Temperature sensitive *Mycobacterium tuberculosis* as a potential vaccine candidate

by

Crystal Tina Pinto
M.Sc. in Biotechnology, Mumbai University, 2010

A Thesis Submitted in Partial Fulfillment
of the Requirements for the Degree of

MASTER OF SCIENCE

in the Department of Biochemistry and Microbiology

©Crystal Tina Pinto, 2015
University of Victoria

All rights reserved. This thesis may not be reproduced in whole or in part, by photocopy
or other means, without the permission of the author.

Supervisory Committee

<u>Temperature sensitive *Mycobacterium tuberculosis* as a potential vaccine candidate</u>

by

Crystal Tina Pinto
M.Sc. in Biotechnology, Mumbai University, 2010

Supervisory Committee

Dr. Francis E. Nano (Department of Biochemistry and Microbiology)
Supervisor

Dr. Caroline Cameron (Department of Biochemistry and Microbiology)
Departmental Member

Dr. Perry Howard (Department of Biology)
Outside Member

Abstract

Supervisory Committee
Dr. Francis E. Nano (Department of Biochemistry and Microbiology)
Supervisor
Dr. Caroline Cameron (Department of Biochemistry and Microbiology)
Departmental Member
Dr. Perry Howard (Department of Biology)
Outside Member

Mycobacterium tuberculosis remains one of the most common worldwide causes of illness and death due to an infectious disease. The emergence of multiple and extreme-drug resistant strains has increased the need to find an effective vaccine for tuberculosis. The goal of our research group is to engineer a temperature-sensitive (TS) *M. tuberculosis* strain that can be used as a tool in vaccine development. One approach to create TS *M. tuberculosis* involves the integration of the essential gene *ligA* encoding a TS NAD+ dependent DNA ligase, which was taken from the psychrophilic organism *Pseudoalteromonas haloplanktis*. The integration and functioning of *ligA* was demonstrated in the fast-growing organism *Mycobacterium smegmatis*. This strain had a TS phenotype with growth limited to below 37°C. The strain was found to have a stable TS phenotype and did not mutate to a temperature-resistant form at a detectable level. Following experiments with the fast growing *M. smegmatis*, the integration of the *ligA* gene was attempted in slow-growing *M. tuberculosis*. Merodiploids of *M. tuberculosis* containing both the psychrophilic and the WT *ligA* gene in its chromosome were obtained.

The second approach used for the development of TS *M. tuberculosis* was the directed evolution of native *M. tuberculosis* essential genes. An advantage of this approach is that the gene encoding the essential protein will resemble the native

M. tuberculosis gene and thus will closely match the native transcriptional and translational rates. A system to screen and select for TS essential genes engineered by directed evolution was designed, where the essential gene on the chromosome of *E. coli* was knocked out and this gene was supplied on a conditionally replicating plasmid. As a first step in developing this directed evolution approach, a family of conditionally replicating plasmids were created and tested in an essential gene knock-out strain of *E. coli*.

Table of Contents

Supervisory Committee ii
Abstract iii
Table of Contents v
List of Tables viii
List of Figures ix
Acknowledgments x
Dedication xi
List of Abbreviations xii
Chapter 1- Introduction 1
1.1. *Mycobacterium tuberculosis* 1
1.2. Pathogenesis of TB 1
1.2.1. Latent and Active TB 1
1.2.2. Overview of events following infection with *M. tuberculosis* 2
1.3. Host immune response against TB 4
1.3.1. Innate immune response 5
1.3.2. Adaptive immune response 7
1.4. Burden of TB disease 9
1.5. Diagnosis and TB treatment 10
1.5.1. Tuberculin skin test (TST) 10
1.5.2. Microbiological studies 11
1.5.3. Radiography 11
1.5.4. TB Treatment 11
1.6. Current vaccines against TB 12
1.6.1. The current BCG vaccine 12
1.6.2. Need for a new vaccine 13
1.6.3. New vaccines in the pipeline 14
1.7. Temperature sensitive (TS) viral vaccines 17
1.7.1. Temperature distribution in humans 18
1.7.2. TS vaccines target mucosal immunity 20
1.7.3. TS bacterial vaccines 21
1.8. Essential genes in bacteria 21
1.9. Essential genes chosen for this study 22
Project Overview 23
Chapter 2- Stable temperature sensitive *M. smegmatis* engineered by incorporating the Arctic *ligA* gene 25
2.1. Introduction 25
2.1.1. Objective of this research 25
2.1.2. Gene replacement in *M. tuberculosis* 27
2.2. Materials and Methods 29
2.2.1. Bacterial strains and growth conditions 29
A) Experiments to engineer TS *M. smegmatis* 29
2.2.2. Codon harmonization of the psychrophilic *ligA* gene 29

2.2.3. Recombining *M. smegmatis* flanking regions with *ligAPh* 30
2.2.4. Cloning the assembled construct into the p2NIL-*sacB* mycobacterial suicide vector.. 33
2.2.5. Restriction digestion, ligation and transformation.. 34
2.2.6. Preparing electrocompetent *M. smegmatis* cells 35
2.2.7. Construction of TS *M. smegmatis* strains 35
2.2.8. Determining the restrictive temperature of TS *M. smegmatis* 36
2.2.9. Mutation rate of TS *M. smegmatis*... 36
2.2.10. Growth curves of WT and TS *M. smegmatis*.. 37
2.2.11. Genetic drift experiments with TS *M. smegmatis*....................................... 37
B) Experiments to engineer TS *M. tuberculosis* 37
2.2.12. Assembling three different psychrophilic *ligA* variants with the *M.TB* flanking regions 37
2.2.13. Constructs designed for allelic exchange using mycobacteriophage recombineering (Hatfull method) 39
2.2.14. Preparing *M.TB* electrocompetent cells and transformation using Hatfull method.. 41
2.2.15. Constructing three different variants of *ligAPh* and cloning into mycobacterial suicide vector (Pavelka method) 42
2.2.16. Pavelka method for preparing electrocompetent *M. tuberculosis* 43
2.2.17. Transforming the *ligAPh* constructs into *M.TB* Erdman and H37Rv strains 44
2.2.18. Screening for primary *M.TB* recombinants.. 44
2.2.19. Future Steps to screen for TS mutants in the level 3 facility 45
2.3. Results 46
2.3.1. Codon harmonization using the Anaconda software package 46
2.3.2. Recombining the *M. smegmatis* chromosomal flanking regions with the Arctic *ligA* gene 47
2.3.3. Cloning the assembled constructs into the mycobacterial suicide vector....... 47
2.3.4. Screening primary and secondary recombinants in *M. smegmatis* 48
2.3.5. Determining restrictive temperature of TS strains.. 50
2.3.6. Growth of TS *M. smegmatis* in broth.. 51
2.3.7. Reversion rate of TS *M. smegmatis* 52
2.3.8. Genetic Drift experiments... 53
2.3.9. *ligA* variants constructed for gene replacement in *M.TB*.............................. 53
2.3.10. Gene replacement in *M.TB* 54
2.4. Discussion and future directions 58
Chapter 3- Strains and genetic elements for directed evolution of essential genes 62
3.1. Introduction 62
3.1.1. Directed Evolution 62
3.1.2. The need to engineer TS strains by directed evolution................................. 64
3.1.3. Objective of this research.. 65
3.1.4. Method designed to mutagenize, screen and select for functional TS *M.TB* mutants 65
3.2. Materials and Methods 69
3.2.1. Designing constructs to create a conditionally replicating plasmid................ 69
3.2.2. Creating the pSEVA-kan plasmid... 71

3.2.3. Cloning the assembled constructs into the pSEVA-kan plasmid.................... 72
3.2.4. Testing the conditionally replicating plasmids 73
3.2.5. Preparing pREDi-pSEVA electrocompetent cells 73
3.2.6. Creating knockout strains of *E. coli* using pREDi system.............................. 74
3.2.7. Testing the system using *E. coli prsA* essential gene..................................... 75
3.2.8. Cloning the *Salmonella prsA* into the pBBR-MCS5 plasmid and transforming into the knockout strain.. 76
3.3. Results... 77
3.3.1. *EssG* constructs assembled using yeast mediated recombination.................. 77
3.3.2. pSEVA514-kan plasmid with *prsA* construct dilutes out efficiently............. 78
3.3.3. *prsA* essential gene knockout created 78
3.3.4. *Salmonella* WT *prsA* supports growth of the *E. coli* strain 80
3.4. Discussion and future directions.. 81
Conclusions.. 84
Bibliography 85
Appendix.. 94

List of Tables

Table 1: Primers designed to amplify the *ligAPh* gene and the *M. smegmatis* flanking regions.. 31
Table 2: Primers designed to screen yeast transformants having correctly assembled constructs. 33
Table 3: Primers designed to amplify the *ligAPh* gene and clone it into the p2NIL vector. 34
Table 4: Primers used to amplify three variants of *ligAPh* gene and the *M. TB* flanking regions 39
Table 5: Primers designed to clone the *M.TB ligA* constructs into the pSTKO-*galK* plasmid. 43
Table 6: The CAI value and the GC content of the *ligA* gene from *M.TB*, *M. smegmatis* and *P. haloplanktis* 46
Table 7: Table representing different tetO-containing promoters with their strengths, and the genes whose expression is controlled by these promoters. 70
Table 8: Primers designed to create the *prsA* construct with the *prsA essG,* the *pir* gene and tetO-containing promoters. 71
Table 9: Primers designed to amplify the *cat* gene and the *E. coli prsA* flanking regions to create the linear chloramphenicol cassette 75
Table 10: Primers designed to clone the the *E. coli* WT *prsA* gene into the pBBR-MCS5 plasmid. 76

List of Figures

Figure 1: Overview of *M. tuberculosis* infection. 4
Figure 2: TB incidence in the world in 2012. 10
Figure 3: Overview of the different types of vaccine candidates for different target populations and stage of vaccine administration. 15
Figure 4: Temperature distribution in the human body from the surface to the core. 19
Figure 5: Genes upstream and downstream of the WT *ligA* gene in the *M. smegmatis* and *M.TB* genome. 26
Figure 6: Allelic gene replacement of the native *M. smegmatis (M.smeg) ligA* gene with the psychrophilic *ligA* using a mycobacterial suicide vector 28
Figure 7: Recombineering using mycobacteriophage proteins. 28
Figure 8: Three different *ligAPh* variants constructed for allelic gene replacement.. 38
Figure 9: Schematic representation of the linear constructs designed for mycobacterial recombineering. 40
Figure 10: Diagrammatic representation of the pSTKO-*galK* mycobacterial suicide vector. 43
Figure 11: *ligAPh* gene recombined with *M. smegmatis* flanking regions. 47
Figure 12: p2NIL clone with the assembled *ligA* construct. 48
Figure 13: *M. smegmatis* merodiploid streaked on M7H9-kan plates with ADC enrichment media and colony PCR of the *M. smegmatis* merodiploid. 49
Figure 14: *M. smegmatis* secondary recombinants replica plated on M7H10-ADC agar plates at 30ºC and 44ºC.. 50
Figure 15: TS and WT *M. smegmatis* streaked on M7H10-ADC plates to determine the restrictive temperature of the TS strain. 51
Figure 16: Growth of TS *M. smegmatis Vs.* WT *M. smegmatis* in broth. 52
Figure 17: Two-week-old streak of *M.TB hygR* colonies on M7H11-OADC-hyg plates resulting from transforming recombineering strain of Erdman (Erdman/pJV53) with the assembled construct. 54
Figure 18: *M.TB-ligA* construct assembled using Geneious.. 55
Figure 19: Sequencing results aligned with the *M.TB ligA* gene using Geneious software 55
Figure 20: Screening primary merodiploids obtained by Pavelka method. 57
Figure 21: Schematic representation of directed evolution experiment. 64
Figure 22: Directed evolution approach for creating TS essential strains 68
Figure 23: Constructs designed to create a conditionally replicating plasmid. 70
Figure 24: pSEVA514-kan plasmid with the R6K ori which depends upon the Pi protein, a product of the *pir* gene, to initiate replication 72
Figure 25: Screening putative clones containing correctly assembled for directed evolution 77
Figure 26: Screening for *prsA essG* gene knockouts. 79

Acknowledgments

First and foremost, I would like to thank my supervisor Dr. Francis Nano for giving me the opportunity to work in his lab. Thank you for your guidance and encouragement throughout my time in Victoria. This is truly invaluable.

I also want to thank my committee members Dr. Caroline Cameron and Dr. Perry Howard for their valuable feedback and support which has helped me progress in my research work.

I want to thank all the members of the Nano Lab especially Daniel Kemp, Stephanie Puckett, Barry Duplantis, Clara McDonald, Stephanie Pearce, Jarek Pankowski and Ralph McWhinnie, for their assistance and support in the lab. A special thanks to Sheila Potter for her assistance with the research carried out in the level-3 BSL facility.

A special thanks to all my friends for their support in the lab, during my writing and coursework. Wish you all the very best in accomplishing your career goals.

Dedication

I would like to dedicate this work to my parents who have encouraged me and supported me always. Thank you Mom and Dad!

List of Abbreviations

ADC	Albumin-dextrose-sodium chloride
APCs	Antigen presenting cells
BCG	Bacillus Calmette Guérin
CAI	Codon adaptation index
DCs	Dendritic cells
2-DOG	2-deoxygalactose (counterselective marker for the *galK* gene)
essG	Essential gene
GC	Guanine-cytosine
HIV	Human immunodeficiency virus
IFN-γ	Interferon-gamma
IL	Interleukin
iNOS	Inducible nitric oxide synthase
LAVs	Live attenuated vaccines
ligACp	*ligA* gene from *Colwellia psychrerythreae*
ligAPh	*ligA* gene from *Pseudoalteromonas haloplanktis*
MCS	Multiple cloning site
MHC	Major Histocompatibility complex
MDR TB	Multiple drug-resistant TB
M.TB	*Mycobacterium tuberculosis*
NK cells	Natural killer cells
NO	Nitric oxide
PAMPs	Pathogen-associated molecular patterns
PRRs	Pathogen recognition receptors
RD1	Region of difference-1
RNS	Reactive nitrogen species
ROI	Reactive oxygen intermediates
ROS	Reactive oxygen species
rBCG	recombinant BCG

TB	Tuberculosis
tetR	Tetracycline repressor
tetO	Tetracycline operator
TL buffer	Tris lysis buffer
TLRs	Toll-like receptors
TNF-α	Tumour necrosis factor- alpha
TS	Temperature-sensitive
URA-DO	Uracil drop-out
WHO	World Health Organization
WT	Wild type
XDR TB	Extreme drug-resistant TB

Chapter 1

Introduction

1.1. *Mycobacterium tuberculosis*

Mycobacterium tuberculosis (M.TB) is a gram-positive acid-fast bacterium from the family *Mycobacteriaceae* and is the main causative agent of tuberculosis (TB). Robert Koch was awarded the Nobel Prize in Medicine or Physiology in the year 1905 for discovering *M.TB* as a pathogen and elucidating the etiology of tuberculosis (T.M. Daniel, 2006). *M.TB* is a slow growing organism that divides every 15-20 hours and takes around 21 days to grow on agar plates. A unique characteristic of this bacterium is its waxy mycolic acid cell wall which makes it impermeable to certain dyes and stains, resistant to antibiotics and killing by host defense mechanisms, and which allows the organism to survive in a dry state for weeks (Murray *et al*, 2005).

1.2. Pathogenesis of TB

1.2.1. Latent and Active TB

TB is primarily a lung disease. However, in 30% of cases it can also lead to extrapulmonary disease (O'Garra *et al*, 2013). TB infection results when the contaminated aerosols harboring TB bacilli inhaled by the healthy individual bypass the bronchial defense mechanisms and enter into the alveoli of the lungs. In the alveoli, the bacilli are rapidly phagocytosed by the alveolar macrophages, neutrophils (Eum *et al*, 2010) and dendritic cells (DCs) (Mihret *et al*, 2012; WHO TB Report 2013). TB infection in the lungs can be presented in either an active or latent form.

Upon initial TB infection, host defense mechanisms prevent the multiplication of bacteria in 90% of the cases. These individuals are usually asymptomatic and referred to as being latently infected. This group can only be diagnosed with a positive tuberculin skin test (Brändli, 1998). In about 5% of the infected cases, *M.TB* overcomes these defense mechanisms and the disease progresses to the active state within several weeks or months. This type of infection is often referred to as primary progressive or symptomatic form of tuberculosis. The infection reaches the regional lymph nodes and is disseminated via the blood stream into other parts of the body such as the liver, kidneys, meninges and other body organs. The symptoms of active TB disease may include cough, fever, weight loss, night sweats, hemoptysis, thoracic lymphadenopathy and lung cavities (O'Garra *et al*, 2013).

1.2.2. Overview of events following infection with *M. tuberculosis*

Once in the lungs, *M.TB* entry into the macrophages is mediated by a diverse array of receptors such as pattern recognition receptors (PRRs) like toll-like receptors (TLRs) which produce proinflammatory cytokines and chemokines (Kleinnijenhuis *et al*, 2011). Neutrophils and monocytes are the first to arrive at the site of infection; they phagocytose the bacteria and release more cytokines and chemokines which drive the recruitment of more leukocytes and dendritic cells (DCs) to the site of infection (Kleinnijenhuis *et al*, 2011; Sakamoto, 2012). DCs are the main antigen presenting cells (APCs) that migrate to the nearest lymph nodes and present *M.TB* antigens to the naive T-cells, thus stimulating infected macrophage killing.

Viable and virulent *M.TB* bacilli escape this killing by preventing phagolysosomal fusion and acidification of the phagosomal compartment (Sturgill-Koszycki *et al*, 1994),

adapting themselves to the intracellular environment and creating a niche to reside inside the macrophages. Finally, a well-organized granuloma develops which consists of a clustering of immune cells (mainly lymphocytes) surrounding a core containing macrophages harboring live *M.TB* bacilli. This form of TB, where the bacilli are contained inside granulomas and kept in check by the host's immune system is called latent TB infection. The granulomas may remain in a dormant state for a lifetime provided the individual's immune system remains healthy (O'Garra *et al*, 2013; Weiner *et al*, 2014).

