Construction of A Molecular Assay for Detection of Drug Resistant Tuberculosis Using Improved Linear-After-The-Exponential PCR

Master's Thesis

Presented to

The Faculty of the Graduate School of Arts and Sciences
Brandeis University
Department of Biology
Lawrence J. Wangh, Advisor

In Partial Fulfillment
of the Requirements for the Degree

Master of Science
in
Biology

by
Ruth-Love Damoah

February 2017
Acknowledgements

I would like to thank God for leading me to this great opportunity and for giving me the strength to do this work. I am eternally grateful to Professor Lawrence Wangh for giving me this project and for guiding me through it. I am thankful that he also provided both the financial and moral support to make this possible and for all the great thoughts and encouragement. I am also grateful to John Rice for working with me to complete this project successfully. I am grateful for all the time spent analyzing data and contemplating solutions. I am also grateful to Aquiles J. Sanchez for taking time to explain difficult concepts from the very beginning, and also for constant support. I am grateful also to Kenneth Pierce, Adam Osborne and the rest of the Wangh lab for their constant support and for being very accommodating. I am truly grateful for all the help I received throughout this project Brandeis biology department and from friends and family. It is because of you all that I made it this far. Thank you all for teaching me that impossible just means no one has found a way to do it just yet. I truly appreciate you all. God bless you.
Abstract

Construction of A Molecular Assay for Detection of Drug Resistant Tuberculosis Using Improved Linear-After-The-Exponential PCR

A thesis presented to the Department of Biology

Graduate School of Arts and Sciences
Brandeis University
Waltham, Massachusetts

By Ruth-Love Damoah

Tuberculosis continues to plague mankind about three centuries after the first epidemic in modern history in the 17th-18th century. In spite of the advances made in understanding the disease and in diagnostics and treatment options since then, eradicating TB is proving difficult. Cases of antibiotic resistance are complicating both the diagnostic and treatment processes of the disease. A cheap, reliable, sensitive assay using Linear-After-The-Exponential PCR (LATE-PCR was invented in the Wangh laboratory) has been developed to meet the Point-Of-Care (POC) diagnosis needed to effectively combat TB. The nature of the assay and the method of amplification make it applicable to all sample types using less costly technology. Using this Assay, as few as five, rpoB gene copies from Mycobacterium tuberculosis can be selectively amplified in samples containing as many as 58,000 copies of human genomic DNA; PCR efficiency was 105.58%. In addition, the Assay can correctly distinguish thirty-five out of thirty-eight different mutations that are known to cause resistance to rifampicin and all thirty-eight strains are can be distinguished from wild-type drug sensitive rpoB gene sequences using Lights-Off Only Probe technology (also invented in the Wangh Laboratory). Although no mutant strain was scored as wild type, not all mutations could be distinguished from each
other. Thus, the resolving power of Lights-Off Only Probes will have to be enhanced if it is clinically important to uniquely distinguish all types of mutations.
Table of Contents

List of Tables ........................................................................................................... ix
List of Figures ........................................................................................................... x
Overview .................................................................................................................. 1
In the Beginning ........................................................................................................ 4
  About Mycobacterium tuberculosis: ...................................................................... 4
    History .................................................................................................................. 4
  Ending the First Epidemic ...................................................................................... 5
  The Recent Epidemic .............................................................................................. 6
  The Horror Today .................................................................................................... 6
  Back with a Vengeance ......................................................................................... 8
The Mycobacterium tuberculosis Complex (MTBC) .............................................. 10
  Basic Morphology ................................................................................................ 10
    The Capsule ......................................................................................................... 10
  Complex cell wall .................................................................................................. 11
  Long Generation Time ......................................................................................... 11
  Genetic component and adaptability .................................................................. 12
Coevolution .............................................................................................................. 14
  The Theory of the Coevolution of TB and Host .................................................. 14
    Cell Wall Specialization ..................................................................................... 15
    Host Related Adaptations .................................................................................. 16
    Major Means of Evolution ............................................................................... 16
  Why TB Has Not Been Eradicated ...................................................................... 17
Tuberculosis as a Disease ....................................................................................... 18
  Disease Cycle ....................................................................................................... 18
Immunobiology ........................................................................................................ 20
  Cellular Response ............................................................................................... 20
  Involved Immune Cells and Killing Mechanisms ............................................... 20
  Survival Strategy .................................................................................................. 21
  Pathogenicity and Virulence .............................................................................. 22
Treatment of TB ....................................................................................................... 25
  Available and Approved Drugs ........................................................................... 25
  General Guidelines .............................................................................................. 26
  Drug Susceptibility and Resistance ................................................................... 27
  New Drugs ........................................................................................................... 27
  Beyond Drug Therapy ......................................................................................... 28
TB Diagnosis ........................................................................................................... 29
  History ................................................................................................................... 29
  Recent Introduction of Molecular Diagnostics ................................................... 29
  What Has Not Been Achieved ............................................................................ 30
Aim of This Research .............................................................................................. 33
  Gene Targets ....................................................................................................... 34
    RNA polymerase Subunit Beta (rpoB) and Region of Difference 9 (rd9) ......... 34
  Sample Types ...................................................................................................... 34
  PCR Methods ...................................................................................................... 35
    Novel Amplification Methods: Linear-Expo-Linear ......................................... 36
    Improved Linear-After-The-Exponential PCR (LATE-PCR) ........................... 38
Detection Method: Lights-On/Lights-Off .............................................................................. 39

Materials and Methods ..................................................................................................... 41

General Setup Procedure ................................................................................................. 41

PCR Reagents .................................................................................................................. 42

General Reagents ............................................................................................................. 43

Other Reagents ................................................................................................................ 43

Primers .............................................................................................................................. 43

Taq Polymerase ................................................................................................................ 46

Magnesium Concentration ............................................................................................. 46

SYBR Concentration ....................................................................................................... 46

ThermaStop/PrimeSafe ................................................................................................. 47

TheraGo ............................................................................................................................. 47

Target .................................................................................................................................. 48

Probe Design and Concentration ...................................................................................... 51

PCR Profile ......................................................................................................................... 52

Summary ............................................................................................................................ 52

LEL Profile ......................................................................................................................... 53

Melt Profiles and Equilibration ....................................................................................... 53

Intermediate Melts ............................................................................................................ 54

Final PCR Reagents and Profile ...................................................................................... 55

Reagents ............................................................................................................................ 55

Protocol .............................................................................................................................. 55

Organization of Results ................................................................................................... 56

Results ................................................................................................................................. 59

Optimized Amplification with LATE-PCR and Lights-Off Only Detection .................... 59

Experiment 1: Dilution Series with and without IC ......................................................... 59

Experiment 2: Dilution Series with IC down to Single Copy Mtb ................................... 60

Summary Results .............................................................................................................. 61

Experiments Leading to Optimized Assay ..................................................................... 64

Experiment 3: LEL with High TB Copy Number ............................................................ 64

Experiment 4: LEL in Lower Copy TB Number (TheraGo 3 Introduced) ....................... 65

Experiment 5: LEL vs. LATE-PCR ................................................................................ 66

Experiment 6: rpoB and rd9 Amplification with LATE-PCR ........................................ 67

Experiment 7: Human Genomic Background ............................................................... 68

Experiment 8: ThermaGo 3 Concentration at 50nM .................................................... 69

Experiment 9: Multiple Reads Between Amplification Cycles ..................................... 70

Experiment 10: Duplex with Internal Control (IC) ......................................................... 71

Experiment 11: Standard BCG Samples ....................................................................... 72

Experiment 12: Platinum Taq Polymerase ................................................................. 73

Experiment 13: New Internal Control (IC) Limiting Primer ........................................ 73

Experiment 14: New Taq Polymerase ........................................................................... 74

Experiment 15: ThermaGo 86-21 and PrimeSafe Combinations ................................. 75

Experiment 16: 3mM vs. 4.5mM Magnesium ............................................................... 76

Experiment 17: 100 vs. 150 ThermaGo 86-21 ............................................................. 77

Experiment 18: 0.24x vs. 0.48x SYBR Concentrations ............................................... 78

Experiment 19: 100nM and 300nM Probe Concentrations ......................................... 79

Experiment 20: 50nM and 100nM rpoB LP Concentration .......................................... 79

Achieving Higher resolution with Lights-Off Only ........................................................ 80
List of Tables

Table 1. Some notable disease-causing members of the Mycobacterium family ........................................13
Table 2. Table of primer sequences and IC target sequence .................................................................45
Table 3. Labels and mutations of all 38 *Mtb* samples ....................................................................50
Table 4. List of biological replicates within the 38 samples tested ....................................................51
Table 5. List of allelic replicates within the 38 samples ......................................................................51
Table 6. Table of dilution series result summary ...............................................................................62
List of Figures

Figure 1. Illustration of LEL PCR .................................................................37
Figure 2. Illustration of amplification steps in improved version of LATE-PCR ........................................38
Figure 3. Schematic representation of Lights-On/Lights-Off and Lights-Off Only technologies ..................39
Figure 4. Illustration of PCR setup ................................................................42
Figure 5. Primer design ...............................................................................44
Figure 6. Structure of ThermaGo and ThermaStop/PrimeSafe .........................................................48
Figure 7. Description of rpoB amplicon ..........................................................................................49
Figure 8. Example of amplification plot showing the four indicators of PCR efficiency ......................57
Figure 9. Example of FAM dissociation curve with On probes ..........................................................57
Figure 10. Example of FAM dissociation curve and signatures with Lights-Off Only .........................58
Figure 11. Example of Quasar dissociation curve .............................................................................58
Figure 12. Dilution series (5,000 copies to 5 copies) of Mtb with and without the IC template ..........60
Figure 13. Dilution series (1,000 copies to single copy) of wild type Mtb DNA with IC present ..........61
Figure 14. Obtaining Cq (Quantification Cycle) values .................................................................63
Figure 15. Determination of Slope for Efficiency Calculation .........................................................63
Figure 16. Amplification and melt curves of rpoB using LEL-PCR protocol ..................................65
Figure 17. LEL-PCR with lower Mtb copy numbers ........................................................................66
Figure 18. Comparison of LATE-PCR and LEL-PCR .................................................................67
Figure 19. New primers for rpoB and rd9 ......................................................................................68
Figure 20. Amplification with background human genomic DNA .....................................................69
Figure 21. Increased ThermaGo concentrations ............................................................................70
Figure 22. Performing melts in between amplification cycles ..........................................................71
Figure 23. Duplex with an amplifiable control ...............................................................................72
Figure 24. Amplification of standard BCG samples .......................................................................72
Figure 25. Comparison of Platinum Taq and normal Taq (with PrimeSafe) ......................................73
Figure 26. Comparison of old and new internal control limiting primer ...........................................74
Figure 27. A comparison of Invitrogen Taq with its 1x buffer and Syd Taq with its 1x buffer ..........75
Figure 28. A comparison of varying PrimeSafe and ThermaGo 86-21 concentrations ...................76
Figure 29. A comparison of 3mM (top) and 4.5mM (bottom) magnesium concentrations ..............77
Figure 30. A comparison of 100nM and 150nM TG86-21 .............................................................78
Figure 31. FAM dissociation of 0.24x and 0.48x SYBR concentrations ............................................79
Figure 32. A comparison of 100nM and 300nM concentrations of all six rpoB off probes ..............79
Figure 33. A comparison of 50nM and 100nM of the rpoB Limiting Primer (LP) ...............................80
Figure 32. Demonstration of improved resolution using sample R1810 ........................................81
Figure 33. Fluorescent signatures of 23 different mutational types as well as plots of biological and allelic replicates ...........................................................................................................91
Figure 34. Mixed Samples of wild type and mutant (rpoB) Mtb samples ..........................................92
Overview

A story is told of a frail little boy, plagued with chesty, hearty coughs, who as an adult spits and coughs up blood until he finally drowns in his own blood. Such is the tragic story of the satirical writer George Orwell (Eric Arthur Blair) author of Animal Farm: diagnosed with tuberculosis (TB) at the age of 35. He died 12 years later before finishing his next novel. A greater portion of his short immiserated life was spent battling this vicious disease. George Orwell died in 1950, just short of the dawn of a new era in TB diagnostics and treatment. During his lifetime, it took one to two months, at the earliest, to positively diagnose the disease using bacterial cultures. Unfortunately, this precious time was something many afflicted with the disease could not afford and still cannot afford.

Since that time major improvements have been made in the diagnostics, treatment, and cure of TB. Due to increasing knowledge of Mycobacterium tuberculosis (Mtb), the process of infection and the availability of antibiotics. Instead of the months for diagnosis, today's technology can render a diagnosis in under two hours in some cases. Also there are many more drugs available and approved for treatment of TB. Additionally there are promising arrays of new and repurposed old drugs in various stages of clinical trial for the TB treatment. Despite all of these advancements there are still many casualties attributed to the disease.

---

a Slight reference to text from book: “Let's face it: our lives are miserable, laborious, and short.” — George Orwell, Animal Farm
b Causative agent of TB
*Mtb* remains a deadly pathogen, earning the title of world’s most successful pathogen\(^1\). This title takes into account its steadfast and efficient morphology as well as its mechanisms to evade host immune system. *Mtb* turns the bodies defensive macrophages into its home and replicating factory\(^2\). Moreover, *Mtb* continues to evolve into a deadlier pathogen\(^3\). Numerous drug resistant strains have arisen around the world: multi-drug resistant MDR, extensively drug resistant – XDR, and even totally drug-resistant – TDR TB strains\(^4\).

In addition to the challenge of drug resistance, *Mtb* is difficult to detect in some high-risk patients, including children and people with HIV. In particular, children cannot cough out a sputum sample, nor can patients with other diseases that cause symptoms of dry mouth. Sputum samples from HIV positive patients are frequently scored as TB negative because their white cell count is low. Failure to diagnose the disease in these patients, especially drug resistant *Mtb* results in improper treatment, which, in turn, can lead to the spread of the disease and the increased incidence of drug resistance.

Thus the need for an extremely sensitive assay that can be applied to various sample types; including the minimally invasive blood samples and less invasive oral sample has been proposed. This paper will discuss an effort to better detect *Mtb* in such samples. The chosen gene targets include the *rpoB* gene\(^5\), the region of difference 9 (*rd9*) which is unique to *Mtb*, and insertion element, *IS6110*. An improved Linear-After-The-Exponential (LATE) PCR technology (developed in the Wangh Lab), has been utilized in these studies. This method has improved sensitivity and holds the promise of less costly technology which would making TB diagnostics

\(^{\text{c}}\) Particularly the RRDR region with known drug resistance (DR) associated mutations
available to a greater number of afflicted individuals who include many of the world’s most improvished people.
In the Beginning

About Mycobacterium tuberculosis:

History

_Mycobacteria_ are a genus of free living soil-dwelling bacteria that infect a wide array of animal hosts. The oldest species _Mycobacterium ulcerans_ traced back as far as ~150mya\(^5\) whereas the younger _Mycobacterium tuberculosis_ (Mtb) species were believed to have existed for a thousandth\(^d\) of that time\(^6\). Conclusive evidence of Mtb as a hosted infectious disease agent was procured in 17,000 year old bison remains\(^7\)\(^8\). However the exact time when Mtb emerged as a human specific pathogen is not clear. Estimates are given between 20,000 to 15,000 years ago when it was postulated to have evolved with the human hosts as they emerged out of Africa\(^9\). However weak connections have been made to 500,000 year-old _Homo erectus_ skeletal remains\(^10\). It is certain that the bacteria existed as a human host 9,000 years ago during the Neolithic period\(^11\). Contrary to original cattle zoonosis theory\(^12\) the host dependent association of mycobacteria to cattle(livestock) and human evolved independently but in the same time space\(^e\)\(^13\),\(^14\). with _Mycobacterium canetti_ as possibly the most closely related to last common ancestor (LCA)\(^17\). _Mtb_ since then affected human health and medicine, politics, economics, civil society and even literature and arts through TB

\(^d\) 150,000 years
\(^e\) Shown by sequencing and genomic studies focusing on single nucleotide polymorphisms (SNPs),
TB as an ancient human disease is evidenced by ancient writings of the Assyrians, who inscribed the symptoms on clay artifacts and also by Hippocrates in describing consumption- “phthisis” 15. Evidence that the disease has been present in its human host for at least 4-5,000 years can be found in mummified remains originating from ancient Egypt and regions in Middle East. These unearthed discoveries include the TB-scarred remains of King Tutankhamen (Egyptian pharaoh—circa1358-1340 B.C.)16. The disease affected many well known historical and prominent figures; including German dictator Adolf Hitler (1889-1945) and revolutionary Nelson Mandela (1918-2013). Others like famous American writer and poet Edgar Allan Poe (1809-1849), Elizabeth Darwin (1841-1851, Charles Darwin’s favorite child), Thomas Daniel Lincoln (1853–1871, youngest son of Abraham Lincoln), and first lady Eleanor Roosevelt (1884-1962) were killed by the disease17 18 19.

