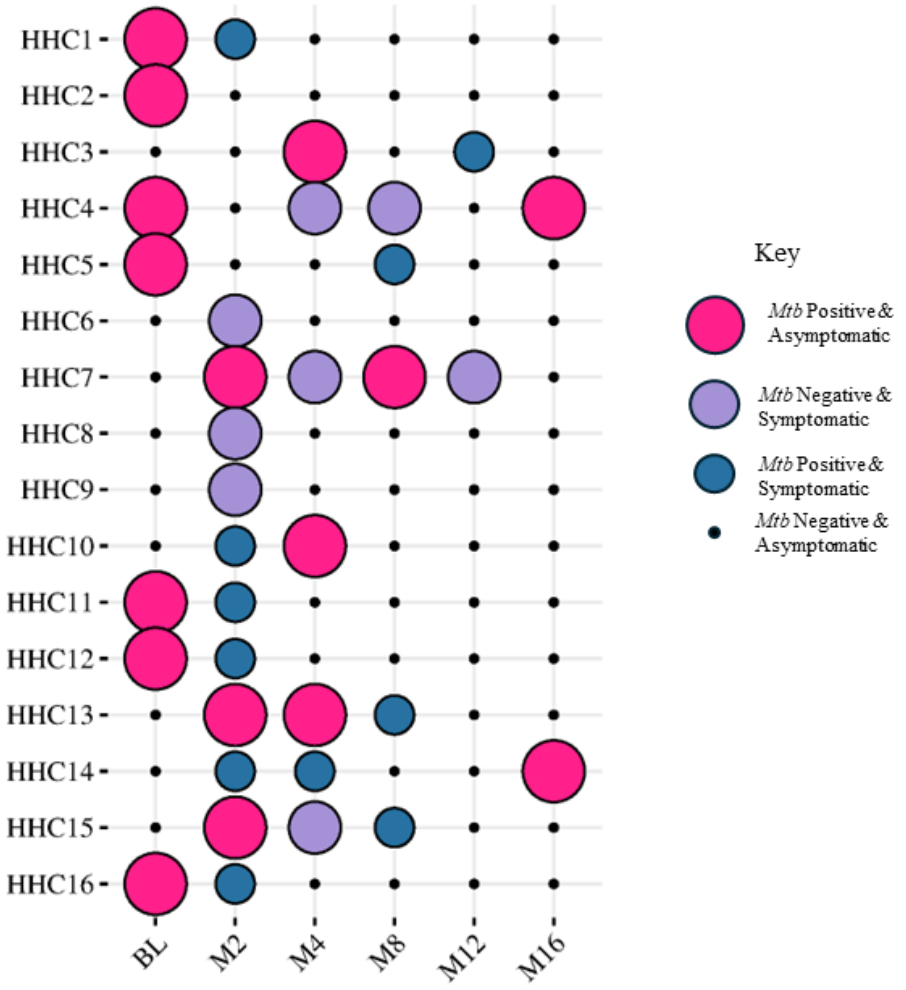


Figure 35: Longitudinal Trends of *Mtb* Positivity and Symptom Reporting in HHCs. Balloon plot showing the progression of *Mtb*-positive HHCs to symptomatic disease over the 16-month time period (BL = baseline, M2-M16 = months post baseline). The large, pink circles indicate *Mtb* positivity and no reported symptoms, the medium, purple circles signify *Mtb*-negative results but at least 1 reported symptom (either cough, weight loss, fatigue, pleuritic chest pains or night sweats), while the small teal circles indicate *Mtb* positivity and symptomatic individuals. The small black circles denote absence of TB detection and no reported symptoms.

5. Discussion

The gold standard for routine *Mtb* diagnosis is still on solid and liquid media. These diagnostic methods require the bacteria to exist in an actively replicating state thereby, failing to detect bacteria in the viable but non-culturable state. In 1996, Paul and colleagues showed that after phagocytosis by human macrophages within the lung granuloma, *Mtb*'s ability to grow on solid media becomes impaired (117, 118). This non-culturable or dormant state of *Mtb* is a result of metabolic remodelling induced by several factors during phagocytosis, including stress

conditions such as host immune pressure, oxygen and carbon starvation, and antibiotic exposure. More than a decade ago *rpf*s were identified and implicated in the stimulation of non-culturable or dormant *Mtb.* The role of *rpfs* was first reported in a cohort of 25 patients, where *Mtb* was detected in 80% of the individuals only after stimulation with *rpf*-containing CF (119). Since the discovery of this phenomenon, several studies have investigated the detection of unculturable bacteria in clinical samples (42). The same group later studied the prevalence of DCTB in TB diagnosed patients undergoing TB treatment and found that 74% still harboured DCTB populations after treatment completion (44). Additionally, in DR-TB cases, CF-supplementation was required to detect highest levels of DCTB, which remained prevalent despite the extended treatment period (45). Given the significance of DCTB, there is an urgent need to develop these methodologies as improved diagnostics capable of identifying all mycobacterial populations during infection.

5.1 Optimization of Experimental Techniques and Study Design

We first needed to streamline culture procedures to ensure all the techniques applied were able to deliver robust, reliable and accurate results. As a starting point, the spot and overnight sputum samples were individually processed to assess the quantum of bacteria detected. As expected, the overnight sample was marginally superior to the spot sample, but the difference was not significant hence, the two samples were pooled and processed as a single sample. Similarly, a study comparing the MGIT TTP for 3 sputum sample types from TB patients which included: (1) overnight pooled sputum samples collected over the course of the night, (2) single morning sputum samples and (3) spot sputum samples which were collected at any time during the day. showed that the single pooled overnight sputum sample was the superior sample for faster detection and bacterial yield (120).

5.2 Longitudinal Prevalence of Culturable and Non-Culturable *Mtb* in Index Participants

To assess the prevalence of *Mtb* infection in index participants after treatment, the bacterial load was evaluated at baseline, during treatment and after treatment for a total period of 16 months. At baseline, 94% of index participants were *Mtb*-positive by routine culture or MPN/CFU assays. CF supplementation significantly increased bacterial yield compared to CF-independent conditions ($p=0.0005$) and CFU assays ($p=0.0004$). Additionally, CF-dependent DCTB populations were significantly higher than bacteria identified in the absence of CF (CF-independent) ($p<0.0001$). A significant reduction in bacterial load was observed at months 2 and 4 in CF-supplemented MPN assays ($p<0.0001$), while MGIT TTP increased significantly

from month 2 to month 16 compared to baseline ($p<0.0001$), consistent with the expected treatment effect. No significant differences were observed between month 4 (pre-treatment completion) and post-treatment timepoints (months 8, 12, and 16). However, the detection of MGIT-positive cases post-treatment suggests possible treatment failure, reinfection, or bacterial persistence due to drug resistance.

MPN assay assessment showed that 89% of index participants harboured DCTB populations at baseline, indicating that these bacteria were present before treatment and likely transitioned to a dormant or non-culturable state due to stressors. Over the 16-month period, culturable bacteria demonstrated a declining trend (slope =-0.5825), as did CF-independent non-culturable bacteria (slope = -0.145). Conversely, CF-supplemented non-culturable bacteria increased over time (slope =0.4235), reinforcing the importance of CF supplementation in detecting *Mtb* populations.

This analysis allowed us to investigate the relationship between bacterial load in index participants and their propensity to transmit the disease to their respective HHCs. These data corroborate well with findings from Peters et al., in which treatment naïve TB participants were assessed for their response to treatment using MPN/CFU assays (44). The study showed that at baseline, CF-supplemented MPNs yielded higher bacterial loads ($p=0.0047$) compared to unsupplemented MPNs. Herein, 36 index cases required stimulation by CF for DCTB populations to emerge, 3 cases emerged in CF-unsupplemented media, while 43 cases emerged in both supplemented and unsupplemented MPNs (44). Over the 14 day follow up period, the participants displayed different treatment responses which were classified into four categories: (1) classic bi-phasic bacterial clearance, (2) early non-responders with slower clearance, (3) paradoxical worsening with an increase in bacterial count upon treatment initiation; and (4) non-responders with no change in bacterial load (44). There is clear evidence that even post treatment, the DCTB population persists as these bacteria most likely did not respond to the treatment regimen, or the first line drugs were ineffective at clearing the infection.

While standard TB chemotherapy is expected to eradicate all susceptible *Mtb* bacteria within the 6-month treatment period, this is often not the case. This treatment failure has mainly been attributed to the presence of the DCTB populations (41, 42, 121). As treatment progresses, the amount of viable bacteria detectable by CFU decreases, while CF-dependent populations detected by CF-supplemented MPNs increase (41), aligning with the trend observed in our cohort.

5.2.1 *Mtb* Drug Resistance and Persistence in Index Participants

Bacteria may adopt a non-culturable state due to various stressors, including stress induced by chemotherapy, host immune response, host health, comorbid diseases and the development of mutations related to drug resistance. To understand why treatment was not fully effective in these participants, we investigated whether they had developed mutations in the genes responsible for drug resistance. Of the 26 index participants with persisting bacteria post treatment, only 3 (11.5%) showed mutations in the genes coding for drug resistance. An additional 3 (11.5%) converted from drug susceptible to DR, however, these participants showed no presence of *Mtb* post treatment completion, indicating treatment success. A mutation in a bacterial drug resistance gene results in a single nucleotide change, leading to the substitution of one amino acid for another in the encoded protein. This change can alter protein function, often diminishing the effectiveness of antibiotics. The stress induced within the host during TB treatment likely contributed to the development of these mutations, thereby, creating resistance to TB drugs (see Table S4 for a list of mutations, corresponding amino acid changes, and their relation to drug resistance in the 6 participants). This data suggests that the persistence of bacterial populations post-treatment may not only be due to the emergence of drug resistance but could also be a result of other stressors forcing the bacteria into a non-replicating state.

5.2.2 TB-HIV Coinfection and Bacterial Load

Next, we examined the effect of TB-HIV coinfection on bacterial load and the longitudinal persistence of bacteria. Of the 26 index participants with persistent bacteria post treatment, 11 (42%) were HIV-positive. Comparison of transmissibility of TB from HIV-TB coinfected individuals show differing results, where som9e studies indicate HIV/TB coinfection increases transmissibility due to higher bacterial loads, while others find no difference (122, 123). To further explore the impact of HIV co-infection, we compared our findings to those of Chengalroyen et al. (2016), who assessed for *Mtb* in sputum with and without CF. The findings of this study aligned with our study, showing that HIV-positive individuals yielded lower levels of CF-supplemented ($p=0.0260$) and unsupplemented ($p=0.0015$) MPN values compared to HIV-negative individuals (42). However, in the study by Chengalroyen and colleagues showed a significant difference in bacterial load by the CFU assay between HIV-positive and HIV-negative index participants, while in our cohort there was no difference. Further stratification of our study participants by CD4 counts (<200 CD4 cells/µL) showed no differences in bacterial loads as assessed by the MPN and CFU assays whilst the study by Chengalroyen and colleagues showed significantly lower recovery of bacteria. However, the HIV status of the index participants did not influence the TTP in the MGIT assay in either study.

Together, these findings highlight the persistence of non-culturable *Mtb* after treatment completion and host factors, such as HIV status, influence bacterial culturability. These persistent bacteria have major implications for TB transmission and relapse disease and therefore, a deeper understanding of this population of bacteria is required for improved TB diagnostics and more importantly their role in TB transmission within HHCs exposed to an infected individual within a household. Hence, we explored the utility of culturomics and molecular techniques to improve detection of all possible *Mtb* populations in the HHCs of infected index participants.