The latently infected individual can remain asymptomatic for years unless the granuloma containment breaks open due to conditions that deplete the immune system such as HIV, old age, malignant disease or malnutrition. This can give rise to reactivation and the spread of TB (Verver *et al*, 2005). This secondary TB occurs in 5-10% of latently infected individuals for reasons not well understood and is largely responsible for the pervasiveness of *M.TB* as a pathogen.

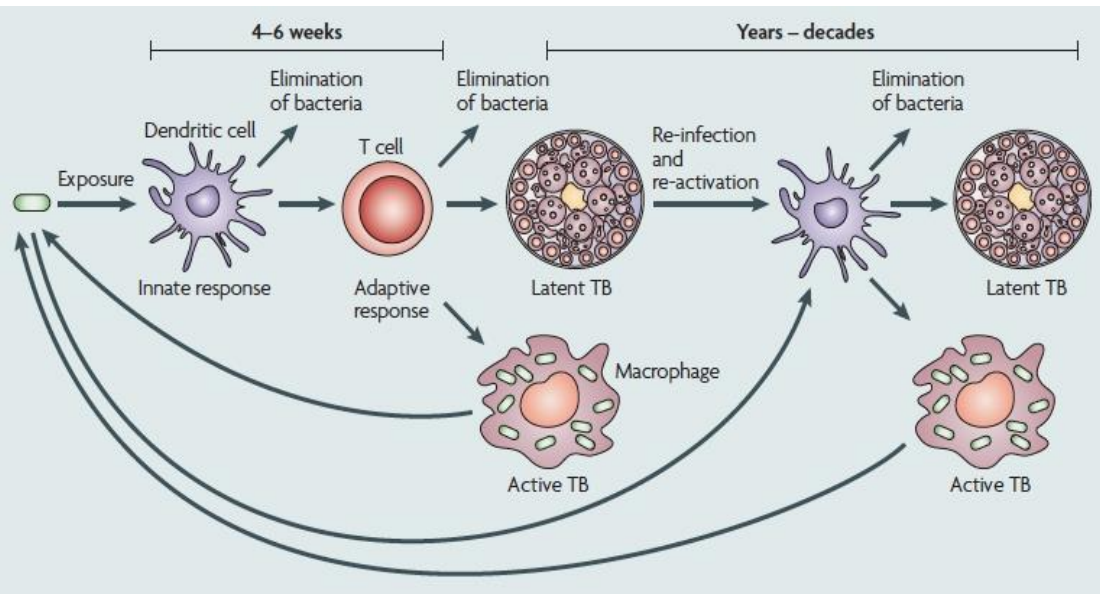

Figure 1: Overview of *M. tuberculosis* infection.
***M.TB* is mainly transmitted through aerosols from individuals with active disease. Once in the alveoli of the lungs, *M.TB* is rapidly phagocytosed by the macrophages and transferred to the lysosome for degradation. However, some bacilli can escape this killing and reside within the macrophages. Dendritic cells engulf the bacteria and present them to T-cells in the draining lymph node. These primed T-cells further activate the macrophages by releasing various cytokines, which generally results in the clearance of infection. If the T-cell response is unable to control the initial infection, symptoms develop within a year of infection and this form of disease is called primary progressive TB. In most individuals, the TB bacilli cannot be completely cleared by the immune system which gives rise to latent TB or asymptomatic TB. In this case, the bacilli are contained inside a mass of cells called granulomas. Latent TB carries a risk of secondary disease following re-infection or reactivation of the initial infection (Adapted from Young *et al*, 2008).**

1.3. Host immune response against TB

M.TB infection elicits both an innate and adaptive immune response. This immune response is directed towards the control and clearance of the pathogen. This includes induction of IFN-γ and TNF-α and the production of reactive oxygen intermediates (ROI) and reactive nitrogen intermediates (RNI). Both $CD4^+$ and $CD8^+$ T-cell response is required for bacterial clearance and the control of *M.TB* infection (O'Garra *et al*, 2013). Moreover, the role of B-cells is also being determined during TB infection (O'Garra *et al*, 2013).

1.3.1. Innate immune response

The *M.TB*-macrophage interactions, role of macrophages and *M.TB* defense mechanisms against host immune response can be summarized as follows:

1.3.1.1. Binding of *M.TB* to macrophages

The innate immune response against *M.TB* is characterized by multiple pattern recognition receptors (PRRs) expressed on the macrophages and DCs which recognize the pathogen-associated molecular patterns (PAMPs) on the bacteria. These PRRs are considered as the first line of defense in response to an invading pathogen. One of the best characterized classes of PRRs, TLR-2 binds to *M.TB* and releases proinflammatory cytokines and chemokines in response to mycobacterial cell wall components such as lipoarabinomannan (LAM), lipomannan (LM) and phosphatidylinositol mannosides (PIMs) (Sakamoto, 2012). These cytokines and chemokines further activate the macrophages, DCs and neutrophils which destroy the TB bacilli and control the spread of infection. They regulate the anti-mycobacterial defense mechanisms of macrophages such as the production of ROI and RNI (Raja, 2004). Two major cytokines IFN-γ and TNF-α upregulate nitric oxide synthase 2 (NOS2) expression and induce the production of RNI within the phagolysosome which results in intracellular killing of *M.TB* (Flynn *et al*, 2001). Cytokines also elicit an adaptive immune response and arrest bacterial growth further controlling the spread of TB infection (Hossain *et al*, 2013).

1.3.1.2. Phagolysosomal fusion

The fate of *M.TB* in the macrophage is one of the most interesting aspects of mycobacterial pathogenesis. After *M.TB* is phagocytosed by the macrophages and the phagosome is fused to the lysosome, it is subject to degradation by acidic hydrolases in the lysosomal compartment. Macrophages activated by cytokines like TNF-α and IFN-γ

can protect themselves from mycobacterial defense mechanisms due to stimulation of inducible NOS (iNOS) which in turn produces nitric oxide and is toxic for *M.TB* (Chan *et al*, 1992). However, studies show that *M.TB* has evolved various strategies to prevent its destruction by the host defense mechanisms. Viable bacilli can prevent the phagolysosomal fusion and survive inside the macrophages (Flynn *et al*, 2001; Raja, 2004). Mycobacterial sulfatides, derivatives of multiacylated trehalose 2- sulfate as well as large amounts of ammonia generated in culture are known to inhibit this fusion. However, the role that ammonia plays in preventing fusion is still debated (Flynn *et al*, 2001; Raja, 2004). *M.TB* is also known to prevent phagosome acidification (Sturgill-Koszycki *et al*, 1994), and maturation (Ehrt *et al*, 2009). Phagosomes containing mycobacteria are also prevented from associating with iNOS limiting exposure to nitrogen radicals, all of which prevents effective antigen processing (Flynn *et al*, 2001; O'Garra *et al*, 2013).

1.3.1.3. Macrophage apoptosis

In addition to ROI and RNI, another mechanism involved in macrophage defense against *M.TB* is apoptosis. This programmed cell death is mediated through the downregulation of the apoptosis inhibitor bcl-2 (Klingler *et al*, 1997). Apoptosis plays a role in host immune response by eliminating the niche for *M.TB* growth. It has direct anti-microbial effects on TB bacilli and packages them in apoptotic bodies. These apoptotic bodies are engulfed by the newly recruited macrophages and DCs, which help in eradicating *M.TB* and stimulating an adaptive immune response (Lee *et al*, 2009).

1.3.1.4. Nramp, neutrophils and NK cells

Phagocytosis and subsequent cytokine production are initiated in the absence of prior exposure to antigens and thus form components of innate immunity. The other innate immunity components include neutrophils, natural resistance associated macrophage proteins (*nramp)*, and natural killer cells (NK cells) (Raja, 2004). Neutrophils are the first cells to arrive at the site of infection and can kill the *M.TB* bacilli using anti-microbial molecules enclosed in their granules such as defensins, lactoferrin and lysozymes (Korbel *et al*, 2008). *Nramp* functions in transporting nitrate from the intracellular sites to the more acidic environment like the phagolysosome where it is converted to nitric oxide (NO). NK cells can directly lyse the *M.TB* infected macrophages due to their cytotoxic functions exerted through perforin and granzyme or granulysin (Korbel *et al*, 2008).

1.3.2. Adaptive immune response

Since it resides within the macrophages, *M.TB* is a classic example of a pathogen that is cleared due to cell mediated immune response. The role of $CD4^+$ T-cells is best understood compared to the role of $CD8^+$ T-cells or B-cells during *M.TB* infection.

$CD4^+$ T helper type 1 cells (Th1 cells) play an important role in protective immune response against TB as *M.TB* is an intracellular pathogen. These cells recognize *M.TB* antigens presented by MHC class II molecules on DCs and macrophages and are initially primed in the draining lymph nodes of the infected lung (Ottenhoff *et al*, 2012). They release proinflammatory cytokines including IFN-γ which activate macrophages and kill intracellular *M.TB* through production of NO and ROS. One of the main characteristics upon *M.TB* infection is the delayed initiation of the T-cell immune response. Some of the reasons for this delay include the inhibition of macrophage apoptosis (Blander *et al*,

2007) which delays the antigen presentation of DCs and the induction of IL-10, which in turn prevents the production of cytokines IFN-γ and IL-17 by $CD4^+$ T-cells (Redford *et al*, 2011).

Infected DCs and macrophages can also present *M.TB* antigens to $CD8^+$ T-cells via class I MHC molecules. $CD8^+$ T-cell activation can lead to the release of various cytokines such as TNF-α and IFN-γ which can activate macrophages to kill intracellular bacteria. Other pathways to mediate bacterial killing are the perforin and granulysin-mediated pathway or by induction of macrophages apoptosis by expression of Fas ligand (Weerdenburg *et al*, 2009).

Antigens from intracellular pathogens are usually present in the cytosol of the infected cells and presented to $CD8^+$ T-cells via class I MHC molecules. For a very long time *M.TB* was thought to reside in the phagosome, giving rise to a question as to how the *M.TB* antigens can be presented via MHC Class I molecules to $CD8^+$ T-cells. However, recent studies indicate that *M.TB* can escape the phagosome and reside in the host cell cytosol (Weerdenburg *et al*, 2009). This knowledge has been used to design a recombinant BCG vaccine against TB which expresses perfringolysin, allowing BCG to escape into the cytosol and provide protection though Class I MHC presentation (Sun *et al*, 2009). There are many questions which still remain unanswered regarding the role that $CD8^+$ T-cells play during *M.TB* infection.

Although B-cell immune response is essential against a broad range of pathogens, the role of B-cells in *M.TB* infection is unclear. Researchers have dismissed their importance for a long time as *M.TB* is an intracellular pathogen. However, recent evidence of B-cell aggregates found in the lungs of TB patients and granulomas of *M.TB*

infected mice suggest a role of B-cells in TB infection which still needs to be understood (Maglione *et al*, 2009).

1.4. Burden of TB disease

TB is primarily a lung disease and a leading cause of illness worldwide, second only to human immunodeficiency virus (HIV). According to the World Health Organisation (WHO) global health consensus, this disease accounted for 8.6 million new cases and 1.3 million deaths in the year 2012, with most of the estimated TB cases in Asia and Africa (Figure 2) (WHO Global TB Report, 2013).

One of the major threats to the control of this disease is drug-resistant TB (DR-TB). *M.TB* has acquired multiple types of drug resistance, including multiple drug-resistant TB (MDR-TB), which is resistant to first line drugs like rifampicin and isoniazid, and XDR-TB, which in addition to first line drugs is also resistant to second line drugs like amikacin and kanamycin. In 2012, there were an estimated 300,000 MDR-TB cases and the diagnosis and treatment of these strains has been very challenging (WHO Global TB Report, 2013).

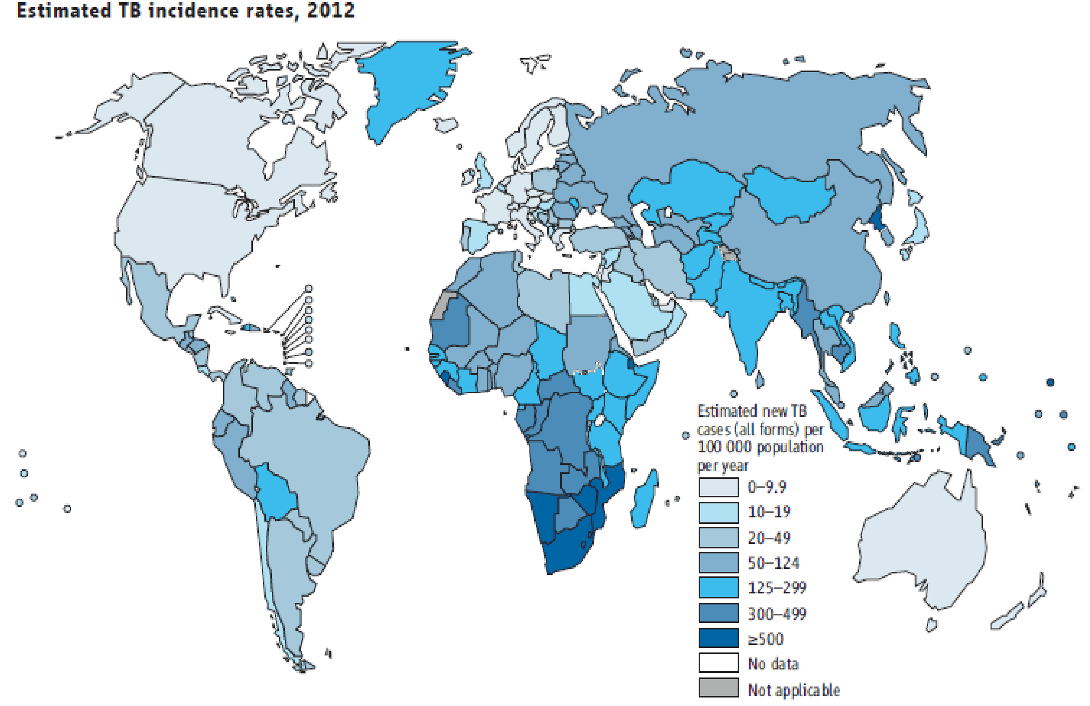

Figure 2: TB incidence in the world in 2012.
The estimated incidence rate of TB is much lower in developed countries like USA, Europe and Canada with ~10 cases per 100,000 people as compared to South Africa and Zimbabwe which have an incidence rate of above 500 cases per 100,000 people (Adapted from WHO Global Tuberculosis Report, 2013).

1.5. Diagnosis and TB treatment

The conventional tests used for screening and diagnosis of TB are the Tuberculin skin test (TST), sputum smear microscopy, a chest X-ray and bacterial culture.

1.5.1. Tuberculin skin test (TST)

The tuberculin skin test (TST) is a safe and inexpensive way of crudely indicating whether the person is infected with TB. It is used to detect TB in infected individuals and screen high-TB risk populations to guide TB control efforts (O'Garra *et al*, 2013).

A tuberculin solution is a glycerol extract of the "tubercle bacillus". In the TST, the tuberculin solution is intradermally injected into the forearm and the induration is detected within 48-72 hours. The person exposed to the bacteria is expected to mount an

immune response in the skin due to the infiltration of macrophages, lymphocytes and local edema at the site of infection. The effectiveness of this test is limited as it cannot exclude TB disease if it is suspected. False negative tuberculin tests are found in 15-25% of patients suffering from active tuberculosis due to reasons such as viral infections or HIV (Korzeniewska-Kosela *et al*, 1994). Additionally, the size of the induration does not help to distinguish between active or latent TB infection (O'Garra *et al*, 2013).

1.5.2. Microbiological studies

Another method for the diagnosis of TB is by culturing *M.TB* from a specimen such as sputum and biopsy tissue collected from the patient. This sample is then stained and observed under a microscope. This test however, does not distinguish between tuberculous and non-tuberculous mycobacteria. Furthermore, the lengthy cultivation time required to grow *M.TB*, (up to 6 weeks) can lead to a delay in diagnosis (Steingart *et al*, 2006).

1.5.3. Radiography

Chest X-ray and CT scans are radiographic methods used for the detection of chest abnormalities. Active TB is characterized by infiltrates or cavities which may appear anywhere in the lungs (O'Garra *et al*, 2013). The presence of these abnormalities can suggest TB infection and help rule out the possibility of pulmonary TB disease in an individual with a positive TST. However, this test must be used in combination with other microbiological tests to definitively diagnose TB.

1.5.4. TB Treatment

Due to the slow growth of the disease-causing organism, active TB disease needs to be treated with first line antibiotics such as rifampicin and isoniazid for a minimum of six

to nine months. TB treatment can be divided into two phases: the first intensive killing phase that destroys actively replicating bacteria, followed by the second phase which targets the persisting bacteria (O'Garra *et al*, 2013). Moreover, most drugs used for TB treatment are toxic to humans and have unpleasant side effects. These side effects in addition to the long treatment duration can lead patients to discontinue their treatment prematurely, leading to the emergence of multiple drug-resistant (MDR) and extensively drug-resistant (XDR) TB. These strains require the use of multiple second line drugs and the treatment to be continued for 1-2 years.

1.6. Current vaccines against TB

Vaccines are defined as biological preparations that elicit immune response against an infectious disease (Clem, 2011). A vaccine is mostly a weakened or killed organism or toxin/surface protein that resembles the disease-causing organism. It stimulates the body to recognize a foreign agent, kill it and remember it so that upon later encounter with the pathogen, it can recognize and destroy it. Vaccines come in a variety of forms and induce different types of immune protection. These forms include whole-cell killed, subunit, toxoid, DNA and live attenuated vaccines (Clem, 2011). When designing a vaccine, one must take into account the lifestyle of the pathogen (intracellular/ extracellular) and the kind of immune response being generated.

1.6.1. The current BCG vaccine

A major breakthrough in the fight against TB was the development of the Bacillus Calmette Guérin (BCG) vaccine by Albert Calmette and Camille Guérin (Liu *et al*, 2009). This strain is a live attenuated form of *Mycobacterium bovis*, which causes cattle TB.