Ending the First Epidemic

The first outbreak reaching epidemic proportions devastated Europe in the 17th to 18th centuries and was thought to be introduced by Indo-Europeans14 20. At this time one in every seven people of the world’s population died to the disease. Various contributions made by early microbiologists, epidemiologist, scientists and physicians helped curb the disease. Robert Koch in 1882 was the first to initially stain and culture the bacterium on solid medium, identifying it as the disease causing agent21 22 23. The rod-shaped bacterium was named a year later. Before he had identified it, Jean-Antoine Villemin had shown the disease could be transmitted from ailing to healthy persons in various studies disproving earlier theories that the disease was inherited or spontaneous24 25. Koch not only identified the bacterium but also showed that it was the disease
causing agent using what we now refer to as Koch’s postulate, providing further evidence for transmittability of the disease. Using the new information many projects begun in efforts to halt the epidemic. Leon Charles Albert Calmette cultured the tuberculin protein produced by Koch and applied it to a diagnostic test: the tuberculin skin test. Calmette together with Camille Guérin discovered *Mycobacterium bovis* with reduced virulence and introduced the BCG vaccination. All these collectively reduced the severity of the epidemic but the decisive blow was dealt with the discovery of streptomycin by Selman Waksman and Albert Schatz. The initial epidemic steadily dropped in the 20th century especially in developed countries with improving healthcare practices, available and effective first-line and second-line antibiotic treatment, BCG vaccine and lastly and questionably natural selection.

The Recent Epidemic

**The Horror Today**

The advent of AIDS in the 80’s and 90’s coupled with economic factors refueled the disease back to epidemic status. The current epidemic was declared in 1993 at the Geneva convention when it was projected that TB would claim about 30 million lives in the next ten years. In efforts to halt the epidemic and reduce the estimated mortality, various global targets were set in the previous *Stop TB Strategy (2006-2015)* and the *End TB Strategy (2016-2035)*. These goals fuel research and projects all over the world. A target set for the *Stop TB Strategy*, of the Millenium Developmental Goals (MDGs), was to halve the TB prevalence and associated

---

This involved three steps: after 1) the presence of the agent in diseased tissue was shown, 2) it was isolated and cultured in vitro and after 3) reintroduction into a non-infected host led to disease.

Part of the Millennium Developmental Goals (MDGs)

Part of the Sustainable Developmental Goals (SDGs)
death: the reference baseline for the target was the 1990 statistics. Halting and reversing TB incidence was also another target of the MDGs. According to the 2015 Global TB report, since the year 2000, TB incidence fell at an annual rate of 1.5% globally with a 18% cumulative reduction in incidence. Yet prevalence\(^i\) and mortality\(^i\) were not reduced equally across all six WHO regions, and TB continues to plague earth’s inhabitants: especially those in underprivileged economically challenged overpopulated regions. WHO estimated 9.6 million new cases worldwide in the year 2014 alone, with 1.5 million deaths. About 400,000 of those deaths were attributed to patients with HIV (1.2 million HIV/TB cases). The majority of cases – 58% – are in the South-East Asia and Western Pacific regions.

With 28% of the total cases, the African region had the highest burden to population ratio (281\(^k\) cases for every 100,000 people). The highest number of cases were recorded in India, Indonesia and China (23%, 10% and 10% respectively of total cases). In the Americas, there were 228,476 total cases in 2014, with 60,468 total Rifampicin resistant (RR) or MDR TB\(^40\). Today, it is estimated that a third of the world’s population is infected with latent TB; this causes 7-8 million cases of active TB with a death toll of about 1.3-1.6 million annually\(^41\). These alarming statistics underpin the newly established and ongoing *End TB Strategy*, approved by all Member States at the May 2014 World Health. It targets a 95% reduction in TB deaths and a 90% reduction in TB incidence and a target of zero catastrophic costs for TB-affected families by 2020. At the current

---

\(^i\) Achieved in three (3) WHO regions – the Americas, the South-East Asia and the Western Pacific Regions – and in nine high-burden countries (Brazil, Cambodia, China, Ethiopia, India, Myanmar, the Philippines, Uganda and Viet Nam)

\(^j\) Achieved in four WHO regions – the Americas, the Eastern Mediterranean, the South-East Asia and the Western Pacific Regions – and in 11 high-burden countries (Brazil, Cambodia, China, Ethiopia, India, Myanmar, Pakistan, the Philippines, Uganda, Viet Nam and Zimbabwe)

\(^k\) Average of 133
rate of improvements and TB incidence, prevalence and mortality decline, these goals will not be met in a timely fashion. Also the incidence and mortality attributed to DR-TB only increases with time. If this goal is to be achieved, easy, fast, reliable diagnosis of TB and drug resistance is a key component in this overarching goal to stop TB\textsuperscript{42}.

**Back with a Vengeance**

The instances of multi- and extensively- drug resistant TB (MDR-TB, XDR-TB\textsuperscript{1} respectively) are rising\textsuperscript{43, 44}. There are even some documented cases of “totally” drug resistant TB \textsuperscript{3, 45}. Cases of “totally drug-resistant” TB showing resistance to all 11 frontline drugs are being reported and these could be due to evolving impenetrable cell wall of the bacteria. To combat these more robust new lineages and strains, new drugs will need to be developed\textsuperscript{46}. Drug resistance, and even worse extensive drug resistance, has been shown to further reduce treatment success. Thus the availability, accuracy and reproducibility of current drug susceptibility testing methods is crucial to these outcomes\textsuperscript{47}. The long term treatment required for cases of XDR TB in areas without the necessary infrastructure further complicates the situation as released patients with failed treatments further spread the disease\textsuperscript{48}. This has led to multiple debates on the return of santoriums and their possible advantages and effects\textsuperscript{49, 50}. This infection is greatly affecting the population and human force by literally reducing it; it affects the economy and even our social lives with the stigma it induces\textsuperscript{51, 52}. To combat this disease we need to discover even more about the activity and strategy of the bacterium, especially from an evolutionary viewpoint. Learning

\textsuperscript{1} MDR-TB: TB caused by Mycobacterium tuberculosis bacilli that are resistant to at least isoniazid and rifampicin. XDR-TB: TB caused by Mycobacterium tuberculosis bacilli that are resistant to rifampicin, isoniazid, plus any fluoroquinolone and at least one of the three injectable second-line drugs: amikacin, kanamycin and capreomycin.
more about the evolutionary mechanisms and tactics of the bacterium will enable us to understand its revived and increased potency and virulence\textsuperscript{53}. 
The *Mycobacterium tuberculosis* Complex (MTBC)

MTBC is used to refer to the members of the *Mycobacterium* sp. that can cause TB in a variety of hosts. The species are *M. tuberculosis*, *M. bovis*, *M. bovis BCG*, *M. mungi*, *M. microti*, *M. africanum*, *M. pinnipedii*, *M. caprae*, *M. canetti*, and *M. surricatae*. **Table 1** lists common MTBC members and Non-Tuberculous *Mycobacterium* (NTM) and the associated diseases. The small (0.2-0.4μm in length) members of the MTBC are characterized as bacillus (rod-shaped bacteria) with cell wall, central DNA (nucleoid lacking an envelope), the cell membrane, a capsule, and ribosomes as cytoplasmic components. Though the microbes lack flagella and other structures for motility, they are not immobile, as was first believed. New research suggests that *Mtb* may move through a “sliding” motion over a surface\(^5\). As mentioned before, they are minute and can easily infiltrate the lungs traveling in a small aerosolysed droplets from an infected person. The success of the disease traces back, largely, to its morphology and basic structure.

**Basic Morphology**

**The Capsule**

The capsule of *Mycobacterium* species that enclose the cell wall have been shown to have glycopeptidolipids (GPL) and phenolic peptidolipids that are used in the “sliding motion”. These amphoteric substances are specific to the *Mycobacterium* species involved. This is believed to
enable them to more effectively inhabit organisms\textsuperscript{54}, as they are able to spread over larger surfaces.

**Complex cell wall**

This gram-positive bacterium has a fluid cell wall, which unlike other regular bacillus is very thick and is composed of three layers: the mycolic layer (comprised of alphamycolate, methoxymycolate and ketomycolate), the peptidoglycan layer and the lipid bilayer\textsuperscript{55}. The cell wall is another component believed to aid the microbe in its infectious ability as well as antibiotic resistance. It has a complex array of 528 total identified proteins that aid in motility and some are even known antigens of T and B cells\textsuperscript{56,57}. A main component of the cell wall believed to be involved is the oxygenated composites of the mycolic layer, which for *Mtb* includes methoxymycolate and ketomycolate components. Though these compounds are sparsely distributed over the cell surface, they impair the microbe at normal temperatures (stable at 37°C) and make the microbe hypersensitive to ampicillin and rifampicin\textsuperscript{58}. This also reduces permeability significantly and makes it hard for macrophages to digest the bacteria. Additionally, the thick cell wall prevents macrophages from digesting the microbe by virtue of its size and also through signaling: signaling prevents the formation of phagolysosome. The cell wall size has also been linked to resistance: thickness of cell wall increases from susceptible to MDR and XDR strains respectively\textsuperscript{59}.

**Long Generation Time**

The thickness of the cell wall confers a long reproducing time, about seventy two times the normal rate for bacteria\textsuperscript{m}. This has been attributed to the complexity of its cell division:

\textsuperscript{m} About 20 mins
coupled with the fact that the cell has to synthesize portions of its complex cell wall, it has missing some key proteins for bacterial cell division and hence must compensate using other complex mechanism\textsuperscript{60}. This can be viewed as a tradeoff between virulence and survival. The bacterium’s virulence is reduced as it is not rapidly dividing and overtaking host cells. However this is an advantage as it reduces its chances of being noticed and tagged as hostile by the host cells defences. Controlling its resources and virulence by lengthening division time is hence critical to its success as pathogen.

**Genetic component and adaptability**

The members of the MTBC’s ability to adapt and survive extreme environments is largely due to protein production. A large percentage of the many genes that have been identified are related to the production of glycoproteins, cell wall components, and other sugars and proteins that are largely found on the cell surface\textsuperscript{61}. The MTBC genome has been found to be gradually deleting the less useful genes to conserve the more useful ones. Thus they constantly assess their genome to be even more efficient\textsuperscript{62}. The efficiency of the members of the MTBC in managing their genome can be traced back to their evolution, specifically their coevolution with the human host. This reflects on the natural struggle in the coevolution between the host and pathogen.
Table 1. Some notable disease-causing members of the Mycobacterium family

<table>
<thead>
<tr>
<th>Species</th>
<th>Disease</th>
<th>MTBC/NTM</th>
<th>Human host</th>
</tr>
</thead>
<tbody>
<tr>
<td>M. kansasii</td>
<td>Pulmonary, disseminated disease</td>
<td>NTM</td>
<td></td>
</tr>
<tr>
<td>M. marinum</td>
<td>Cutaneous disease</td>
<td>NTM</td>
<td></td>
</tr>
<tr>
<td>M. avium</td>
<td>Lymphadenitis</td>
<td>NTM</td>
<td></td>
</tr>
<tr>
<td>M. abscessus</td>
<td>Pulmonary, cutaneous disease</td>
<td>NTM</td>
<td></td>
</tr>
<tr>
<td>M. xenopi</td>
<td>Cutaneous disease</td>
<td>NTM</td>
<td></td>
</tr>
<tr>
<td>M. malmoense</td>
<td>Pulmonary, Lymphadenitis</td>
<td>NTM</td>
<td></td>
</tr>
<tr>
<td>M. tuberculosis</td>
<td>TB</td>
<td>MTBC</td>
<td>Frequently</td>
</tr>
<tr>
<td>M. caprae</td>
<td>TB</td>
<td>MTBC</td>
<td>Occasionally</td>
</tr>
<tr>
<td>M. microti</td>
<td>TB</td>
<td>MTBC</td>
<td>Rarely</td>
</tr>
<tr>
<td>M. canetti</td>
<td>TB</td>
<td>MTBC</td>
<td>Rarely</td>
</tr>
</tbody>
</table>

There are over 125 recognized NTM species and others yet to be identified. The above table shows the 18 common NTMs and 8 common MTBC members.
Coevolution

The Theory of the Coevolution of TB and Host

Despite the long history of members of the MTBC with numerous hosts, there seems to be a preference for human hosts. The human-hosted bacterium does not cross-infect other animal species. This is believed to be a result of the coevolution of the microbe and the host, allowing the microbe to increase its virulence in the human landscape while losing its choices in possible hosts and the benefits of cross species transmission. This tradeoff is one that may be more beneficial as humans are numerous and mobile so the coverage of the microbe is not as limited\textsuperscript{64, 65, 66, 67}. This coevolution is remarkable as some species of MTBC show specific adaptions to hosts in specific regions: that is they may be drug resistant in some regions but still remain drug susceptible in other regions\textsuperscript{68}. There are believed to be 7 major lineages existing today\textsuperscript{69}. All seven are found in East African regions but exist in various combinations across the globe. Of these seven lineages 2-4 are the modern forms while lineage 1, 5 and 6 are thought to be the more ancient forms\textsuperscript{70}. The modern forms are known to have entire sections of genes deleted, TbD1 in lineages 2-4, which are present the ancestral forms.

The evolution of these microbes is not autonomous and is largely affected by the host. The ability of the host to migrate largely affects its evolution. Not only does host migration affect the spread of the microbe but it brings it to a new environment which leads to natural selection, increased adaptations and hence evolution. For example, migration into largely populated, poor
areas give the microbe a chance to spread and diverge, increasing its virulence. Major movements of populations from Africa and regions in Asia played a significant role here. In addition to the geography, there are several human host-specific events that have enabled the evolution of the microbe. There are several events that have been traced to coincide with the increase in strains of MTBC. These include wars: the First and Second World Wars and the Cold Wars. George Orwell was described to be affected by this as he stayed long periods in poverty-stricken, war ravished areas with poor aeration. Also the HIV explosion in the late 19th to early 20th century increased the bacteria’s virulence as it provided a weaker immune system, in which the bacterium can be active. Additionally, the virus affects the detection of the bacteria. Since this make it hard to diagnose TB in HIV infected individuals, the treatment is in turn very difficult. The explosion in HIV made the microbe even more virulent. These and other host-specific characteristics like poverty, poor infrastructure and overpopulation, make it a preferred host and in turn adds to the evolvability of the microbe.

**Cell Wall Specialization**

One aspect of this evolution is the adapted cell wall. Though other members of the *Actinobacteria* family have mycolic cell walls, some are not covalently linked to the peptidoglycan cell layer as in *Mycobacteria*. Also they have been shown to have different properties including a higher melting temperature and different stress response. Furthermore, there is evidence that the evolution of *Mycobacteria* in relation to its close *Actinobacteria* relatives places importance on the production and regulation of lipids. There is an observed difference in the profile of genes relating to metabolism of fatty acid. As already noted the cell wall plays an
important role in its virulence, thus it falls in line that the cell evolve to maximize the efficiency of production and modification of its components.

**Host Related Adaptations**

The cell wall of MTBC is specially adapted to the host as well. It is able to evade the host immune system response by preventing the repair of broken cell membrane\(^75\). Also, the pathogen does not reduce its immunogenicity but rather remolds the tissue in the area it affects capitalizing on the delayed response time of the immune system. This remodeling includes the exclusion of lymphocytes and the influx of lipids disregarding lipid homeostasis in the area\(^76\). In addition to interfering with the host \(Mtb\) recognition pathway, the cell wall is also adapted to utilize host lipids and toxic waste. The bacteria are able to utilize the toxic waste, succinyl-coA, created by the host during respiration and incorporate it into its highly tough and complex cell wall. This is used in the manufacture of cholesterol to strengthen the peptidoglycan wall. This is relevant in that it shows more of a mutual symbiosis than a parasitic relationship. The host is able to rid the cells of toxic material without having to rely on energy costing pathways and the bacterial in turn gain materials for its development. This may influence the cell’s reaction to the microbe and provide a possible explanation for the long latent phase as the bacteria relives the cell of the stress of converting the harmful substance into waste or the process of recycling\(^77\).

**Major Means of Evolution**

Major pathways that are recognized to drive the Mycobacteria evolution include horizontal gene transfer (HGT) and the deletion of genes\(^78\). Though duplications are generally credited for the most significant evolutionary process as they lead to co-option and
specializations, they are shown to not significantly affect the evolution of the microbe. In bacteria, HGT and deletions play a more important role. MTBC is known to delete dormant and non-useful portions of its genome, ranging from gene libraries to bases, to create and maintain a compact library of useful genes. Most conserved non-transcribed regions have been predicted to be promoter regulatory regions and the like. Mtb’s use and maintenance of its gene libraries is one of the reasons why the disease is so difficult to eradicate today.