5.3 Culturomics and Molecular Techniques to Identify *Mtb* infection in HHCs

A major factor contributing to the TB epidemic is the population of undiagnosed individuals that subsequently remain untreated due to the limitations of routine diagnostic tests. This population continues to contribute to the spread of TB. Detecting the DCTB population has added great value in addressing this diagnostic gap. In household settings, MPN/CFU assays can identify a subset of the population that would otherwise have been missed by routine diagnostic methods. By leveraging the enhanced detection capabilities of MPN/CFU assays, researchers can obtain a more accurate estimate of TB prevalence within households. This will allow for a comprehensive assessment of transmission dynamics within the household and more importantly, the identification of both symptomatic and asymptomatic household carriers, enabling early initiation of treatment to break the chain of transmission. MPN/CFU assays can further identify *Mtb* in patients despite having completed TB treatment regimens and are considered clinically cured, possibly continuing to transmit the disease.

The primary aim of this study was to identify missing cases of *Mtb* infection within HHCs using culturomics and molecular assays and to map transmission links within these individuals. We first assessed the prevalence of *Mtb* in HHCs at baseline to determine the extent of disease spread that had already occurred prior to inception of the study, either within the household or within the community.

Several studies highlight the underestimation of household TB burden. A systematic review evaluating TB prevalence in South African households, measured by routine diagnostic assays such as GeneXpert, MGIT culture and smear microscopy reported rates below 10% between the different cohorts, depending on the methodology applied for detecting *Mtb* in sputum samples (124, 125). In a study analysing sputum using smear microscopy and MGIT from 282 HHCs from 130 households of TB diagnosed index participants, reported a 3.9% TB prevalence (126). Martinson and colleagues (2021) used GeneXpert and MGIT culture on 2166 HHCs and

reported an overall TB prevalence of 3.2%. In another study, Kigozi et al., 2019 conducted an investigation for HHC TB wherein, HHCs of patients with pulmonary TB were screened for TB infection using standard diagnostic testing. The data from this study cohort showed a 6.6% detection rate of new TB cases (127).

In contrast, our study identified microbiologically confirmed TB in 112 (45%) HHCs over the 16 months. This substantial increase in *Mtb*-positive cases may be attributed to our study design wherein, all HHCs were subjected to culture testing, regardless of symptoms. Similarly, a study investigating TB transmission dynamics within households over 24 months showed the presence of stable bacterial loads in participants post-treatment. The extended follow-up underscores the importance of continued monitoring of HHCs even after the completion of the standard treatment regimen in index participants. These findings emphasize the need for long-term surveillance and possibly extending preventive measures for HHCs to mitigate the risk of TB resurgence (128).

The MPN assay is emerging as a novel and useful approach for identifying missing cases of TB within households. Our data showed that at baseline, the MPN assay increases the detection of *Mtb* in HHCs from 11% to 18%, with the CFU assay adding another 2%. Over the 16-month period, the MPN assay identified 55 cases of *Mtb* positivity in HHCs that were missed by the conventional MGIT culture assay.

The second objective of this study was to assess whether molecular based assay, specifically qPCR, could serve as early predictors for the progression of TB disease. Additionally, we aimed to determine whether qPCR could detect *Mtb* DNA in the sputum of HHCs that were unculturable. While several studies have explored the potential of qPCR-based methods for TB diagnosis, most have focused on evaluating sensitivity and specificity in individuals with active TB disease, with limited application in the household setting (129-131). A recent study evaluating in-house qPCR detection of TB from tongue swabs in adults, showed that qPCR had a high level of sensitivity and specificity (89.6 % and 96.2%, respectively) compared to GeneXpert (87.4% and 98%) (132). In our study, we used qPCR to detect *Mtb* in sputum samples obtained from HHCs of infected individuals. The HHCs displayed no clinical evidence of TB, however, qPCR detected *Mtb* DNA in an additional 6.9% of individuals that would otherwise remain undetected if using culture-based methods only. Finally, we next sought to map transmission links between HHCs and their respective index participants to distinguish between household and community acquired infections.

5.4 Transmission Dynamics within Households

Our WGS data analysis to determine the number of households transmission events showed that 27 (39%) of the transmission events occurred within the household, with 18 (26%) of these cases detected by MPN/CFU assays. Historically, WGS was not the primary method for inferring transmission in contact tracing; instead, more crude methods such as spoligotyping and RFLP were commonly used. While most household transmission studies primarily focus on individuals with diagnosed TB, fewer studies combine active contact tracing within households with detailed transmission dynamics.

A surveillance study conducted in British Columbia enrolled culture positive children who were strain typed using MIRU-VNTR genotyping to identify clusters between the group. The results were then compared with strain types from a previous study conducted on adult patients within the region. The clustered samples were then analysed by WGS to determine relatedness. WGS showed that household transmission was the source of infection in 16% of enrolled children (133). A similar study conducted in England enrolled participants from the same household as previously reported cases and genotyped each sample by MIRU-VNTR. Herein, 3.9% of cases were reported to have occurred within the household (134). In contrast, our findings are significantly different in comparison to these two studies as we identified transmission in 24 (24%) households within the 100-household cohort. Statistical analysis confirmed that the increase in detection in our study was significant ($p < 0.0001$). These differential observations are likely due to methodological differences between studies as we applied several culture conditions to enhance the detection of genetically diverse *Mtb* populations within samples from a single individual. By utilizing varied culture conditions, the likelihood of identifying the full spectrum of bacterial populations within samples increases substantially, allowing for the detection of transmission events that may be overlooked when relying solely on standard culture methods.

Genetic heterogeneity also plays a crucial role in household transmission dynamics, as variations in *Mtb* genotypes influence bacterial behaviour, including response to treatment, virulence, transmissibility and the ability to persist in dormant states. This genetic heterogeneity extends to the variations of strains found within a host, which can have significant implications for both the diagnosis of the strain and the ability for the drug to eradicate all bacterial populations. Understanding these variations is essential for improving diagnostic accuracy and treatment strategies, as certain strains may be more resistant to therapeutic interventions or

capable of evading detection under standard diagnostic methods. A recent study which investigated the differentially culturable state (DCS) of different clinical *Mtb* genotype families showed that *Mtb* strains exhibit distinct propensities to adopt a DCS characterized by a reversible state of dormancy where the bacteria are viable but not readily culturable under standard laboratory conditions (135). Moreover, the study highlighted that different *Mtb* genotype families have variable dependency on CF-supplementation for resuscitation from the DCS. For example, LAM (lineage 4) strains required CF, while Beijing (lineage 2) strains did not. In our study, we observed the phenomenon described by Gordhan et al., within HH32, wherein, a mixed infection of LAM and Beijing were detected in the MGIT culture, CFU assay and the first dilutions of the MPN assay, containing the most concentrated cells. However, the serial dilution of the sputum in the MPN assay facilitated the separation of the mixed infection. LAM, which requires CF, was recovered in the diluted CF-supplemented MPN well, while Beijing, which does not require CF, was separated in the diluted CF-unsupplemented MPN well. Collectively, these findings emphasize the importance of incorporating multiple culture conditions to better understand *Mtb* transmission dynamics within households. The detection of DCTB populations in this study clearly uncovered greater genetic variability and strengthened our understanding of transmission dynamics in households with TB confirmed individuals.

5.5 Overall Conclusions

Herein, we demonstrated that the application of a combination of culturomics and molecular methods resulted in higher positivity rates compared to MGIT culture alone. Seven percent of HHCs that were culture negative, were detected by qPCR at baseline. The MPN assay detected additional *Mtb* positivity within HHCs and index participants that would have otherwise remained undetected under routine diagnostic assays. The detection of TB within HHCs in this cohort at baseline increased from 11% to 20% when a combination of MGIT culture and MPN/CFU assays were used. More importantly, by applying different culture conditions to a single sputum sample, we were able to uncover greater genotypic heterogeneity and map more transmission inferences within the household. Fifty-seven percent of transmission events were identified by MPN/CFU assays that would have been missed with routine diagnostic tests only. These data clearly highlight that application of a combination of culture and molecular techniques allows detection of many more *Mtb* populations in sputum which, is not only useful at the diagnostic level but have major implications for understanding transmission within households.

5.6 Limitations and Prospective Research

Some of the limitations of this study include:

1. Exclusion of participants across the study timeline due to terminations and missed visits reduced the sample size.
2. Due to the limited availability of funding, GeneXpert testing was not included as a diagnostic comparison. GeneXpert analysis may have provided a more direct comparative view of TB detection at clinics.
3. HHCs that become *Mtb*-positive at a later stage could not be followed up for 16 months due to the fixed timeline of the study.
4. Further research into the directionality of transmission inferences within households may provide a different view of transmission dynamics within households. Additionally, mapping transmission of individuals within communities by incorporating their daily habits may contribute to understanding larger transmission networks and highlight possible ways of restricting transmission within the communities.

6. Appendices

6.1 Appendix A – HoTT DCTB Ethics Approval

Human Research Ethics Committee: (Medical)
FWA Registered No IRB 00001223

SECRETARIAT: Suite 189, Private Bag x2600, Houghton 2041, South Africa Tel: +27-11-274 9200 Fax: +27-11-274 9281

04 June 2020 **EMAILED**

Miss P Bhana

Regulatory Consultant
Perinatal HIV Research Unit
PO Box 114
Diepkloof
1864

Email: bhanap@phru.co.za

Dear Miss Bhana,

PROTOCOL: HOTT-DCTB - ASSESSING HOUSEHOLD TUBERCULOSIS TRANSMISSION USING DETECTION OF DIFFERENTIALLY CULTURABLE TUBERCLE BACTERIA (HOTT-DCTB)

ETHICS REFERENCE NO: 190910 **DAIDS Risk/Benefit Category: 45 CFR 46.404**

RE : NOTIFICATION OF RESUMPTION OF STUDY ACTIVITIES

We acknowledge receipt of your letter dated 27 May 2020 with the following documentation pertaining to the above-captioned trial.

The Chairperson has noted and approved the following:

"In anticipation of the national lockdown, we suspended all HOTT DCTB study activities. The study was in the preparation phase prior to lockdown. No participants have yet been enrolled. Since the president's announcement that South Africa will be transitioning to Level 4 of the lockdown, we would like to inform the committee of our intention of restart study activities and begin participant enrolment on 10 June 2020.

The following steps have been taken to ensure the safety of our teams and participants:
* The study team has been divided into two groups and will work on a rotational schedule. Members of the different groups will not interact with each other to limit contact and potential spread of COVID-19.
* Study personnel who have respiratory symptoms or fever will not come to work and all staff are being screened daily for symptoms.
* Personal Protective Equipment has been purchased to be used by staff when interacting with participants. All staff are required to wear cloth masks at work and N95 or FFP masks when interacting with patients.
* Study team members have been trained on the severity of COVID-19 and PPE best practices.
* Our teams will follow national protocol on number of personnel in a single vehicle and restriction of movement in high risk areas.
* As this study is largerly household-based all staff members will be provided with adequate PPE when visits households. Interactions at the household level will be limited and where possible study procedures will be performed outside. Household members will be offerred screening for COVID-19 and referred for appropriate testing if they report symptoms.

Additionally, the HOTT DCTB study processes will not interfere with hospital or clinic responses to COVID-19 and will not utilise any resources that could place the health system under undue strain."

The above has been noted for the Ethics Committee information and records.