The BCG vaccine was first created by subculturing the bacteria nearly 230 times in ox-bile detergent and glycerol-soaked potato slices until the *M. bovis* strain lost its virulence properties (Liu *et al*, 2009). These cultures were then distributed to different laboratories around the world to manufacture the vaccine in different countries. As BCG is a live vaccine, there was need to passage the strain in fresh media every few weeks. Different passaging conditions in different laboratories gave rise to different strains of BCG. Four BCG strains that are majorly used are the BCG-Pasteur, BCG-Japan, BCG-Danish and BCG-Glaxo (Liu *et al*, 2009; Behr, 2002). Very little is understood about the attenuation of the vaccine leading to a safe strain. However, studies have identified a deletion of the Region of Difference-1 (RD1) in all strains of BCG which is present in both *M. bovis* and *M.TB* (Mahairas *et al*, 1996).

1.6.2. Need for a new vaccine

The BCG vaccine was developed to prevent serious forms of TB in infants. This vaccine was administered to all neonates in areas with high prevalence of TB. Although BCG protects 80% of the children against severe forms of tuberculosis like miliary and meningeal TB, it imposes a high risk to immunodeficient individuals (Hesseling *et al*, 2006).

The protection provided by BCG vaccine is transient and the vaccine is ineffective against pulmonary TB in adolescents and adults, which is the most prevalent form of TB today (O'Garra *et al*, 2013). It also causes BCGosis which is a disseminated form of BCG infection, in HIV and other immunocompromised individuals (Mansoor *et al*, 2010). Although TB can be treated with combination of antibiotics, treating TB cases cannot prevent disease transmission in highly endemic populations. Thus, there is a need

to develop a safe and effective vaccine against tuberculosis that prevents the establishment of disease in a susceptible host and controls TB progression to an active state (WHO Global TB Report, 2013).

1.6.3. New vaccines in the pipeline

In the early 1990s, WHO declared TB as a global emergency leading to significant progress in the study of this disease. The development of techniques for genetic manipulation of mycobacteria, the sequencing of the *M.TB* genome and the progress in understanding the immunology of the disease has provided us with an opportunity to develop much more effective TB vaccines (WHO Global TB Report, 2013).

In the past decade, two strategies for vaccine development against TB have been used. The first strategy is to develop a vaccine that would be more highly efficacious than the current BCG vaccine and replace it – such as the live attenuated *M.TB* strain or the recombinant version of BCG. The second method is called the “prime boost” strategy in which the current BCG vaccine is first administered to infants and a new vaccine will then be given as a booster dose, with the aim to improve the efficacy of the current BCG vaccine and provide long lasting immunity (Romano *et al*, 2012).

Vaccines developed against TB are divided into four broad categories: i) recombinant viral vaccines encoding *M.TB* putative protective antigens (Tameris *et al*, 2013), ii) subunit vaccines containing putative protective antigens encapsulated in liposomes or T-cell stimulating adjuvants (Day *et al*, 2013) iii) live attenuated vaccines (LAVs) that include recombinant BCG (Grode *et al*, 2013) or genetically attenuated *M.TB* vaccines (Arbues *et al*, 2013) and iii) therapeutic vaccines (Yang *et al*, 2011) administered with the purpose to eradicate *M.TB* organisms from the human body and

prevent relapse or re-infection. At present, there are 12 vaccine candidates in different phases of clinical trials (Weiner *et al*, 2014) and their status as of July 2013 is summarized in Figure 3. Majority of these vaccines aim to prevent TB disease, either by blocking TB infection upon exposure to TB bacilli (pre-exposure) or interfering with the reactivation of latent TB (post-exposure).

Target populations	Infection/Disease	Vaccine type	Advanced Candidates
Infant	Uninfected	Preexposure/Preventive BCG replacement	rBCG: VPM1002 r-Mtb: MTBVAC
Infant	Uninfected BCG	Preexposure/Preventive Prime-boost	Viral vectored: MVA85A/Aeras-485 Protein/adjuvant: H4:IC-31
Adolescent/ Adult	LTBI/BCG (TST^{+})	Postexposure/Preventive Prime-boost	Viral vectored: MVA85A/Aeras-485 Protein/adjuvant: M72:AS01$_{E}$ H56:IC-31 ID93:GLA-SE
Adolescent/ Adult	Active TB	Therapeutic	Killed mycobacteria: *M. indicus pranii* *M. vaccae* RUTI

Figure 3: Overview of the different types of vaccine candidates for different target populations and stage of vaccine administration (Adapted from Weiner *et al*, 2014).

1.6.3.1. Recombinant viral vaccines and subunit vaccines

A major pathway in the development of improved TB vaccines is the development of subunit vaccines. Subunit vaccines are non-live or non-replicating in case of viral vectors, and are either delivered through viral vector systems or as recombinant proteins mixed with T-cell stimulating adjuvants. These vaccines are mostly used as booster vaccines after being initially primed with BCG or recombinant BCG (rBCG) or attenuated *M.TB* vaccines and are aimed at providing long-lasting protective immunity. These vaccines are considered safe to be delivered to immunocompromised individuals.

An example of a recombinant vector based vaccine is MVA85A (Modified Vaccinia Ankara, MVA) developed by the University of Oxford (Tameris *et al*, 2013). This vaccine expresses the *M.TB* antigen 85A and has completed the phase IIb clinical trial. However, this vaccine did not show better vaccine efficacy as compared to the current BCG vaccine (Tameris *et al*, 2013). M72+AS01$_E$ is an example of a subunit vaccine which is a fusion protein of *M.TB* antigens 32A and 39A encapsulated in the adjuvant AS01$_E$. This vaccine has completed the phase IIa clinical trials in South Africa and is found to provide good safety and immunogenicity (Day *et al*, 2013).

1.6.3.2 Therapeutic vaccines

In addition to designing vaccines that would induce protective immunity and prevent new TB infections, a vaccine that would completely eradicate *M.TB* and prevent reactivation in latently infected individuals is also required. RUTI and *Mycobacterium vaccae* are the therapeutic vaccines undergoing clinical trials. RUTI is used to complete latent TB treatment after a short duration of antimicrobial therapy and uses constitutes of detoxified liposomal fragments of *M.TB* (Montagnani *et al*, 2014). *Mycobacterium vaccae* is a non-living preparation of the organism that is also aimed at being used with antimicrobial therapy (Yang *et al*, 2011).

1.6.3.3. Live attenuated vaccines (LAVs)

LAVs were among the first vaccines developed to induce immunity against intracellular pathogens such as smallpox (cowpox vaccine, 1796) and *M.TB* (BCG vaccine, 1908-1920) (Baxby, 1977; Calmette, 1931). These attenuated strains can replicate within the host but are incapable of causing disease due to the mutations incorporated. As LAVs are live bacteria, they can mimic a natural infection and elicit a

strong immune response which includes $CD4^+$ and $CD8^+$ T-cell activation (Behar *et al*, 2007). Although LAVs are the best vaccines available against intracellular pathogens, they have some disadvantages. Particularly, the vaccine being a live pathogen poses high risk to immunocompromised individuals and the regulatory requirements are very strict due to safety concerns.

LAVs against TB are aimed at replacing the current BCG vaccines by rBCG or genetically attenuated *M.TB* vaccines. The rBCG vaccines are aimed at being safe, highly immunogenic and giving long lasting protection, even against the highly virulent MDR and XDR strains. Examples of live attenuated vaccine candidates currently in clinical trials are VPM1002 and MTBVAC (WHO TB Report, 2013).

VPM1002 is a live recombinant BCG strain that expresses listeriolysin of *Listeria monocytogenes* and contains the hygromycin marker replacing the urease gene (BCG ΔureC::hly HmR, VPM1002). This rBCG strain aims at improved release of BCG-derived antigens into the cytosol and increased apoptosis of infected host cells *in vitro* (Grode *et al*, 2013). The MTBVAC is a double deletion mutant of *M.TB* with the phoP and fad26 genes deleted. PhoP/PhoR is a transcriptional regulator that regulates the transcription of *M.TB* virulence genes and fad26 gene is involved in synthesis of cell wall lipid phthocerol dimycocerosates (DIM) which plays a role in *M.TB* virulence. This is the first live attenuated *M.TB* vaccine to enter Phase I clinical trials (Arbues *et al*, 2013).

1.7. Temperature sensitive (TS) viral vaccines

In addition to engineering attenuated forms of an organism, another approach to create live vaccines is by engineering bacteria that are sensitive to small increases in temperature. Several TS viral vaccines were engineered by passaging the virus repeatedly

in tissue culture at cool temperatures and making them cold-adapted (Dubes *et al* 1956, Maassab *et al*, 1999). These cold adapted viruses grew better when compared to their parent strains at cooler temperatures (Dubes *et al* 1956). These strains make use of the body temperature gradient described below and are capable of limited replication in the host (White *et al*, 2011).

TS viral vaccines have shown to be effective through the development of both the Sabin polio virus vaccine (Dubes *et al*, 1956) and the live attenuated influenza vaccine, also known as FluMist (Maassab *et al*, 1999). In particular, the FluMist vaccine is a live attenuated cold adapted viral vaccine and provides better protection as compared to the traditional inactivated influenza vaccines (Belshe *et al*, 2007). This virus fails to grow at 38°C, suggesting a good inactivation temperature for a TS vaccine (White *et al*, 2011). If a TS strain with low inactivation temperature (<33°C) is engineered, the vaccine will be very restricted in its ability to disseminate in the body. This low temperature may prevent replication at cool sites in the body thus, limiting the induction of immunity (Duplantis *et al*, 2011).

1.7.1. Temperature distribution in humans

To understand the potential use of TS vaccines, it is important to understand the temperature distribution in mammals. This is best understood using the two compartment model of body temperatures (Figure 4).

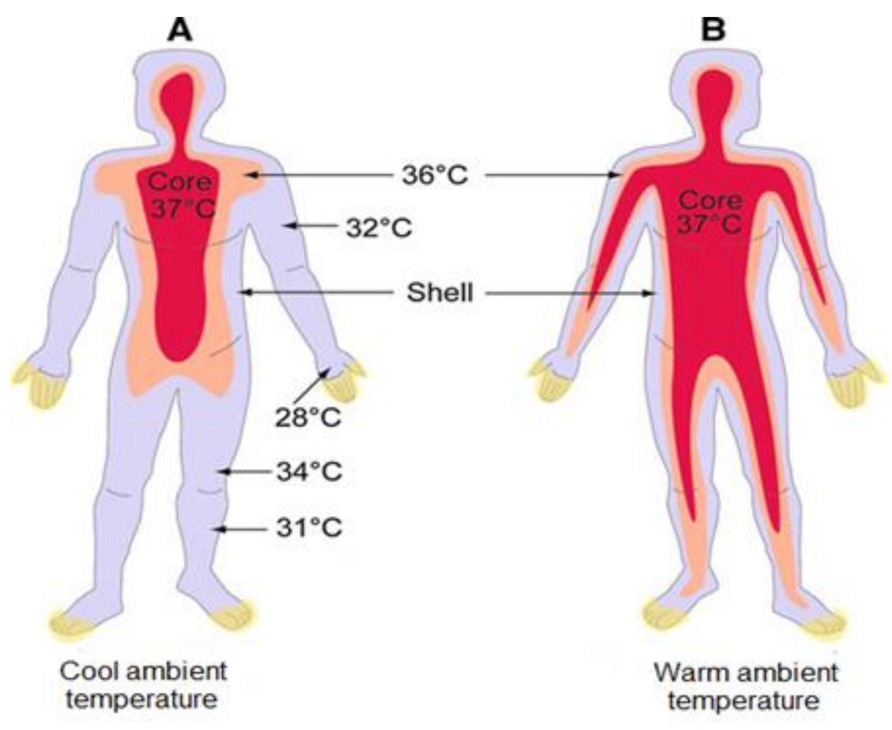

Figure 4: Temperature distribution in the human body from the surface to the core. The temperature in the periphery is lower compared to the core body temperature which is maintained at 37°C. The size of the core is reduced considerably at cool ambient temperature (A) in comparison to the size at warm ambient temperature (B). The yellow areas are the acral regions that control the body temperature through heat loss or heat gain (Adapted from White *et al*, 2011).

In this model, the body compartment is divided into two parts: the shell and the core. The boundaries of the two body compartments change based on the temperature of the surroundings. At a cool ambient temperature, the outermost shell is maintained at lower temperatures (28-31°C) compared to the core temperature which is maintained at 37°C (Figure 3A). At a warm ambient temperature, there are two main changes that occur in the body compartments. Firstly, the size of the core increases to include both the arms and legs. Secondly, the large temperature gradient that exists in the shell at cold ambient temperature is reduced considerably (Figure 3B). This change in the compartmental size is mainly due to vasoconstriction at cool ambient temperatures and vasodilation at warm

ambient temperatures (White *et al*, 2011). At a comfortable room temperature, the skin is usually found to be in the range of 32-35°C. Although the body temperature can vary significantly due to variations caused by clothing, gender, illnesses and repeated exposure to different climates, the core body temperature is regulated at approximately 37°C. There are some variations in the human body temperature that need to be considered when designing TS vaccines. However, we cannot forget the success of TS viral vaccines and also the fact that as long as an individual is alive, the core body temperature cannot dip below 37°C, allowing a gradient to exist from the body surface to the core (White *et al*, 2011).

1.7.2. TS vaccines target mucosal immunity

A common method for delivering a live vaccine is to deposit the vaccine material in the skin via scarification and then puncturing the skin to introduce the vaccine into different layers of the skin. The BCG vaccine and the small pox vaccine are examples of live vaccines that have generated successful immunity upon being vaccinated at the skin sites (Duplantis *et al*, 2011; White *et al*, 2011).

TS vaccines are generally introduced subcutaneously or intramuscularly at the skin sites or via droplets in the nose. This gives the TS strain an opportunity to replicate in the skin before it encounters the warm core body temperature. While replicating in the skin, the TS strain primes the dermal dendritic cells (DCs) which are the main APCs patrolling the dermal compartments. These DCs present the foreign antigen to the T-cells in the lymph node within 18 hours of exposure, allowing a quick and efficient Th1-cell mediated immune response to be generated before the strain reaches the body core. This Th1-mediated immune response is useful for defense against intracellular pathogens like

M. tuberculosis, S. typhi and *F. tularensis.* Thus, a TS vaccine generated against an intracellular pathogen like *M.TB* will target the dermal tissues rich in dermal dendritic cells (White *et al*, 2011).

1.7.3. TS bacterial vaccines

There are no examples of TS bacterial vaccines generated as human vaccines. However, TS bacterial vaccines have been used in veterinary medicine since the late 1900's and these vaccines were made TS by random chemical mutagenesis. Examples of TS bacterial vaccines generated are *Mycoplasma synoviae* (Markham *et al*, 1998) and *Bordetella avium* (White *et al*, 2011). However, as these vaccines were generated before rapid genome sequencing technology was available, the mutations that contributed to virulence and temperature sensitivity could not be determined (White *et al*, 2011).

1.8. Essential genes in bacteria

Due to advances in microbial genetics and molecular biology, a number of bacterial genomes have been sequenced. This has led to the identification of a number of essential genes and provides insight into the roles and functions of these essential genes in the bacterial genome. The exact definition of an essential gene is still under debate. However, for our research purposes, a gene will be considered essential if it is required for the survival of the organism under all growth conditions.

In the last few decades, researchers have developed interest in understanding more about the essential genes in different bacterial species. Understanding the genes that are necessary for survival provides insights into the basic elements of life (Glass *et al*, 2006) and can be used as potential targets to develop new antibiotics (Wilkinson *et al*, 2001).

Identifying genes that are essential for an organism has been difficult as current methodologies identify only non-essential genes, permitting us to infer that the remaining genes are essential (Gerdes *et al*, 2003). The most effective method to study the essentiality of genes involves targeted deletion of every gene within the genome and testing for cell viability (Ji *et al*, 2001; Baba *et al*, 2006). However, most of the methods have relied on identifying "missed hits" by transposon mutagenesis and outgrowth selection (Gerdes *et al*, 2003). Comparative genomics has allowed the identification of 127 genes that are conserved across different bacterial species. Good examples of essential genes conserved across different bacterial species are those which play a role in key cellular processes like DNA replication and protein synthesis.

1.9. Essential genes chosen for this study

Previous work by Barry Duplantis in the Nano lab demonstrated that substituting *ligA* gene that encodes the NAD+ dependent DNA ligase from psychrophilic organisms imparts temperature sensitivity to mesophilic organisms and the inactivation temperature of these TS strains range from 33-37°C (Duplantis *et al*, 2011). Thus, the *ligA* gene which plays an important role in DNA replication and repair has been chosen for this study. In addition to this gene, three other essential genes which play an important role in different metabolic pathways have been chosen and are summarised in the Appendix Table A1. The genes chosen have different functions in different metabolic pathways and we think that these genes will generate a range of TS phenotypes. This will give us an option to select from a large pool, the strains which have a desired inactivation temperature, and thus test their suitability as potential vaccine candidates.

Project Overview

Mycobacterium tuberculosis is the main causative agent of TB, which is the reason for approximately 1.3 million deaths worldwide in 2012. BCG, the only vaccine against TB and is ineffective against pulmonary TB, which is the most prevalent form of TB in adults today. A vaccine is needed that can prevent TB disease and control the progression of latent TB disease to an active state.

Research carried out by Barry Duplantis in the Nano lab showed that substituting psychrophilic essential genes into mesophilic organisms makes them TS and induces good protective immunity in mice (Duplantis *et al*, 2010). We therefore hypothesize that substituting TS essential genes into *M.TB* will make them TS and this TS strain will be able to induce a strong Th1 T-cell immune response. The applications of this TS *M.TB* strain might be its use as a potential vaccine candidate, in diagnostic testing and drug development.

In addition to the psychrophilic essential gene approach to create TS *M.TB*, we have attempted to engineer TS essential genes by using the directed evolution approach. The hypothesis that drives our desire to use directed evolution to the temperature-sensitivity of essential genes is that this approach will allow us to create a greater variety of genes encoding essential proteins than we would be able to discover using bio-prospecting of psychrophilic bacteria. We reason that directed evolution should allow us to make a wide variety of essential proteins temperature-sensitive and, for any one protein, we should be able to create a set of protein variants with a range of inactivation temperatures.

The first chapter provides background information on *M.TB*, including the current knowledge about host immune response against *M.TB* and the various strategies that this

pathogen has evolved to prevent its killing by the host. It also provides insight into the current TB vaccines in the pipeline and the background information necessary to understand the hypothesis. The second chapter describes the engineering of TS *M. smegmatis* strain by allelic gene replacement of the *ligA* essential gene with its naturally occurring homolog from the psychrophile *P. haloplanktis*. The third chapter describes the method developed to screen and select TS essential genes created by directed evolution.