Why TB Has Not Been Eradicated

Because the first epidemic was successfully curbed one would imagine that there are treatments and vaccines available. That is indeed true as the disease is treatable and curable. However, this new wave of TB has gained momentum because of various factors acting in unity. Poverty is a well-recognized factor that not only affects the quality of life of patients but also access to proper treatment and diagnosis. Although advancements have been made in diagnostics, there are still challenges and restraints in various focus groups like infants and patients that cannot produce sputum for testing. In addition to this, the HIV pandemic (since 1990) further complicates diagnosis and increases the rate of active TB. There is also the emergence of drug resistant TB that reduces the efficacy of drug regimens and limits the available drugs for treatment. The low efficacy of available useful vaccines renders them ineffective. Thus it is essential that diagnostic testing which provide resistance information to allow quicker, more directed treatment is prioritized. The nature of the disease and its life cycle provides insight into where and how to find the microbe and diagnose the disease.
Tuberculosis as a Disease

Disease Cycle

Once an uninfected person inhales aerosolized droplets of *Mtb*, primary infection, the disease can take one of two pathways: progress to the latent or active stage. In the primary infection, the microbe gains access to the subpleural cavities of the right lung (usually the lower lobe) through the mucosal lining and are phagocytosed by macrophages but may also infect type II alveolar epithelial pneumocytes-AEP\(^n\). This, however, does not lead to the clearance of the pathogen as it evades the cell’s killing mechanisms in ways detailed below. The antigen presenting cells-APCs- (Dendritic cells –DC- more so than macrophages) are activated to recruit T-Lymphocytes to the site of infection where they surround infected macrophages, forming a granuloma: referred to as “caseation necrosis” translating to cheesy death, named after its resemblance in appearance to cheese. The Ghon Focus (primary lung lesion and infected lymph node-LN) is formed by now as the presentation of microbe introduced the infection to the LN. The battleground, the granuloma, at this point consists of both dead microbes, host cells and some live *Mtb* pathogens centered in the foamy encasement created for survival.

The infection may be stalled and the microbe may remain dormant for many years, undetected in the body, in the latent phase and little is known about it. This phase may last a few

\(^n\) More abundant in alveoli
months or sometimes longer (up to 11 years with a 5-15% lifetime risk of developing the active disease\textsuperscript{90}. In this case, the infection progresses to the active pulmonary disease (occurring in the lungs) where multiple lesions are formed (milliary TB), possibly leading to progressive lung disease (usually wrought by HIV, and malnutrition) and severe destruction of lung tissue. The primary infection may however be effectively stalled and reactivated through a second reintroduction of the virus, termed secondary infection. This usually leads to extrapulmonary tuberculosis occurring in areas of the body other than the lungs. Extrapulmonary forms of TB include tuberculous lymphadenitis, tuberculous peritonitis, urogenital, musculoskeletal, cardiovascular, gastrointestinal, ocular, cutaneous, miliary, endocrine and metabolic tuberculosis. Tuberculosis may also affect the central nervous system, liver, biliary tract, and pancreas. Despite the ubiquity of this bacteria, locating it in its latent phase is very difficult. There have been reports of the microbe existing in the bone marrow in the latent phase\textsuperscript{91}. There are a few tests that have been developed using the transcriptional activity of the host as an indicator of the presence of the latent form of \textit{Mtb}. However, it has been discovered that proteins synthesized in the latent and active phases are not entirely separate\textsuperscript{92}. Most of the transcripts that are targeted for diagnosis of the latent phase are also present in the active phase of the disease.
Immunobiology

Cellular Response

After gaining entry to the subpleural cavities of the lungs through alveolar passages, *Mtb* may infect constitutively expressed macrophages (primary host cells) and type II AEPs. Microbes gain entry to these cells through various means including macrophage mannose receptors, complement receptors and toll-like receptor 2 (TLR2). Surfactant protein A, on alveolar surfaces, up-regulates mannose receptor activity and facilitates this process. As stated earlier, for antigen presentation, circulating DCs are more efficient. They introduce *Mtb* antigens (includes lipoarabinomannan - LAM, heat shock proteins - HspX, Esat6/CF) to T-cells to initiate their activation: thereby spreading the infection to the LN. Differentiated Th1 CD4+ T-cells are usually the main cells responsible for the adaptive immune response. They participate in granuloma formation by walling off areas with recruited macrophages: macrophages typically fuse and become foamy due to imported low-density lipoprotein (LDL) and other cholesteroloxides.

Involved Immune Cells and Killing Mechanisms

Once engulfed in the endosomal phagosome, if the process of normal maturation is not interfered with, a phagolysosome is formed after fusion with lysosome where the microbe is killed employing reactive oxygen intermediates (ROI), reactive nitrogen intermediates (RNIs), lysosomal enzymes, and toxic peptides in an acidic environment. Th1 cells produce
IFN-γ and CD40 ligand activate macrophages to destroyed phagocytosed bacteria; they also produce 1) IL-2, 2) IL3 and GM-CSF, TNF-α, and 3) LT-α and CXCL2. They induce apoptosis through the Fas (CD95) ligand system and LT-beta. Cytotoxic T-lymphocytes (CTL) contribute slightly to cell lysis and microbe killing. Recruited bactericidal neutrophils employ methods such as phagocytosis, antimicrobial peptides, ROI, RNI but also only contribute slightly to this infection. Natural killer cells that secrete IFN-γ, IL-15, and IL-18, to regulate CTLs contribute slightly in the absence of APCs. Humoral immunity may play a minor role. B-cells produce antibodies that opsonize microbes making them better targets for phagocytosis and subsequent destruction. Opsonization also improves antigen processing and presentation by neutrophils, macrophages and DCs. However, in some cases it was shown to have no effect or elevate success rates of eliminating Mtb: that is with antibodies specific to unique Mtb antigens such as lipoprotein, α-crystallin, and LAM. Aspects of innate immunity such as granulocytes (most phagocytes with the exception of macrophages) and aspects of adaptive immunity (CTL and humoral immunity) play a non-pivotal role in controlling the infection.

Survival Strategy

The mechanisms mentioned above will clear the microbes if not inhibited. However Mtb has several devices to inhibit many of these mechanisms including down-regulation of IL12, MHC II and host cell apoptosis. One major mechanism is the prevention of the formation of the phagolysosome by inhibition of phagosomal maturation. The exact mechanism that leads to

---

o For T-cell proliferation and differentiation
p For induced macrophage differentiation in the bone marrow
q Aid macrophage recruitment to infected site
this outcome is not known though several explanations have been proposed. Prominent among these include phagosomal conditions like very weak acidic environment, absence of proton ATPases, inhibited Ca\(^{2+}\), and absence of Rab 7 protein (small GTPase involved in endosome trafficking of late phagosome). This becomes an advantage as the microbe is not killed but rather utilizes the cell as a production factory. This mechanism is made even more effective when *Mtb* prevents the apoptosis of infected cells employing anti-apoptosis factor like *nuoG* and NDH-1 and other mechanisms. Foamy macrophages also express high levels of the antiapoptotic Bcl2 molecule. The macrophage’s failure to commit suicide not only blocks the cell’s defense mechanism but also provides security for microbes that are shielded from lymphocytes in these environments with low MHC-II presentation. Other survival mechanisms include the induced apoptosis of Th1 cells through the Fas/Fas ligand mechanism by foamy macrophages. Also some strains of *Mtb* show resistance to RNIs. Additionally some cell induced immune mechanisms aid the progression of the bacteria. An example beside phagocytosis is the induced Inflammation and production of TH1 cytokines such as TNF-\(\alpha\), which causes extensive tissue and early death damage usually a result of severe lung inflammation.

Pathogenicity and Virulence

*Mtb* virulence is dependent on morbidity and mortality rate and hence the gene encoded virulence factors are different in comparison to other bacterial pathogens. While most phagocytosis and subsequent killing of microbes or antibody-opsonized *Mtb* may be inhibited as elevated levels of Ca\(^{2+}\) are associated with phagosome formation. Also Ca\(^{2+}\) associated to many host immune responses such as respiratory burst as well as nitrous oxide (NO) and cytokine production.
bacterial pathogens have virulence factors linked to rapid colonization of host cells and toxin production, *Mtb* in contrast has virulence factors related to its survival under anoxic, nutrient deprived, hostile environments of macrophages: all features of its slow colonization rate and latent phase\textsuperscript{134}. Some of the recognized genes contributing to virulence include culture filtrate proteins (CFPs) like KatG (catalase-peroxidase) and SodA (superoxide dismutase) that degrade ROIs. Additional genes, including *Mas*, *FadD26*, *MmpL7*, *MmaA4*, and *OmpA*\textsuperscript{135,136,137}, code for cell surface protein unique to pathogenic mycobacteria. The genes also code enzymes involved in host fat metabolism and utilization, after the microbe switches from consuming carbohydrates to fatty acids. This category encompasses more than 200 genes\textsuperscript{138} including *Icl*, *LipF*, *FadD33*, and mce4 transporter\textsuperscript{139}. Some of the virulence genes, such as *MgtC* and *MbtB*, are involved in metal intake. Some of the genes code for enzymes that help the bacteria in the anoxic phagosome such as *Nitrate reductase, KatG*, and *AhpC*. Some Transcriptional Regulators turn on a range of other genes associated with the pathogens survival in the unfavorable environment inside the macrophage. Notable among these are the sigma factors (*Sigma A*\textsuperscript{aa}, *Sigma F*\textsuperscript{140}, *Sigma E*\textsuperscript{141}, and

---

\textsuperscript{s} Catalyzes the synthesis of long-chain, methylated branched fatty acids
\textsuperscript{t} Creates pores in liposomes
\textsuperscript{u} Allows bacteria and plants to grow on acetate or fatty acids as sole carbon source
\textsuperscript{v} Lipase/esterase that may function in lipid degradation
\textsuperscript{w} 36 fatty acid metabolic genes
\textsuperscript{x} Enzymes whose function is to synthesize mycobactin and carboxymycobactin, the major siderophores in *M. tuberculosis*
\textsuperscript{y} A catalase: peroxidase that degrades H₂O₂ and organic peroxides
\textsuperscript{z} An alkyl hydroperoxide reductase, and enzymes of this type function to detoxify organic hydroxyperoxides
\textsuperscript{aa} Essential principal mycobacterial sigma factor: presumably necessary for most mycobacterial housekeeping gene transcription
Sigma $H^{142}$ and Pho response regulator ($PhoP^{143}$). Some other transcriptional regulators like IdeR, RelA$^{144}$, Hsp$R$ and Whi$B^3^{bb}$ contribute to the pathogen’s virulence.

Knowing these virulence factors is key in the treatment of TB. Often they are made targets the aspects of antibiotics. The available antibiotics target the cell wall and other key metabolic functions conferred by some virulence factors. Streptomycin, the initial drug of choice, worked by inhibiting protein synthesis and hence blocking many of the important pathways for the microbes virulence.

---

$bb$ Responsible for sporulation and cell septation: the persistence or latent state of $Mtb$ is analogous to bacterial sporulation
Treatment of TB

Since the introduction of streptomycin to treat TB in 1964, many other antibiotics have been used and many have become useless because the bacteria are so steadfast and have various mechanisms for evading the host immune system. Currently, multidrug treatment is the only means of fighting the disease. This involves taking several medications for at least half a year. To ensure that the right medications are taken for the right amount of time, there are global guidelines to the treatment of the disease.

Available and Approved Drugs

There are several antibiotic drugs approved for the treatment of TB and are grouped into five classes\textsuperscript{145, 146}. Group one, the first-line oral drugs, consist of pyrazinamide, ethambutol, isoniazid and rifampicin\textsuperscript{cc}. Apart from streptomycin, which is also a first-line drug although it is a group two injectable drug, all other drugs are second-line drugs. Other group two injectable drugs are kanamycin, amikacin, and capreomycin. The fluoroquinolones are group three and include levofloxacin, moxifloxacin, and ofloxacin. Group four oral bacteriostatic second-line drugs include para–aminosalicylic acid, cycloserine, terizidone, thionamide, and protonamide. The role of group five drugs in treatment of drug resistant TB is still unclear and as such, they are rarely used in treatment regimens; they include clofazimine, linezolid, amoxicillin/clavulanate,

\textsuperscript{cc} Also known as rifampin
thioacetazone, imipenem/cilastatin, high dose isoniazid, and clarithromycin. First-line drugs are used in the treatment of drug susceptible TB. For drug resistant TB, second-line drugs are used. With the wide array of first and second-line drugs, the WHO has issued global guidelines in the use of these drugs in the treatment of TB\textsuperscript{147}.

**General Guidelines**

Although TB is curable, the task of treatment is long and unpleasant. To ensure the success of treatment the WHO has set global standardized guidelines for treatment. Multiple-drug therapy is the main method to combat the disease\textsuperscript{148}. As the name suggests, this involves giving patients several drugs over the course of treatment. All drugs have different targets and this increases the rate of success. That is, where one drug may fail to eliminate the bacteria, another drug will be active. In addition, this reduces the possibility of emergence of drug resistance. The six-month treatment schedule is divided into two parts. The first part consists of an antibiotic regimen of four first-line drugs: isoniazid (INH), rifampicin (RIF), pyrazinamide (PZA), and ethambutol (EMB) for two months. This is followed by four months of INH and RIF. This cocktail of drugs is not absolute. For example, HIV patients on antiretroviral drugs cannot take RIF due to documented drug interactions. Age, pregnancy and drug susceptibility are also considered when choosing the drugs to be taken for the treatment period. Directly observed therapy (DOT) is the method that is employed in the treatment of TB. To prevent prolonged treatment and the development of drug resistance patients are observed when taking their medication in order to avoid issues of compliance and to ensure that the regimen is successfully completed.
Drug Susceptibility and Resistance

Despite efforts made to halt TB drug resistance, resistance to antibiotic agents is on the rise. Cases of MDR-TB are increasing and in 2014, WHO reported about 190,000 lives out of the estimated 480,000 who develop the disease were lost\textsuperscript{40}. In case of MDR-TB, second-line drugs are used in the treatment, but it is often not successful. The average success rate for MDR-TB is 50\% and the drugs are more expensive with greater side effects. In addition, treatment regimens normally last longer, increasing the possibility of noncompliance. For XDR TB when some second-line drugs cannot be used the rate of success is even lower, about 30\%. Due to these discouraging statistics, there is a quest for new TB antibiotics, and especially combination of antibiotics that are feasible in the frame of multiple-drug therapy\textsuperscript{149}.

New Drugs

Although there are many types of antibiotics available for the treatment of TB, there exists a persisting and ever-growing need for new antibiotics and antibiotic combinations. Issues such as the high cost, high toxicity and low efficacy of second line drug resistant TB drugs are a major driving factor for this need\textsuperscript{150}. Including the required long period for drug administration, unfavorable drug interactions (especially with antiretroviral drugs) and in some cases shortage of drugs also play a role\textsuperscript{151}. To answer this need there is various ongoing clinical trials\textsuperscript{152} or repurposed old antibiotics, new antibiotics and new versions of TB antibiotics\textsuperscript{153,154}. Promising among these is sirturo whose active substance is bedaquiline. Currently it is available when other drugs are in low supply. In addition, nitroimidazole OPC-67683, or delamanid is being developed by the Otsuka Company and is promising\textsuperscript{155,156,157,158}. The company began phase 3 trials of the
drug in 2011. Another potential drug, AZD5847, by AstraZeneca is undergoing phase 2a trials in South Africa, which started in Dec 2012\textsuperscript{159}. Bayer and the TB Alliance are also developing promising fluoroquinolones to add to those already available in the second line TB drugs\textsuperscript{160}. With all these developments, treatment of TB and drug resistant TB should improve in coming years.

**Beyond Drug Therapy**

Although drug development offers hope for the future of TB treatment, data on drug resistant TB casualties suggest this alone is not enough. Of the 480,000 people estimated to have developed MDR-TB, almost 10\% of them have XDR-TB strains. WHO estimated that MDR-TB cases are composed of about 20\% of previously treated cases and 3.3\% of new cases. This data suggests the diagnostics and treatment are failing to fully capture drug resistance cases. In the 2014 WHO report, of the people estimated to have developed MDR-TB, only 123,000 of them were diagnosed. Furthermore, about 10\% of the diagnosed cases had not started MDR treatment\textsuperscript{40}. Almost three quarters of drug resistance cases go unnoticed. It is clear that currently not all diagnosed cases of drug resistant TB are treated. However, when we fail to identify such a vast majority of cases, hope for reducing and even dealing with issues of drug resistance dwindles. Diagnostic tools have improved and the number of cases detected globally has increased. TB diagnostics has come a long way but there are existing and new concerns not addressed; these include but are not limited to making test, more assessable to children and developing fast, cheap reliable tests for a point-of-care system.
TB Diagnosis

History

Initially, very invasive methods, sometimes surgery, were used to diagnose TB. These were replaced by culturing methods, or skin antigen tests. These increased the tuberculin sensitivity but had low sensitivity and specificity and symptoms could be attributed to preexisting respiratory conditions, as in the case of George Orwell. The introduction of drug susceptibility testing set higher standards. It allowed the association of mutations in corresponding gene targets of drugs and made possible the design of molecular tests for drug susceptibility. Though drug susceptibility testing is still used today, molecular diagnostic test are the frontrunners in TB diagnostics.

Recent Introduction of Molecular Diagnostics

Advances made in the diagnostics of TB from classical culturing method has made way for molecular testing techniques such as the Gene Xpert MTB/RIF assay and Genotype MTBDRplus: both approved by WHO for global use. They are effective and useful tools for the detection of TB in adults\textsuperscript{161}. The Xpert MTB/RIF (Xpert; Cepheid; Sunnyvale, CA) assay is a novel diagnostic tool that uses the rpoB as a gene target, specifically the 81bp rifampicin resistance-determining region (RRDR), and is able to detect TB and provide drug susceptibility information in about 2hours\textsuperscript{162}. The GenoType MTBDRplus assay (MTBDR; Hain Lifescience GmbH, Nehren, Germany) has probes two more targets in addition to rpoB (katG and inhA genes associated with isoniazid (INH))
resistance) and hence is able to provide information on MDR\textsuperscript{163}. These new assay have been shown to be useful in TB burdened areas in Africa and many new cases the rate of diagnosing the disease has increased globally.