KINDLY FORWARD TO THE RELEVANT INVESTIGATORS / CRA / SPONSOR / STUDY CO-ORDINATORS - WHERE APPLICABLE

Regards,

Thashinreddy

MR THASHIN REDDY

For and on behalf of the Human Research Ethics Committee: (Medical)

6.2 Appendix B – Media

Table S1: Culturing Media

Media	Reagents	Method
Tween®80 (25%)	10 mL Tween80	Tween80 was dissolved in 40 mL of distilled water and filter sterilized using a 0.2 µM filter and 50 mL syringe.
7H9	4.7 g of 7H9 BD Middlebrook 100 mL BD OADC, 2 mL Glycerol 2 mL 25% Tween®80	7H9 media was dissolved in 900 mL of distilled water, with glycerol. Once autoclaved, the temperature of the media was allowed to reduce to 55^{o}C before adding Tween®80 and OADC, followed by sterilising using a 0.2 µM stericup filter.
7H11	20.5 g of 7H11 Sigma Middlebrook media 100 mL OADC 5 mL Glycerol.	The 7H11 media was dissolved in 900 mL of distilled water and glycerol. After autoclaving and cooling, Tween®80 and OADC were incorporated. The media was dispensed into petri dishes in 25 mL aliquots.
PBS (pH 7.4)	PBS Tablets	Five tablets dissolved in 1 L of distilled water and autoclaved. Once autoclaved, the pH was corrected to 7.4. The media was filtered in a 0.2 µM stericup filter.

Table S2: Reagents for DNA Extraction using CTAB.

Reagent	Components	Method
TE buffer	0.6057 g of 10 mM Tris-HCl 1.8612 g of 10 mM EDTA	A homogenous solution was prepared using 500 mL of distilled water and autoclaved. Adjust the pH of Tris-HCl to 8 using HCl or NaOH.
CTAB/NaCl	4.1% NaCl, 10% N-cetl-N,N,N-trimetyl ammonium bromide	CTAB was added to 300 mL of distilled water while stirring on a heating block before filter sterilizing using a 0.22 µM filter and a 50 mL syringe.

Table S3: Buffers and Solutions used for Spoligotyping

Buffer/Solution	Reagents	Method
Stock solutions		
10X SSPE	105.19 g sodium chloride 13.7 g sodium di-hydrogen phosphate 3.36 g EDTA Distilled water	All 3 reagents were dissolved in 900 mL of distilled water and adjusted to a pH of 7.4 using NaOH. The volume was made up to 1 L using distilled water.
10% SDS	50 g sodium dodecyl sulphate Distilled water	Dissolved in 500 mL of distilled water and filtered using a 500 mL stericup with a filter of 0.2 µM.
0.5 M Ethylenediaminetetracetic acid (EDTA)	93 g EDTA Distilled water	The powder was homogenized in 500 mL of distilled water.

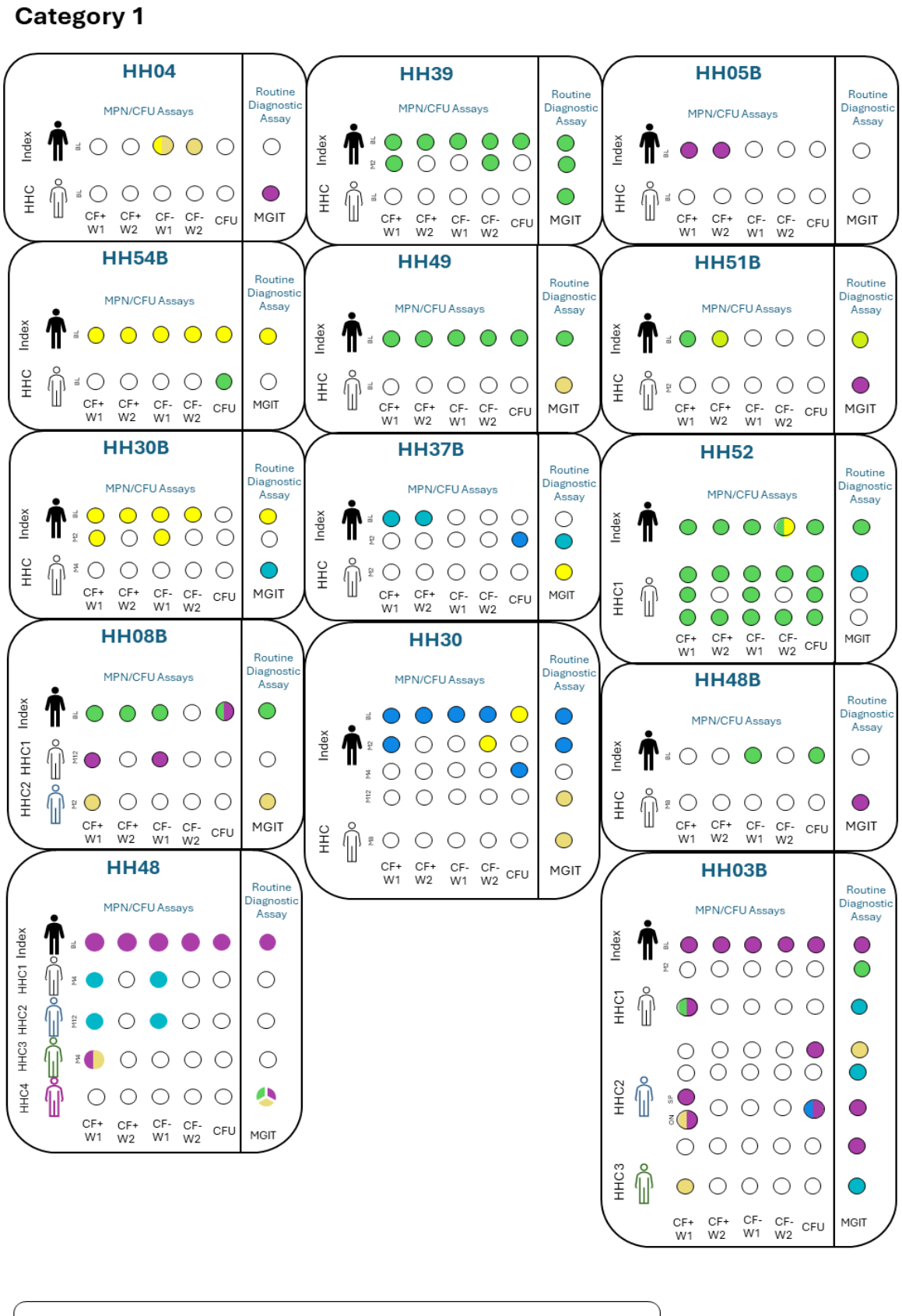
Working solutions		
2x SSPE/0.1%SDS	200 mL 10X SSPE 10 mL 10% SDS Distilled water	The stock solutions were added to a Schott bottle and the volume was made up to 1 L with distilled water
2x SSPE/0.5%SDS	200 mL 10X SSPE 20 mL 10% SDS Distilled water	The stock solutions were added to a Schott bottle and the volume was made up to 1 L with distilled water
1% SDS	10 mL 10% SDS Distilled water	The stock solution was added to a Schott bottle and the volume was made up to 1 L with distilled water
20 mM EDTA	20 mL 0.5M EDTA Distilled water	The stock solution was added to a Schott bottle and the volume was made up to 500 mL with distilled water

6.3 Appendix C – Software and Bioinformatics Tools used for Data Analysis

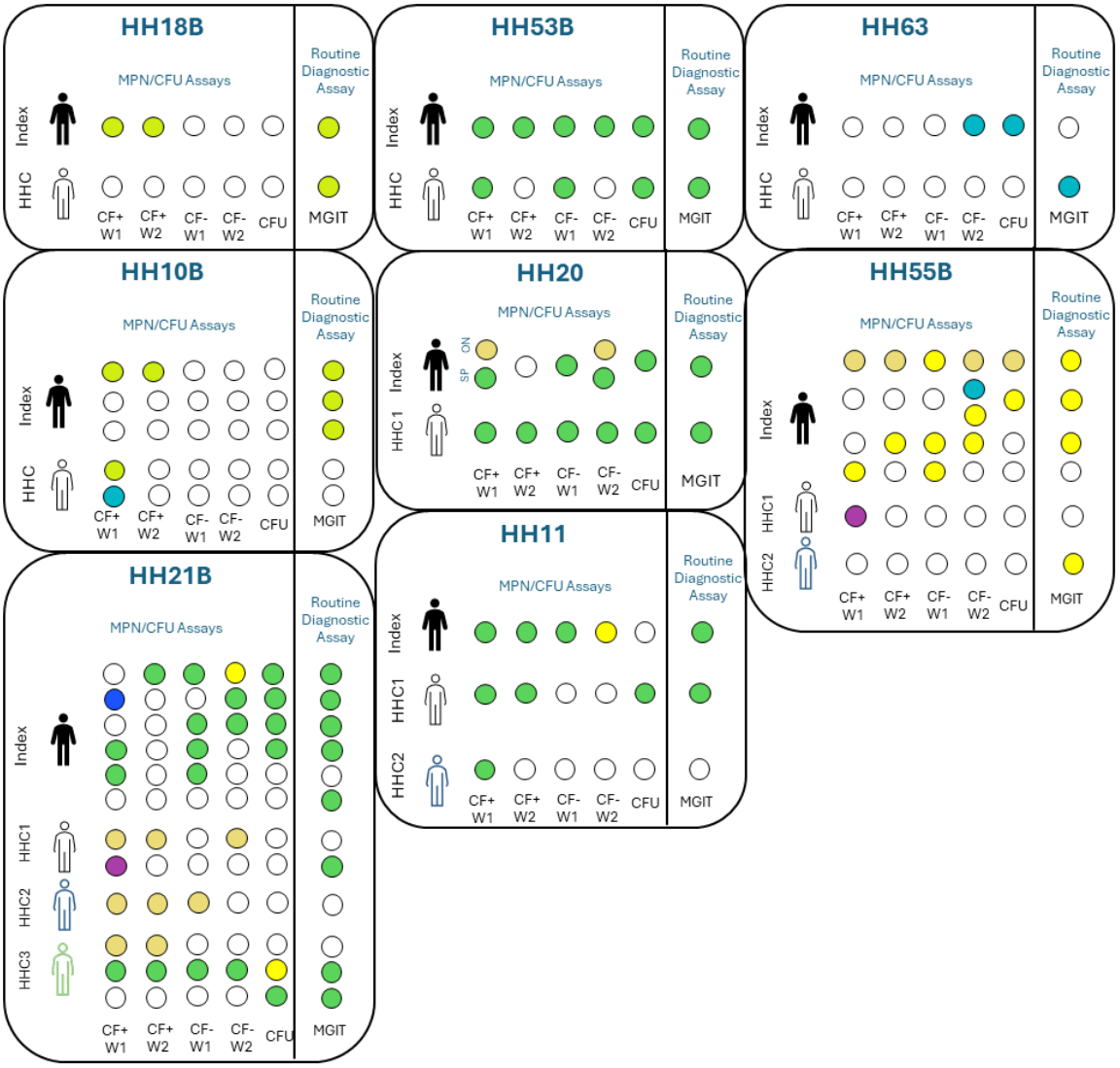
1. **Software**
    a. **GraphPad Prism Software** (Version 10) is a statistical and graphing software used to construct the graphs and determine statistical significance in this study.
    b. **SpotClust** – TB Insight is an open access online bioinformatics tool used to determine the genotypic family of samples based on spoligotyping patterns. (https://tbinsight.cs.rpi.edu/about_spotclust.html)
    c. **MPN Calculator** (version 2) is an Excel program created based on a Bayesian mathematical model to enumerate the organisms from the liquid limiting dilution assay. (http://www.wiwiss.fu.berlin.de/fachbereich/vwl/iso/ehemalige/wilrich/index.html)
    d. **Microsoft Excel** (version 2501, Build 16.0.18429.20132) was used as a database for all raw data. Additionally, Excel was used for graphing and statistical analysis.
    e. **SRplot** is an open access platform for graphing statistical analysis and was used to create the Balloon plot. (https://www.bioinformatics.com.cn/srplot)
2. **Bioinformatics Tools**
    a. **Linux** is an operating system used to run the pipeline codes.
    b. **Kraken2** to assign taxonomic labels to DNA sequences using a k-mer approach.
    c. **PRINSEQ** to filter, reformat and trim sequences for quality control.
    d. **Python** – Programming language used for scripting and facilitating Kraken2 and Phylip execution.
    e. **BWA-MEM** to align DNA sequences to the reference genome.
    f. **Picard 2.8.0** - A toolkit for manipulating high-throughput sequencing data, including duplicate marking, sorting, and indexing BAM files.
    g. **SAMTools** - A suite of programs for interacting with and processing high-throughput sequencing alignment data.
    h. **Pilon** - Improves genome assembly by identifying and correcting sequencing errors.