In addition to the intellectual organization of this thesis, the reader should also recognize a practical constraint on the logic of the experiments that I performed. *M.TB* is a Level 3 human pathogen and all work with live bacteria has to be carried out in a Level 3 containment facility. Financial, regulatory and facility access constraints led to a division of labour. My work was confined to the Level 2 laboratory, and my role in a group effort was to support the goal of creating TS *M. tuberculosis* strains. My work, as described in this thesis, contributed in three areas: (1.) I created a TS *M. smegmatis* strain, demonstrating that the psychrophilic *ligA* gene could be used in mycobacterial species. (2.) I made several genetic constructs that were used to transform *M.TB* (done by Sheila Potter in the Level 3 facility) (3.) I developed the *E. coli* strains and plasmids that will allow for the next generation of experiments leading to the creation of TS strains of *M. tuberculosis*.

Chapter 2- Stable temperature sensitive *M. smegmatis* engineered by incorporating the Arctic *ligA* gene

2.1. Introduction

There are more than 100 different essential genes that are conserved across the domain bacteria (Gerdes *et al*, 2003). Introducing mutations into an essential gene that makes its product temperature-sensitive makes the entire organism temperature-sensitive (Maassab *et al*, 1985, Duplantis *et al*, 2010). Using this principle, Barry Duplantis from the Nano lab substituted psychrophilic essential genes into different mesophilic organisms which made them TS. A TS *M. smegmatis* strain with an inactivation temperature of 33°C was generated by substituting the psychrophilic *ligA* gene from *Colwellia psychrerythreae (ligACp)* into the mesophile *M. smegmatis.* This strain had 54% of the native *M. smegmatis ligA* gene deleted from its chromosome and was created by introducing the *ligACp* gene on a plasmid (Duplantis *et al*, 2010).

2.1.1. Objective of this research

Work by Barry Duplantis showed that substituting the *ligA* gene from the psychrophile *Pseudoalteromonas haloplanktis* into the mesophile *F. novicida* made the organism TS with a restrictive temperature of 37°C and provided a good immune response suggesting a good inactivation temperature for a successful vaccine (Duplantis *et al*, 2010). The objective of this experimental line of research was therefore, to create a TS *M. tuberculosis* strain with a restrictive temperature of 37°C by substituting the *ligAPh* gene into its chromosome. This TS strain generated could be used as a safe level-2 laboratory strain, a potential vaccine candidate and might be used in diagnostic testing and drug development.

The slow growth and infectious nature of *M.TB* make it necessary to first confirm the validity of the experiments in its research surrogate *M. smegmatis*. Therefore, our objective was to study whether the *ligAPh* gene can impart temperature sensitivity in *M. smegmatis. M. smegmatis* is commonly used as a model organism for *M.TB* as it is a fast-growing organism and non-pathogenic. Moreover, this organism has a waxy mycolic acid cell wall like other mycobacterial species, allowing transformations to be carried out in a similar way as we would in *M. tuberculosis* (Etienne *et al,* 2005).

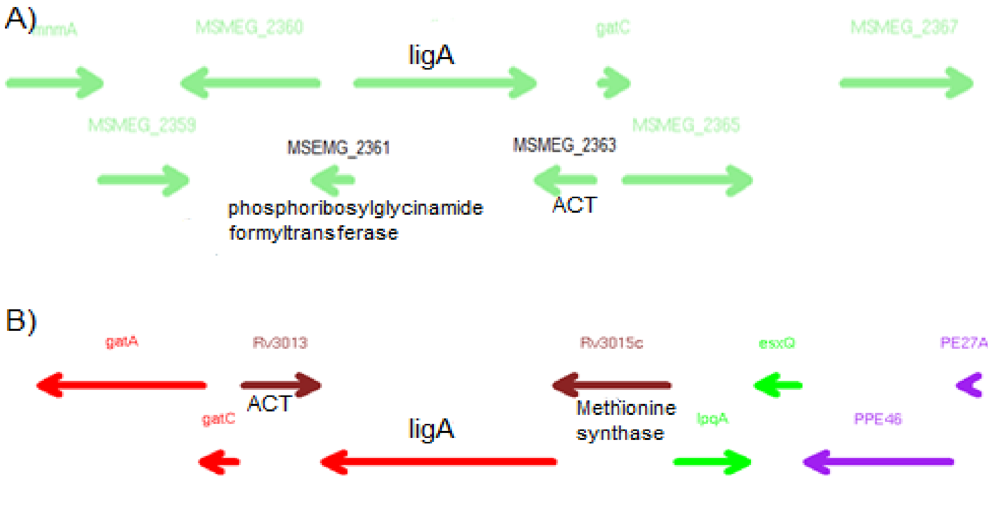

Figure 5: Genes upstream and downstream of the WT *ligA* gene in the *M. smegmatis* and *M.TB* genome. A) Genes upstream and downstream of the *ligA* gene in the *M. smegmatis* genome (Adapted from Smegmalist; Kapopoulou *et al*, 2010) B) Genes upstream and downstream of the *ligA* gene in the *M. tuberculosis* genome (Adapted from Tuberculist; Lew *et al*, 2010)

In the *M. smegmatis* genome, the *ligA* gene is flanked by the amino acid binding ACT domain-containing protein and the phosphoribosyl glycinamide formyltransferase both of which are non-essential (Figure 5A). In the *M.TB* genome, the *ligA* gene is flanked by the ACT domain-containing protein on one side and is in an operon with the methionine synthase gene on the other side (Figure 5B), both the genes being

non-essential. These differences were to be considered when substituting the *ligAPh* gene into the *M.TB* genome.

2.1.2. Gene replacement in *M. tuberculosis*

Allelic exchange in *M.TB* is very complex due to the high rate of illegitimate recombination and slow growth of the organism (Kalpana *et al*, 1991). Different gene replacement strategies in *M.TB* have been developed which include use of non-replicating vectors (Husson *et al*, 1990) and incompatible plasmids (Balasubramanian *et al*. 1996). However, these methods require large amounts of DNA (1-10μg) and yield low number of mutants. Counterselectable markers like *sacB* help in identifying the mutants, but require multiple steps of transformation and selection (Pavelka and Jacobs, 1999).

Recombineering using mycobacteriophage encoded-recombination proteins enhances the recombination frequencies in both *M. smegmatis* and *M.TB* (Van Kessel *et al*, 2007). This method makes use of mycobacteriophage Che9c, which encodes gp60 and gp61 proteins. These proteins are homologs of RecE and RecT proteins found in the Rac prophage. The RecE protein functions as 5'-3' dsDNA-dependent exonuclease and the RecT protein is a ssDNA binding protein, which promotes annealing of complementary DNA strands, strand invasion and strand exchange (Noirot *et al*, 1998).

We attempted allelic exchange of the native *M.TB ligA* gene with the Arctic *ligA* gene in the *M. tuberculosis* genome using two different methods: (1.) The Pavelka method which makes use of a mycobacterial suicide vector (Figure 6; Pavelka and Jacobs, 1999) and (2.) Hatfull's method which uses mycobacteriophage proteins (Figure 7; Kessel and Hatfull, 2007).

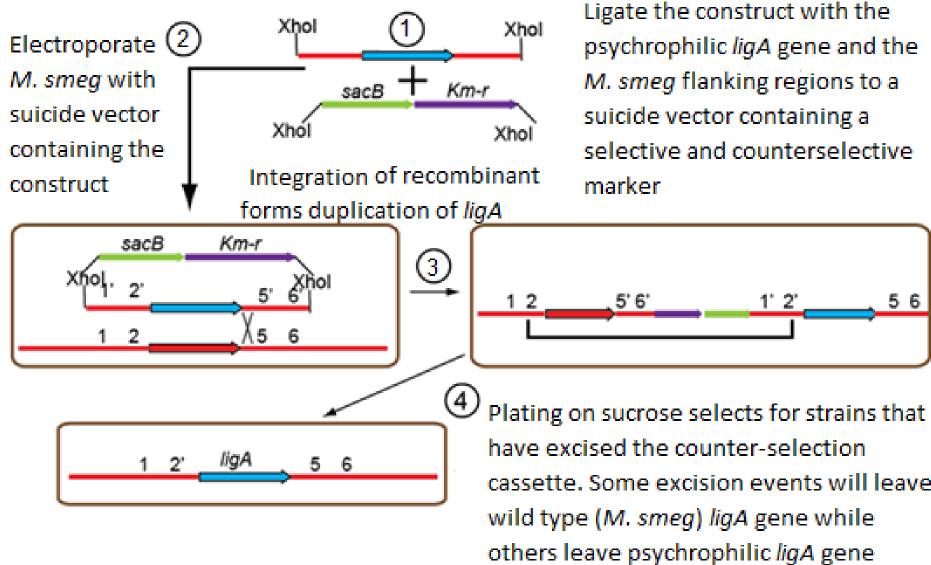

Figure 6: Allelic gene replacement of the native *M. smegmatis (M.smeg) ligA* gene with the psychrophilic *ligA* using a mycobacterial suicide vector (Figure credits- Dr. Francis E. Nano).

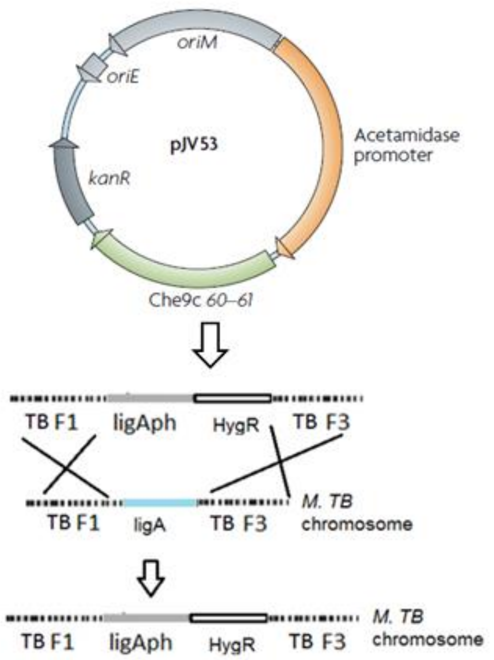

Figure 7: Recombineering using mycobacteriophage proteins.
The pJV53 plasmid contains the Che9c mycobacteriophage encoded proteins gp60 and gp61 under the control of an inducible acetamidase promoter. These gp60 and gp61 proteins are homologs of RecE and RecT proteins in Rac prophage which facilitate allelic exchange at regions of homology as depicted in the figure (The pJV53 plasmid was adapted from Van Kessel *et al*, 2008).

2.2. Materials and Methods

2.2.1. Bacterial strains and growth conditions

All enzymes used in this study were purchased from NEB or Fermentas unless otherwise noted. All reagents and chemicals were purchased from Sigma unless otherwise noted. The M7H10 and M7H11 agar media are products of BD biosciences. The strains used in this study were *M. smegmatis* mc^2 155 (Snapper et al, 1990), *M. tuberculosis* H37Rv and Erdman strains, *Saccharomyces cerevesiae* MYA3666 and *Escherichia coli* MG1655. The *M. smegmatis* strain was cultured in Middlebrook 7H9 media supplemented with albumin-dextrose-sodium chloride (M7H9-ADC) enrichment medium and the *M. tuberculosis* H37Rv and Erdman strains were cultured in Middlebrook 7H9 medium supplemented with oleic acid-albumin-dextrose-catalase (M7H9-OADC) enrichment medium. 0.05% Tween 80 was added to the Middlebrook medium as *M. smegmatis* and *M. tuberculosis* tend to form clumps in liquid medium. When required, kanamycin (Km) and hygromycin (*hyg*) antibiotics were added to a final concentration of 15µg/ml and 100µg/ml respectively.

A) Experiments to engineer TS *M. smegmatis*

2.2.2. Codon harmonization of the psychrophilic *ligA* gene

The nucleotide sequence of the *ligA* gene from *Pseudoalteromonas haloplaktis* (*ligAPh*) was obtained from NCBI (GenBank Accession no. AF126866) and was codon harmonized for *M. smegmatis* using the program Anaconda (Moura *et al*, 2005; http://bioinformatics.ua.pt/software/anaconda/). The nucleotide sequence was submitted to Integrated DNA technologies (http://www.idtdna.com/site) for gene synthesis. Codon harmonization was necessary as the G+C content of *ligAPh* is 40% as compared to that of *M. smegmatis*, which has 70% G+C content.

2.2.3. Recombining *M. smegmatis* flanking regions with *ligAPh*

2.2.3.1. Primer design to assemble the codon harmonized Arctic *ligA* gene and the *M. smegmatis* flanking regions

The ~1kb regions flanking the *ligA* gene in the *M. smegmatis* chromosome were PCR amplified using standard PCR protocols. One end of the primer sequence was designed to overlap with the yeast cloning vector pRS416 and the other end of the primer overlapped the Arctic *ligA* gene sequence. The codon harmonized *ligAPh* gene was PCR amplified using primers designed to overlap the *M. smegmatis* flanking regions (Table 1). The *M. smegmatis* genomic DNA was isolated using the standard protocol and was used as a template for amplifying the flanking regions. All the constructs were amplified using the Q5® HF DNA polymerase system purchased from NEB. The PCR mixture contained the following reagents in their final concentration equivalents: 1X Q5® Buffer, 1X GC enhancer, 200µM dNTPs, 0.5µM primers, 0.02Units/µL Q5® HF DNA polymerase and 10ng template DNA and was brought up to volume with nuclease free water. The reactions were run in a Techne® Endurance TC-512 Gradient Thermal cycler using the standard protocol as per Q5 DNA polymerase protocol specifications (initial denaturation: 98°C for 5 min, denaturation: 98°C for 1 min, annealing and extension: 72°C for 1 min followed by a final extension for 5 min). The resultant PCR products were visualized on a 0.7% agarose gel containing Gel Red and the nucleotide fragment of the correct size was purified using the QIAquick PCR purification protocol.

Table 1: Primers designed to amplify the *ligAPh* gene and the *M. smegmatis* flanking regions. Overlaps to adjacent regions are indicated in the lower case font and enzyme cut site (XhoI) is underlined.

	Size of constructs	Forward (5'-3')	Reverse (5'-3')
M. smegmatis left flanking region	~1000bp	gtgagcgcgcgtaatacgactcactCTCGAGCAGATGGGTGGCGGGGTC	gagatgctgctggccatTCTGGCAGGCTAGCCGAGCG
M. smegmatis right flanking region	~1000bp	agcagcataacggctgaACGGGACCTCGGCGGTGT	taaccctcactaaagggaacaaaagctggaCTCGAGGAGGGGGTTGTCGGTCGGCT
ligAPh gene	~2100bp	ctcggctagcctgccagaATGGCCAGCAGCATCTCGGA	acaccgccgaggtcccgttCAGCCGTTATGCTGCTCCA

2.2.3.2. Assembling the psychrophilic *ligA* gene with *M. smegmatis* flanking regions

The *M. smegmatis* flanking regions and the codon harmonized psychrophilic *ligA* gene were assembled together using yeast mediated recombination. To make chemically competent yeast cells, a frozen mid-log phase culture of *S. cerevesiae* MYA3666 stored at -80°C was thawed and added to 4 ml of Yeast Peptone Dextrose Adenine (YPD-A) broth a day before the transformation. From this stock, the cultures were serial diluted to a final ratio of 1:128 and grown overnight at 30°C. Next day, the cultures in the mid-log phase were chosen and 1ml of this culture was added to 25ml YPD-A broth. The flasks were grown at 30°C with shaking for 2-4 hours until the A_{600} was ≤ 0.6-0.8. The cultures were centrifuged for 5 min at 3000 rpm. The pellet was resuspended in 10 ml of TL buffer and centrifuged again. The supernatant was discarded and the cells were resuspended in 1ml of fresh TL and incubated at 30°C for 45 min, inverting occasionally to make fresh competent cells. 100ng of pRS416 linear plasmid (cut with XhoI, XbaI and

SalI), 100ng of each of PCR amplified *M. smegmatis* flanking regions, *ligA* gene, and 5µl of boiled herring sperm DNA were added to the competent *S. cerevesiae* cells and mixed gently to evenly distribute the DNA. The control transformation mixture contained only the pRS416 plasmid. 400µl of TLP was added and the mixture was incubated at 30ºC for 1 hour. Cells were heat shocked in a 42ºC water bath for 20 min and transferred to ice for 3 min. The transformation mixture was plated on Uracil-dropout (URA-DO) plates and incubated at 30ºC for 48 hours to select for cells with the pRS416 plasmid containing the URA3 marker.

2.2.3.3. Screening yeast transformants having correctly assembled constructs

To isolate plasmid DNA, the yeast transformants obtained were inoculated into 3 ml of URA-DO broth (with glucose) and incubated for 16 hours at 30ºC in a spinning wheel incubator. The cultures were pelleted and suspended in P1 buffer. 20 units of zymolase was added and the tubes were incubated at 37ºC for 15 min. P1 buffer containing washed glass beads was added and the tubes were agitated on a BioSpec Products Mini-Beadbeater™ for 100 seconds. The tubes were reincubated at 37ºC for 20 min and then transferred to ice. P2 buffer was added to the tubes and the supernatant was added to P3 buffer and spun at 13000 rpm for 10 min. The supernatant was applied to EZ-10 spin columns, washed with PB and PE buffers and the plasmid DNA was eluted in EB. The buffer recipes can be found at http://openwetware.org/wiki/Qiagen_Buffers. The SIGMA 1-15 micro-centrifuge was used for all small volume spins unless otherwise noted. The putative positive recombinants were determined by PCR amplification using primers specific to the *M. smegmatis* flanking regions (Table 2) and running the PCR

products on an agarose gel. These clones were gel purified using standard PCR purification protocol and analyzed by sequencing.

Table 2: Primers designed to screen yeast transformants having correctly assembled constructs. Underline indicates the restriction site (XhoI) added.