However, there are still disadvantaged groups like HIV infected TB patients, kids, infants and sick people that are unable to produce enough sputum for testing. Other tests including the urine antigen assays for HIV patients\textsuperscript{164}. Recent advances of molecular assays for the detection of TB (including the Xpert \textit{MTB/RIF}) have improved the detection rate, however there is still the need to reduce the approximately 3 million undetected cases\textsuperscript{40}. Rapid diagnostics of respiratory tract infections (RTIs), including TB, are crucial elements in overcoming this mountain of undetected cases. They provide direct information on the pathogen causing the disease which enables more directed care instead of trial and error with a broad range of antibiotics which may contribute to increased resistance\textsuperscript{165}.

What Has Not Been Achieved

Though alot has been achieved in diagnosing TB, the available techniques are far from perfect. Smear tests have been shown to have low sensitivity and specificity in addition to temperature and treatment requirements that further affect the results\textsuperscript{166,167}. Though culturing methods are better, detecting TB by culturing traditionally takes about 4-8 weeks and requires additional testing\textsuperscript{168}. New culturing techniques have been able to reduce this time to about 5-14 days but this still falls short of rapid detection rates\textsuperscript{169,170,171}. WHO has deemed serological tests inconsistent and imprecise and as such do no recommend them\textsuperscript{172}. 
Multiplex assays like the multiplex ligation-dependent probe amplification (MLPA), that simultaneously detect TB and give information on multiple drug resistance use many gene targets including the rpoB, katG, inhA, embB, 16S rRNA, IS6110 are cost and material ineffective\textsuperscript{173}. Blood based PCR techniques\textsuperscript{174} have been developed some using multiple targets(IS6110 and MPB64) and backlight cards. Though this is a cheaper and fast alternative, it still does not provide information on drug susceptibility. A multi-probe PCR assay using blood culture broth has been developed to identify \textit{Mycobacterium} species in about 3hrs\textsuperscript{175}. Though the cost induced is lower there is still no drug resistance info provided due to the chosen gene target. Though there are assays to differentiate the MTBC members from the NTM members, they do not provide information on drug susceptibility and are not effective clinical tools\textsuperscript{176}.

One may discount improving diagnosis in search of other ways to alleviate the TB burden. There are other ways to tackle the epidemic. An analysis of data sets on TB biomarkers may provide possible targets for new TB diagnostic tools\textsuperscript{177}. However these still require more research and analysis and ultimately will be more time consuming. Recent advances in TB drug discovery and development includes the submission of new drugs like delamanid and bedaquiline, and new and updated TB trials and trial designs, TB biomarkers and host-directed therapies\textsuperscript{178}. A possible new vaccine for infants, MVA85A which was modestly tolerated but lacked efficacy against tuberculosis or \textit{M tuberculosis} infection is one of the attempts at reducing the number of affected infants\textsuperscript{179}. This and other similar approaches need further research and will take more time to perfect and become readily available to the public for use and consumption. In the mean time,
simpler, less invasive and dependable diagnostic methods can alleviate the toll of TB on kids and infants.
Aim of This Research

One person dies every 3 minutes in India due to TB\textsuperscript{180}. This number has decreased significantly and to keep the downward trend, fast, cheap, reliable and comprehensive diagnosis is key. The standard for diagnostic tools set in the 3\textsuperscript{rd} edition International Standards for TB care emphasize the need for reduced diagnostic delay and the use of rapid molecular testing as primary means for diagnostics: for both pulmonary and extrapulmonary TB in children and adults\textsuperscript{181}. Isolating and identifying the causing agent is one of the major aims in TB diagnostics\textsuperscript{182}. Genome instability leading to multiple strains is one of the major challenges to this aim that should be addressed by all molecular assays. The sensitive and specific assay proposed will tackle that challenge. Also cost analysis and impact must be considered in the design of new diagnostic tools. Diagnostic tools now seek to identify specific \textit{Mycobacterium} species through nucleic acid testing and provide drug susceptibility testing through mutation analysis\textsuperscript{183}. There is a reported increase in the incidence of MDR TB which can be attributed largely to incomplete and improper diagnostics\textsuperscript{184}. Thus diagnostic tools need to include drug susceptibility information.

What new diagnostic tools must achieve in real world setting have been discussed and researched\textsuperscript{185}. As expected of every excellent diagnostic tool, an assay must be molecular-based, correctly amplify the bacterial DNA even in low copy numbers (single copy even), must not rely on sputum samples and if possible must provide information on drug resistance. Thus, a new molecular assay for the detection of \textit{Mtb} using the \textit{rpoB} gene as a target is proposed to meet
these needs.

Gene Targets

One of the key needs of diagnostic tests is their applicability in a point-of-care (POC) system. With this in mind, three genes were targeted in this study. These genes are the RNA polymerase subunit beta (rpoB) gene—particularly the RRDR region with known drug resistance (DR) associated mutations—the exclusive region of difference 9 (rd9). These targets were chosen because they indicate that Mtb DNA is present in the sample and because in the case of rpoB the target can inform about the presence/absence of resistance to rifampicin, a key first line drug.

RNA polymerase Subunit Beta (rpoB) and Region of Difference 9 (rd9)

The Mtb rpoB region is extensively studied target because the RRDR sequence within this gene is a hotspot for many mutations which cause rifampicin resistance. TB strains that exhibit rifampicin drug resistance are classified as MDR-TB, and XDR-TB depending on how many other drug resistant genes are also present in these strains.

The rpoB gene is not unique Mtb. In contrast, the rd9 is specific to the Mtb species and therefore identifies the presence of Mtb.

Sample Types

Another concern for diagnostic tests is the limits in sample types. As stated before, sputum samples, which are currently used, are not always available. They are also invasive and present a health risk for medical personnel taking these samples. Therefore, our Assay is being developed for use with all types of samples including, oral samples and blood plasma samples. Samples from these sources also contain high levels of human DNA from the patient. For this
reason, our Assay is being developed to retain sensitivity and specificity for \textit{Mtb} in the presence of a vast excess of human DNA. The possibility of finding at least low levels of \textit{Mtb} in blood makes sense given the route of infection and the progression of the disease, see above. The presence of low levels of \textit{Mtb} in oral samples also makes sense in light of the fact that tuberculosis is a pulmonary disease associated with coughing and sneezing which could deliver bacteria to oral cavity. In fact, multiple studies indicate that low levels of \textit{Mtb} can be amplified from blood and oral samples of infected patients\textsuperscript{189 190 191 192 193 194 195 196 197}. Our Assay seeks to increase the sensitivity and specificity of \textit{Mtb} detection in these samples. In order to mimic conditions of oral and blood plasma samples\textsuperscript{198} my experiments utilized low levels of non-living \textit{Mtb} genomes in the presence of high backgrounds of human DNA and other oral microbiome\textsuperscript{199 200 201 208}.

**PCR Methods**

The amplification methods being used are an improved version of Linear-After-The-Exponential (LATE) and Linear-Expo-Linear (LEL-PCR). All these technologies developed in the Wangh Lab end with asymmetric PCR after symmetric PCR has greatly increased the number of starting DNA. This results in an abundance of single stranded targets that can be probed in a single tube reaction. This setup saves time and resources as the amplification and detection reaction can be run in a single tube. The added advantage of the version of LATE-PCR performed for this project is that owing to the unique and stringent primer design (LEL-PCR primers) and PCR reagents (TheraStop/PrimeSafe and ThermaGo) there is increased sensitivity. LEL-PCR and LATE-PCR are explained in more detail below.
Novel Amplification Methods: Linear-Expo-Linear

Linear-Expo-Linear (LEL) PCR is a novel PCR method developed to increase selectivity for a rare event. LEL PCR uses tailed-primers to achieve three different consecutive amplification profiles; first a linear amplification of DNA, followed by an exponential and then linear amplification. Like LATE-PCR, two primers are used in unequal quantities; one primer is used at a lower concentration (limiting primer) than the other (excess primer). The excess primer can be up to twenty times more than the limiting primer.

In LEL-PCR, a limiting primer (LP) is initially extended and its overhang sequence is incorporated into the amplicon (Figure 1, Step 1). This first round of linear amplification is repeated for about ten rounds. This is followed by a drop in temperature to allow a lower Tm excess primer (XP) to bind; it extends and incorporates its overhang sequence into the amplicon also (Figure 1, Step 2). This second step is done only once to prevent formation of non-specific product at lower temperatures. Following this, the temperature is raised to select for amplicons with both tails incorporated. Since the primers are now fully complementary to the amplicon with both tails incorporated into its sequence, their melting temperature (Tm) also increases. Thus they are active at higher temperatures and amplify only products that were initially amplified. Since both primers are active, the amplification is exponential in this phase.

As in LATE-PCR, the LP runs out after a while as exponential amplification comes to an end. At this point there are still many more excess primers and they continue to make single stranded DNA product. This results in an abundance of single stranded DNA of our target region that can be probed in a single tube reaction. LEL-PCR is illustrated in Figure 1 below.
Figure 1. Illustration of LEL PCR

**Step 1**
After denaturation of template strands, the limiting primer (LP) binds and is extended at 70°C. The excess primer (XP) is not active as it cannot bind to the template (low Tm).

**Step 2**
After about 10 rounds of linear amplification, the temperature is dropped.

**Step 3**
Elevating the temperature allows the reaction to select the product with both LP and XP tails incorporated. Since both primers have Tm's lower than 80°C when the tail does not hybridize to the target, only the amplicon with both tails incorporated is used as a template at 80°C.

**Key**
- DNA Template
- 5’
- Excess Primer (XP)
- Tails
- Limiting Primer (LP)
- Polymerase
- G-C/A-T base pairs

This product is carried through about 50 rounds of exponential, and then linear amplification (LATE-PCR) due to unequal primer concentrations. This results in double strand (ds) production followed by single strand (ss) production.
**Improved Linear-After-The-Exponential PCR (LATE-PCR)**

LEL-PCR was designed to begin with linear amplification, but it was discovered that the excess primer was active during the initial linear phase of amplification. Thus the true amplification profile was exponential amplification followed by linear amplification as in Liner-After-The-Exponential (LATE) PCR. As such, it was concluded that step 2 of the process (Figure 1) where the temperature was reduced to allow the excess primer to bind was not necessary. As such, that step was eliminated. This amplification was however not true LATE-PCR as the primer design follows LEL-PCR and does not meet the requirement of LATE-PCR. Figure 2 illustrates this improved version of LATE-PCR.

---

**Figure 2. Illustration of amplification steps in improved version of LATE-PCR**
Detection Method: Lights-On/Lights-Off

After single strands have been produced in excess in a reaction, they are available to bind probes in the same reaction. Lights-On/Lights-Off (Figure 3) uses short probes (oligos that hybridize to amplicon) modified with fluorophores (On probes) and quenchers (Off probes). When the On probes bind to the amplicon, they are not in close proximity to the quenchers and as such fluoresce. As more On probes bind, the fluorescence increases and this happens as a function of temperature. Thus fluorescence increases as the temperature near then probe Tm; the derivative of the increase in fluorescence produces a peak. This can be seen in the rpoB and rd9 on probes used in this project.

![Figure 3. I) Schematic representation of Lights-On/Lights-Off and II) Lights-Off Only technologies: A) Single stranded (ss) DNA product is available to B) bind off probes in a specific order, depending on the target sequence, that C) gives rise to a fluorescent signature reflective of the sequence.]

When the quenchers bind the amplicon however, they reduce the fluorescence. They do this in two ways, when they hybridize to the amplicon with the quencher modification adjacent
to the fluorophore modification, the fluorescence of the On probe is reduced. However, the Off probe also hybridizes to the amplicon forming a double stranded DNA with quenchers at the end; this configuration is able to “steal” SYBR from the reaction, causing a reduction in fluorescence. This results in a valley in the FAM/SYBR derivative plot. The Lights-Off Only\textsuperscript{dd} probing technology that is used in this project operates based on this principle. Figure 3 (Part II) illustrates Lights-Off only as used in this project.

\textsuperscript{dd} Developed in the Wangh lab
Materials and Methods

Various experiments were conducted using improved LATE-PCR and LEL-PCR profile. To optimize PCR conditions, the reagents used and their respective concentrations changed often. Also the PCR profile used in amplification was changed depending on the goals of the experiment conducted. As such, there is no universal PCR mix and profile for all the experiments discussed below. The general setup procedure and reagents that were used consistently will be discussed briefly, followed by a more detailed discussion of the PCR reagents and profile that was changed in the course of the project.

General Setup Procedure

PCR amplification was the main method used in this research. Samples were aliquoted in volumes of 16μL per PCR reaction tube. The reaction master mix for all the experiments were made in a clean PCR room with the appropriate protective garment; sterile techniques were observed. Reagents were added in labeled tubes; following this human genome, IC template, and Mtb templates were added in that order. There were times when to achieve uniformity among replicates this order was changed. All Mtb templates were added last outside the PCR room to create sub-mixtures (sub-mixes). Mixtures were shaken with a vortex mixer for about 10s and spun for about 3s before aliquots were taken for PCR reaction tubes. When new dilutions of reagents were made, the time spent vortexing and spinning was increased. Taq and genomic samples were always kept on ice when not in use. For no template controls (NTCs), aliquots of
the mixture without the IC templates or *Mtb* templates were used. The set up process scheme is illustrated in Figure 4.

**Figure 4. Illustration of PCR setup**

**PCR Reagents**

General master mixes were prepared using water, 10x PCR buffer, dNTP’s, Mg$^{2+}$, 10x SYBR, PCR reagents similar to hot start reagents (ThermaStop/PrimeSafe and ThermaGo), the determined sets of probes and primers, and Taq polymerase (normal or platinum). The volume of the mixes was then adjusted with water so that each reaction tube would have a total volume of 16μL. PCR reagents had to be changed constantly to achieve optimal results. Some reagents like the concentration of hot start reagents had to be adjusted constantly as other regents were added or changed. An amplifiable internal control (IC) was added to establish the success of amplification and to ensure that all reagents were functioning properly.
**General Reagents**

Though some reagents were varied in experiments, some were always present at fixed concentrations. The master mix for all 23 experiments to be discussed below contained 1x PCR buffer and 0.3 mM dNTPs. The concentration of 1.25 U of Taq polymerase was used in all experiments. However, the brand of Taq polymerase used was varied between Invitrogen (normal and platinum Taq) and Syd Taq. Apart from these reagents, all other reagents were varied in the optimization process.

**Other Reagents**

*Primers*

**Primer Design**

LEL primers were generally designed to follow specifications illustrated in Figure 5 below. LEL primers for rpoB, and the internal control (IC) were adapted from existing efficient LATE PCR primers\(^{ff}\). Additional LEL primers for rd9 were designed using designed rd9 LATE PCR primers. The suggested concentration ratio for LP to XP LATE PCR primers (1:20) was used in most experiments\(^ {203}\). Other experiments have a lower ratio (1:40). Unlike in LATE-PCR, the melting temperature of the amplicon (Tm\(^A\)) is more than 25°C higher than the starting excess primer melting temperature (Tm\(^X\)). However, the differing between the final melting temperature of the limiting primer and excess primer (Tm\(^L\) and Tm\(^X\) respectively) is greater than or equal to 0°C as LATE-PCR dictates. The primer design is illustrated below.

---

\(^{ee}\) Refers to the Taq polymerase enzymes without the antibody

\(^{ff}\) Designed by John Rice
Primers for rpoB

Primers for rpoB were designed to match the initial LEL-PCR profile; primer set three (03), which was the most efficient was used in experiments 3-5. This initial LEL primer had low efficiency and the Tm during the LATE-PCR phase did not meet LATE-PCR specifications completely; as a result, it was replaced with primer set five (05) in subsequent experiments. The sequences for primer sets 03 and 05 are shown below in Table 2.

Primers for rd9

Primers that were used to amplify the rd9 region are shown in the table below. These primers also conformed to the LEL-PCR design: used in experiment 6. This was the second set of primers derived from the 3rd set of rd9 LATE-PCR primers. Although the efficiency of this primer set is low - delayed amplification - in comparison to the rpoB primers, the probe signal is still strong.

Primers for The Amplifiable Internal Control (IC)

Initial experiments (3-9) were done without an amplifiable control. The primer for the IC were LEL-PCR primers as well. There were setbacks in designing primers for the amplifiable control as the IC excess primer (XP) interfered with the amplification of rpoB target and its probe hybridization. After trying variations of the IC primers with no success, a new IC template was designed to mimic the primer portions of the rpoB amplicon. The target was designed such that
rpoB and the IC shared the same XP. This eliminated the incompatibility issues between the two XPs. The IC LEL-PCR limiting primer (LP) was then designed to match the new target. The primer was made to mimic the 5’ end of the rpoB LP, ICrpoB LP 2 (used in experiments 10-12). The initial 8 base pairs of the sequence were changed to make the IC primer selective for the IC control target only and to preserve the sensitivity and selectivity of the rpoB primers for the rpoB target.

A derivative of this new limiting primer was designed - ICrpoB LP 2_1 – to eliminate secondary structure in the IC amplicon and allow probes to bind more easily (used in experiments 1, 2, 13-17, 19-23).