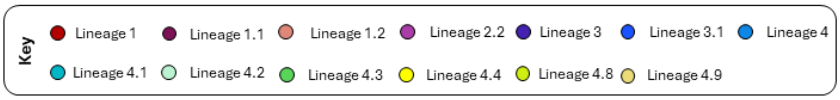
i. **BCFtools** - Tools for variant calling and manipulating VCF and BCF (Binary Call Format) files.
j. **Phylip** - A package for inferring phylogenetic trees from molecular sequence data.
k. **IQ Tree** - Software for phylogenetic inference using maximum likelihood and to generate phylogenetic trees.
l. **SnpEff** - Annotates and predicts the effects of genetic variants on genes and proteins.
m. **QuantTB** - A bioinformatics tool designed for quantifying *Mtb* from sequencing data.

6.4 Appendix D – Supplementary Figures and Tables

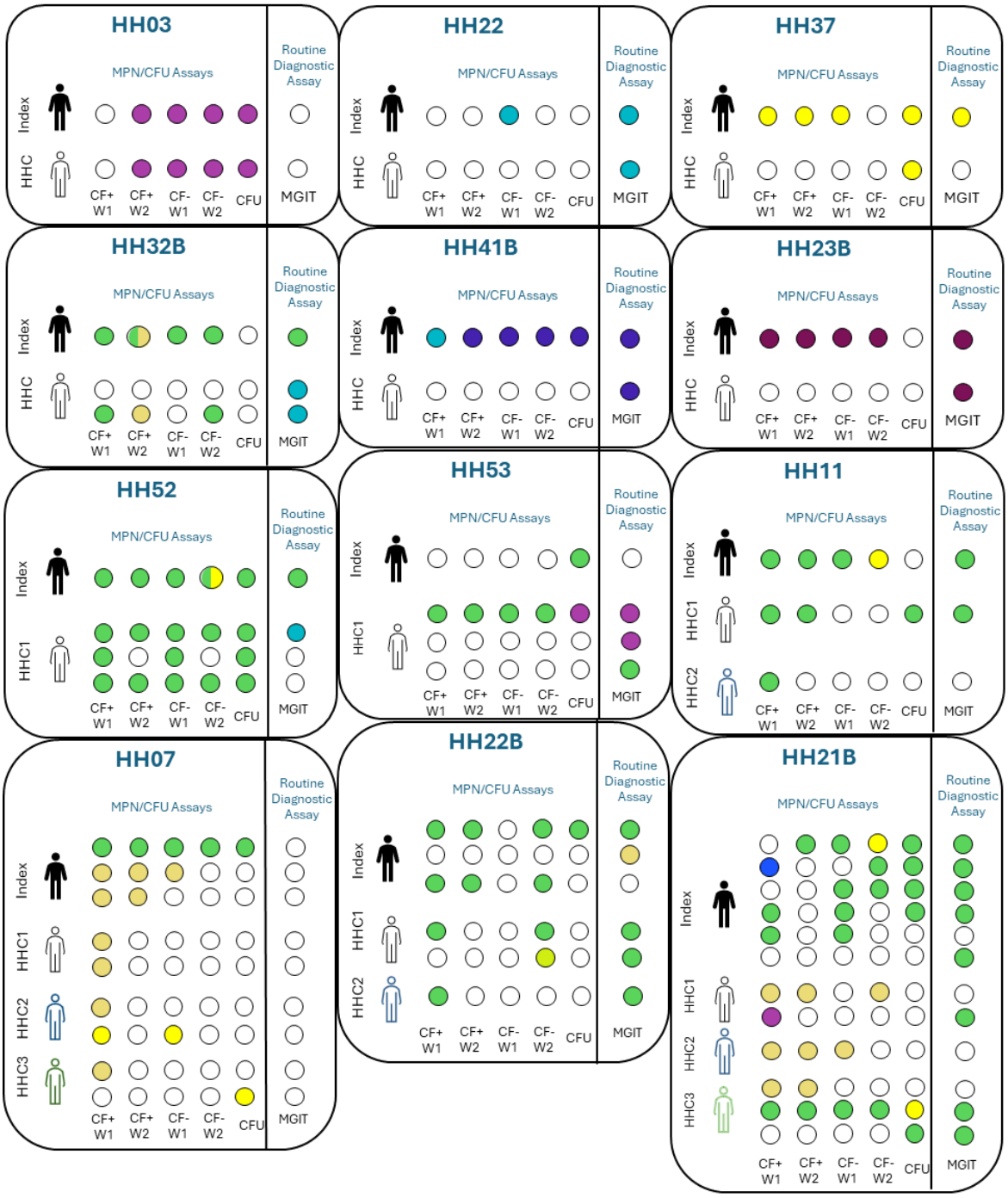


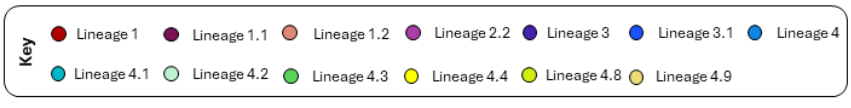
Category 2


HH18B
MPN/CFU Assays
Routine Diagnostic Assay
Index
HHC
CF+ W1
CF+ W2
CF- W1
CF- W2
CFU
MGIT
HH53B
MPN/CFU Assays
Routine Diagnostic Assay
Index
HHC
CF+ W1
CF+ W2
CF- W1
CF- W2
CFU
MGIT
HH63
MPN/CFU Assays
Routine Diagnostic Assay
Index
HHC
CF+ W1
CF+ W2
CF- W1
CF- W2
CFU
MGIT
HH10B
MPN/CFU Assays
Routine Diagnostic Assay
Index
HHC
CF+ W1
CF+ W2
CF- W1
CF- W2
CFU
MGIT
HH20
MPN/CFU Assays
Routine Diagnostic Assay
Index
SP
ON
HHC1
CF+ W1
CF+ W2
CF- W1
CF- W2
CFU
MGIT
HH55B
MPN/CFU Assays
Routine Diagnostic Assay
Index
HHC1
HHC2
CF+ W1
CF+ W2
CF- W1
CF- W2
CFU
MGIT
HH11
MPN/CFU Assays
Routine Diagnostic Assay
Index
HHC1
HHC2
CF+ W1
CF+ W2
CF- W1
CF- W2
CFU
MGIT
HH21B
MPN/CFU Assays
Routine Diagnostic Assay
Index
HHC1
HHC2
HHC3
CF+ W1
CF+ W2
CF- W1
CF- W2
CFU
MGIT

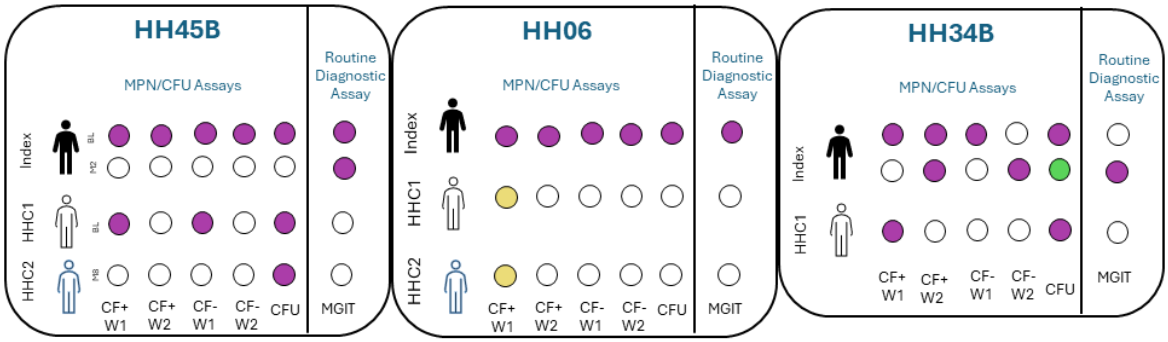

Key
Lineage 1
Lineage 1.1
Lineage 1.2
Lineage 2.2
Lineage 3
Lineage 3.1
Lineage 4
Lineage 4.1
Lineage 4.2
Lineage 4.3
Lineage 4.4
Lineage 4.8
Lineage 4.9

Category 3


HH03
HH22
HH37
HH32B
HH41B
HH23B
HH52
HH53
HH11
HH07
HH22B
HH21B
MPN/CFU Assays
Routine Diagnostic Assay
Index
HHC
HHC1
HHC2
HHC3
CF+ W1
CF+ W2
CF- W1
CF- W2
CFU
MGIT