	Expected size of the construct	Forward (5'-3')	Reverse (5'-3')
M. smegmatis flanking regions assembled with the *ligA* gene	~4100bp	gtgagcgcgcgtaatacgactcactCTCGAGCAGATGGGTGGCGGGGTC	taaccctcactaaagggaacaaaagctggaCTCGAGGAGGGGTTGTCGGTCGGCT

2.2.4. Cloning the assembled construct into the p2NIL-*sacB* mycobacterial suicide vector

To carry out the gene replacement of the psychrophilic *ligA* with the *M. smegmatis ligA* gene, we needed to introduce the *ligAPh* gene with the *M. smegmatis* flanking regions into a cloning vector that can replicate in *E. coli* but not in *M. smegmatis* (a "suicide vector"). The p2NILpGOAL*ΔligA* (Korycka-Machala *et al*, 2007) is a mycobacterial suicide vector containing the kanamycin resistant marker, the *hsp60-sacB*-Ag85-*lacZ* cassette derived from the pGOAL17 vector and a deleted version of the *ligA* gene from *M. smegmatis*. Initial cloning attempts with this vector were not successful due to its large size (~15kb) and the presence of inverted repeats. Thus, PCR amplifying the p2NILpGOAL*ΔligA* vector was a difficult goal to achieve. To reduce the size, the vector was first digested with PacI to remove the *lacZ-sacB* cassette, and PstI to get rid of the deleted *ligA* portion. This reduced the size of the p2NIL vector containing the kanamycin resistance marker to 4.7kb. The assembled construct was amplified using primers containing HindIII and BamHI restriction sites (Table 3) and cloned into the MCS of this p2NIL vector. The putative positive clones were determined by digestion with restriction enzymes unique to the assembled *ligA* construct and further confirmed by sequencing.

The 2.2kb *sacB* fragment with the hsp60 promoter was next cloned into the PacI site of this p2NIL-*ligA* clone to give a complete vector with the assembled construct, the kanamycin marker and the hsp60-*sacB* counterselectable marker.

Table 3: Primers designed to amplify the *ligAPh* gene and clone it into the p2NIL vector. Underline indicates the restriction sites added (BamHI restriction site was added to the forward primer and HindIII restriction site was added to the reverse primer).

	Expected size of construct	Forward (5'-3')	Reverse (5'-3')
ligAPh assembled construct to be cloned into the p2NIL vector	~4100bp	ctcactGGATCCCAGATGGGTGGCGGGGTC	agctggAAGCTTGAGGGGGTTGTCGGTCGGCT

2.2.5. Restriction digestion, ligation and transformation

The restriction digestion and ligations were carried out using the standard protocols unless otherwise noted. The ligation was carried out using the standard NEB protocol and the restriction enzymes were either Fermentas fast digest or NEB enzymes. The concentrations were estimated by measuring the absorbance at 260nm using the ND 1000 Nanodrop.

For transformation, chemically competent DH10B cells were incubated on ice for 3-5 min. 2-3μl of the ligation product was added and the competent cells were left for additional 15 min on ice. The mixture was heat shocked at 42°C for 1 min and transferred to ice. 250μl of Luria Bertani (LB) broth was added and the culture was plated on LB agar plate supplemented with ampicillin (250μg/ml) and incubated at 37°C overnight.

2.2.6. Preparing electrocompetent *M. smegmatis* cells

A single colony of *M. smegmatis* was inoculated in M7H9 broth containing 0.05% Tween-80 and ADC enrichment (M7H9-ADC) and grown with shaking at 37°C for 48 hours. This culture was inoculated into fresh 100ml broth and grown overnight at 37°C with shaking until the A_{600} was 0.8-1.0. The cells were kept on ice for 1 hour and spun at 5000 rpm for 20 min. The supernatant was discarded immediately and all remaining steps were performed on ice. The cells were washed two times with ice-cold 10% glycerol, resuspended in glycerol and immediately used for transformation. The transformation efficiency of the electrocompetent cells was ~ $1x10^5$ transformants/μg.

2.2.7. Construction of TS *M. smegmatis* strains

The p2NIL-*sacB* vector containing the *ligAPh* gene was transformed into *M. smegmatis* mc^2 155 electrocompetent cells to obtain merodiploids containing both the psychrophilic and the WT *ligA* gene in the *M. smegmatis* chromosome. The merodiploids were selected on M7H10-Km agar plates and tested for the presence of the *ligAPh* gene in the *M. smegmatis* chromosome by PCR amplifying the *ligAPh* gene using primers in Table 1. The merodiploids were grown in M7H9-ADC broth containing 10% sucrose at 30°C to force the excision of the *Km-SacB* cassette along with the targeted essential gene. The cultures were plated on M7H10-ADC plates containing 10% sucrose and incubated at 30°C. The sucrose resistant colonies obtained were replica plated on M7H10-ADC agar plates and incubated at 30°C and 42°C to select for TS phenotypes. The presence of the psychrophilic gene in *M. smegmatis* genome was confirmed by amplifying the *ligAPh* gene and the region surrounding the psychrophilic *ligA* gene using primers in Appendix Table A2.

2.2.8. Determining the restrictive temperature of TS *M. smegmatis*

A "restrictive temperature" is one that "restricts" or does not allow the growth of an organism. For our TS strains we have defined the restrictive temperature as 1°C higher than the highest temperature that supports the growth of individual colonies on agar medium. To determine the maximum growth temperature of the TS strain, the *M. smegmatis* colonies were streaked on M7H10-ADC agar plates and incubated at different temperatures. The plates were buried in aluminum pellets to reduce the temperature fluctuations in the incubator. A temperature probe connected to the Fluke 53II thermometer was inserted into the pellets to measure the minimum and the maximum temperatures while incubating the plates. The restrictive temperature assigned to the strains should be considered accurate to ±0.5°C.

2.2.9. Mutation rate of TS *M. smegmatis*

Two colonies of TS *M. smegmatis* were inoculated into 5 ml of M7H10-ADC broth and the cultures were grown for 18-20 hours until the A_{600} at 30°C reached 0.9-1.0. Approximately 1 x10^{10} cells were plated on 150mm diameter M7H10-ADC plates and incubated at 3°C above the restrictive temperature. The actual CFU plated was determined by plating serial dilutions of the original cultures incubated at 30°C. The mutation rate was calculated by dividing the number of colonies arising on plates incubated above restrictive temperature to the number of viable cells plated. If too many bacterial cells were plated or if the plates were incubated close to restrictive temperature, a lawn of bacterial cells appeared and this masked the growth of temperature-resistant mutants.

2.2.10. Growth curves of WT and TS *M. smegmatis*

The WT and TS *M. smegmatis* cultures were grown until late exponential phase at 30°C in M7H9-ADC broth with 0.05% Tween-80. The cultures were then diluted 1:20 and incubated in a water bath with shaking at 30°C and 38°C. The readings were monitored using a Klett-Summerson photoelectric calorimeter as clumping prevented the use of the microplate reader. The readings were taken every 2 hours over a 24 hour time period.

2.2.11. Genetic drift experiments with TS *M. smegmatis*

It is possible that mutations can accumulate in the *ligAPh* gene over time due to genetic drift and render the TS strain temperature-resistant. To test this, the TS *M. smegmatis* strain was grown for 36-48 hours at 30°C in M7H9-ADC broth and subcultured subsequently for 10 passages over 3 weeks. At the end of the 10th passage, the cultures from the first and the last passage were plated onto M7H10-ADC plates and incubated at 30°C and 42°C to check for temperature sensitivity. Glycerol stocks were prepared of the TS *M. smegmatis* strain at each passage and stored in the -80°C freezer. The *ligAPh* gene was sequenced to identify any mutations that may have accumulated over time.

B) Experiments to engineer TS *M. tuberculosis*

2.2.12. Assembling three different psychrophilic *ligA* variants with the *M.TB* flanking regions

When engineering TS *M. smegmatis*, the *ligAPh* gene had none of the codons derived from the WT *M. smegmatis ligA*. However, three different constructs of *ligAPh* that differed in the strength of their ribosome binding site (RBS) were constructed when attempting to create TS *M.TB*. This was done so that the psychrophilic *ligA* gene would

be expressed at a level close to that of the native *ligA* gene. The three different TS-*ligA* variants with the strength of the RBS are depicted in Figure 8. The primers designed to create the different *ligA* gene variants and amplify the *M.TB* flanking regions are shown in Table 4. The *M.TB* flanking regions were assembled with the *ligAPh* gene using yeast mediated recombination as described previously. The *ligA* constructs were further confirmed by sequencing using primers designed in Table 4.

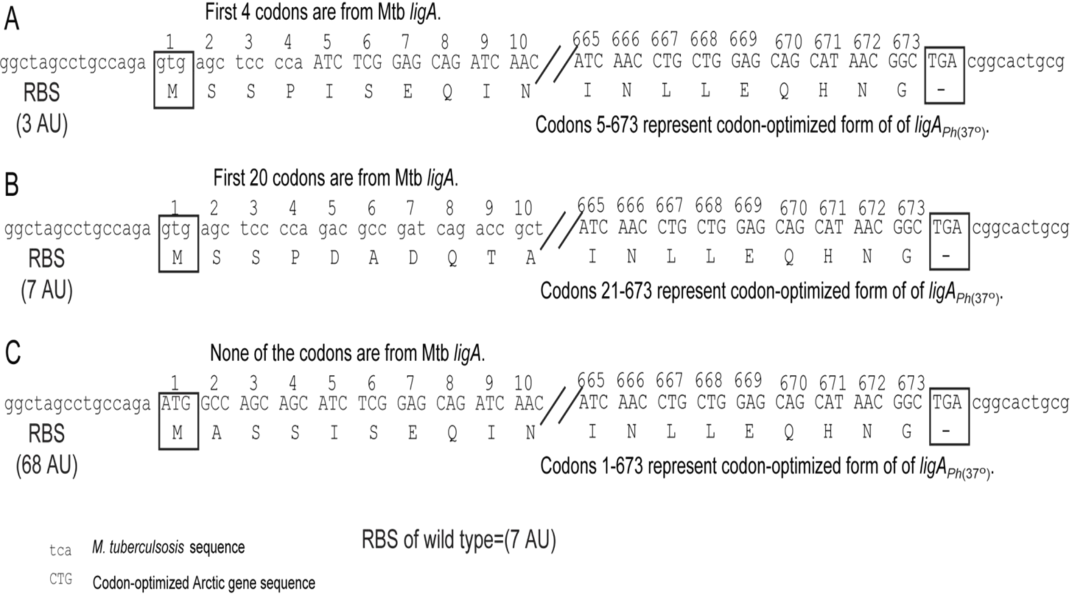

Figure 8: Three different *ligAPh* variants constructed for allelic gene replacement. The *ligA* variants designed differ in the number of codons derived from the native *M.TB ligA* gene and the strength of the ribosome binding site (RBS) .The lowercase letters represent the *M.TB* sequence and the uppercase letters represent the codon-harmonized psychrophilic *ligA* gene sequence (Figure credits- Dr. Francis E. Nano).

Table 4: Primers used to amplify three variants of *ligAPh* gene and the *M. TB* flanking regions. Overlaps to adjacent regions are indicated in the lower case font and enzyme cut sites are underlined. The codons derived from the WT *M.TB ligA* gene are indicated in red color. (HindIII restriction site was added to the forward and reverse primer).

	Size of constructs	Forward (5'-3')	Reverse (5'-3')
M. TB right flanking region	~1000bp	gtgagcgcgcgtaatacgactcactAAGCTTTGACGTTGACGTCCTTGAGC	TggagcagcataacggctgaCGGCACTGCGAATCTCCG
M. TB left flanking region	~1000bp	tccgagatgctgctggccatTCTGGCAGGCTAGCCGAG	taaccctcactaaagggaacaaaagctggAAAGCTTGCACCGCCGCGCGAGAGG
M. TB left flanking region (recombined with the *ligAPh* gene having 20 codons derived from WT *ligA*)	~1000bp	tcggcgtctggggagctcacTCTGGCAGGCTAGCCGAG	taaccctcactaaagggaacaaaagctggAAAGCTTGCACCGCCGCGCGAGAGG
M. TB left flanking region (recombined with the *ligAPh* gene having 4 codons derived from WT *ligA*)	~1000bp	ccgagattggggagctcacTCTGGCAGGCTAGCCGAGG	taaccctcactaaagggaacaaaagctggAAAGCTTGCACCGCCGCGCGAGAGG
ligAPh gene	~2100 bp	cggagattcgcagtgccgTCAGCCGTTATGCTGCTCCA	CtcggctagcctgccagaATGGCCAGCAGCATCTCGGA
ligAPh gene (20WT codons derived from *M.TB*)	~2100 bp	cggagattcgcagtgccgTCAGCCGTTATGCTGCTCCA	gtgagctccccagacgccgatcagaccgctcccgaggtgttgcggcagtggcaggcactgTACAACTACTACGTGCTCGATACGC
ligAPh gene (4WT codons derived from *M.TB*)	~2100 bp	cggagattcgcagtgccgTCAGCCGTTATGCTGCTCCA	CCTCGGCTAGCCTGCCAGAgtgagctccccaATCTCGGAGCAGATCAACCATC

2.2.13. Constructs designed for allelic exchange using mycobacteriophage recombineering (Hatfull method)

The hygromycin resistant gene (*hygR*) with its promoter (~1.5kb) was amplified from the pST-KO plasmid (Parikh *et al*, 2013) and cloned into the *ligA-M.TB* assembled

constructs in the pRS416 vector (as described above). Eight constructs were designed that differed in the orientation of the *hygR* marker (forward or reverse with respect to the *ligA* gene), the position where the cassette was inserted (immediately downstream of the *ligA* gene or 200bp in the left and the right *M.TB* flanking regions) (Figure 9) and in the number of codons that were derived from WT *M.TB ligA* gene. The primers used for amplifying the *hygR* gene and the pRS416 vector containing the *ligA-M.TB* assembled constructs are summarised in Appendix Table A3. The gene organization of the constructs was confirmed by DNA sequence analysis. The recombinant plasmids were linearized by digestion with ScaI prior to their use in electroporation experiments with *M.TB* cells.

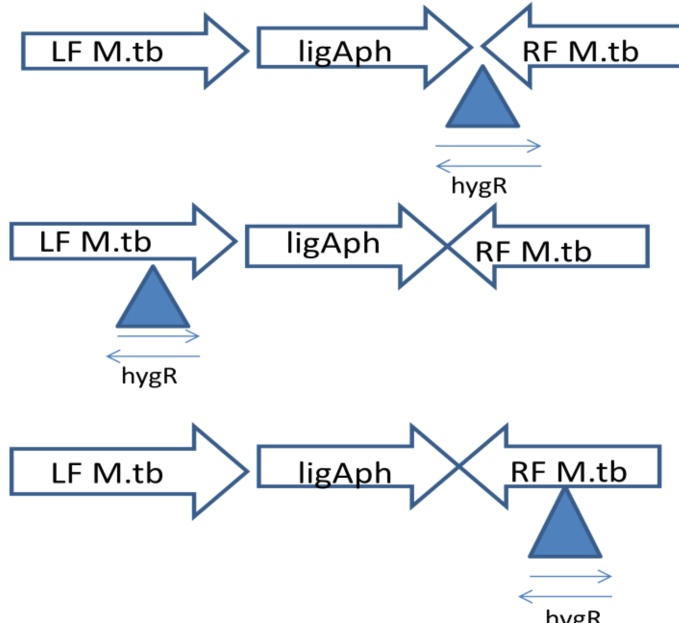

Figure 9: Schematic representation of the three linear constructs designed for mycobacterial recombineering. The constructs designed for recombineering using mycobacteriophage proteins differed in the orientation of the *hygR* cassette (forward or reverse with respect to the *ligA* gene), the position where the cassette was inserted (immediately downstream of the *ligA* gene or 200bp in the left and the right *M.TB* flanking regions) and in the number of codons that were derived from the native *M.TB ligA* gene.

2.2.14. Preparing *M.TB* electrocompetent cells and transformation using Hatfull method

(These CL3 experiments were carried out by Sheila Potter)

The mycobacterial recombineering strains were generated using the pJV53 plasmid (Kessel and Hatfull, 2006; obtained from Addgene #26904) which contains the kanamycin marker and the Che9c 60 and 61 genes expressed under the control of the inducible acetamidase promoter. To create electrocompetent *M.TB* cells, the H37Rv and Erdman strains were grown to a mid-log phase, until the A_{600} was 0.8-1.0 .The cells were transformed with 50ng of pJV53 plasmid DNA and plated on M7H11-OADC agar plates containing kanamycin and incubated for three weeks. The colonies that grew on the plates were inoculated in 3ml of M7H9-OADC broth containing Tween 80 and grown at 37°C in a 100 rpm shaker. This was then subcultured into 50ml of M7H9 induction medium containing succinate and kanamycin. After the *M.TB* culture reached an A_{600} of 0.5, acetamide was added to a final concentration of 0.2% and the culture was grown overnight. The cells were pelleted at 2000 x g for 5 min, washed three times 10% glycerol and spun at 2000 x g for 5 min. The cells were next resuspended in 10% glycerol and used for transformation directly without freezing. The linearized plasmids with the constructs containing the *hygR* cassette were transformed into electrocompetent cells as described in the Pavelka method (Pavelka and Jacobs, 1999). The pMXI–*hygR* plasmid transformed into pJV53/Erdman cells was used as a control.

The *hygR* recombinant colonies were recovered by plating on M7H11-OADC plates containing hygromycin and incubating at 30°C for nine weeks. The colonies that were *hygR* were replica plated at 31°C, 36°C and 41°C and tested for temperature

sensitivity. The presence of the psychrophilic *ligA* gene in the *M.TB* chromosome was tested by PCR amplification and sequencing.

2.2.15. Constructing three different variants of *ligAPh* and cloning into mycobacterial suicide vector (Pavelka method)

The Pavelka method for gene replacement makes use of the mycobacterial suicide vector similar to the method used to create a TS *M. smegmatis* strain. *SacB* is a widely used counterseletable marker in *Mycobacteria.* However, it can be spontaneously inactivated which complicates it use (Pavelka and Jacobs, 1999). Further, research from the Barkan lab showed that selection of *M. smegmatis* recombinants obtained by allelic exchange improved 98% - 100% using the the *sacB-galK* double counterselection system (Barkan *et al*, 2010). Therefore, a new pSTKO-*galK* vector containing two counterselectable markers was created. This vector was created by cloning the *galK* gene with its mycobacterial optimized promoter (MOP) (derived from pDB88 vector, Barkan *et al*, 2010) downstream of the *sacB* gene, with the hsp-60 promoter driving the expression of both the *sacB* and the *galK* gene (Figure 10). The primers used for amplifying the *galK* gene and the pST-KO vector are shown in the Appendix Table A4. The *ligA-M.TB* constructs were cloned into the MCS of the pSTKO-*galK* plasmid using primers as shown in Table 5.