Table 2. Table of primer sequences and IC target sequence

<table>
<thead>
<tr>
<th>Primer Name</th>
<th>Sequence (5’-3’)</th>
<th>Size (bp)</th>
<th>Tm Before* (°C)</th>
<th>Tm After** (°C)</th>
</tr>
</thead>
<tbody>
<tr>
<td>rpoB LP 03</td>
<td>TCGTGAATACCTCCAGCTCGGACCTCACGTGACAGCCG</td>
<td>40</td>
<td>71</td>
<td>81</td>
</tr>
<tr>
<td>rpoB XP 03</td>
<td>CGAGTCCATCACCTCGGATCACACCCACAGAGTT</td>
<td>35</td>
<td>61</td>
<td>81</td>
</tr>
<tr>
<td>rpoBL LP05</td>
<td>AACACAGCCAGGCAGCTCAGTGACAGACCG</td>
<td>31</td>
<td>72</td>
<td>80</td>
</tr>
<tr>
<td>rpoBL XP05</td>
<td>AAACCTACCAACACCTTGACACTGGAGATCATACCCGAGCTGT</td>
<td>43</td>
<td>60</td>
<td>79</td>
</tr>
<tr>
<td>rd9 LP 3_1</td>
<td>TCAGCATTCTGACGGTGCCGGGAGGAATTTGTGACATC</td>
<td>43</td>
<td>69</td>
<td>81</td>
</tr>
<tr>
<td>rd9 XP 3_2</td>
<td>AAACCTACCAGCTACCCCGAGTCATTCCAGCTCAAA</td>
<td>38</td>
<td>59</td>
<td>79</td>
</tr>
<tr>
<td>ICrpoB LP 2</td>
<td>AGGGCTGACGCACAGGCCAGCTCAGTGACACTG</td>
<td>35</td>
<td>68</td>
<td>80</td>
</tr>
<tr>
<td>ICrpoB LP 2_1</td>
<td>AAGGAGAGGGAGGCCACGTGCTAGAGTTTATATATCATCAT</td>
<td>48</td>
<td>66</td>
<td>81</td>
</tr>
<tr>
<td>ICrpoB 2 (IC Target)</td>
<td>AATGCGTTCACCACAAACCCAGAAATAGGCACGGCTCAGTGATACACTGTTATATCATCATCTACACTACTACTATATTACTTTATATATATAATATACAGCGCAGCGCGGTGTTAATAGTGACAAGCAAGAGAGGCGAGCTCAGTGATACACTGTTATATCATCATCTACACTACTACTATATTACTTTATATATATAATATACAGCGCAGCGCGGTGTTAATAGTGACAAGCAAGAGAGGCGAGCTCAGTGATACACTGTTATATCATCATCTACACTACTACTATATTACTTTATATATATAATATACAGCGCAGCGCGGTGTTAATAGTGAC</td>
<td>144</td>
<td>N/A</td>
<td>81</td>
</tr>
</tbody>
</table>

Blue letters refer to the primer tail sequence, while red letters refer to mismatches to the rpoB template.

*refers to the melting temperature of primer to the original target

**refers to the melting temperature of the primer after the tails are incorporated into the amplicon sequence

Primer concentration

Primer concentrations for monoplexes with only one primer set generally followed suggested LATE-PCR concentrations; that is a 1:20 ratio of limiting primer (LP) to excess primer (XP): 50nM of LP and 1μM of XP. This is the ratio used in experiments 3-11, 18 and 19. However, when the rpoB and IC ampicons shared a common XP, the LP-XP ratio was reduced to 1:40 as the final XP concentration was doubled to 2μM. This is true for experiment 9-15. To optimize the
amplification of the IC, the IC LP concentration was also doubled to 100nM in experiments 20 returning the ratio to 1:20. The same logic was applied in experiments 1, 2, 20-23 after testing the rpoB LP concentration at 50nM and 100nM, double the concentration in experiment 18; this enabled successful amplification of lower \textit{Mtb} copies (5 copies and below). This prevents competition with the IC targets for the shared XP.

\textit{Taq Polymerase}

Invitrogen Taq polymerase was used in most experiments (3-11, 13-14) The effects of platinum Invitrogen Taq was also tested in experiment 12. All other experiments were done using the comparatively cheaper Syd Taq. Taq polymerase concentrations remained constant (1.25U) for all experiments. Each polymerase was used with its respective 10x buffer at a final concentration of 1x.

\textit{Magnesium Concentration}

Experiments that were run before the introduction of Syd Taq (experiments 3-13) had a 3mM Magnesium concentration. However, in experiment 14 where Syd and Invitrogen Taq were compared in their respective buffers, the replicates with Syd Taq and buffer had a final magnesium concentration of 4.5mM; this is because the Syd buffer includes 1.5mM magnesium. In experiment 16, 3mM and 4.5mM magnesium concentrations were compared. Based on the results of that experiment, all following experiments used 4.5mM magnesium.

\textit{SYBR Concentration}

SYBR green was used at 0.24x final concentration for experiments 3-17. However, this amount was increased to enable proper resolution between the wild type and mutants when using Off probes only. SYBR concentrations of 0.24x and 0.48x were compared in experiment 16.
Following this experiment, SYBR was used at 0.48x. Also care was taken when making working stock dilutions from the general stock to ensure that the quality of SYBR was preserved for the best results.

**ThermaStop/PrimeSafe**

ThermaStop and PrimeSafe were used hot start reagents; they have the same sequence but different modifications. The structure of ThermaStop and PrimeSafe are shown in Figure 6 below; they are both single strands with quencher modifications at the end of the hairpin formed by the 3′ and 5′ end. They have a conformation that sequesters Taq at low temperature where the hairpin is formed. In so doing they reduce Taq reactivity and increase faithfulness. Thus, ThermaStop and PrimeSafe prevent mis-priming and formation of non-specific product. However, unlike other hot start reagents, they are DNA and as such are not degraded with high temperatures but present throughout the reaction and resume their activities one the hairpin forms at low temperatures (about 45°C).

In experiments (3-6), ThermaStop was used at a final concentration of 600nM. However, owing to availability, this was switched to 1μM PrimeSafe as background human genome was added (experiment 7); the concentration was further increased to 1.2μM (experiments 8-14). In experiment 15 the PrimeSafe concentrations needed for a clean reaction with Syd Taq was determined from 0nM, 300nM, 600nM and 1.2μM. Following this experiment, PrimeSafe was used at 600nM (experiments 1, 2, 16-23).

**ThermaGo**

ThermaGo (TG) is a hot start reagent similar to ThermaStop and PrimeSafe in its mode of activity. However, it has a different structure; it is composed of two separate strands that
hybridize at higher temperatures. It is therefore present for longer periods in the reaction than ThermaStop/PrimeSafe and increases the faithfulness of Taq to prevent mis-priming at high temperatures. It is especially important in preventing secondary rise during the linear amplification phase of LATE-PCR (also known as product evolution).

Initial experiment (3) did not have this reagent. In experiments 4-7, ThermaGo 3D (TG3D) was used at 35nM. This was then increased to 50nM in experiments 8-14. However, more effective, shorter version - ThermaGo 86-23 (TG86-23) - was used in experiment 1 and 2, and 15-23. In experiment 15, 0nM, 100nM, 200nM, and 400nM concentration of TG86-23 were tested. In experiment 17 100nM and 150nM TG86-23 were tested. Based on the results of this experiment, TG86-23 was used at a final concentration of 150nM.

![Figure 6. Structure of ThermaGo and ThermaStop/PrimeSafe](image)

Target

Due to difficulty in designing an LEL-PCR excess primer for the existing IC template that was compatible with the rpoB excess primer, a new IC template was designed (used in experiments 1,2, 10-23). The new template was designed to mimic the primer portions of the rpoB amplicon. The target was designed to share the rpoB XP but with introduced mismatch in the 5’tail region to reduce the amplification efficiency of the IC. Also the new IC target mimicked the 5’ end of the rpoB LP. Cultured Mtb samples were used in the experiments. Samples were
initially tested in a nanodrop to determine their concentration. The initial stock concentration (usually a multiple $10^6$ copies) was diluted to 50,000 or 5,000 working stock. All lower *Mtb* copies used were made from these higher concentrations. Standardized BCG samples: tested concentrations of $10^{-4}$ units – $10^{-7}$ units (experiment 9). Initially, human genome was not added in reaction mixtures; it was added to the background of reactions in increasing concentration across various experiments.

**Figure 7.** Description of rpoB amplicon (5'-3'): Blue and green highlights refer to the primer regions (green highlight refers to the primer tails); codons 506-539 are labeled; boxed sequences refer to probe hybridization region, probes sequences are highlighted in gray-pink letters indicate intentional mismatch to the amplicon.

**Target Concentrations**

*Mtb* genomes were tested from as many as 5 million copies to single copies per 1μL for different experiment; the specific copies in each replicate will be specified in the results section for each experiment. The IC DNA was tested at 1000 copies or 5000 copies per reaction (experiment 10); following this experiment, 5,000 copies (experiments 1, 2, 11-16, 18-23) was used. In experiment 17, 1,000 copies IC DNA was used. Background human genomic DNA was
added at 17,500 copies (experiments 7-8). It was removed once again to optimize the concentration of other reagents (experiments 10-13) and reintroduced at 46,000 (experiments 9, 14-16), and finally 58,000 copies (experiments 1, 2, 18-23).

**Target Types Tested**

Various *Mtb* strains were tested. Among the 38 stains that were tested (Table 3), 5 were wild type for rpoB (WT) and 35 of them had were mutated rpoB genes (Mut). All samples were tested in replicates; for the purposes of this paper three kinds of replicates will be discussed – technical replicates, biological replicates and allelic replicates. Technical replicates are those that are completely identical; that is, they were aliquoted from the same sub-mix and contain the same reagents and the same *Mtb* genomic DNA (or human and IC DNA) at the same concentration. Biological replicates (Table 4) are the samples that contained different strains of *Mtb* samples but these strains either have the mutation or WT status. Allelic Replicates (Table 5) will be used to refer to the samples that contained different *Mtb* strains with different mutations in the same codon position.

### Table 3. Labels and mutations of all 38 *Mtb* samples: Drug resistance (DR) key: R implies rifampicin resistance and X implies extensive drug resistance.

<table>
<thead>
<tr>
<th>Sample #</th>
<th>DR</th>
<th>ID</th>
<th>Mutation</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>R</td>
<td>296</td>
<td>510 CCG</td>
</tr>
<tr>
<td>2</td>
<td>R</td>
<td>334</td>
<td>511 CCG</td>
</tr>
<tr>
<td>3</td>
<td>R</td>
<td>9339</td>
<td>512 AAA</td>
</tr>
<tr>
<td>4</td>
<td>R</td>
<td>602</td>
<td>513 AAA</td>
</tr>
<tr>
<td>7</td>
<td>R</td>
<td>3728</td>
<td>516 GGC + 533 CCG</td>
</tr>
<tr>
<td>5</td>
<td>R</td>
<td>4939</td>
<td>516 GGC</td>
</tr>
<tr>
<td>6</td>
<td>R</td>
<td>16074</td>
<td>516 GGC</td>
</tr>
<tr>
<td>8</td>
<td>R</td>
<td>761</td>
<td>516 GTC</td>
</tr>
<tr>
<td>9</td>
<td>R</td>
<td>1873</td>
<td>516 GTC</td>
</tr>
<tr>
<td>10</td>
<td>R</td>
<td>10212</td>
<td>516 GTC</td>
</tr>
<tr>
<td>11</td>
<td>R</td>
<td>4440</td>
<td>516 TAC</td>
</tr>
<tr>
<td>12</td>
<td>R</td>
<td>2686</td>
<td>526 AGC</td>
</tr>
<tr>
<td>13</td>
<td>R</td>
<td>8498</td>
<td>526 AAC</td>
</tr>
<tr>
<td>14</td>
<td>R</td>
<td>892</td>
<td>526 CGC</td>
</tr>
<tr>
<td>15</td>
<td>R</td>
<td>851</td>
<td>526 CTC</td>
</tr>
<tr>
<td>16</td>
<td>R</td>
<td>145</td>
<td>526 GAC</td>
</tr>
<tr>
<td>17</td>
<td>R</td>
<td>2504</td>
<td>526 GAC</td>
</tr>
<tr>
<td>18</td>
<td>R</td>
<td>3487</td>
<td>526 GAC</td>
</tr>
<tr>
<td>19</td>
<td>R</td>
<td>4330</td>
<td>526 GAC</td>
</tr>
<tr>
<td>20</td>
<td>R</td>
<td>709</td>
<td>526 GGC</td>
</tr>
<tr>
<td>21</td>
<td>R</td>
<td>531</td>
<td>526 TAC</td>
</tr>
<tr>
<td>22</td>
<td>R</td>
<td>987</td>
<td>526 TAC</td>
</tr>
<tr>
<td>23</td>
<td>R</td>
<td>2493</td>
<td>526 TAC</td>
</tr>
<tr>
<td>24</td>
<td>R</td>
<td>511</td>
<td>531 TGG</td>
</tr>
<tr>
<td>25</td>
<td>R</td>
<td>3640</td>
<td>531 TTG</td>
</tr>
<tr>
<td>26</td>
<td>R</td>
<td>3889</td>
<td>531 TTG</td>
</tr>
<tr>
<td>27</td>
<td>R</td>
<td>7113</td>
<td>531 TTG</td>
</tr>
<tr>
<td>28</td>
<td>X</td>
<td>104</td>
<td>533 CCG</td>
</tr>
<tr>
<td>29</td>
<td>R</td>
<td>965</td>
<td>533 CCG</td>
</tr>
<tr>
<td>30</td>
<td>R</td>
<td>1870</td>
<td>533 CCG</td>
</tr>
<tr>
<td>31</td>
<td>R</td>
<td>5247</td>
<td>Del (512-514)</td>
</tr>
<tr>
<td>32</td>
<td>R</td>
<td>4381</td>
<td>Del (517)</td>
</tr>
<tr>
<td>33</td>
<td>R</td>
<td>1810</td>
<td>Del (517-519)</td>
</tr>
<tr>
<td>34</td>
<td>R</td>
<td>201</td>
<td>Del (526-527)</td>
</tr>
<tr>
<td>35</td>
<td>R</td>
<td>238</td>
<td>Del (526-527)</td>
</tr>
<tr>
<td>36</td>
<td>X</td>
<td>164</td>
<td>Wild Type</td>
</tr>
<tr>
<td>37</td>
<td>R</td>
<td>1277</td>
<td>Wild Type</td>
</tr>
<tr>
<td>38</td>
<td>R</td>
<td>2613</td>
<td>Wild Type</td>
</tr>
</tbody>
</table>
**Probe Design and Concentration**

A set of On and Off probes were used for the IC, rpoB and rd9 targets. Probe design followed the Lights-On-Lights-Off probing technology developed in the Wangh lab. These probes are modified with fluorophore (On) and quenchers (Off) and produce peaks and valleys respectively at the temperatures which they bind. The rd9 probes were viewed in the CAL red channel while the IC and rpoB probes were viewed in the quasar channel. Lights-Off only, a new technology developed in the Wangh Lab was also used as means of detection. In this method,
only probes modified with quenchers (Off probes) are used in detection. Carbon linkers were used to prevent amplification at the 3’ end of probes. A total of 6 probes that hybridize to the target between 50-80°C were used. These probes steal SYBR when they hybridize to the target causing a valley in the SYBR channel. The presence of the dip indicates the presence of *Mtb* genome whereas shifted or altered valleys indicated a mutated rpoB gene. The *IC On* probe and *IC Off* probe were used at 50-100nM. The *rpoB On* probe was used at 90-400nM. The six rpoB Off probes were used at 100-300nM.

**PCR Profile**

**Summary**

Initially the temperature profile for LEL-PCR was used but this was later changed to that of LATE-PCR. Generally, samples were tested using 16μL aliquots. All samples were tested in technical replicates comprising 2-10 tubes based on the number of *Mtb* copies added to the tube. In other words, when very low numbers of genomes were used many replicates were used in order to establish reproducibility.

The temperature profile included an initial denaturation step at 97°C for 1 minute. This was followed by cycles of denaturation at 97°C for 7s, and a subsequent anneal and extension at 69°C for 45s. This initial denaturation and extension was repeated for a total 10 cycles (experiments 1-4, 7-23) or 20 cycles (experiments 5-6). In early experiments (experiments 3-4), there was a single cycle where the temperature was dropped down to 60°C for 20s and 40s. This step was eliminated in later experiments. There was also 40-70 cycles of a second denaturation and extension; 97°C for 7s, and 75/78°C for 30s-45s. During each cycle of this step, the fluorescence was read three times and the average was recorded for amplification plots.
The penultimate step was the equilibration step; this step enabled the reaction to cool down and probe hybridization to occur. It also allowed SYBR to be distributed among the double stranded DNA. This was followed by a melt procedure from 25°C to 100°C. Two different Strategene machines were used to run PCR experiments. It was noticed that the fluorescence scale on one of the machines was about half of the other one so some results may have a smaller scale than others. These experiments will be marked with a symbol (†) showing that the decrease in peak heights and fluorescence values is not reflective of a decrease in efficiency but difference in fluorescence scales of different machines.