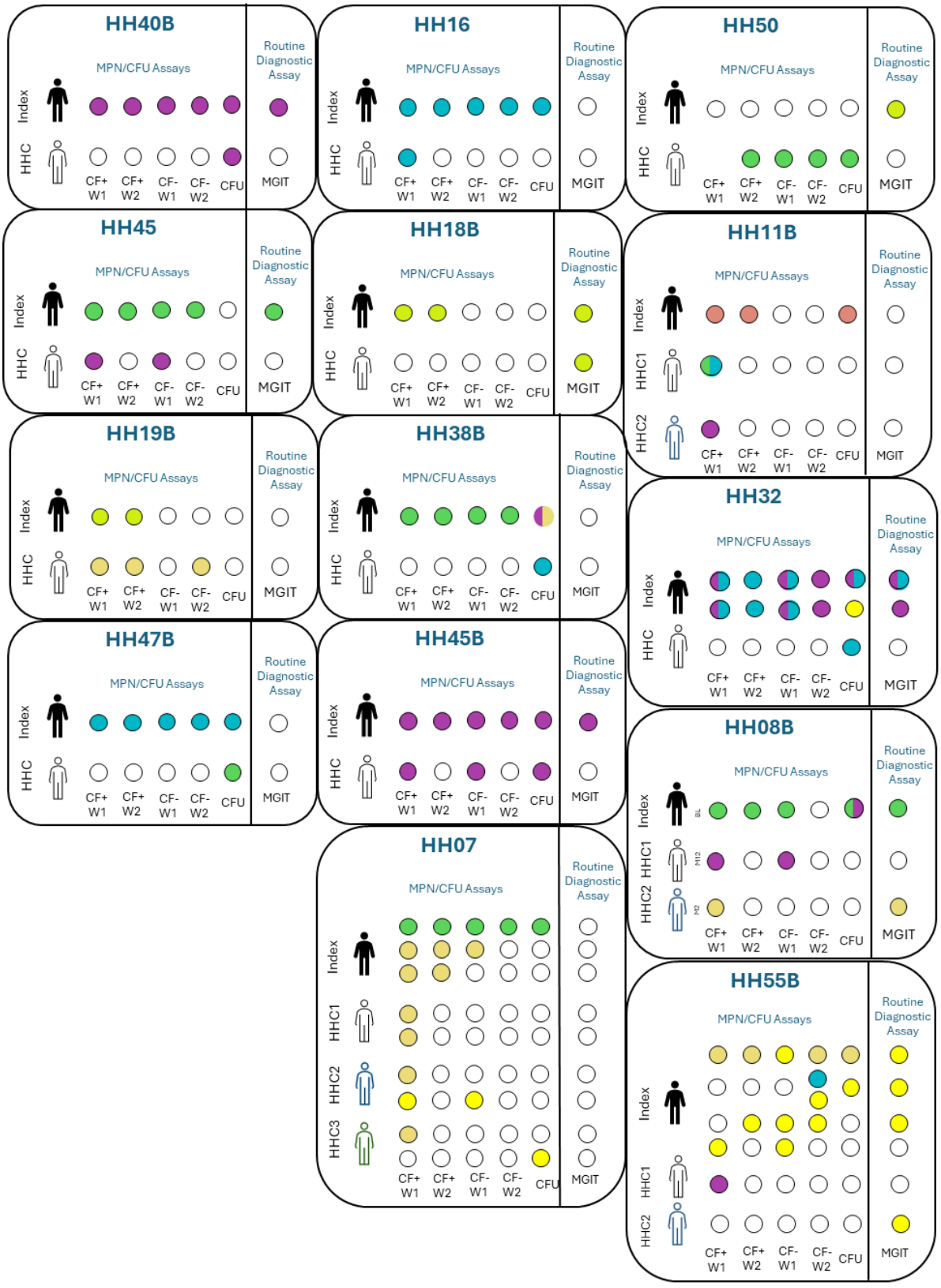

Key
Lineage 1
Lineage 1.1
Lineage 1.2
Lineage 2.2
Lineage 3
Lineage 3.1
Lineage 4
Lineage 4.1
Lineage 4.2
Lineage 4.3
Lineage 4.4
Lineage 4.8
Lineage 4.9

Category 3


HH45B
MPN/CFU Assays
Routine Diagnostic Assay
Index
HHC1
HHC2
BL
M2
M8
CF+ W1
CF+ W2
CF- W1
CF- W2
CFU
MGIT
HH06
MPN/CFU Assays
Routine Diagnostic Assay
Index
HHC1
HHC2
CF+ W1
CF+ W2
CF- W1
CF- W2
CFU
MGIT
HH34B
MPN/CFU Assays
Routine Diagnostic Assay
Index
HHC1
CF+ W1
CF+ W2
CF- W1
CF- W2
CFU
MGIT


Key
Lineage 1
Lineage 1.1
Lineage 1.2
Lineage 2.2
Lineage 3
Lineage 3.1
Lineage 4
Lineage 4.1
Lineage 4.2
Lineage 4.3
Lineage 4.4
Lineage 4.8
Lineage 4.9

Category 4


HH40B
HH16
HH50
HH45
HH18B
HH11B
HH19B
HH38B
HH32
HH47B
HH45B
HH08B
HH07
HH55B
MPN/CFU Assays
Routine Diagnostic Assay
Index
HHC
HHC1
HHC2
HHC3
CF+ W1
CF+ W2
CF- W1
CF- W2
CFU
MGIT
Key
Lineage 1
Lineage 1.1
Lineage 1.2
Lineage 2.2
Lineage 3
Lineage 3.1
Lineage 4
Lineage 4.1
Lineage 4.2
Lineage 4.3
Lineage 4.4
Lineage 4.8
Lineage 4.9

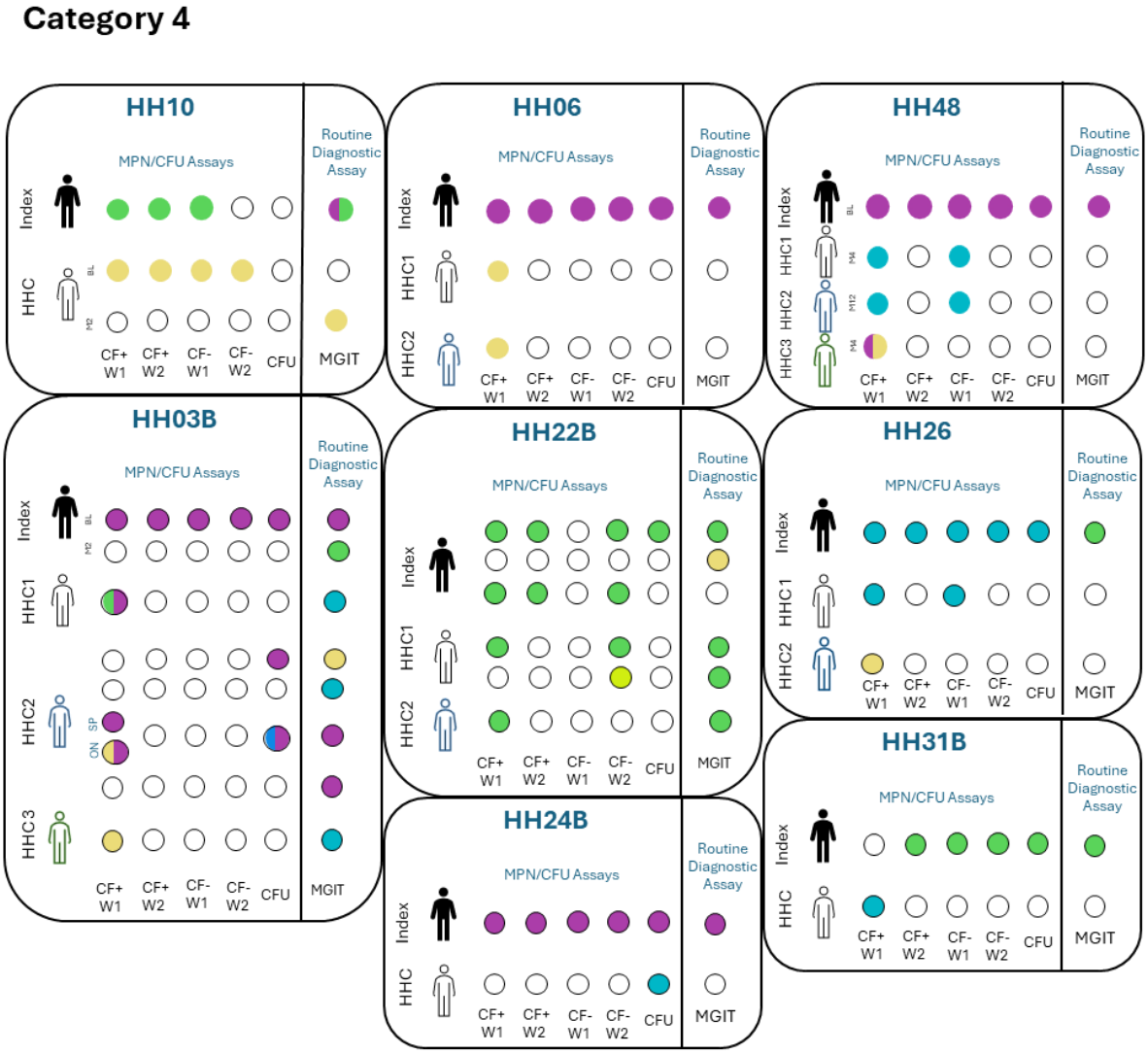


Figure S1: Graphical Illustration of *Mtb* Strain Distribution within Households. All households with at least one positive HHC were categorized based on whether a transmission link can be made within the household and which assay allows for identification of the link. gDNA for sequencing was obtained from 6 different culture conditions: First well of growth in the CF-supplemented MPN (CF+ W1), last well of growth in the CF-supplemented MPN (CF+ W2), First well of growth in the CF-unsupplemented MPN (CF- W1), last well of growth in the CF-unsupplemented MPN (CF- W2), CFU and MGIT (maybe put the figure from the results again at the beginning to show this). The colours in each circle represent the lineage identified, uncolourd circles represent no growth or unavailability of DNA. The particpants were placed in one of four categories of transmission. Category 1 consists of households wherein a transmission link cannot be made within the household (SNP>13); HHCs identified by the MGIT assay. Category 2 consists of households wherein transmission links can be inferred within the household (SNP<13); transmission links were identified by MGIT and MPN/CFU assays. Category 3 consists of HHs wherein, transmission links can be inferred within the HH (SNP<13); transmission links were identified by MPN/CFU assays only. Category 4 consists of HHs wherein transmission links cannot be made within the HH (SNP>13); transmission links identified by MPN/CFU assays only.

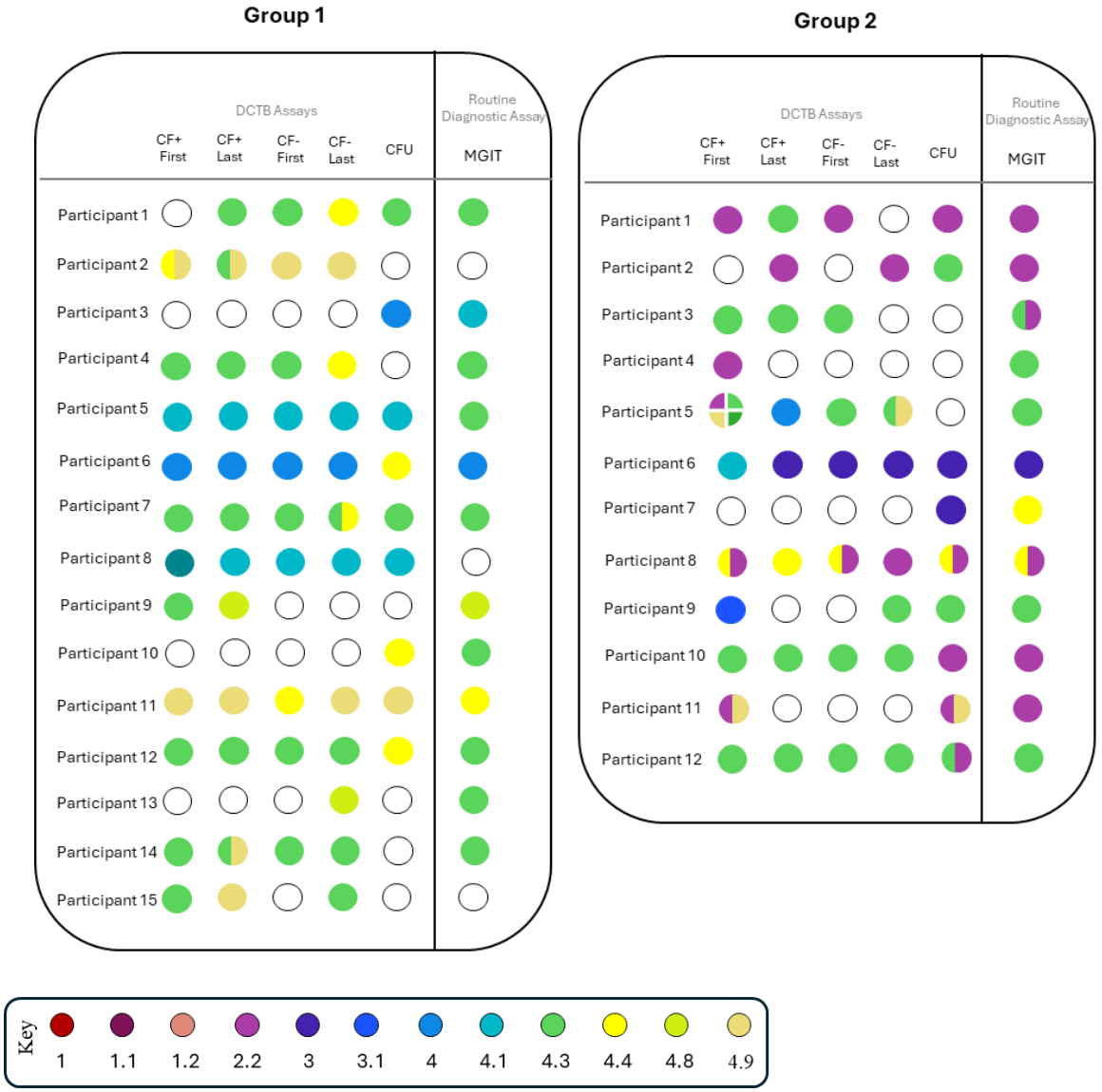


Figure S2: Display of Mycobacterial Genetic Heterogeneity in Infected Individuals Illustrative depiction of the different strains within a single individual at a single time point. Different culture conditions are able to detect different strains. Each color represents a different genotype family or genotype lineage.