Table 5: Primers designed to clone the *M.TB ligA* constructs into the pSTKO-*galK* plasmid. The enzyme cut sites are underlined. (BamHI restriction site was added to the forward primer and XbaI was added to the reverse primer).

	Expected size of constructs	Forward (5'-3')	Reverse (5'-3')
PSTKO-*galK* plasmid	~5500bp	ataagtGGATCCATGGCTAGTTCCA	ataagtTCTAGATTAATTAACTAGTGTTAACGG
Assembled *M.TB ligA* constructs	~4100bp	AAGATCGGATCCTGACGTTGACGTCCTTGAGC	AAGATCTCTAGAGCACCGCCGCGCGAGAGG

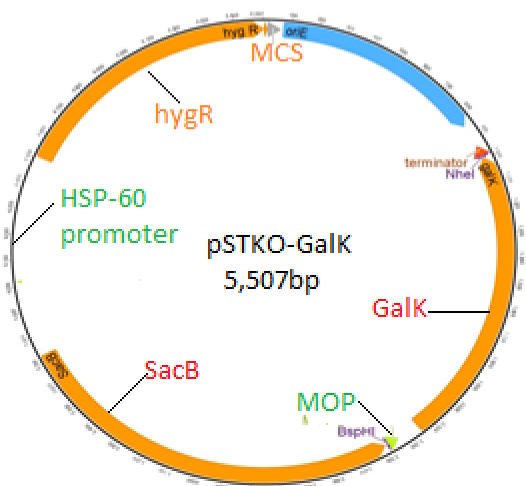

Figure 10: Diagrammatic representation of the pSTKO-*galK* mycobacterial suicide vector. The pSTKO-*galK* vector contains the *hygR* gene and two counterselectable markers *sacB* and *galK* cloned in front of the hsp60 and MOP promoter. The three *ligA* variants with the *M.TB* flanking regions were cloned into the multiple cloning site (MCS) of this vector. (This figure was designed using Geneious version 6, http://www.geneious.com, Kearse *et al*, 2012; Figure credits: Sheila Potter).

2.2.16. Pavelka method for preparing electrocompetent *M. tuberculosis*

One ml aliquots of *M.TB* H37Rv and Erdman strain were thawed and transferred to 10ml M7H9-OADC broth and incubated at 37°C with slow shaking (100 rpm) until saturated (~5 days). 1ml of this well grown culture was subcultured 1:50 into 100ml of M7H9 media in a roller bottle and incubated at 37°C on a roller apparatus at 100 rpm for 5-7 days until A_{600} was 0.8-1.0. The density of the cells was checked two days before the

electroporation experiment to test if the culture was expanding as expected. The cells were transferred into two 50ml conical flasks and pelleted at 3000 rpm for 5 min at room temperature. The cells were washed three times with 10% glycerol and spun at 3000 rpm for 5 min. After the washes, the cells were resuspended in 1ml of 10% glycerol and used for transformation directly without freezing.

2.2.17. Transforming the *ligAPh* constructs into *M.TB* Erdman and H37Rv strains

The PSTKO-*galK* plasmid DNA with the assembled construct (~ 1 µg) was added to 100µl *M.TB* electrocompetent cells using a 0.2 cm gap electroporation cuvette. Electroporation was carried out using the Biorad Gene pulser for 2500 mV, 25 µF, and 1000 Ω. 1 ml of M7H9 media was added to the cuvette and the cell suspension was transferred to a 15ml conical tube. The electroporation was repeated 5-10 times using the suicide plasmid with different constructs. Two control electroporations were carried out; one without DNA and the other with plasmid DNA bearing a mycobacterial origin of replication and hygromycin resistant marker (pMXI vector). The tubes were incubated at 37°C in a shaker at 100 rpm overnight and plated on M7H11-OADC plates containing hygromycin. Negative control was used to test for spontaneous hygromycin resistance. Plates were wrapped and incubated for 3-4 weeks in a 37°C CO_2 incubator.

2.2.18. Screening for primary *M.TB* recombinants

The primary recombinants obtained on M7H11-OADC plates were tested for the presence of *ligAPh* in the chromosome of *M.TB* by PCR amplification and restriction digestion using enzymes specific to the *P.haloplanktis ligA*. The primary recombinants were also grown in M7H9-OADC broth to make glycerol stocks and stored at -80°C.

2.2.19. Future Steps to screen for TS mutants in the level 3 facility

Due to time and financial constraints, the screening for TS mutants in the level-3 laboratory was not completed. The future steps to screen for TS mutants would be as follows: the primary recombinants obtained on M7H11-OADC plates with hygromycin will be inoculated into M7H9-OADC broth containing 2% sucrose and 0.5% 2-deoxygalactose (2-DOG); the counterselective marker for *galK*, to force the excision of the counterselection cassette. The sucrose and 2-DOG resistant colonies obtained will be replica plated on M7H11-OADC plates and incubated at 30°C and 42°C for eight weeks and three weeks respectively, to identify the strains that are TS. TS colonies if any will be streaked on M7H11-OADC plates containing sucrose-2DOG and hygromycin. This will be done to test whether the *M.TB* colonies are sucrose-2DOG resistant and hygromycin sensitive, confirming loss of the selection and counterselection cassette.

2.3. Results

Mycobacterium smegmatis, a research surrogate for *M.TB* was initially chosen as a host organism to test for temperature sensitivity imparted by the Arctic *ligAPh* gene.

2.3.1. Codon harmonization using the Anaconda software package

As the G+C content of the *ligA* gene from *M. smegmatis* was ~70% as compared to the *ligAPh* gene which was 44%, the *ligAPh* gene needed to be first codon harmonized for *Mycobacterium* species. To achieve this, the psychrophilic gene was added to the *M. smegmatis* ORFome and codon harmonized using the anaconda software package (Moura *et al*, 2005; http://bioinformatics.ua.pt/software/anaconda).The *ligAPh* gene was codon harmonized to have a G+C content of 62% and CAI value of 0.719, which is very close to the CAI value of native *ligA* gene (0.776). Codon harmonization was done to ensure that codon usage frequency of *ligAPh* gene would closely match the codon usage frequency of the host organism.

Table 6: The CAI value and the GC content of the *ligA* gene from *M.TB*, *M. smegmatis* and *P. haloplanktis* is depicted below. The *ligAPh* was codon harmonized to have a CAI value of 0.719 and GC content of 62%, very close to that of native *ligA* gene in *M. smegmatis*.

Gene and Organism	CAI value	% GC content
ligA from *M.TB*	0.774	69
ligA from *M. smegmatis*	0.776	69.7
Codon harmonized *ligA* from *P. haloplanktis*	0.719	62
ligA from *P. haloplanktis*	0.196	44

2.3.2. Recombining the *M. smegmatis* chromosomal flanking regions with the Arctic *ligA* gene

The regions flanking the *M. smegmatis ligA* (~1000bp) were PCR amplified and joined with the psychrophilic *ligA* gene using yeast mediated recombination. The putative positive clones were determined by isolating the pRS416 plasmid DNA from the yeast colonies that grew on URA-DO plates, amplifying the constructs using primers specific to the constructs and running the PCR products on an agarose gel (Figure 11). The positive clones were confirmed by sequencing.

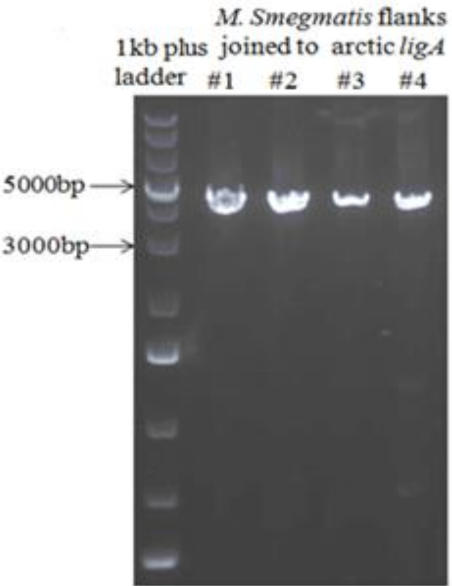

Figure 11: The *ligAPh* gene recombined with the *M. smegmatis* flanking regions. Four putative positive clones (of size ~4kb) were obtained that contained the *ligAPh* gene recombined with the *M. smegmatis* flanking region. The putative clones were determined by PCR amplifying the plasmid DNA from the yeast recombinants obtained and running the products on an agarose gel. The #3 and #4 putative positive clones were further confirmed by sequencing.

2.3.3. Cloning the assembled constructs into the mycobacterial suicide vector

The *M. smegmatis-ligA* assembled construct was cloned into the BamHI-HindIII site of the p2NIL suicide vector. One positive clone was obtained (Figure 12) which contained the *M. smegmatis-ligA* construct cloned into the MCS of the p2NIL suicide

vector. The *sacB* counterselectable marker was cloned into the PacI site of this vector as described in the methods section.

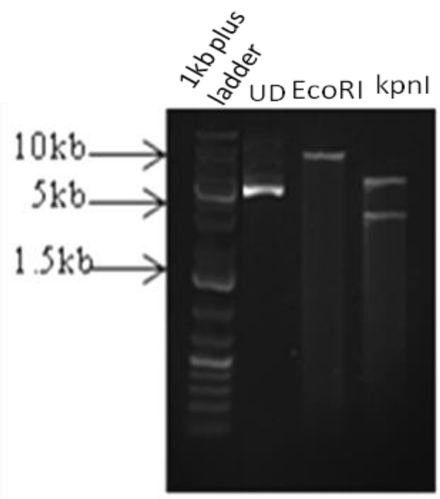

Figure 12: p2NIL clone with the assembled *ligA* construct.
The putative positive p2NIL clones were determined by digesting the plasmid DNA with EcoRI and KpnI restriction enzymes, which are specific to the psychrophilic *ligA* gene, and running the samples on an agarose gel. One positive clone was obtained as shown above. The first lane in the gel represents the DNA ladder. The undigested sample is run as a control followed by digestion with the restriction enzymes EcoRI and KpnI. This clone gave a correct sized fragment of ~8.8 kb upon digestion with EcoRI and two fragments of size ~ 5.6kb and ~3.1kb upon digestion with KpnI indicating the presence of the *ligA* construct in the p2NIL vector.

2.3.4. Screening primary and secondary recombinants in *M. smegmatis*

The mycobacterial suicide vector with the assembled *ligA* construct was transformed into electrocompetent *M. smegmatis* mc^2 155 and the primary recombinants were determined by selecting on M7H10-ADC-kan plates. One kanamycin resistant colony was obtained (Figure 13A), indicating a putative merodiploid with the psychrophilic gene and the kanamycin marker inserted in the *M. smegmatis* chromosome. PCR amplification using primers specific to the psychrophilic *ligA* gene gave products of expected size, further suggesting the presence of *ligAPh* in the *M. smegmatis* chromosome (Figure 13B).

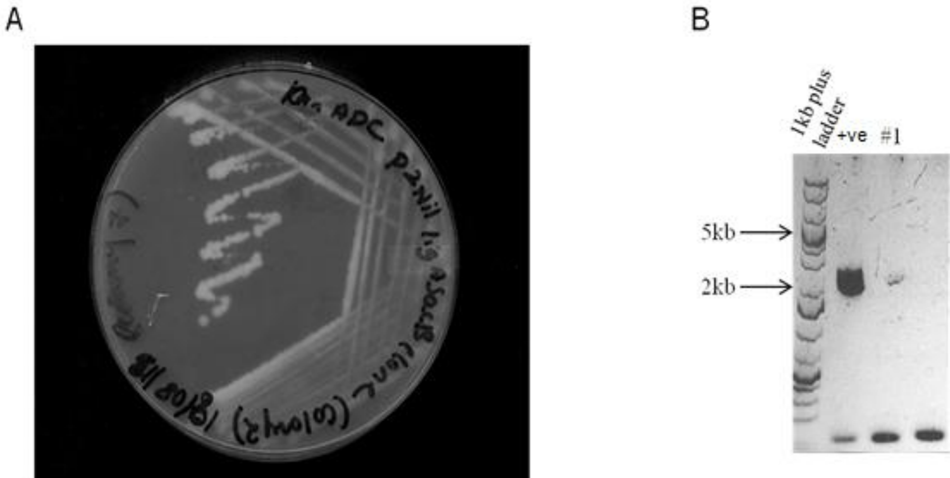

Figure 13: A) *M. smegmatis* merodiploid streaked on M7H9-kan plates with ADC enrichment media. B) Colony PCR of the *M. smegmatis* colony obtained on the Km plate with primers specific to the psychrophilic *ligA*, gave a correct sized fragment of 2.1 kb on an agarose gel. This indicated the presence of *ligAPh* in the *M. smegmatis* chromosome. The *ligAPh* gene cloned into the p2NIL vector was used as a positive control (lane 2).

Secondary recombinants were obtained by growing and plating the merodiploid on M7H9-sucrose agar plates, to force the excision of the counterselectable cassette (Figure 6). These sucrose resistant colonies were replica plated at low and high temperatures to select for a TS phenotype. The strains which grew at 30°C but failed to grow at 44°C were considered TS (Figure 14A and 14B). 50% of the strains failed to grow at 44°C, indicating that the strains had lost the WT *ligA* gene and retained the psychrophilic *ligA* gene in the chromosome. The colonies were simultaneously picked on M7H9-ADC-Km agar plates to confirm the loss of the kanamycin marker (Figure 14C). The presence of the psychrophilic *ligA* gene in the *M. smegmatis* chromosome was further confirmed by sequencing.

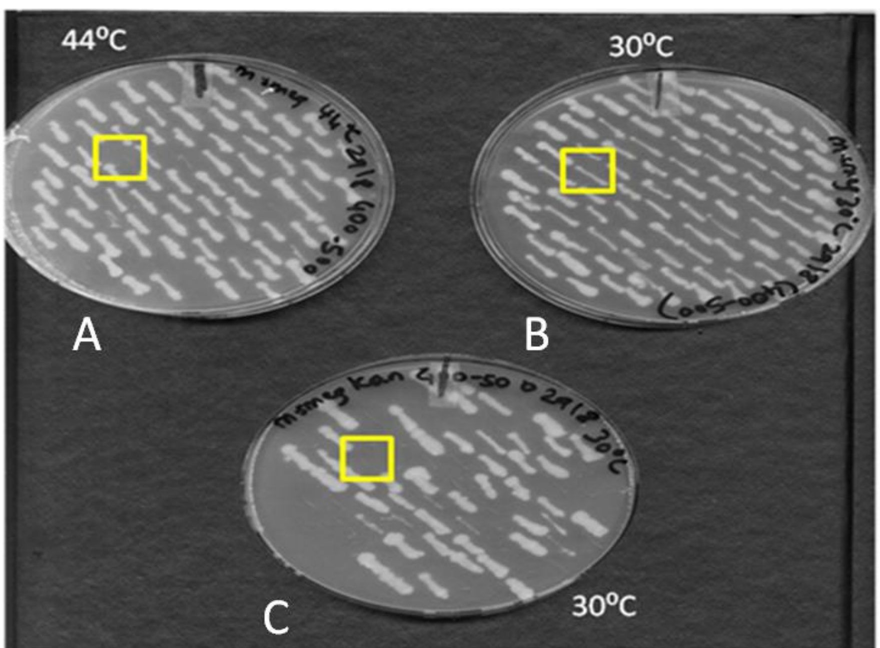

Figure 14: ***M. smegmatis*** **secondary recombinants replica plated on M7H10-ADC agar plates at A) 44°C and B) 30°C. Some colonies grew at 30°C but failed to grow at 44°C indicating the loss of the WT** ***M. smegmatis ligA*** **gene and the presence of the psychrophilic gene in the chromosome (One such TS strain is highlighted in yellow). C)** ***M. smegmatis*** **colonies replica plated on M7H10-ADC-kan plates to confirm the loss of the kanamycin selectable marker.**

2.3.5. Determining restrictive temperature of TS strains

The TS colonies were streaked on M7H10-ADC plates and incubated at 33°C, 35°C, 37°C and 39°C to determine restrictive temperature of the TS strain. We define the restrictive temperature of the TS organism as 1°C above the lowest temperature that allows the formation of individual colonies on agar medium. The TS strains grew well at 33°C and 35°C but did not form individual colonies at 37°C and did not grow above 39°C when compared to the wild type, which grew well at all temperatures (Figure 15). Therefore, the restrictive temperature of the TS strain was determined to be 37°C, which is our desired restrictive temperature to induce a good protective immunity.

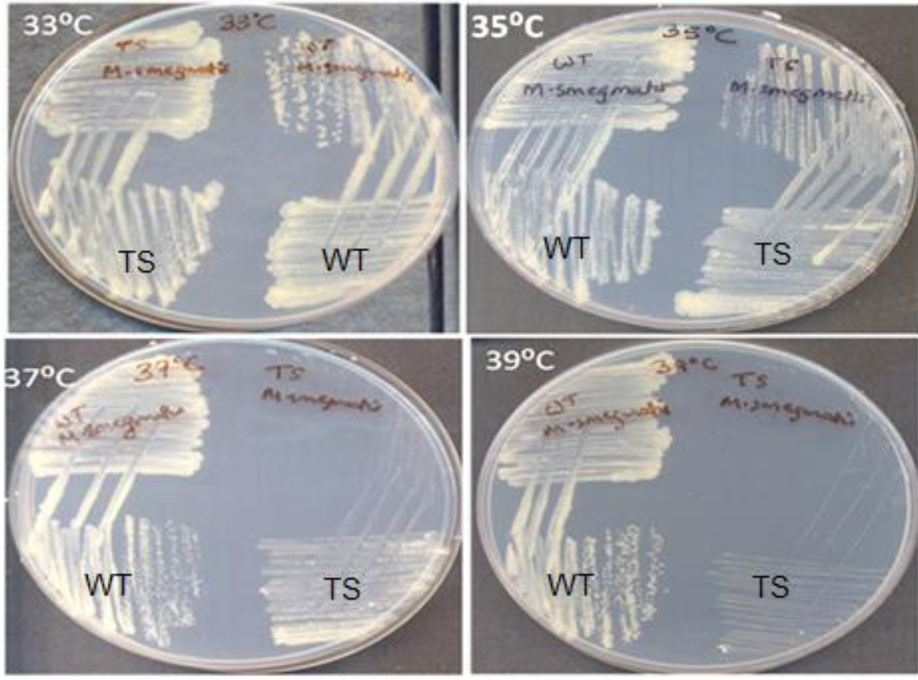

Figure 15: TS and WT *M. smegmatis* streaked on M7H10-ADC plates to determine the restrictive temperature of the TS strain. The plates were incubated at 33°C, 35°C, 37°C and 39°C. TS *M. smegmatis* grows like the WT at 33°C and 35°C. However, the TS strain fails to form individual colonies above 37°C and does not grow at 39°C.