**LEL Profile**

The stages for LEL-PCR are illustrated in Figure 1. The temperature profile used to achieve these stages includes an initial denaturation step at 97°C for 1 minute; followed by 10 or 20 cycles of 97°C for 7s, and 69°C for 45s; a dropdown to 60°C for 20-45s; 40-70 cycles of 97°C for 7s, and 75°C or 78°C for 45s with fluorescence reading each cycle; a degree drop in temperature per cycle for 45 cycles (70°C - 25°C), and 45s per cycle; followed by a melt and probe hybridization procedure from 25°C to 100°C. This profile was used in experiments 3-4.

**LATE-PCR profile**

The PCR protocol for experiments 1 and 2, and 5-23 were similar to the LEL profile but eliminated the dropdown step. The resulting temperature profile is as follows: 97°C for 1 minute; 10 cycles of 97°C for 7s, and 69°C for 45s; 40-70 cycles of 97°C for 7s, and 75°C or 78°C for 45s with fluorescence reading each cycle; a degree drop per cycle for 45 cycles (70°C - 25°C), 45s per cycle (a single step at 25°C for 10 minutes was also done alternatively); lastly, a melt and probe hybridization procedure from 25°C to 100°C was done.
Melt Profiles and Equilibration

The equilibration step before a melt was performed was done in two ways. Initially, the reaction was cooled down slowly from 70°C to 25°C by dropping the temperature 1°C each cycle and staying at that temperature for 45s. Thus this step lasted for 45 cycles. This was used in experiments 3-12. In experiments 1 and 2, and 13-23, the equilibration step was done as a single step when the temperature was quickly dropped to 25°C after the final extension; the temperature was then kept at 25°C for 10 minutes. Melts were performed after the equilibration step. The temperature was increased by 1°C per cycle. Initially, each cycle lasted 45s and 3 reads were taken for a recorded average (experiments 3-17). Fast melts lasted either 6s (with a single read) or 18s (with three reads). Longer melts lasting 1 minute per cycle were performed in experiments 1 and 2, and 18-23 where 10 reads were taken each cycle.

Intermediate Melts

In some experiments, the second extension step of 40-70 cycles were separated by melts. Usually 25-30 cycles of denaturing, annealing, and extension were performed; this was followed by a melt procedure. The extension procedure was then continued in 10-15 more cycles and a second melt was taken. A final extension step of 15-20 cycles was performed followed by a final melt. The first and second melts were partial melts that covered a smaller temperature range within 25°C-100°C; it usually encompassed a temperature region of interest (probe hybridization regions). This process of interrupting amplification to perform temperature melts was used in experiment 9.
Final PCR Reagents and Profile

Reagents

After optimization tests, the \textit{Mtb} assay currently has the following reagents at given concentrations: 1x Syd PCR buffer, 0.3mM dNTP’s, 1.25U of Syd Taq polymerase, 600nM PrimeSafe, 150nM ThermaGo 86-23, 300nM \textit{rpoB} Off probes 1-6, 100nM \textit{ICrpoBLP 2_1}, 100nM \textit{rpoB LP 05}, 2μM \textit{rpoB XP 05}, 0.48 SYBR, 5000 IC template copies, 58,000 copies of human genomic DNA and 1μL of \textit{Mtb} at desired copy number. Water is also added; the amount added brings the final volume of the mix to 16μL per reaction.

Protocol

The final amplification profile is as follow: initial denaturation at 97°C for 1 minute; followed by 10 cycles of 97°C for 7s, and 69°C for 45s; 50 cycles of 97°C for 7s, and 75°C for 45s with florescence reading each cycle; equilibration at 25°C for 10 minutes; then a melt from 25°C to 100°C for 75 cycles, 1 minute per cycle.
Organization of Results

The following plots are examples of results that will be shown and discussed below. Important labels are shown and also some visual guides to understanding the efficiency and significance of the results are given. Figure 8 shows an example of an amplification plot and also indicates the four visual measures of amplification efficiency; the quantification cycle (Cₚ), the steepness of the slope of exponential amplification, the distance between technical replicates, and the absence/presence of secondary rise during the linear phase. Figure 9 shows an example of a FAM dissociation curve and the characteristic peaks with On probes whereas Figure 10 shows the FAM dissociation curve with Lights-Off Only technology. Figure 11 shows the Quasar dissociation curves. The peak for the internal control (IC) is not always present as the amplifiable internal control (IC) was only added in later experiments. Also, initially, there was secondary structure in the single stranded IC product. A hairpin formed at about 40°C (labeled in Figure 9) prevented the IC On probe from binding and thus significantly reduced the IC On probe peak size. As such it is not clear in early experiments with the IC. Results will be briefly discussed as well as the specific changes made in the experiments, before the plots are shown. Also experiments shown will be accompanied by a table (or description), briefly listing the total technical replicates.

---

86 Cycle number where DNA is amplified enough for SYBR fluorescence to exceed threshold levels
88 Indicate the rate of doubling/exponential increase
89 Indicates reproducibility
90 Indicates stability of single stranded product; secondary rise indicates product evolution or the conversion of single stranded product to double stranded product.
(TR) tested under each condition and the number out of those that were successfully amplified and identified, the number that resembled the IC signal and the number that did not amplify at all. The examples of results are shown below in Figures 8-11.

**Figure 8.** Example of amplification plot showing the four indictors of PCR efficiency

**Figure 9.** Example of FAM dissociation curve with On probes
Figure 10. Example of FAM dissociation curve and signatures with Lights-Off Only

Figure 11. Example of Quasar dissociation curve
Results

Optimized Amplification with LATE-PCR and Lights-Off Only Detection

PCR conditions have been optimized to yield highly reproducible results with high sensitivity and specificity. Samples with few *Mtb* genomic DNA (5-1 copies) are efficiently detected with high efficiency in as much as 58,000 human genomes. Below are two results from *Mtb* dilution experiments using the optimized conditions. The first experiment shows a dilution series from 5,000-5 copies of *Mtb* with and without the internal control to assess the effects of the IC on the amplification; the second experiment shows the Lights-Off Only detection of 1,000 copies to single copy *Mtb* genomes to determine the detection limits of the assay. Experiments discussed below are not numbered in the order they were performed but the order in which they are presented. The summary data of dilutions performed and the derivation of the PCR efficiency are shown also.

**Experiment 1: Dilution Series with and without IC**

The optimized assay was tested with 58,000 copies of human genomic DNA in the background and was shown to have high reproducibility and efficient amplification (steep slope during the exponential phase). All technical replicates in an *Mtb* dilution series ranging from 5,000 copies to 5 copies were detected with and without the IC template (5,000 copies). The components of the finalized assay included the following: 1X Syd PCR buffer, 4.5mM Mg$^{2+}$, 300nM dNTPs, 0.48x SYBR-Green, 600nM PrimeSafe, 150nM ThermaGo86-21, 1.25U Syd Taq DNA
polymerase, 100nM rpoB LEL-PCR limiting primer (rpoBL LP_05), 2µM rpoB LEL-PCR excess primer (rpoBL XP_05), 300nM rpoB Off probe 1-6, 100nM IC limiting primer (ICrpoBLP 2_1). This results of the amplification (both individual technical replicates and averages) and melt are shown below.

**Figure 12. Dilution series (5,000 copies to 5 copies) of Mtb with and without the IC template:** A) shows amplification Plots of all the technical replicates (100% of all replicates tested were amplified and detected) B) &C) show the amplification and FAM dissociation, respectively, of the averages of the technical replicates

**Experiment 2: Dilution Series with IC down to Single Copy Mtb**

The sensitivity of the assay was demonstrated. Technical replicates of a dilution series ranging from 1,000 Mtb copies to 1 Mtb copy were detected efficiently in 58,000 copies of human genomic DNA in the background with Lights-Off Only technology. All replicates with 1,000-10 copies of Mtb were detected. Eight out of ten of the technical replicates with a single Mtb genome were successfully amplified. The remaining two had signal similar to the amplifiable control indicating that they may not have any Mtb DNA. The results of the melt in the FAM channel (Lights-Off only) are shown below.
Figure 13. Dilution series (1,000 copies to single copy) of wild type Mtb DNA with IC present: 100% of all technical replicates of 10-1,000 copies of Mtb were amplified and detected, while 20% of the single copy samples were not amplified and are suspected to have no Mtb genome as the resemble the IC in the plot of the averages.

Summary Results

From the above data the PCR efficiency was calculated. The delta CT (ΔCT) values of dilution series as well as detection rate were also calculated. The theoretical ΔCT values for tenfold dilution series with 100% is 3.32. The calculated ΔCT values for higher copy number dilutions were 3.2 which is close to this value. Generally, the accepted range for PCR efficiency (calculation shown below) falls within the range of 90-110% as ΔCT values are expected to be slightly above or below the expected 3.32. The efficiency for amplification reactions with the IC (107.74%) and without the IC (105.58%) fall within this range. The A_f values refer to the amplification factor: that is the factor by which the DNA copy number increases each time. The theoretical value showing 100% efficiency is 2; that implies the DNA doubles every cycle. Experimentally obtained values were close to this value: 2.07 and 2.06 for replicates with and without the IC respectively. This reaffirms the efficiency of the assay. Summary results for
experiments 1 and 2 and efficiency information are show in the table below. Dilutions 1 and 3 show data from experiment 1 while Dilution 4 shows data from experiment 2.

**Table 6. Table of dilution series result summary:** values used in dilution 1 & 3 are from experiment 1; values from dilution 2 are from an experiment not shown; This experiment was chosen because larger amounts of *Mtb* are needed to offset the effects of the IC in efficiency calculations. TR refers to technical replicates. *Efficiency was not calculated for these dilutions because the Mtb copy numbers used were too low and as such was heavily ΔCq affected by the IC.*

<table>
<thead>
<tr>
<th>Mtb Copy Number</th>
<th>Log (Copy Number)</th>
<th>Cq</th>
<th>ΔCq</th>
<th>Number of TR tested</th>
<th>Number of TR detected</th>
<th>Number of TR like IC</th>
<th>Detection Rate (%)</th>
<th>PCR Efficiency (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Dilution 1: No Internal Control (IC) added for 5,000-5 Mtb copies</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5,000</td>
<td>3.699</td>
<td>18.1</td>
<td>-</td>
<td>3</td>
<td>3</td>
<td>0</td>
<td>100</td>
<td>105.5 (AF=2.06)</td>
</tr>
<tr>
<td>500</td>
<td>2.699</td>
<td>21.31</td>
<td>3.21</td>
<td>3</td>
<td>3</td>
<td>0</td>
<td>100</td>
<td></td>
</tr>
<tr>
<td>50</td>
<td>1.699</td>
<td>24.49</td>
<td>3.18</td>
<td>3</td>
<td>3</td>
<td>0</td>
<td>100</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>0.699</td>
<td>26.49</td>
<td>2</td>
<td>8</td>
<td>8</td>
<td>0</td>
<td>100</td>
<td></td>
</tr>
<tr>
<td>0 (5000 IC copies)</td>
<td>-</td>
<td>24.06</td>
<td>-</td>
<td>2</td>
<td>-</td>
<td>-</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Dilution 2: IC added for 50,000-500 Mtb copies</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>50,000</td>
<td>4.699</td>
<td>15.65</td>
<td>-</td>
<td>3</td>
<td>3</td>
<td>0</td>
<td>100</td>
<td>107.7 (AF=2.07)</td>
</tr>
<tr>
<td>5,000</td>
<td>3.699</td>
<td>18.83</td>
<td>3.18</td>
<td>3</td>
<td>3</td>
<td>0</td>
<td>100</td>
<td></td>
</tr>
<tr>
<td>500</td>
<td>2.699</td>
<td>21.96</td>
<td>3.13</td>
<td>3</td>
<td>3</td>
<td>0</td>
<td>100</td>
<td></td>
</tr>
<tr>
<td>0 (5000 IC copies)</td>
<td>-</td>
<td>23.49</td>
<td>-</td>
<td>10.3</td>
<td>-</td>
<td>-</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Dilution 3: Internal Control (IC) added for 5,000-5 Mtb copies</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5,000</td>
<td>4</td>
<td>17.6</td>
<td>-</td>
<td>3</td>
<td>3</td>
<td>0</td>
<td>100</td>
<td>*</td>
</tr>
<tr>
<td>500</td>
<td>3.74</td>
<td>20.92</td>
<td>3.32</td>
<td>3</td>
<td>3</td>
<td>0</td>
<td>100</td>
<td></td>
</tr>
<tr>
<td>50</td>
<td>3.703</td>
<td>23.35</td>
<td>2.43</td>
<td>3</td>
<td>3</td>
<td>0</td>
<td>100</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>3.699</td>
<td>24.61</td>
<td>1.26</td>
<td>8</td>
<td>8</td>
<td>0</td>
<td>100</td>
<td></td>
</tr>
<tr>
<td>0 (5000 IC copies)</td>
<td>-</td>
<td>24.06</td>
<td>-</td>
<td>2</td>
<td>-</td>
<td>-</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Dilution 4: IC added for 1,000-1 Mtb copies</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1,000</td>
<td>3.778</td>
<td>-</td>
<td>-</td>
<td>3</td>
<td>3</td>
<td>0</td>
<td>100</td>
<td>*</td>
</tr>
<tr>
<td>100</td>
<td>3.708</td>
<td>-</td>
<td>-</td>
<td>3</td>
<td>3</td>
<td>0</td>
<td>100</td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>3.7</td>
<td>-</td>
<td>-</td>
<td>8</td>
<td>8</td>
<td>0</td>
<td>100</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>3.699</td>
<td>-</td>
<td>-</td>
<td>10</td>
<td>8</td>
<td>2</td>
<td>80</td>
<td></td>
</tr>
<tr>
<td>0 (5000 IC copies)</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>3</td>
<td>-</td>
<td>-</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
**Figure 14. Obtaining C_q (Quantification Cycle) values:** The Y-axis of amplification plots (Fluorescence) was scaled using the Log function and the automatic machine Threshold was calculated; generated C_q (shown in red box) values were recorded for various curves.

**Figure 15. Determination of Slope for Efficiency Calculation**

The equations used in calculating PCR efficiency and amplification factor ($A_F$) are shown below:

**Equation 1 (PCR Efficiency)**

$$Efficiency = \left(-1 + 10^{-\frac{1}{m}}\right) \times 100$$

**Equation 2 (Amplification factor, $A_F$)**

$$A_F = 10^\frac{1}{m}$$
Experiments Leading to Optimized Assay

Results from many experiments shaped the assay into its optimized form. For the purposes of this paper, the key experiments that illustrate the assay optimization process are shown below. Experiments are discussed in the order they were performed; however, they may be other experiments in between the experiments that are not shown here.

**Experiment 3: LEL with High TB Copy Number**

The initial LEL-PCR protocol was used to amplify large copy *Mtb* genomes (5,000,000 copies and 10,000 copies) successfully with very low efficiency. Although all technical replicates tested (three-3- for each *Mtb* concentration) made the right products, there was very little reproducibility among replicates even at high *Mtb* concentrations.

The temperature profile included 10 initial rounds of extension at 69°C, a drop down to 60°C for 20s and 50 more rounds of extension at 78°C. The reagents used were 1X Invitrogen PCR buffer, 3mM MgCl2, 300nM dNTPs, 0.24X SYBR-Green, 600nM *ThermaStop*, 1.25U Invitrogen Taq DNA polymerase, 50nM LEL-PCR limiting primer (*rpoBL LP 03*), 1 µM LEL-PCR excess primer (*rpoBL XP 03*), 100nM On probe (*rpoB On probe 2*). These reagents served as the baseline and were modified in later experiments. The results of the amplification and melts (in Quasar 670 and FAM channels) are shown below.
Experiment 4: LEL in Lower Copy TB Number (ThermaGo 3 Introduced)

A new hot start agent, *ThermaGo 3*- TG3, was added to the baseline reagents to provide a cleaner reaction. Technical replicates showed more reproducibility and the no template controls (NTCs) were delayed by about 4 cycles. The addition of 35nM of TG3 allowed the time spent at the drop down to 60°C to be doubled (40s) without making too much nonspecific product; this provided more time for uniform amplification among replicates. All other reagents were kept the same. A dilution series, 1000 copies -1 copy, for *Mtb* genome was performed. Though the results are more reproducible among technical replicates(TR) samples with high *Mtb* copies, not all replicates were successfully amplified or even produced the right product. All three (3) TR of 1,000 *Mtb* copies samples were amplified. Two (2) of three (3) TR of 100 *Mtb* copies amplified were the right product, while the other made non-specific product. Four (4) of six (6) TR of 10 *Mtb* copies the right product although they were not reproducible; one (1) of the six (6) was extremely delayed and made little rpoB product while the other made non-specific product. Two (2) of ten (10) TR of single copy *Mtb* were successfully amplified, one (1) made non-specific product while the other seven (7) did not begin amplification at all. The results are shown in the plots below.
Figure 17. LEL-PCR with lower Mtb copy numbers: LEL-PCR worked after further optimization with fewer copies TB template (black-1000 copies, purple-100 copies, green-10 copies, orange- single copy, blue- no template control (NTC))

Experiment 5: LEL vs. LATE-PCR

The need for a drop down to 60°C was tested. This experiment demonstrated that true LEL-PCR was not occurring as amplification was possible even in the absence of a drop in temperature to engage the excess primer (XP); implying that the XP is active even during the initial amplification. Performing PCR without a drop in temperature (LATE-PCR profile) increased the reproducibility between technical replicates; three replicates of Mtb samples and two of the no template control were tested. Additionally, the ΔCq values were reduced indicating earlier amplification for lower copy numbers and increased efficiency.