Table S4: Drug Resistance Mutations in *Mtb* from Participants who Converted from Drug-Susceptible TB

Participant	Gene	Mutation	Amino Acid Change	Drug Resistance	Genotype Lineage(s)
1	*rpoB*	L443S	Leucine → Serine	Rifampin	4.3
2	*katG*	S315T	Serine → Threonine	Isoniazid	4.1.1.3
3	*katG*	S315N	Serine → Asparagine	Isoniazid	4.3.4.2.1 & 4.4.1
4	*katG*	S315T	Serine → Threonine	Isoniazid	4.1.1.3
	katG	S315T	Serine → Threonine	Isoniazid	
5	*embB*	M306V	Methionine → Valine	Ethambutol	4.4.1.1
	rpoB	D435V	Aspartic Acid → Valine	Rifampin	
6	*rpoB*	S450L	Serine → Leucine	Rifampin	2.2.1

6.5 Appendix E – Turn-it-in Report

turnitin

Digital Receipt

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

The first page of your submissions is displayed below.

Submission author:	Astika Sewcharran
Assignment title:	Final submission (all, regardless of course/programme)
Submission title:	A.Sewcharran_MSc Dissertation_1822059_final_26 March 2025...
File name:	A.Sewcharran_MSc_Dissertation_1822059_final_26_March_202...
File size:	20.17M
Page count:	132
Word count:	33,190
Character count:	185,187
Submission date:	26-Mar-2025 11:04AM (UTC+0200)
Submission ID:	2625668359

Molecular Analysis of Tuberculosis Transmission in Household Contacts Using a Culturomics Approach

UNIVERSITY OF THE WITWATERSRAND JOHANNESBURG

Astika Sewcharran (1822059)

7. References

1. Glaziou P, Floyd K, Raviglione MC. Global Epidemiology of Tuberculosis. *Semin Respir Crit Care Med*. 2018;39(3):271–85.

2. World Health Organization. Global tuberculosis report 2024. Geneva: World Health Organization; 2024. Contract No.: 2024. Licence: CC BY-NC-SA 3.0 IGO.

3. Tomlinson C. In-depth: What new WHO TB numbers mean for SA SpotLight2023 [Available from: https://www.spotlightnsp.co.za/2023/11/10/in-depth-what-new-who-tb-numbers-mean-for-sa/.

4. Kanabus A. Information about Tuberculosis. High burden TB countries. TBFACTS.ORG2022.

5. World Health Organization. Global Tuberculosis Report. Geneva; 2023.

6. Lonnroth K, Jaramillo E, Williams BG, Dye C, Raviglione M. Drivers of tuberculosis epidemics: the role of risk factors and social determinants. Soc Sci Med. 2009;68(12):2240–6.

7. Lönnroth K, Castro KG, Chakaya JM, Chauhan LS, Floyd K, Glaziou P, et al. Tuberculosis control and elimination 2010-50: cure, care, and social development. Lancet. 2010;375(9728):1814–29.

8. Bagcchi S. WHO's Global Tuberculosis Report 2022. Elsevier BV; 2023. p. e20.

9. Tiruviluamala P, Reichman LB. Tuberculosis. Annu Rev Public Health. 2002;23:403–26.

10. Smith I, . *Mycobacterium tuberculosis* pathogenesis and molecular determinants of virulence. 2003. p. 463–96.

11. Davenne T, McShane H. Why don't we have an effective tuberculosis vaccine yet? Expert Rev Vaccines. 2016;15(8):1009–13.

12. Converse SE, Cox JS. A protein secretion pathway critical for *Mycobacterium tuberculosis* virulence is conserved and functional in *Mycobacterium smegmatis*. J Bacteriol. 2005;187(4):1238–45.

13. Chakraborty S, Rhee KY. Tuberculosis Drug Development: History and Evolution of the Mechanism-Based Paradigm. Cold Spring Harb Perspect Med. 2015;5(1).

14. Murray S. Challenges of tuberculosis control. CMAJ. 2006;174(1):33–4.

15. Cha J, , Ho YL, , Kyung SL, , et al. Radiological findings of extensively drug-resistant pulmonary tuberculosis in non-AIDS adults: Comparisons with findings of multidrug-resistant and drug-sensitive tuberculosis. 2009;10:207–16.

16. Johansen IS, Roen A, Kraef C, Martín-Iguacel R, Nemeth J, Fenner L, et al. Risk of tuberculosis after initiation of antiretroviral therapy among persons with HIV in Europe. 2024(1878-3511 (Electronic)).

17. Adeyemo S, Sangotola A, Korosteleva O. Modeling Transmission Dynamics of Tuberculosis–HIV Co-Infection in South Africa. Epidemiologia. 2023;4(4):408–19.

18. Dheda K, Perumal T, Moultrie H, Perumal R, Esmail A, Scott AJ, et al. The intersecting pandemics of tuberculosis and COVID-19: population-level and patient-level impact, clinical presentation, and corrective interventions. Lancet Respir Med. 2022;10(6):603–22.

19. South African National AIDS Council (SANAC). National Strategic Plan on HIV, STIs and TB 2012–2016. Pretoria, South Africa: South African National AIDS Council; 2011.

20. Van der Walt MM, S. The First National TB Prevalence Survey. 2018. p. 25.

21. World Health Organization. Companion handbook to the WHO guidelines for the programmatic management of drug-resistant tuberculosis. Geneva: World Health Organization; 2014.

22. Zhang H, Liu M, Fan W, Sun S, Fan X. The impact of *Mycobacterium tuberculosis* complex in the environment on one health approach. Front Public Health. 2022;10:994745.

23. Rustad TR, Minch KJ, Brabant W, Winkler JK, Reiss DJ, Baliga NS, et al. Global analysis of mRNA stability in *Mycobacterium tuberculosis*. Nucleic Acids Research. 2012;41(1):509–17.

24. Kiazyk S, , Ball T, . Latent tuberculosis infection: An overview. 2017;43:62–6.

25. Brites D, Gagneux S. Co-evolution of *Mycobacterium tuberculosis* and Homo sapiens. 2015(1600-065X (Electronic)).

26. Bottai D, Frigui W, Sayes F, Di Luca M, Spadoni D, Pawlik A, et al. TbD1 deletion as a driver of the evolutionary success of modern epidemic *Mycobacterium tuberculosis* lineages. Nat Commun. 2020;11(1):684.

27. Coscolla M, Gagneux S, Menardo F, Loiseau C, Ruiz-Rodriguez P, Borrell S, et al. Phylogenomics of *Mycobacterium africanum* reveals a new lineage and a complex evolutionary history. Microb Genom. 2021;7(2).

28. Ngabonziza JCS, Loiseau C, Marceau M, Jouet A, Menardo F, Tzfadia O, et al. A sister lineage of the *Mycobacterium tuberculosis* complex discovered in the African Great Lakes region. Nat Commun. 2020;11(1):2917.

29. Brynildsrud OB, Pepperell CS, Suffys P, Grandjean L, Monteserin J, Debech N, et al. Global expansion of *Mycobacterium tuberculosis* lineage 4 shaped by colonial migration and local adaptation. Sci Adv. 2018;4(10):eaat5869.

30. Silva ML, Ca B, Osorio NS, Rodrigues PNS, Maceiras AR, Saraiva M. Tuberculosis caused by *Mycobacterium africanum*: Knowns and unknowns. PLoS Pathog. 2022;18(5):e1010490.

31. Said H, Ratabane J, Erasmus L, Gardee Y, Omar S, Dreyer A, et al. Distribution and Clonality of drug-resistant tuberculosis in South Africa. BMC Microbiol. 2021;21(1):157.

32. Comas I, Gagneux S. The past and future of tuberculosis research. 2009(1553-7374 (Electronic)).

33. World Health Organization. WHO consolidated guidelines on tuberculosis. Module 3: Diagnosis – rapid diagnostics for tuberculosis detection. Geneva; 2024.

34. Kobashi Y. Current status and future landscape of diagnosing tuberculosis infection. Respiratory Investigation. 2023;61(5):563–78.

35. Gago G, , Diacovich L, , Gramajo H, . Lipid metabolism and its implication in mycobacteria–host interaction. 2018;41:36–42.

36. Abrahams KA, , Besra GS, . Mycobacterial cell wall biosynthesis: A multifaceted antibiotic target. 2018;145:116–33.

37. Lumpur K, . The manual MGIT system for the detection of M tuberculosis in respiratory specimens : an experience in the University Malaya. 2009;31:93–7.

38. Data AW-gS, . Phylogenetic Analysis of *Mycobacterium tuberculosis* Strains in Wales by Use of Core Genome Multilocus Sequence Typing To Analyze Whole-Genome Sequencing Data. 2019. p. 1–11.

39. Fadzilah MN, Ng KP, Ngeow YF. The manual MGIT system for the detection of M. tuberculosis in respiratory specimens: an experience in the University Malaya Medical Centre. Malays J Pathol. 2009;31(2):93–7.

40. Kana BD, Gordhan BG, Downing KJ, Sung N, Vostroktunova G, Machowski EE, et al. The resuscitation-promoting factors of *Mycobacterium tuberculosis* are required for virulence and resuscitation from dormancy but are collectively dispensable for growth in vitro. *Proc Natl Acad Sci U S A*. 2008;105(33):11915–20.

41. Mukamolova GV, Turapov O, Malkin J, Woltmann G, Barer MR. Resuscitation-promoting factors reveal an occult population of tubercle Bacilli in Sputum. Am J Respir Crit Care Med. 2010;181(2):174–80.

42. Chengalroyen MD, Beukes GM, Gordhan BG, Streicher EM, Churchyard G, Hafner R, et al. Detection and quantification of differentially culturable tubercle bacteria in sputum from patients with tuberculosis. *Am J Respir Crit Care Med*. 2016;194(12):1532–40.

43. Shi J, , Zheng D, , Zhu Y, , et al. Role of MIRU-VNTR and spoligotyping in assessing the genetic diversity of Mycobacterium tuberculosis in Henan. 2018:1–12.

44. Peters JS, McIvor A, Papadopoulos AO, Masangana T, Gordhan BG, Waja Z, et al. Differentially culturable tubercle bacteria as a measure of tuberculosis treatment response. *Front Cell Infect Microbiol*. 2022;12:1064148.

45. Gordhan BGS, A. Letsoalo, M. Chinappa, T. Yende-Zuma, N. Padayatchi, N. Naidoo, K. Kana, B. D. Detection of differentially culturable tubercle bacteria in sputum from drug-resistant tuberculosis patients. Front Cell Infect Microbiol. 2022;12:949370.

46. Kamariza M, Shieh P, Ealand CS, Peters JS, Chu B, Rodriguez-Rivera FP, et al. Rapid detection of *Mycobacterium tuberculosis* in sputum with a solvatochromic trehalose probe. Sci Transl Med. 2018;10(430).