2.3.6. Growth of TS *M. smegmatis* in broth

Growth of the WT and TS *M. smegmatis* strains were monitored for 24 hours by inoculating the colonies in M7H9-ADC broth with Tween 80 at 30°C and 38°C. Both the strains grew well at a permissive temperature of 30°C (Figure 16A). However, at a restrictive temperature of 38°C, the TS *M. smegmatis* strain initially grew like the WT and ceased to grow after 5 hours, as determined by measuring the optical density (Figure 16B).

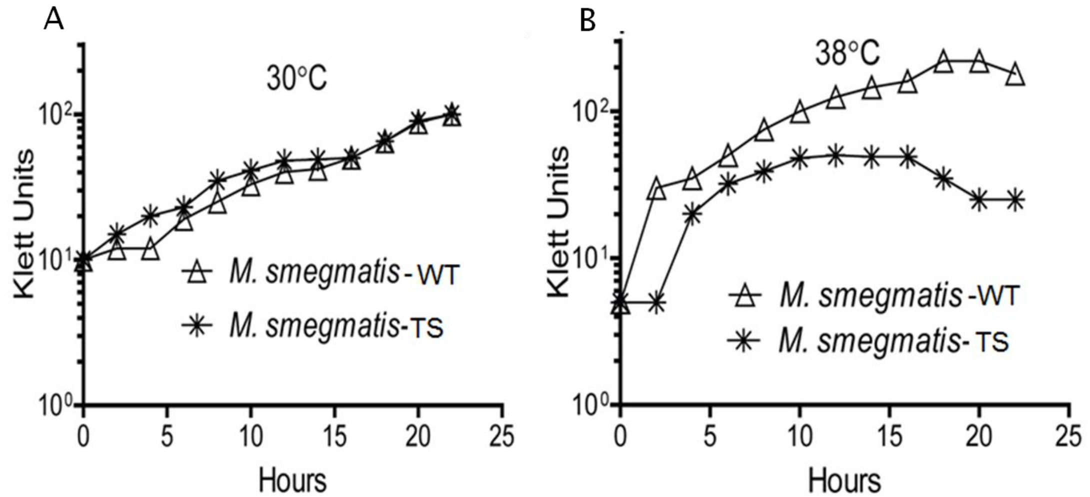

Figure 16: Growth of TS *M. smegmatis Vs.* WT *M. smegmatis* in broth. On comparing the growth of the TS *M. smegmatis* with the WT *M. smegmatis*, the TS strain grows well at permissive temperature of 30ºC (A) but fails to grow at a rate equal to that of the WT at a restrictive temperature of 38ºC (B).

2.3.7. Reversion rate of TS *M. smegmatis*

Although the psychrophilic *ligA* gene has adapted itself to a cold environment from millions of years, the gene may gain its stability at high temperature due to single point mutations and become temperature-resistant. To test this possibility, 10^{10} cells were plated on M7H10 agar plates and the plates were incubated 3ºC higher than the restrictive temperature. There were no colonies present on the plates, indicating that the TS strain has a very stable phenotype. These experiments were repeated three times with similar results. It is possible that the growth of any temperature-resistant mutant could be masked among the pool of dying cells. To test this, 10^3 cells of WT *M. smegmatis* was added to the TS *M. smegmatis* culture and the plates were incubated at 39ºC. Colonies were observed on the plates in numbers consistent to the WT culture added, confirming the validity of the reversion experiments performed.

2.3.8. Genetic Drift experiments

Even though we did not detect any revertants at the restrictive temperature, it might happen that the TS *M. smegmatis* strain could revert to a temperature-resistant form over time due to selection pressure or through genetic drift. To test this possibility, the TS strain was subcultured in M7H9-ADC broth over a period of three weeks at 30°C and at the end of the tenth subculture experiment, the strains were plated on M7H10-ADC plates and incubated at 30°C and 38°C. No reversions were noted on incubating the plates at 38°C indicating that the *M. smegmatis* strain did not lose its TS phenotype. Further, sequencing results confirmed that there were no mutations present in the psychrophilic *ligA* gene indicating that the TS strain was stable.

2.3.9. *ligA* variants constructed for gene replacement in *M.TB*

After confirming temperature sensitivity imparted by the *ligAPh* gene in *M. smegmatis*, we attempted to replace this *ligAPh* gene into the *M.TB* chromosome. The *ligAPh* gene substituted into the *M. smegmatis* genome had none of its codons derived from the WT *ligA* gene. To maximise the chances that the psychrophilic gene would be expressed very close to the native homolog, the *ligAPh* gene was integrated into the *M.TB* chromosome with transcription driven by the host promoter and the translation of the mRNA directed by the host ribosome binding site (RBS). Three different variants of the *ligAPh* gene were designed which differed in the strength of the RBS (Figure 8). These *ligA* variants were successfully assembled with the *M.TB* flanking regions using yeast mediated recombination and were confirmed by sequencing. For gene replacement using mycobacteriophage recombineering, eight constructs were created with the *hygR* gene inserted into the *M.TB* assembled constructs as described in the methods section (Figure 9, Page 42).

2.3.10. Gene replacement in *M.TB*

2.3.10.1. Mycobacteriophage recombineering to create TS *M.TB* (Hatfull method)

Each of the linear constructs containing the *ligAPh* gene, the *hygR* gene and the *M.TB* flanking regions were transformed separately into the H37Rv and Erdman strains containing the pJV53 plasmid, as described in the methods section. *M.TB* Erdman *hygR* colonies were obtained on incubating the M7H11-OADC-hyg plates at 30°C. The colonies were streaked on M7H11-hyg plates and incubated at 31°C, 36°C and 41°C to test for TS phenotype. Four *M.TB* colonies grew well at 31°C and 36°C but did not grow at 41°C indicating they were TS (Figure 17).

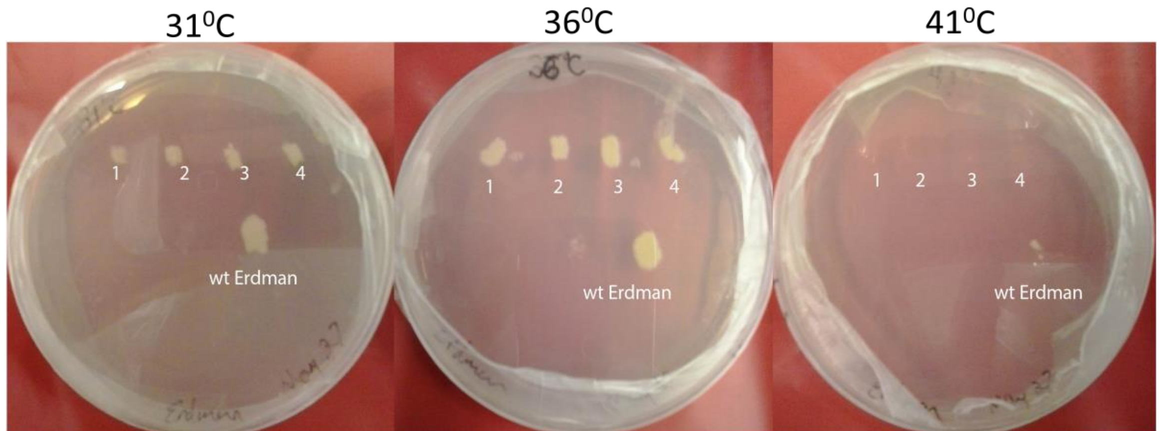

Figure 17: Two-week-old streak of *M.TB hygR* colonies on M7H11-OADC-hyg plates resulting from transforming recombineering strain of Erdman (Erdman/pJV53) with one of the assembled construct (*ligAPh* gene flanked by *hygR* gene and *M.TB* flanking regions). All four *M.TB* Erdman colonies (1, 2, 3 and 4) grow well at 31°C and 36°C but fail to grow at 41°C indicating a TS phenotype. WT Erdman is the Erdman/pJV53 strain with pMXI plasmid which can replicate in mycobacteria and confers *hygR*. This plasmid served as a control. (Figure credits-Sheila Potter).

The construct that resulted in TS *hygR* colonies is shown in Figure 18. This construct had the *hygR* gene cloned 45bp downstream of the *ligA* gene and was in reverse

orientation with respect to the *ligA* gene. The *ligAPh* gene had four codons derived from the native *M.TB ligA* gene.

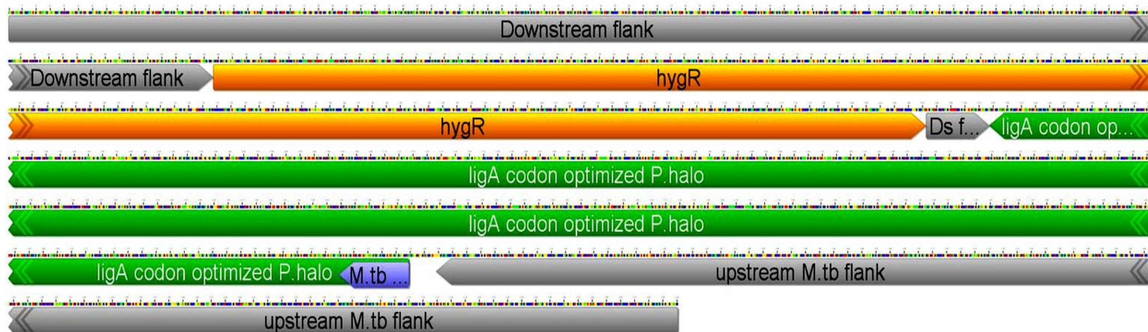

Figure 18: *M.TB-ligA* construct assembled using Geneious. The *M.TB* Erdman *hygR* colonies resulted from transforming the construct which has the *hygR* gene cloned downstream of the *ligA* gene and is in reverse orientation with respect to the *ligA* gene (shown in orange). This *ligAPh* gene (shown in green) has four codons derived from the *M.TB ligA* gene. The arrow in blue represents the four codons derived from *M.TB ligA*. (This figure was designed using Geneious version 6, http://www.geneious.com, Kearse *et al*, 2012; Figure credits: Sheila Potter).

2.3.10.2. TS phenotype is not due to allelic gene replacement

Two *hygR M.TB* strains were sequenced to test whether the TS phenotype resulted from the presence of the *ligAPh* gene in the *M.TB* chromosome. The sequencing results revealed that the TS *M.TB* strain had the native *ligA* gene retained in its chromosome and the *hygR* gene was not present downstream of the native *M.TB ligA* gene (Figure 19). These results suggested that the TS *M.TB* clones were spontaneously *hygR* or the *hygR* gene was inserted at another position in the chromosome.

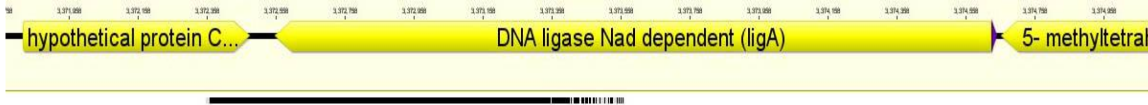

Figure 19: Sequencing results aligned with the *M.TB ligA* gene using Geneious software. The yellow arrows represent the *M.TB ligA* gene flanked by the upstream and downstream genes in its chromosome. The black bars represent the results obtained by sequencing the *ligA* region in one of the *hygR M.TB* strain. This demonstrates that the *M.TB* clone aligns perfectly with the WT *ligA* gene, and the psychrophilic *ligA* gene is not integrated near the WT *ligA* gene in the *M.TB* clone. (Figure credits: Sheila Potter).

2.3.10.3. Gene replacement using suicide vector (Pavelka method)

The *ligAPh-M.TB* assembled constructs (Figure 8) were cloned into the PSTKO-*galK* vector as described in the methods section and confirmed by sequencing. The vector containing constructs were separately transformed into *M.TB* H37Rv and Erdman electrocompetent cells and plated on M7H10-hyg plates. The plates were incubated at 37°C to obtain primary recombinants. After 4-5 weeks, colonies with two different morphologies (smooth and rough colonies) were present on the plates. The smooth and the rough *M.TB* H37Rv colonies were screened for the presence of *ligAPh* gene using primers specific to the psychrophilic gene. Around 50% of the smooth colonies and 20% of the rough large colonies were found to have the *ligAPh* gene suggesting primary merodiploids (Figure 20).

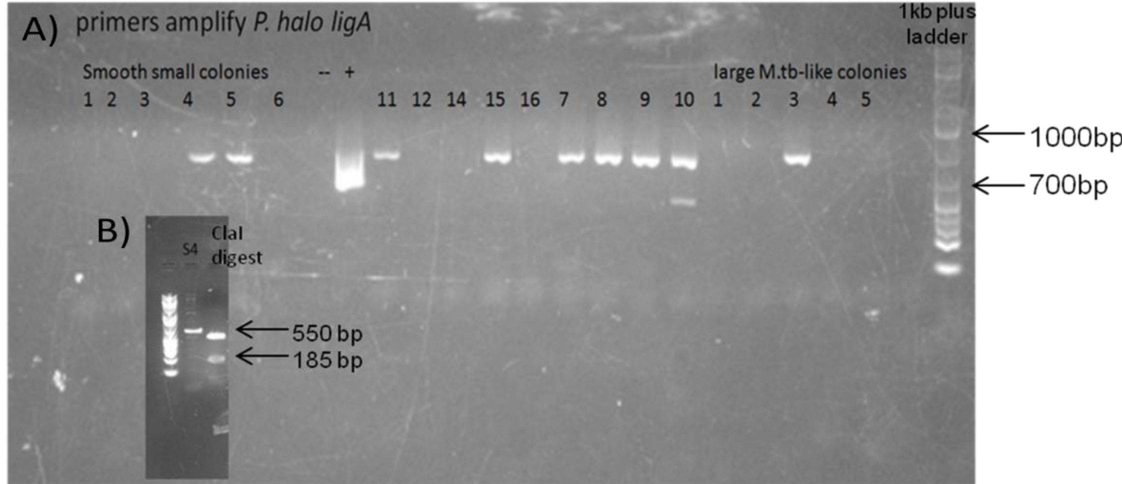

Figure 20: Screening primary merodiploids obtained by Pavelka method.
A) The H37Rv primary merodiploids were screened for the presence of the psychrophilic *ligA* gene in the *M.TB* chromosome using primers specific to *ligAPh*. The size of the expected PCR product is ~720bp. PCR amplifying smooth small colonies 1 to 6 and 10 to 16 gave eight putative positive merodiploids. These products were obtained by amplifying the *ligAPh* gene from *M.TB* colonies having a smooth morphology. The remaining putative positive merodiploid was obtained by amplifying the *ligAPh* from *M.TB* colonies having a rough morphology (large *M.TB* like colonies 1 to 5). The *ligAPh* gene amplified from TS *M. smegmatis* was used as a positive control and the *M.TB H37Rv* strain was used as a negative control. B) The figure in the inset is the *ligAPh* gene amplified from H37Rv merodiploid clone 4 (S4). This clone upon digestion with ClaI gave correct size products of 535bp and 185bp. This indicates the presence of *ligAPh* in the chromosome of *M.TB*.

2.4. Discussion and future directions

2.4.1. The TS *M. smegmatis* strain is a good potential vaccine candidate

One of the World Health Organisation (WHO) Millennium Development Goals (MDGs) is to reduce the burden of TB disease by 2015. In addition to advanced diagnostic and therapeutic strategies, new vaccine development is urgently needed to limit initial TB infection, progression of TB disease and reactivation of TB (WHO Global Report 2013). As a step towards achieving this goal, we engineered a TS *M. smegmatis* strain by replacing the native *ligA* gene with the Arctic *ligA* gene in its chromosome.

T-cell mediated immune response is required for the clearance of intracellular pathogens like *M.TB*. The TS strain functions by multiplying in the body shell as the temperature in the shell is lower, thus inducing a good Th1-cell immune response (Duplantis *et al*, 2010). The live Flu-Mist vaccine against influenza is a good example of a cold-adapted vaccine that induces good protection immunity and has an inactivation temperature of 38°C (Belshe *et al*, 2007). At a core body temperature of 37°C, the TS strain is unable to grow and to cause disease (White *et al*, 2011). Our TS strain had an inactivation temperature of 37°C which was our desired inactivation temperature. A restrictive temperature of 33°C or lower would restrict the growth of the TS strain at the cool body sites, thus limiting its ability to mount a good immune response.

A successful TS vaccine is one that does not revert to a temperature-resistant form. A vaccine that reverts to wild type can have adverse consequences. One such example is the type-3 Oral Sabin Poliovirus attenuated vaccine. A single point mutation in the genome of this attenuated strain resulted in reversion to wild type, giving rise to vaccine-associated poliomyelitis (Cann *et al*, 1984). The TS strain that we have created is stable as it does not revert to a temperature-resistant form at a detectable level.

To test the possibility of the TS strain to revert to a temperature-resistant form over time due to genetic drift or selection pressure, we did a rigorous examination of the stability of the TS strain by passaging it 10 times in-vitro in broth medium. The TS strain did not lose its temperature sensitivity suggesting the strain is stable. Sequencing the *ligAPh* gene confirmed that no mutations accumulated through several generations of growth, confirming its stability.