Two experiments were run simultaneously using the new baseline reagents with 35nM TG3. Sub-mixes (containing no Mtb, 100 copies and 1,000 copies of Mtb DNA) created from a
master mix was split for two separate reactions for the two machines. The amplification protocol of one of the machines skipped the drop down step (Labeled “LATE-PCR”). All other aspects of the PCR protocol were the same for the two reactions. The results of the experiment from the two amplification protocols are shown below.

![Amplification Plots](image1.png)  ![FAM Dissociation Curve](image2.png)  ![Quasar Dissociation Curve](image3.png)

**Figure 18. Comparison of LATE-PCR and LEL-PCR:** LATE-PCR (top) was compared to LEL PCR (bottom) using 100 and 1000 Tb copies (red and green respectively). Most remarkable difference was the increased efficiency and reproducibility: evidenced by smaller ΔCq.

**Experiment 6: rpoB and rd9 Amplification with LATE-PCR**

New primers increased the efficiency of amplification by reducing the Cq. Keeping the baseline reagents, the rpoB LEL-PCR primers (*rpoBL XP 03.* and *LP 03.*) were switched for more efficient ones (*rpoBL XP 05* and *LP 05*). New rd9 LEL primers (*rd9 LP 3-2* and *XP 3-1*) were also used to amplify the rd9 gene. Both primer sets were used to amplify 5,000 copies of *Mtb* genome. LATE-PCR, without a drop down to 60°C, was used. Replicates with *Mtb.* were tested in groups of three (3) whiles the no template controls (NTCs) were tested in groups of (2). The results of this experiment are shown below.
Figure 19. New primers for \textit{rpoB} and \textit{rd9}:
\textit{rpoB} (grey) and \textit{rd9} (turquoise) were amplified with 5,000 copies of TB using LATE-PCR. Differences in peak height (probe hybridization plots-middle) due to arbitrary fluorescence in 2 channels.

\textbf{Experiment 7: Human Genomic Background}

It was demonstrated that PrimeSafe (PS) and ThermaGo (TG) amounts were not enough to prevent product evolution and scattering when human genomic DNA was added. Hence, higher concentrations were needed in reactions. Using the two hot start reagents - 1µM PS and 35nM TG3 - the assay was tested with human genome in the background. Two sub-mixes were made; one containing human genomic DNA at 17,500 copies and another without any background DNA. An \textit{Mtb} genome dilution (500-5 copies) was done with both sub-mixes. Two (2) replicates of the no template control (NTC) were tested; they did not make any product without background human genomic DNA. Three (3) replicates of at 500 copies were \textit{Mtb} tested, and eight replicates of \textit{Mtb} at 50 and 5 copies were tested. Without human genome in the background, there was no product evolution and \textit{rpoB} signal in the FAM channel was well resolved; whereas product evolution occurred when human genomes were added and the \textit{rpoB} FAM signal could not be resolved. Although replicates without the human genome added were better, they also exhibited low efficiency and reproducibility at lower copy \textit{Mtb}; one of the replicates tested with 50 \textit{Mtb} copies produced little \textit{rpoB} and as such had a small product peak in the Quasar channel. Although the Quasar channel showed little difference when human genomes were added, the FAM channel takes precedence since the end goal is to read results in the FAM channel. As such, PS and TG3 concentrations were increased to obtain cleaner, better resolved signals.
Figure 20. Amplification with background human genomic DNA: 17,500 copies of human genomic DNA was added in the background of samples (bottom, top is the control). Though amplification is not as clean, the same number of replicates comes up with or without human genomic DNA in the background.

**Experiment 8: ThermaGo 3 Concentration at 50nM**

Increased ThermaGo (TG3) concentration, 50nM, was shown to delay the formation of non-specific product (delayed no template controls-NTCs) and the onset of product evolution. The PS was used at 1.2µMkk. Thus the master mix of future experiments contained 1.2µM PS, 50nM TG3. Samples with 17,500 human genomes were tested; two (2) technical replicates of the NTC and eight (8) of samples containing 50 Mtb copies were tested with and without the 50nM TG3. The results of the reaction are shown below.

---

kk 1.2µM PrimeSafe had been shown to be more effective than 1µM PrimeSafe in a previous experiment.
Figure 21. Increased ThermaGo concentrations: 50nM ThermaGo was added to clean up the amplification reaction (Bottom).

Experiment 9: Multiple Reads Between Amplification Cycles

Quantification analysis was demonstrated as quick melts at different cycles of amplification showed probe signals proportional to the amount to Mtb DNA tested. This showed that experiments can be paused to run various analyses without interfering with the amplification process. Using the reagents concentrations, the same, with 1.2µM PrimeSafe and 50nM ThermaGo 3, 500 (3 replicates), 50 (8 replicates) and 5 copies (8 replicates) of Mtb were amplified with and without a background of 46,000 human genomes. The rpoB On probe concentration was increased to 400nM to be in probe excess. Short melts were taken from 40°C-80°C after 40 and 50 cycles of amplification. A full melt from 25°C-100°C was done after 65 cycles. The averages of samples that produced the rpoB amplicon are shown. All samples with 50-500 Mtb copies had very high reproducibility and all produced the rpoB product. Few replicates with 5 Mtb copies produced rpoB; three (3) in replicates without human genomes and one (1) in replicates with the human genomes added. One (2) replicate without human genomes produced nonspecific product; two (2) replicates with the human genomes added produced non-specific product. All
other replicates of the eight (8) tested did not make any product. The results of this experiment are shown below.

**Figure 22. Performing melts in between amplification cycles:** Fluorescence measurements were taken 40 and 50 cycles after amplification before the reaction ended at 65 cycles. Human genomic DNA (HG) was used (46,000 copies) in the background. No significant differences were observed in the control and samples with background human DNA except for the 5copy TB samples (2 more replicates amplified successfully in the control)

**Experiment 10: Duplex with Internal Control (IC)**

The amplifiable internal control was successfully amplified as a duplex and its signal could be clearly seen in the FAM channel. The probe signal in Quasar was not visible as secondary structure in the single stranded product prevents the On probes from binding. In this experiment, 50nM of the IC limiting primer (ICrpoBLP 2) and 50nM IC On and Off probes were added to the reaction mixture while keeping other reagents the same; the rpoB On probe was used at 100nM.

Three (3) replicates of the IC template (ICrpoB 2) was tested at 1,000 and 5,000 copies; this was duplexed with 500 (three-3- replicates tested) and 50 (eight-8- replicates tested) Mtb copies; no background human DNA was added. All replicates were amplified with the right product and were very reproducible. The average results of this experiment re shown below.
Figure 23. Duplex with an amplifiable control: The IC (probe peak highlighted by red box) and rpoB were amplified successfully in same reaction without a reduction in the rpoB probe signal.

**Experiment 11: Standard BCG Samples**

Detection rate of standard BCG sample concentrations were demonstrated to be equal, if not slightly better than current standard. Currently, the detection limit is sample concentration of 1E-6. Standard BCG samples of concentration 1E-4 – 1E-7 were amplified. The internal control template was added at 5,000 copies; all reagents were kept at the same concentration. No background human DNA was added. Five (5) replicates of concentrations 1E-4 and 1E-5 were tested whiles ten (10) replicates of concentrations 1E-6 and 1E-7 were tested. The results are shown below, fraction of replicates with the rpoB product are also shown.

Figure 24. Amplification of standard BCG samples: BCG samples, concentrations 1E-4 – 1E-7, were tested: 1E-4 (5/5 replicates), 1E-5 (5/5 replicates), 1E-6 (9/10 replicates), and 1E-7 (4/10 replicates)

\[\text{Fraction of replicates that gave the right probe signal.}\]
Experiment 12: Platinum Taq Polymerase

Platinum Invitrogen Taq (with antibody: used without PrimeSafe) was compared to the normal Invitrogen Taq (no antibody: used with 1.2µM PrimeSafe) and was shown to be slightly better. The rpoB On probe concentration was lowered to 90nM to boost the IC signal. The rpoB excess primer (XP) concentration was doubled (2µM) to prevent competition between the IC and rpoB templates and amplicons for the XP. All other reagents were kept the same for all samples tested including 50nM ThermaGo 3. Ten (10) technical replicates of the sub mix with the platinum and normal Invitrogen Taq were amplified and the results are shown below.

Figure 25. Comparison of Platinum Taq and normal Taq (with PrimeSafe): 10/10 samples tested with just platinum Taq produced the right product but 1/10 of the samples with normal Taq and PrimeSafe showed product evolution

Experiment 13: New Internal Control (IC) Limiting Primer

A new IC limiting primer (LP), \textit{ICrpoBLP 2_1}, was introduced to remove the secondary structure in the IC amplicon to allow the probes to bind. The old IC LP, \textit{ICrpoBLP 2}, was compared to the new LP; Invitrogen Taq was used with 1.2µM PrimeSafe and 50nM ThermaGo 3. The shared excess primer was used at 2 µM and the rpoB On probe was at 90nM. The IC target was used at 5,000 copies. Monoplex reactions of the IC with old and new LP’s as well as rpoB and duplexes of rpoB and the old and new IC LP’s were performed and the results are shown below. Two (2) technical replicates of no template controls (NTCs-shown in the same color but with dashed lines) were tested; generally, none of the NTC replicates, with the exception of one NTC replicate of
the rpoB monoplex, were amplified and produced non-specific product. Monoplexes and
duplexes with DNA templates were tested in replicates of eight (8); all of these replicates yielded
the right product however, the products of the new IC product were able to bind the probe and
show a small probe peak in Quasar channel. They also lacked the secondary structure peak in
FAM.

Experiment 14: New Taq Polymerase

The performance of Invitrogen Taq (1.25U) and buffer (1x) was compared to that of the
relatively cheaper Syd Taq (1.25U) and buffer (1x). Syd Taq was shown to be more effective by
reducing C_q values and increasing reproducibility; however, product evolution increased in Syd
Taw and the corresponding FAM signals were not well resolved. The new IC LP was used in a
monoplex reaction (four -4- technical replicates tested): IC target only at 5,000 copies. All other
reagents were used at regular concentrations including the magnesium, shared excess primer
and, rpoB limiting primer. The rpoB on probe, IC On and Off probes were all used at 100nM
concentration. Human genomic DNA was added at 46,000 copies. The results of the amplification and melt are shown below.

Figure 27. A comparison of Invitrogen Taq with its 1x buffer and Syd Taq with its 1x buffer

**Experiment 15: ThermaGo 86-21 and PrimeSafe Combinations**

PrimeSafe concentrations ([PS]) and ThermaGo 86-21 concentrations ([TG86-21]) for a clean reaction with Syd Taq and buffer were determined; 600nM PS and 100nM TG86-21 was chosen as the PS and TG3 concentrations at which the reaction was clean with minimal effects the Cq value and probe signals. Sub mixes with 0nM PS and 400nM TG8-21, 300nM PS and 200nM TG86-21, 600nM PS and 100nM TG86-21, and 1.2µM PS and 50nM TG86-21. Each sub mix was tested with two technical replicates of the IC and three technical replicates of *Mtb* at 500 copies. All other reagents were kept the same including 46,000 copies of human genomic DNA. All replicates were amplified with the right products. The results of the amplification and melt procedures are shown below.
Figure 28. A comparison of varying PrimeSafe and ThermaGo 86-21 concentrations: the amplification plots of all the different combinations, the FAM dissociation curve and the Quasar dissociation curve are shown. Plots are ordered as follows: 0nM PS and 400nM TG86-21, 300nM PS and 200nM TG86-21, 600nM PS and 100nM TG86-21, and 1.2µM PS and 50nM TG86-21.

Experiment 16: 3mM vs. 4.5mM Magnesium

Optimum magnesium concentrations for the new Syd Taq and buffer were determined to be 4.5mM; this concentration increases reproducibility and exponential amplification rate (steeper slopes). 1.5mM and 3mM magnesium were added to different sub mixes of containing Syd Taq and Buffer. This brought the final magnesium concentration to 3mM and 4.5mM respectively (Syd buffer contains 1.5mM magnesium). The 5,000 copies of the internal control template were used (three -3- replicates of the IC only): 500 copies (three -3- replicates) and 50 copies (eight -8- replicates) of Mtb were tested. Human genomic DNA was added at 46,000
copies. The results of the comparison are shown below. Despite noisy signal in the FAM channel, all replicates produced the right IC and rpoB products.

![Graphs showing fluorescence over cycles and temperature](image)

**Figure 29.** A comparison of 3mM (top) and 4.5mM (bottom) magnesium concentrations: plots of the amplification, FAM dissociation and Quasar dissociation are shown in that order.

**Experiment 17: 100 vs. 150 ThermaGo 86-21**

Using 4.5mM magnesium, new baseline, and keeping all other reagents the same, 100nM and 150nM concentrations of ThermaGo 86-21 (TG86-21) were compared; 150nm was shown to delay IC product, and also reduce product evolution, while increasing reproducibility. In this experiment, the IC template was reduced to 1,000 copies; 46,000 copies of human genomic DNA was also added. Two replicates of the IC only and eight of the duplex with 50 copies of *Mtb* were tested. All replicates produced the right product, when product evolution was absent. All other reagents were kept the same. The results of the experiment are shown below.
Figure 30. A comparison of 100nM and 150nM TG86-21

Experiment 18: 0.24x vs. 0.48x SYBR Concentrations

Using 150nM ThermaGo 86-21 (TG86-21) 0.24x SYBR and 0.48X SYBR concentrations were compared; 0.48x SYBR was shown to improve resolution and increase probe hybridization signal. Monoplex reactions of the IC (three -3- replicates), wild type rpoB Mtb DNA (six -6- replicates), and mutant rpoB Mtb DNA (six -6- replicates) at 50 copies were tested; 1 µM of the shared excess primer was used in each Monoplex reaction. The human genomic DNA in the background was increased to 58,000 copies; all other reagents were kept the same. All replicates generated the right product. The results (averages of all replicates) are shown below.
Figure 31. FAM dissociation of 0.24x and 0.48x SYBR concentrations

Experiment 19: 100nM and 300nM Probe Concentrations

Using 0.48x SYBR concentration, wild type (500 copies) and mutant (5,000 copies) rpoB Mtb DNA were amplified using Lights-Off only. Six Off probes for rpoB (rpoB Off 1-6) were compared at 100nM and 300nM; 300nM probe concentrations improved the resolution of signatures of both mutant and wild type rpoB. In setting up this experiment, the shared primer was used at 1µM as the rpoB gene was amplified alone (without the IC). Human genomic DNA (58,000 copies) was added to the master mix. All six replicates of the mutant and wild type DNA generated the right product and were included in the average signatures shown in the results below.

Figure 32. A comparison of 100nM and 300nM concentrations of all six rpoB off probes

Experiment 20: 50nM and 100nM rpoB LP Concentration

Doubling the concentration of the rpoB limiting primer (LP) to 100nM (same as IC LP) was shown to be effective in allowing rpoB to be amplified for lower copies of Mtb (5 and below) in the presence of the IC. Increasing the rpoB LP concentration reduced competition between Mtb and the IC; this allowed enough rpoB product to be generated for a distinct rpoB signature in FAM. This signature is also similar to that seen in higher copy numbers of Mtb. All other reagents
were kept the same and 300nM of the six rpoB Off probes were used. Human genomic DNA was added to the master mix at 58,000 copies. Three (3) technical replicates of the IC only and eight (8) of the IC and wild type (WT) and mutant (Mut) rpoB Mtb genome at 5 copies were tested. All replicates reproducibly generated the right product. The results (averages of FAM dissociation plots) of the amplification of 5 copies of Mtb are shown below.

![Graphs showing FAM dissociation plots for 50nM and 100nM rpoB Limiting Primer (LP)](image)

**Figure 33. A comparison of 50nM and 100nM of the rpoB Limiting Primer (LP): averages of the FAM dissociation plots are shown.**

**Achieving Higher resolution with Lights-Off Only**

Over the course of optimizing the assay, the resolution of signatures has been improved. As shown in experiments 17 and 18. This is demonstrated in experiment 21 below. This is not an actual result of one experiment but a comparison of two separate experiments; one performed before the optimal conditions were determined. This demonstrated improved resolution enabled distinction to be made between the 35 mutant samples tested and the wild type (discussed in experiment 22).

**Demonstrating Improved Resolution**

The resolution of mutant samples was improved in the course of optimizing the assay. An example of this improved resolution is in the detection of mutant sample R1810; three (3)
technical replicates of 500 copies were tested in both experiments to be compared. Wild type signals are obtained from the average of three biological replicates, each tested in triplicates. An earlier experiment was done with 1X Syd PCR buffer, 4.5mM Mg²⁺, 300nM dNTPs, 0.24x SYBR-Green, 600nM PrimeSafe, 150nM ThermaGo 86-21, 1.25U Syd Taq DNA polymerase, 50nM rpoB LEL-PCR Limiting Primer (rpoBL LP_05), 2μM rpoB LEL-PCR Excess Primer (rpoBL XP_05), 100nM IC limiting primer (ICrpoBLP 2_1), 300nM rpoB Off probe 1-6, 5,000 copies of the IC target, 58,000 human genomes. The results of that experiment (Figure 26 A) are compared with that of the results obtained using the finalized assay in experiment 22 above (Figure 26 B). The comparison is shown below.