47. National Guidelines on the Treatment of Tuberculosis Infection. In: Health Do, editor. Knowledge Hub: National Department of Health; 2023.

48. Ngabonziza JC, Ssengooba W, Mutua F, Torrea G, Dushime A, Gasana M, et al. Diagnostic performance of smear microscopy and incremental yield of Xpert in detection of pulmonary tuberculosis in Rwanda. BMC Infect Dis. 2016;16(1):660.

49. Abraham PR, Sharma VD, Shivannavar CT. Diagnosis of TB from smear & culture negative sputum specimens by IS 6110 based PCR. Indian J Med Res. 2012;135(2):249–51.

50. Swai HF, Mugusi FM, Mbwambo JK. Sputum smear negative pulmonary tuberculosis: sensitivity and specificity of diagnostic algorithm. BMC Res Notes. 2011;4:475.

51. Sreevatsan S, Pan X, Stockbauer KE, Connell ND, Kreiswirth BN, Whittam TS, et al. Restricted structural gene polymorphism in the *Mycobacterium tuberculosis* complex indicates evolutionarily recent global dissemination. *Proc Natl Acad Sci U S A*. 1997;94:9869–74.

52. Yimer SA, , Kalayou S, , Homberset H, , et al. Lineage-Specific Proteomic Signatures in the Mycobacterium tuberculosis Complex Reveal Differential Abundance of Proteins Involved in Virulence, DNA Repair, CRISPR-Cas, Bioenergetics and Lipid Metabolism. 2020;11:1–19.

53. Backus KM, Dolan MA, Barry CS, Joe M, McPhie P, Boshoff HI, et al. The three *Mycobacterium tuberculosis* antigen 85 isoforms have unique substrates and activities determined by non-active site regions. J Biol Chem. 2014;289(36):25041–53.

54. Sharp SEL, M. Sierra, S. G. Poniecka, A.Poppiti, Jr, . Löwenstein-Jensen media: No longer necessary for mycobacterial isolation. 2000. p. 770–3.

55. Procedure PT, . BBL™ Mycobactosel ™ L-J Medium. 2015. p. 9–12.

56. Cepheid, . Cepheid GeneXpert MTB/Rif Ref. F August 2019. 2019.

57. Miller LP, Crawford JT, Shinnick TM. The *rpoB* gene of *Mycobacterium tuberculosis*. Antimicrob Agents Chemother. 1994;38(4):805–11.

58. Steingart KRH, MeganNg, Vivienne Hopewell, Philip C Ramsay, Andrew Cunningham, JaneUrbanczik, RichardPerkins, Mark, . Fluorescence versus conventional sputum smear microscopy for tuberculosis : a systematic review. 2006.

59. Sahile HAR, Celine Shapira, Tirosh Andersen, Raymond J.Av-Gay, Yossef, . DMN-Tre Labeling for Detection and High-Content Screening of Compounds against Intracellular Mycobacteria. 2020. p. 3661–9.

60. Mäkinen J, Marttila HJ, Marjamäki M, Viljanen MK, Soini H. Comparison of two commercially available DNA line probe assays for detection of multidrug-resistant *Mycobacterium tuberculosis*. J Clin Microbiol. 2006;44(2):350–2.

61. Pei S, Song Z, Yang W, He W, Ou X, Zhao B, et al. The catalogue of *Mycobacterium tuberculosis* mutations associated with drug resistance to 12 drugs in China from a nationwide survey: a genomic analysis. The Lancet Microbe. 2024;5(11).

62. Achimi P. Hain GenoType Line Probe Assay. TB Diagnostics Workshop; TB Diagnostics Workshop2019.

63. Kim MJ, Nam YS, Cho SY, Park TS, Lee HJ. Comparison of the Xpert MTB/RIF Assay and Real-time PCR for the Detection of *Mycobacterium tuberculosis*. Ann Clin Lab Sci. 2015;45(3):327–32.

64. Campbell JR, Krot J, Elwood K, Cook V, Marra F. A Systematic Review on TST and IGRA Tests Used for Diagnosis of LTBI in Immigrants. *Mol Diagn Ther*. 2015.

65. Chee CBE, , Reves R, , Zhang Y, , et al. Latent tuberculosis infection: Opportunities and challenges. 2018;23:893–900.

66. Müller A. Dignostic tests for active tuberculosis TB Online2011 [Available from: https://www.tbonline.info/posts/2011/8/9/diagnosic-tests-active-tuberculosis/#:~:text=Therefore%20chest%20x%2Drays%20are,a%20chest%20x%2Dray%20alone.

67. Menzies D, Adjobimey M, Ruslami R, Trajman A, Sow O, Kim H, et al. Four Months of Rifampin or Nine Months of Isoniazid for Latent Tuberculosis in Adults. N Engl J Med. 2018;379(5):440–53.

68. Churchyard GJ, Swindells S. Controlling latent TB tuberculosis infection in high-burden countries: A neglected strategy to end TB. PLoS Med. 2019;16(4):e1002787.

69. Balcells ME, Thomas SL, Godfrey-Faussett P, Grant AD. Isoniazid preventive therapy and risk for resistant tuberculosis. Emerg Infect Dis. 2006;12(5):744–51.

70. Carranza C, Pedraza-Sanchez S, de Oyarzabal-Mendez E, Torres M. Diagnosis for Latent Tuberculosis Infection: New Alternatives. Front Immunol. 2020;11:2006.

71. Kunkel A, Crawford FW, Shepherd J, Cohen T. Benefits of continuous isoniazid preventive therapy may outweigh resistance risks in a declining tuberculosis/HIV coepidemic. Aids. 2016;30(17):2715–23.

72. Sterling TR, Villarino ME, Borisov AS, Shang N, Gordin F, Bliven-Sizemore E, et al. Three months of rifapentine and isoniazid for latent tuberculosis infection. N Engl J Med. 2011;365(23):2155–66.

73. Churchyard GJ, Fielding KL, Lewis JJ, Coetzee L, Corbett EL, Godfrey-Faussett P, et al. A trial of mass isoniazid preventive therapy for tuberculosis control. N Engl J Med. 2014;370(4):301–10.

74. Krishnan S, Chaisson RE. US Guidelines Fall Short on Short-Course Tuberculosis-Preventive Therapy. Clinical Infectious Diseases. 2023;78(3):514–7.

75. Goletti D, Delogu G, Matteelli A, Migliori GB. The role of IGRA in the diagnosis of tuberculosis infection, differentiating from active tuberculosis, and decision making for initiating treatment or preventive therapy of tuberculosis infection. Int J Infect Dis. 2022;124 Suppl 1:S12–s9.

76. Alemu A, , Tadesse M, , Seid G, , et al. Does Xpert ® MTB / RIF assay give rifampicin resistance results without identified mutation ? Review of cases from Addis. 2020:1–6.

77. van Cleeff MRAK-N, Lydia E. Meme, Helen Odhiambo, Joseph A. Klatser, Paul R., . The role and performance of chest X-ray for the diagnosis of tuberculosis: A cost-effective analysis in Nairobi, Kenya. 2005. p. 1–9.

78. World Health Organization. WHO consolidated guidelines on tuberculosis. Module 2: Screening – systematic screening for tuberculosis disease. Geneva; 2021.

79. Wekesa C, Sekaggya-Wiltshire C, Muyanja SZ, Lume I, Nabaggala MS, Parkes-Ratanshi R, et al. Comparing adherence to MDR-TB treatment among patients on self-administered therapy and those on directly observed therapy: non-inferiority randomized controlled trial. Trials. 2023;24(1):326.

80. Cerrone M, Bracchi M, Wasserman S, Pozniak A, Meintjes G, Cohen K, et al. Safety implications of combined antiretroviral and anti-tuberculosis drugs. Expert Opin Drug Saf. 2020;19(1):23–41.

81. Organization WH. WHO Expert Consultation on the Definition of Extensively Drug-Resistant Tuberculosis. Geneva: World Health Organization; 2021.

82. Vanino E, Granozzi B, Akkerman OW, Munoz-Torrico M, Palmieri F, Seaworth B, et al. Update of drug-resistant tuberculosis treatment guidelines: A turning point. International Journal of Infectious Diseases. 2023;130:S12–S5.

83. Mirzayev F, Viney K, Linh NN, Gonzalez-Angulo L, Gegia M, Jaramillo E, et al. World Health Organization recommendations on the treatment of drug-resistant tuberculosis, 2020 update. Eur Respir J. 2021;57(6).

84. Ei PW, Aung WW, Lee JS, Choi GE, Chang CL. Molecular Strain Typing of *Mycobacterium tuberculosis*: a Review of Frequently Used Methods. J Korean Med Sci. 2016;31(11):1673–83.

85. Thiel T, Michalek W, Varshney RK, Graner A. Exploiting EST databases for the development and characterization of gene-derived SSR-markers in barley (Hordeum vulgare L.). Theor Appl Genet. 2003;106(3):411–22.

86. Fomukong NG, Tang TH, al-Maamary S, Ibrahim WA, Ramayah S, Yates M, et al. Insertion sequence typing of *Mycobacterium tuberculosis*: characterization of a widespread subtype with a single copy of IS6110. Tuber Lung Dis. 1994;75(6):435–40.

87. Lok KH, Benjamin WH, Jr., Kimerling ME, Pruitt V, Lathan M, Razeq J, et al. Molecular differentiation of *Mycobacterium tuberculosis* strains without IS6110 insertions. Emerg Infect Dis. 2002;8(11):1310–3.

88. Murthy SE, , Chatterjee F, , Crook A, , et al. Pretreatment chest x-ray severity and its relation to bacterial burden in smear positive pulmonary tuberculosis. 2018;16:1–11.

89. Gupta-Wright A, Peters JA, Flach C, Lawn SD. Detection of lipoarabinomannan (LAM) in urine is an independent predictor of mortality risk in patients receiving treatment for HIV-associated tuberculosis in sub-Saharan Africa: A systematic review and meta-analysis. BMC Medicine. 2016;14:1–11.

90. Rufai SBK, Parveen Singh, Amit Prajapati, Suneel Balooni, Veena Singh, Sarman, . Comparison of Xpert MTB / RIF with Line Probe Assay for Detection of Rifampin-Monoresistant *Mycobacterium tuberculosis*. 2014. p. 1846–52.

91. Hoffmann CJV, Ebrahim Rakgokong, Modiehi Masonoke, Katlego van der Watt, Martin Chaisson, Richard E. Martinson, Neil A., . High Prevalence of Pulmonary Tuberculosis but Low Sensitivity of Symptom Screening among HIV-Infected Pregnant Women in South Africa. 2013. p. 8–12.

92. Nikolayevskyy V, Trovato A, Broda A, Borroni E, Cirillo D, Drobniewski F. MIRU-VNTR Genotyping of *Mycobacterium tuberculosis* Strains Using QIAxcel Technology: A Multicentre Evaluation Study. PLoS One. 2016;11(3):e0149435.

93. Kabir S, Uddin MKM, Chisti MJ, Fannana T, Haque ME, Uddin MR, et al. International Journal of Infectious Diseases Role of PCR method using IS6110 primer in detecting *Mycobacterium tuberculosis* among the clinically diagnosed childhood tuberculosis patients at an urban hospital in Dhaka. *Int J Infect Dis*. 2018;68:108–14.

94. Zheng CZ, Yuding Zhu, Guoqiang Li, Song Sun, Honghu Feng, Qin Luo, Mei Wu, Fanzi Li, Xuefeng Hill, Véronique Rastogi, Nalin Sun, Qun, . Suitability of IS 6110 -RFLP and MIRU-VNTR for Differentiating Spoligotyped Drug-Resistant *Mycobacterium tuberculosis* Clinical Isolates from Sichuan in China. 2014.