![Diagram](image)

**Figure 32. Demonstration of improved resolution using sample R1810**: signatures of the wild type and mutant R1810 before the assay was optimized and after the assay was optimized

**Experiment 22: Thirty-eight (38) Mtb Samples of Wild Type and Mutant rpoB**

The assay was used to test 38 Mtb samples at 500 copies: three (3) of the Mtb samples were wild type (WT) for rpoB, while the remaining 35 were biological and allelic replicates of Mtb strains that were mutated (Mut) in the rpoB genome. The specific breakdown of the samples is shown in table 4 above. Three technical replicates of each sample was tested.
Since one machine was too small for all the samples, the samples, prepared from one master mix, were split between two machines. All individual samples were tested as three (3) technical replicates. The three biological replicates of the WT rpoB Mtb, as well as 3 technical replicates of the IC and sample R1870 were tested in both machines. The results from the 3 biological replicates from the two machines were average to generate a general/normalized WT fluorescence signature. The IC curve, and curve of sample R1870 was also averaged in the same manner to generate generalized signatures.

From these results it was shown that the fluorescent signatures generated for the same sample in the different machines were similar; thus, the averages of the results of 3 technical replicates of the other samples were compared to the general WT and IC signatures generated from both machines. The results are shown below: they have been grouped under the headings wild type, individual samples\textsuperscript{mm}, biological replicates, allelic replicates, different signatures\textsuperscript{nn}, double mutation/mixed sample, low quality DNA and indistinguishable.

*Wild Type*

\textsuperscript{mm} Referring to samples that were tested alone without other biological replicated

\textsuperscript{nn} Referring to samples that produced different signatures than others biological replicates
Individual Samples
**Biological Replicates**
Allelic Replicates
Different

Double Mutation/Mixed sample
Low Quality DNA

Indistinguishable

1. Deletion 526-527
   - WT: R 238
   - R 238

2. Mutation 526 GAC (CAC)?
   - WT: X 145
   - X 145

3. Mutation 510 CCG (AGC)
   - WT: R 296
   - R 296

4. Deletion 517
   - WT: R 4381
   - R 4381
Figure 33. Fluorescent signatures of 23 different mutational types as well as plots of biological and allelic replicates: each plot has its own key showing the samples or mutation tested, and next to each plot is a blown up image of the probe hybridization temperature range (50°C-80°C)

Experiment 23: Mixtures of Wild Type (WT) DNA and Mutant (Mut) *Mtb* DNA

Before the assay was finalized, mixed samples of WT and Mut rpoB DNA were tested: 500 *Mtb* copies were tested in each reaction tube. The assay used at the time contained 1X Syd PCR buffer, 4.5mM magnesium, 300nM dNTPs, 0.24x SYBR-Green, 600nM PrimeSafe, 150nM ThermaGo 86-21, 1.25U Syd Taq DNA polymerase, 100nM rpoB LEL-PCR Limiting Primer (*rpoBL LP_05*), 2µM rpoB LEL-PCR Excess Primer (*rpoBL XP_05*), 100nM IC limiting primer (*ICrpoBLP 2_1*), 300nM rpoB Off probe 1-6, 10,000 copies of the IC target; human genomic DNA was present in the background at 58,000 copies. The results are shown below.
Figure 34. Mixed Samples of wild type and mutant (rpoB) \textit{Mtb} samples: \textit{Mtb} samples that are wild type and mutants of the rpoB gene: wild type and mutant DNA were mixed in ratios of and 1:1, and 1:9 wild type: mutant DNA respectively.
Discussion

This thesis describes a new, sensitive, highly specific assay for the detection and characterization of drug sensitive and rifampicin resistant *Mtb*. It was constructed by combining several novel technologies that have been invented in the Wangh laboratory over the past decade. These technologies include LATE-PCR and Lights-Off Only technology; LEL-PCR primer design was also used. This thesis demonstrates in the results that the assay I have constructed has the following characteristics: 1) it is reliable 2) relatively less costly and 3) informative. The reliability of the assay stems from its robustness, efficiency, specificity and sensitivity. The assay utilizes cheaper Taq and Lights-Off Only detection technology; both of these contribute to lowering the cost of production. Finally, rpoB amplicon has been characterized as wild type or mutant and this provides information regarding the rifampicin drug resistant nature of the infection.

The assay is, very robust. *Mt\(b\) can be amplified under very stringent conditions. This makes the PCR protocol flexible; it can be changed to obtain various information such as *Mt\(b\) genome quantification data. This was demonstrated in experiment 9 where amplification was paused several times to run probe hybridization melts. This robustness gives room for the PCR profile to be modified to obtain even more information on the nature of infection. Results from the assay are also consistent and reproducible, further pointing to its robustness.
The assay is also very effective. The efficiency is 105.58%, within the accepted values for high PCR efficiency. This is evident in the steepness of slopes of the exponential phase of amplification plots in experiment 1. Quantification cycle, \( C_q \), values are lower than standard values; for example, 50 \( Mtb \) copies have a \( C_q \) value of about 24, the normal \( C_q \) value for amplification of 1,000 copies. This shows faster doubling time as copies twenty times less that the reference amount (1,000 copies) come up at the same time. This phenomenon is partly due to the ten initial rounds of extension. These rounds as shown in experiment 5, are not linear. It can be inferred that the starting DNA is increased by the time reaction reaches the exponential amplification phase (in the amplification diagram). This also explains why \( \Delta C_q \) values get smaller with decreasing \( Mtb \) copy number. Additional background DNA (IC target and human genomes) also play a role in this phenomenon.

The assay is very sensitive and specific; \( Mtb \) was detected in increasing background human genomic DNA. Detection rate of samples with five (5) or more \( Mtb \) genomes is 100%. Eight (8) of the ten (10) samples that tested with a single copy \( Mtb \) genome were detected. It was determined that the remaining two (2) samples did not have any \( Mtb \) genome as their signature resembled that of the amplifiable control. Also wild type rpoB and mutant rpoB are mostly distinguishable (91% of the samples tested) from each other; majority of the mutations, even those that are allelic replicates, can be distinguished from each other. This detection of low copies of \( Mtb \) and characterization of rpoB gene amplified (wild type or mutant) occurred in the presence of 58,000 human genomic DNA in the background: that is about 41,000,000 times the representative \( Mtb \) genome. This suggests that the assay has a high specificity and sensitivity.
Finally, the PCR protocol can provide all this information in under two (2) hours. Since the assay has high sensitivity and specificity, *Mtb* detection can be done effectively within this time. Together, all these factors contribute to making the assay reliable. Another important quality is the resolution between different mutant fluorescent signature and the of the wild type. This is discussed in more detail below.

Resolution of Mutations Associated with Rifampin Resistance:

The assay measures the SYBR fluorescent values and through that a fluorescent signature is generated. The resolution of this signature is good enough to distinguish between wild type and mutant rpoB. Thirty-three (33) of thirty-six (36) tested samples were correctly identified as wild type or mutants (91% correctly matched). Two of the samples had very poor quality and produced different signatures than their biological replicates (Sample X145) or signatures close to that of the internal control at high *Mtb* concentrations (5,000 copies). Twenty (20) of the twenty-three (23) different mutational types tested were correctly identified as mutants (87%). The remaining three produced fluorescent signatures that were similar to the wild type signature in the probe hybridization region and as such could not be definitively identified as mutants.

Also all biological replicates that were tested gave similar signatures; this includes both the wild type and several mutant samples. Samples that were biological replicate produced fluorescent signatures with the same pattern. The signatures were identical in curve for most portions of the curve under the probe hybridization region; they differed slightly in very few portions. These results were reproducible in several experiments.

---

Originally 38 but two samples were not counted as poor DNA quality affected the results.
The allelic replicates were largely distinguishable from each other. Plots of allelic replicates revealed signatures with different slopes and depth of valleys; these difference can be used to tell different alleles apart. Some allelic replicates however produced similar signatures. This could be because the probe and SYBR concentrations do not provide enough resolution to distinguish them. An alternative explanation could be that the mutations in the alleles have the same thermodynamic effects on the probe, resulting in similar signatures.

Mixtures of wild type and mutant DNA produced fluorescent signatures that became closer to the wild type signature as the wild type DNA concentration was increased. This shows that mutant rpoB DNA present among wild type rpoB DNA in low copies can be identified. The signatures of such mixtures are can also benefit from improved resolution. One sample (R 3487) produced a different fluorescent signature that its biological replicates: the signature it produced was similar to that of mutations in codon 516. The sample will be sequenced to verify the identity of the mutation.

Optimizing the Assay Through Variable Testing

A number of variables, were tested including PCR reagents and amplification protocol. The variables tested to determine optimum levels are summarized below.

1) Design of Primers Their Relative Concentrations
2) Taq Polymerases and Their Buffers
3) Additional Reagents that Suppress Mis-Priming (PrimeSafe and ThermaGo)
4) Concentration and Quality of SYBR Green
5) Probe Concentration
6) PCR Thermal Cycling
7) The Length of Time for Probe Binding

These variables are discussed in more detail below.
Reagents

The reagents used in the assay have changed over the course of the project. It became clear that PCR reagents such as PrimeSafe 2 and ThermaGo were necessary to prevent mis-priming and product evolution. This is especially true in the presence of human genomic background; and when using Syd Taq and buffer.

Syd buffer has increased magnesium content and also generally has different thermodynamic effects on DNA compared to Invitrogen buffer. For example, the internal control peak in FAM cannot be fully and consistently resolved in Syd buffer. The magnesium content alone is not responsible for this phenomenon. In experiment 15 where 3mM and 4.5mM magnesium concentrations were compared, the efficiency of amplification was increased with 4.5mM magnesium (reduced $C_q$ values) as the Tm’s of double strands were increased was increased. The increased stability of strands contributed to the lack of resolution in the FAM signal for the IC; however, it was also evident that even at 3mM magnesium concentrations, the signal was not resolved. Thus the Syd buffer has other factors that affect the thermodynamic stability of DNA strands.

Taq polymerase and buffer concentrations have been constant although the brand used was changed to the relatively cheaper Syd version. These concentrations seem to produce efficient results. Platinum Taq offers the advantage of clean results without creating a dip that might affect the fluorescent signature created by Off probes. Since platinum Taq is quite expensive and goes against the aim of making this assay cheap, an alternative method was used. The PrimeSafe 2 concentrations was reduced to reduce the depth of the valley it produces, while
ThermaGo 3 (TG3) was switched out for the more effective ThermaGo 86-21(TG86-21). The concentration of ThermaGo was also increased to keep the reaction clean.

Primers were designed to fit the LEL-PCR thermal cycling. Thus they seem to be robust and effective. It was noticed that primers (excess primers) were active up to 20°C above their relative predicted Tm’s. This made performing linear amplification with only the limiting primer difficult. Alternatively, this made it possible to increase stringency and provide sensitivity and specificity to reactions. Also, limiting primer concentrations have been shown to affect the results; they boost the production of double stranded product thus preventing the IC signal from overshadowing the rpoB signal at low Mtb copies and also produce a more defined and distinct probe signal (experiment 19).

Together, with increased SYBR and probe concentrations (experiment 17 and 18 respectively), the changes made to the assay have improved the resolution and clarity of signatures produced in the melt procedures. The assay reagent for the amplification of rpoB have been optimized and gives consistent and clear results. The assay in general can be expanded and further improved.

PCR protocol

The PCR protocol has evolved over the experiments shown above. The initial plan of using an LEL-PCR protocol was abandoned as it is difficult to design excess primers that are not active during the initial linear amplification with the limiting primer alone; also primers doing this while keeping the efficiency of the primers high was difficult. Thus, the second step, dropping the temperature to allow the excess primer to be incorporated into the amplicon, was eliminated; it
was unnecessary and only introduced more errors. As shown in experiment 3, eliminating the second step resulted in increased primer efficiency (lowered \( C_q \) values).

The time for equilibration was also shown to improve results and the resolution of fluorescent signatures. Samples that were melted again after being left to equilibrate overnight showed better resolution. This can be used to improve future experiments. Additionally, the time taken to run each experiment is greatly affected by the melt procedure. Melts can be performed in an elaborate fashion (10 reads per cycle, each cycle lasting a minute) or rapidly (3 reads per cycle each cycle lasting 18s). The quick melt produced similar results to the long melt but had reduced definition. In a setting where a quick yes/no answer was needed, the quick melt is sufficient as it is possible to determine if \( Mtb \) is present and also if the rpoB gene amplified is mutant or wild type.

Since the assay reagents have been configured to prevent mis-priming and product evolution, stopping the amplification procedure to run quick melts in between is possible. This was shown in experiment 8. This provides the opportunity to provide quantitative information. Together with the qualitative information provided by the melts, this could provide a lot of information to physicians, simplifying the determination of treatment regimen.

In addition to the timeliness and efficiency of the assay, the assay could be relatively cheaper than available assays. In using relatively cheaper Syd Taq in amplification and Lights-Off Only technology for detection, the assay has a potential for being cheap. A good point of care test (POCT) must be affordable, reliable, robust, rapid, sensitive and specific; the current assay meets these standards.
Challenges

Though the assay currently meets POCT standards of being affordable, robust, reliable, sensitive, specific, and fast it can be further improved. The assay gives reproducible results, and has good resolution in the resulting fluorescent signatures; but as discussed this resolution is not absolute. Some mutations are not distinguishable from the wild type. Also biological replicates are not exactly the same and not all allelic replicates can be told apart from each other. As shown in the comparison made in experiment 24 above, it is possible to improve resolution. In experiments 17 and 18, it was shown that the signature produced is affected by SYBR and Off probes concentrations respectively. Thus, the resolution can be improved to differentiate between these mutants and the wild type. Currently, the assay provides drug resistance information on rifampicin only. Although it is an important first line drug, there are first line drugs that the \textit{Mtb} could have developed resistance to. This points to the need to expand the assay to include other drug targets to provide more information to physicians.

Future Experiments

Immediate experiments would be to reconfigure the SYBR and probe concentrations such that all mutant samples will produce fluorescent signature distinct from the wild type signature. The resolution might also be affected by the melt procedure and equilibration time. The range for the melt can be reduced to include the probe hybridization region and the double stranded products only. This will give more time to equilibrate as samples will not be rapidly cooled to 25°C before a melt. Also equilibration time can be increased to 15 minutes (instead of 10 minutes) to improve resolution and definition of fluorescent signatures.
Following this, LEL-PCR designed primers for katG and inhA will be added to the assay. This multiplex will allow for more information on the drug resistance. The assay reagent concentrations might need to be slightly altered to all this to happen as a single tube reaction. Next, the sensitivity and specificity of the assay will need to be measured in a blinded test. The numerical value for sensitivity and specificity must also be determined in various sample types; that is oral samples and blood samples. Once this is done, the assay will be truly complete.
Conclusion

TB continues to claim more lives as drug resistant strains of *Mtb* increase. It is quickly becoming a serious global issue. A good point of care test (POCT) that is affordable, reliable, robust, rapid, sensitive and specific will be a powerful tool that simplifies the detection and treatment process. An assay was developed to detect *Mtb* at very low copy numbers using LATE/TI-PCR. Currently the assay detects low copies of *Mtb* DNA, even single copy *Mtb*, efficiently. It can also distinguish most mutant rpoB samples from the wild type thus providing drug resistance information. Though this meets POCT standards, the degree to which these standards are met can be increased. The resolution of the fluorescent signatures can be improved to produce more distinct signatures that allow differentiation between wild type rpoB and all mutational types. The assay can also be expanded to provide more information on drug resistance aside rifampin drug resistance. TB is can be treated and cured. It is time to eradicate it.
References


Waksman, Selman Abraham. Streptomycin: background, isolation, properties, and utilization. publisher not identified, 1953.


49 Dheda, K and Migliori, GB. The global rise of extensively drug-resistant tuberculosis: is the time to bring back sanatoria now overdue?. Lancet. 2012; 379: 773–775


76 Russell, David G. "The evolutionary pressures that have molded Mycobacterium tuberculosis into an infectious adjuvant." Current opinion in microbiology 16.1 (2013): 78-84.


78 Alex Mira, Howard Ochman, Nancy A. Moran, Deletional bias and the evolution of bacterial genomes, Trends in Genetics, Volume 17, Issue 10, 1 October 2001, Pages 589-596


86 Kaufmann, SHE, Lange, C, Rao, M et al. Progress in tuberculosis vaccine development and host-directed therapies—a state of the art review. Lancet Respir Med. 2014; (published online March 24.)


(Sakamoto, K. "The pathology of Mycobacterium tuberculosis infection." Veterinary Pathology Online 49.3 (2012): 423-439.)


132 Understanding Clinical Trials”, U . S. National Institutes of Health clinicaltrials.gov/


Teo, Jeanette WP et al. “Clinical Utility of RD1, RD9 and hsp65 Based PCR Assay for the Identification...


204 Probes were previously designed by John Rice