95. Dlamini MT, Lessells R, Iketleng T, de Oliveira T. Whole genome sequencing for drug-resistant tuberculosis management in South Africa: What gaps would this address and what are the challenges to implementation? J Clin Tuberc Other Mycobact Dis. 2019;16:100115.

96. Cox H, Goig GA, Salaam-Dreyer Z, Dippenaar A, Reuter A, Mohr-Holland E, et al. Whole-Genome Sequencing Has the Potential To Improve Treatment for Rifampicin-Resistant Tuberculosis in High-Burden Settings: a Retrospective Cohort Study. J Clin Microbiol. 2022;60(3):e0236221.

97. Kremer KB, Annelies Schouls, Leo van Soolingen, Dick van Embden, Jan, . Spoligotyping protocol. 2002. p. 1–19.

98. Inna Vitol JD, Barry Kreiswirth, Natalia Kurepina, Kristin P. Bennett, . "Identifying Mycobacterium tuberculosis Complex Strain Families using Spoligotypes", Infection, Genetics and Evolution. 2006:pp. 491–504.

99. Martinez L, Shen Y, Mupere E, Kizza A, Hill PC, Whalen CC. Transmission of *Mycobacterium Tuberculosis* in Households and the Community: A Systematic Review and Meta-Analysis. 2017(1476-6256 (Electronic)).

100. Nikolayevskyy V, , Trovato A, , Broda A, , et al. MIRU-VNTR Genotyping of Mycobacterium tuberculosis Strains Using QIAxcel Technology : A Multicentre Evaluation Study. 2016:1–11.

101. Naga Mamo AA-O, Furi Gilo R, Fikadu Tesema AA-O, Fetene Worku N, Teshome Kenea T, Kebede Dibisa D, et al. Household Contact Tuberculosis Screening Adherence and Associated Factors Among Pulmonary Tuberculosis Patients on Follow-Up at Health Facilities in Shashamane Town, Southeast Ethiopia. 2023(1177-889X (Print)).

102. van der Werf MJ, , Ködmön C, . Whole-genome sequencing as tool for investigating international tuberculosis outbreaks: A systematic review. 2019;7:1–9.

103. Achkar JM, Jenny-Avital ER. Incipient and subclinical tuberculosis: defining early disease states in the context of host immune response. J Infect Dis. 2011;204 Suppl 4(Suppl 4):S1179–86.

104. Bekken GK RC, Selvam S, Jesuraj N, Hesseling AC, Doherty TM, Grewal HMS, Vaz M, Jenum S. Identification of subclinical tuberculosis in household contacts using exposure scores and contact investigations. BMC Infect Dis 2020;20.

105. Rickman HMC, S. Lala, S. G. Waja, Z. Salazar-Austin, N. Hoffmann, J. Dooley, K. E.Hoffmann, C. J. Chaisson, R. E. Martinson, N. A., . Subclinical tuberculosis and adverse infant outcomes in pregnant women with HIV. International Union Against Tuberculosis and Lung Disease (The Union); 2020. p. 681–5.

106. Ismail NA, Mvusi L, Nanoo A, Dreyer A, Omar SV, Babatunde S, et al. Prevalence of drug-resistant tuberculosis and imputed burden in South Africa: a national and sub-national cross-sectional survey. Lancet Infect Dis. 2018;18(7):779–87.

107. Saito K, , Warrier T, , Somersan-Karakaya S, , et al. Rifamycin action on RNA polymerase in antibiotictolerant Mycobacterium tuberculosis results in differentially detectable populations. 2017;114:E4832–E40.

108. Bjorn-Mortensen K, , Zallet J, , Lillebaek T, , et al. Direct DNA extraction from Mycobacterium tuberculosis frozen stocks as a reculture-independent approach to whole-genome sequencing. 2015;53:2716–9.

109. Dixit AA-O, Freschi L, Vargas R, Calderon R, Sacchettini J, Drobniewski F, et al. Whole genome sequencing identifies bacterial factors affecting transmission of multidrug-resistant tuberculosis in a high-prevalence setting. 2019(2045-2322 (Electronic)).

110. Drain PA-O, Bajema KL, Dowdy D, Dheda K, Naidoo K, Schumacher SG, et al. Incipient and Subclinical Tuberculosis: a Clinical Review of Early Stages and Progression of Infection. LID - 10.1128/CMR.00021-18 [doi] LID - e00021-18. 2018(1098-6618 (Electronic)).

111. Walker TM, Ip Cl Fau - Harrell RH, Harrell Rh Fau - Evans JT, Evans Jt Fau - Kapatai G, Kapatai G Fau - Dedicoat MJ, Dedicoat Mj Fau - Eyre DW, et al. Whole-genome sequencing to delineate Mycobacterium tuberculosis outbreaks: a retrospective observational study. (1474-4457 (Electronic)).

112. Clark TG, Mallard K, Coll F, Preston M, Assefa S, Harris D, et al. Elucidating emergence and transmission of multidrug-resistant tuberculosis in treatment experienced patients by whole genome sequencing. (1932-6203 (Electronic)).

113. Alaridah N, Hallbäck ET, Tångrot J, Winqvist N, Sturegård E, Florén-Johansson K, et al. Transmission dynamics study of tuberculosis isolates with whole genome sequencing in southern Sweden. (2045-2322 (Electronic)).

114. Xiao YXA-O, Chan THA-O, Liu KH, Jou RA-O. Define SNP thresholds for delineation of tuberculosis transmissions using whole-genome sequencing. (2165-0497 (Electronic)).

115. Lyon Sarah M, Rossman Milton D. Pulmonary Tuberculosis. Microbiology Spectrum. 2017;5(1):10.1128/microbiolspec.tnmi7–0032–2016.

116. Swindells S, Ramchandani R, Gupta A, Benson CA, Leon-Cruz J, Mwelase N, et al. One Month of Rifapentine plus Isoniazid to Prevent HIV-Related Tuberculosis. N Engl J Med. 2019;380(11):1001–11.

117. Paul S, Laochumroonvorapong P Fau - Kaplan G, Kaplan G. Comparable growth of virulent and avirulent *Mycobacterium tuberculosis* in human macrophages in vitro. 1996(0022-1899 (Print)).

118. Biketov S, Mukamolova Gv Fau - Potapov V, Potapov V Fau - Gilenkov E, Gilenkov E Fau - Vostroknutova G, Vostroknutova G Fau - Kell DB, Kell Db Fau - Young M, et al. Culturability of *Mycobacterium tuberculosis* cells isolated from murine macrophages: a bacterial growth factor promotes recovery. FEMS Immunol Med Microbiol. 2000(0928-8244 (Print)).

119. Mukamolova GV, Turapov O Fau - Malkin J, Malkin J Fau - Woltmann G, Woltmann G Fau - Barer MR, Barer MR. Resuscitation-promoting factors reveal an occult population of tubercle Bacilli in Sputum. 2010(1535-4970 (Electronic)).

120. Mpagama SG, Mtabho C, Mwaigwisya S, Mleoh LJ, Boer IM, Heysell SK, et al. Comparison of overnight pooled and standard sputum collection method for patients with suspected pulmonary tuberculosis in northern Tanzania. Tuberc Res Treat. 2012;2012:128057.

121. Peters JS, Ismail N, Dippenaar A, Ma S, Sherman DR, Warren RM, et al. Genetic Diversity in *Mycobacterium tuberculosis* Clinical Isolates and Resulting Outcomes of Tuberculosis Infection and Disease. Annu Rev Genet. 2020;54:511–37.

122. Carvalho AC, DeRiemer K Fau - Nunes ZB, Nunes Zb Fau - Martins M, Martins M Fau - Comelli M, Comelli M Fau - Marinoni A, Marinoni A Fau - Kritski AL, et al. Transmission of Mycobacterium tuberculosis to contacts of HIV-infected tuberculosis patients. 2001(1073-449X (Print)).

123. Windels EA-O, Wampande EM, Joloba ML, Boom WH, Goig GA, Cox H, et al. HIV co-infection is associated with reduced *Mycobacterium tuberculosis* transmissibility in sub-Saharan Africa. 2024(1553-7374 (Electronic)).

124. Carter N, Webb EL, Lebina L, Motsomi K, Bosch Z, Martinson NA, et al. Prevalence of subclinical pulmonary tuberculosis and its association with HIV in household contacts of index tuberculosis patients in two South African provinces: a secondary, cross-sectional analysis of a cluster-randomised trial. 2023.

125. Lungu PS, Kabaso ME, Mihova R, Silumesii A, Chisenga T, Kasapo C, et al. Undernotification and underreporting of tuberculosis in Zambia: a national data quality assessment. BMC Health Services Research. 2022;22(1):1074.

126. Little KMM, R. Martinson, N. Golub, J. Chaisson, R. Dowdy, D. Yield of household contact tracing for tuberculosis in rural South Africa. BMC Infect Dis. 2018;18(1):299.

127. Kigozi NG, Heunis JC, Engelbrecht MC. Yield of systematic household contact investigation for tuberculosis in a high-burden metropolitan district of South Africa. BMC Public Health. 2019;19(1):867.

128. Wood R, Johnstone-Robertson S, Uys P, Hargrove J, Middelkoop K, Lawn SD, et al. Tuberculosis Transmission to Young Children in a South African Community: Modeling Household and Community Infection Risks. Clinical Infectious Diseases. 2010;51(4):401–8.

129. Kafwabulula M, Ahmed K Fau - Nagatake T, Nagatake T Fau - Gotoh J, Gotoh J Fau - Mitarai S, Mitarai S Fau - Oizumi K, Oizumi K Fau - Zumla A, et al. Evaluation of PCR-based methods for the diagnosis of tuberculosis by identification of mycobacterial DNA in urine samples. (1027-3719 (Print)).

130. Wang Y, Cui J, Li Y, Wang M, Han W, Liu A, et al. Rapid quantitative PCR on tongue swabs for pulmonary tuberculosis in adults: a prospective multicentre study. LID - 10.1183/13993003.01493-2024 [doi] LID - 2401493. 2025(1399-3003 (Electronic)).

131. Kay A, Vasiliu A, Carratala-Castro L, Mtafya B, Mendez Reyes JE, Maphalala N, et al. Performance of a stool-based quantitative PCR assay for the diagnosis of tuberculosis in adolescents and adults: a multinational, prospective diagnostic accuracy study. (2666-5247 (Electronic)).

132. Wang Y, Cui J, Li Y, Wang M, Han W, Liu A, et al. Rapid quantitative PCR on tongue swabs for pulmonary tuberculosis in adults: a prospective multicentre study. LID - 10.1183/13993003.01493-2024 [doi] LID - 2401493. (1399-3003 (Electronic)).

133. Guthrie JL, Delli Pizzi A, Roth D, Kong C, Jorgensen D, Rodrigues M, et al. Genotyping and Whole-Genome Sequencing to Identify Tuberculosis Transmission to Pediatric Patients in British Columbia, Canada, 2005-2014. J Infect Dis. 2018;217(10):1613–21.

134. Lalor MK, Anderson LF, Hamblion EL, Burkitt A, Davidson JA, Maguire H, et al. Recent household transmission of tuberculosis in England, 2010–2012: retrospective national cohort study combining epidemiological and molecular strain typing data. BMC Medicine. 2017;15(1):105.

135. Gordhan BG, Padarath K, Sewcharran A, McIvor A, VanNieuwenhze MS, Waja Z, et al. Clinical Strains of *Mycobacterium tuberculosis* Representing Different Genotype Families Exhibit Distinct Propensities to Adopt the Differentially Culturable State. Pathogens. 2024;13(4).