Our growth phenotype data indicated that the TS *M. smegmatis* strain behaves like the parent strain at 30°C, but its growth ceases at a restrictive temperature of 38°C. This TS phenotype matches the growth phenotype of the TS *S. typhimurium* and TS *F. novicida* strains created in our lab, giving us confidence that the strain will behave *in vivo* as desired (Duplantis *et al*, 2010).

2.4.2. Gene substitution in *M.TB*

My work in a group effort was to support the goal of creating a TS *M.TB* strain. As all the work carried out in *M. tuberculosis* was confined to the level-3 containment facility, regulatory and facility access constraints led to a division of labour. My work as described in above, contributed in two areas: 1) Creating a TS *M. smegmatis* strain, demonstrating that the psychrophilic *ligA* gene could be used in mycobacteria and 2) Creating different genetic constructs that were transformed into *M.TB* (done by Sheila Potter in the level-3 facility).

We attempted to replace the native *ligA* gene in *M.TB* with its psychrophilic homolog by allelic gene replacement using two approaches: the mycobacteriophage recombineering approach and the suicide vector approach. We used these two approaches concurrently in order to maximize our research results due to the limited time that we had

access to the level 3 facilities. The *M.TB* strains obtained by phage recombineering were hygromycin resistant and showed a TS phenotype suggesting that the strain was TS. However, sequencing results of the TS strain confirmed the absence of the Arctic *ligA* gene and the *hygR* marker in the *M.TB* chromosome. This allows us to infer that the TS phenotype may have been due to mutations in a different essential gene which made it TS and spontaneously hygromycin resistant. Alternatively, the TS phenotype may have occurred due to the *ligAPh* allele being inserted at another position in the chromosome and being functional and dominant over the WT *ligA* gene. Illegitimate recombination and spontaneous hygromycin resistance occurs at a high rate in *M.TB*, creating problems in isolating knockout mutants (Kalpana *et al*, 1991, Parish *et al*, 1999).

We also attempted to create a TS *M.TB* strain by using a suicide vector method. PCR amplification and restriction digestion suggests the presence of *M.TB* primary merodiploids. As the level 3 lab had to be shut down for repairs, time was insufficient to complete the experiments. Therefore, the future steps will include confirming the presence of this *ligAPh* gene in the *M.TB* merodiploids by sequencing and passaging the strains repeatedly through media containing sucrose and 2-DOG, to select for the exclusion of the counterselection cassette. We will then screen and select the strains that are *TS* and have the psychrophilic *ligA* gene on the *M.TB* chromosome. The presence of this psychrophilic gene will be confirmed by sequencing. We will be wary of false positive results since *M.TB* has a high rate of illegitimate recombination.

2.4.3. Significance of this study

A large proportion of the research work in pathogens is carried out *in vitro*. TB is a human pathogen, thus limiting its work within the level-3 containment facility. Many research labs around the world are unable to work with *M.TB* due to this constraint. The TS *M.TB* strain that we engineer should allow genetic manipulations to be carried out in level-2 laboratories, facilitating TB research. The BCG vaccine, the only vaccine available against TB has side-effects in children and is in-effective in adults. The TS strain that we generate can be tested for its potential use as a live TB vaccine as it might be successful in providing long-term protection by mediating a strong Th1-cell mediated immune response. Another potential use of this TS *M.TB* strain would be in diagnostic testing, as effective diagnostic tests are required to detect TB infection and help control the spread of TB.

Chapter 3- Strains and genetic elements for directed evolution of essential genes

3.1. Introduction

3.1.1. Directed Evolution

Directed evolution is a process to engineer enzymes with improved properties or novel functionalities by using repeated rounds of mutagenesis, screening and selection (Bloom *et al*, 2005; Bloom and Arnold, 2009). Most directed evolution methods focus on introducing small changes in the enzymes, as most enzymes have weak promiscuous functions that can be improved with few mutations (Aharoni *et al*, 2005). In a laboratory setting, directed evolution of a protein involves the following steps:

i) Mutagenesis: Two of the most common approaches for mutating a protein are A) error prone PCR (EP-PCR) and B) DNA shuffling. EP-PCR is a technique that introduces a small number of base pair mutations in the gene by imperfect DNA replication (McCullum *et al*, 2010) while, DNA shuffling generates novel genes by recombination of homologous gene fragments at cross-over points (Giver *et al*, 1998).

ii) Screening and selection of improved mutants: The identification of improved mutants depends on the engineering goal. Sometimes, it is not possible to design an effective screening and selection technique and the mutants obtained must be assayed on a high throughput screen using robotic-based assays. Mutants with the desired properties are generally selected as parents for the next generation. These mutants undergo further rounds of mutagenesis and screening to select for functional proteins with improved properties (Arnold and Bloom, 2009). As we are screening and selecting for essential

genes, we have an advantage as the protein of interest is linked to the survival of the host cell.

In most cases where a high throughput and sensitive screening assay has been designed, mutations have been found that improve the protein properties. The mutations can be classified based on how they affect the target protein as follows: i) Beneficial or adaptive mutations - give rise to a protein with improved properties ii) Neutral mutations- result in a protein with unaffected stability or function and iii) Deleterious mutations- which are harmful for the organism (Arnold and Bloom, 2009). Around 30-50% of random mutations are found to be strongly deleterious, 50-70% of the mutations are approximately neutral and 0.5-0.01% are beneficial (Arnold and Bloom, 2009). Thus, in a laboratory context, it is possible to find significant number of mutations that are neutral and a small number of mutations that can improve the stability or the existing function of a protein (Arnold and Bloom, 2009). Many research groups have been successful in engineering proteins having desired improved properties using directed evolution. One such example is the use of directed evolution to enhance the thermostability of a psychrophilic protease without affecting its catalytic activity at low temperatures (Miyazaki *et al*, 2000).

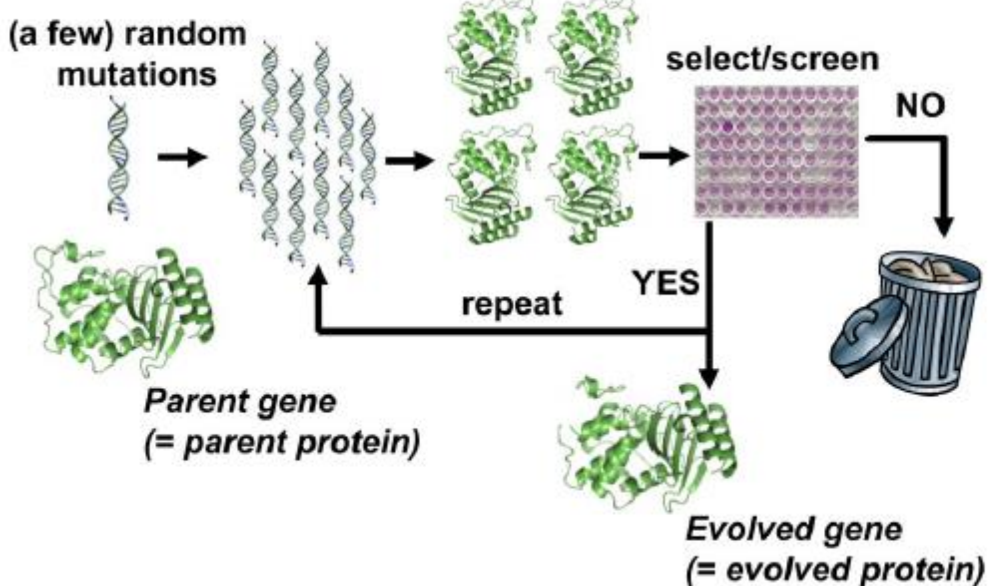

Figure 21: Schematic representation of the directed evolution experiment.
The experiment begins with the parent gene being mutated by random mutagenesis technique like EP-PCR and staggered extension shuffling. The library of mutant genes is cloned into an expression vector to produce mutant proteins and the proteins with target properties are screened and selected. The mutants that do not give rise to a functional protein or improved properties are discarded whereas those with desired properties are passed through further rounds of mutagenesis, screening and selection to give rise to a protein with improved properties (Adapted from Arnold and Bloom, 2009).

3.1.2. The need to engineer TS strains by directed evolution

Barry Duplantis in the Nano lab previously demonstrated that substituting psychrophilic genes into mesophilic pathogens makes them temperature-sensitive and induces protective immunity in mice (Duplantis *et al*, 2010)**.** However, advances in synthetic biology and TS variants of tetracycline repressors engineered in our lab (McWhinnie *et al*, 2014) have given us confidence that TS essential genes derived by directed evolution will be superior to those derived from psychrophiles.

In brief, essential genes can be engineered to be temperature-sensitive by two or three rounds of mutagenesis. These genes derived by directed evolution will have few nucleotide changes as compared to the Arctic essential genes which have hundreds of

nucleotide changes with respect to the native *M.TB* homologue. Fewer nucleotide changes would mean that the native transcription and translation rates will be maintained in a state similar to the native essential gene.

3.1.3. Objective of this research

The main objective of the work described in this chapter is to engineer TS essential genes by mutating the native essential genes using EP-PCR and designing an effective method for screening and selecting these TS strains. We think that TS strains having a range of different inactivation temperatures can be engineered using this approach as different essential genes having different functions can be targeted. This method may be applied to different bacteria and could be used to create different TS variants of organisms like *S. typhi*, *F. tularensis* and *M. tuberculosis.* The strains will then be tested for their suitability as appropriate vaccine candidates.

3.1.4. Method designed to mutagenize, screen and select for functional TS *M.TB* mutants

The aim of this method is to create a TS strain that will have multiple mutations so that it cannot revert to a temperature-resistant form on plating $<10^{10}$ cells of TS strain on agar plates. Making TS essential genes through multiple mutations does not mean that the TS mutants cannot be generated by single amino acid residue changes. However genes with multiple mutations are less likely to revert to their original sequence. Since our aim is to generate TS mutants with multiple amino acid residue changes, it is challenging to use structure based studies as the changes that impart temperature sensitivity can be anywhere in the protein. The method that we have designed to obtain TS mutants by directed evolution is described below:

3.1.4.1. Mutagenesis of native *M.TB* essential genes

The experiments explained below are hypothetical experiments which we will be carried out in the future. To generate a range of TS variants, we will use a combination of two very powerful mutagenesis techniques: i) EP-PCR (Rasila *et al*, 2009, description already mentioned on page 59) and ii) staggered extension PCR (StEP), which is a method of in-vitro mutagenesis and recombination (Zhao *et al*, 1998). We will adjust the EP- PCR to obtain TS variants with 10 nucleotide changes per kb of DNA and we will generate a library of mutants. We will then mix and match this pool of templates by staggered extension PCR, to make a complex bank of mutants and lastly screen and select for functional TS mutants.

3.1.4.2. Screening and selection of functional TS mutants

The mutants obtained must be screened and selected for using a very powerful high-throughput approach which cannot be performed in *Mycobacteria*. These experiments must be carried out in *E. coli* strains as they are much simpler to work with. **Therefore, our starting material to obtain a clone bank of TS mutants is the native *M.TB* essential gene codon harmonized for *E. coli*.**

The screening and selection method that we have designed is dependent upon the wild-type essential gene *(essG)* conditionally expressed on a plasmid in an *E. coli* strain This *E.coli* strain has a copy of the *essG* deleted from its chromosome. A clone bank of mutagenized *essG* will be supplied on the second plasmid. The expression of the wild type essential gene will be repressed and this will allow us to select and screen for 0.05-0.1% of the TS mutants that contain the functional mutagenized *essG* on the second plasmid (Fig 22).

An added advantage to our plan is that the first plasmid is diluted out using a R6K origin of replication. This R6K ori is dependent upon the Pi protein, a product of the *pir* gene to initiate replication (Kittleson *et al*, 2011). We will thus have the *pir* gene cloned downstream of the tetracycline repressor (*tetR*) controlled promoter. After introducing the clone bank of mutagenized essential gene into the *E. coli* strain, we will remove the (tetracycline) *tetR* inducer. Withdrawing the inducer of *tetR* will repress the *pir* gene and the plasmid will be diluted out. Only those *E. coli* strains would survive which have a functional *essG* TS mutant on the second plasmid. These experiments may sound challenging. However, the fact that we are working with essential genes allows us to select for viability and thus a functional TS essential gene.

We will use the Genetic QPix2 robot to pick 50,000 colonies and we expect to obtain 50-100 functional TS mutants. We will sequence these clones to obtain the nucleotide sequence and this will give us information regarding the amino acid sequence of the TS *M.TB* homologue. As we had codon harmonized the *M.TB* gene for expression in *E. coli*, we will now revert the nucleotide sequence of the functional TS mutants for *M.TB* and test if this mutant imparts TS in *M. smegmatis*. Our goal will be to select the best three TS variants and we will purchase the modified TS *M.TB* gene that will be made synthetically.

3.1.4.3. Introducing TS essential genes in *M.TB*

Before we proceed with any experiments in *M.TB,* we will substitute the TS essential genes into the chromosome of *M. smegmatis* and test if a TS phenotype is conferred by the variant *M.TB essG*. As TS *M. smegmatis* grows much faster as compared to *M.TB*, we can quickly identify any problems with the functioning of the TS

mutants. All the gene substitutions will be done with the codon-harmonized version of TS mutants for *M.TB*. We will substitute the TS *essG* into the chromosome of *M. smegmatis* as done previously, where we use mycobacterial suicide vector for integration into the chromosome and counterselectable markers to force excision (Appendix Figure A6).

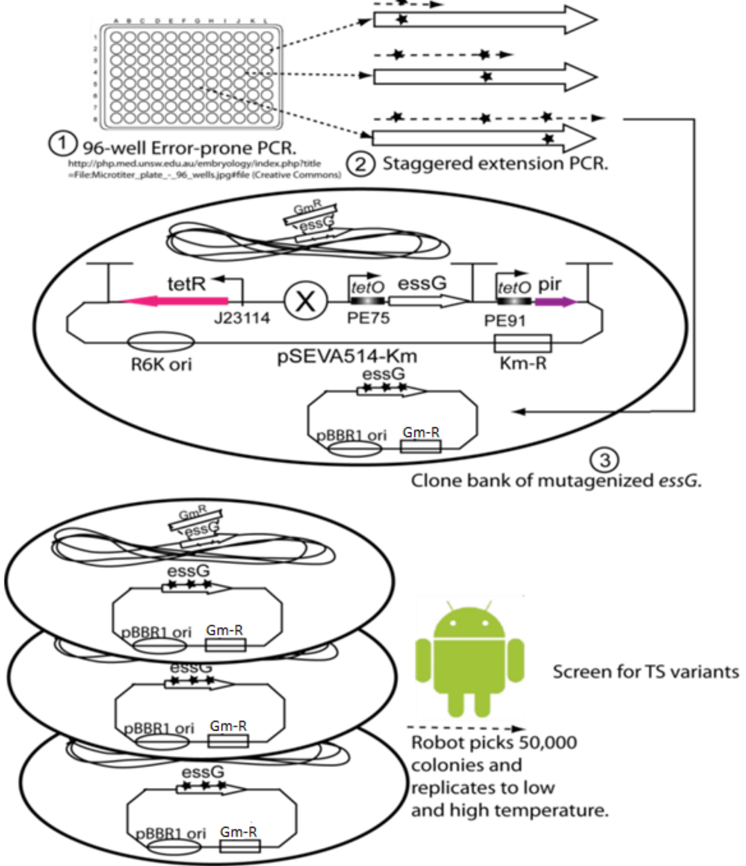

Figure 22: Directed evolution approach for creating TS essential strains (Figure credits- Dr. Francis E. Nano).
Step1: The *essG* is subjected to error prone (EP-PCR) in a 96-well plate to create 96 different mutant templates. After confirming that the mutant templates have 10 nucleotides changes/kb, the templates are mixed together and used as starting template for the staggered extension PCR (Step 2). The amplicons generated are cloned into a low copy Gm-R plasmid and electroporated into a custom-made *E. coli* strain. This strain has the *essG* on its chromosome deleted and the viability of this strain is maintained by the expression of the *essG* supplied on a conditionally replicating plasmid (step3). The expression of the *essG* on this plasmid is dependent upon the *tetR* controlled promoter. The replication of this plasmid is regulated by the expression of the *pi* protein which is also cloned downstream of a *tetR* controlled promoter. When the inducer of *tetR* is removed, the *essG* and the *pir* gene are no longer expressed and the plasmid will no longer replicate. At this point, the viability of the cell will be dependent upon the functional mutant *essG* supplied on the GmR plasmid. This method would allow for a strong selection of the functional mutant essential genes. A robotic colony picker will then pick 50,000 clones into 384-well micro-titer plates and we will replica plate the colonies onto agar plates, which will be incubated at different temperatures. A large number of these clones will be TS and we will screen and select for the clones that are TS at 37^{0}C and do not revert to a temperature-resistant form at a detectable level.

3.2. Materials and Methods

3.2.1. Designing constructs to create a conditionally replicating plasmid

The constructs designed to create a conditionally replicating plasmid were as follows: the tetracycline (*tetR*) *gene* was cloned downstream of the J23114 constitutive promoter (synthetic promoter designed by the Chris Anderson laboratory at the University of California, Berkeley; http://parts.igem.org/Part:BBa_J23114) and the *essG (E. coli MG1655)* and *pir* gene (DIAL strain; Kittleson *et al*, 2011) were cloned downstream of the *tetR*-controlled promoters, which contain an internal tetO sequence. These synthetic promoters were made by Ralph McWhinnie in the Nano lab and have been tested in both *E. coli* and *F. novicida* (McWhinnie *et al*, 2014). Three different constructs were designed that differed in the essential gene (*essG*), the strength of the tetO-containing promoters (Table 7) and RBS controlling the expression of the *essG* and the *pir* gene (Figure 23). Each DNA fragment was amplified using primers shown in Table 8 and assembled together using yeast mediated recombination as described previously. The yeast transformants obtained were screened for putative positive recombinants by PCR amplifying the construct and confirming the clones by sequencing.

Table 7: Table representing different tetO-containing promoters with their strengths, and the genes whose expression is controlled by these promoters.

tetO-Promoters	Strength	Essential genes
PE-75	Weak	*prsA*
PE-91	Medium	*Pir*
PE-100	Medium	*thrS*
PE-101	Medium	*Pir*
PE-123	Weak	*pgsA*
PE-128	Medium	*Pir*

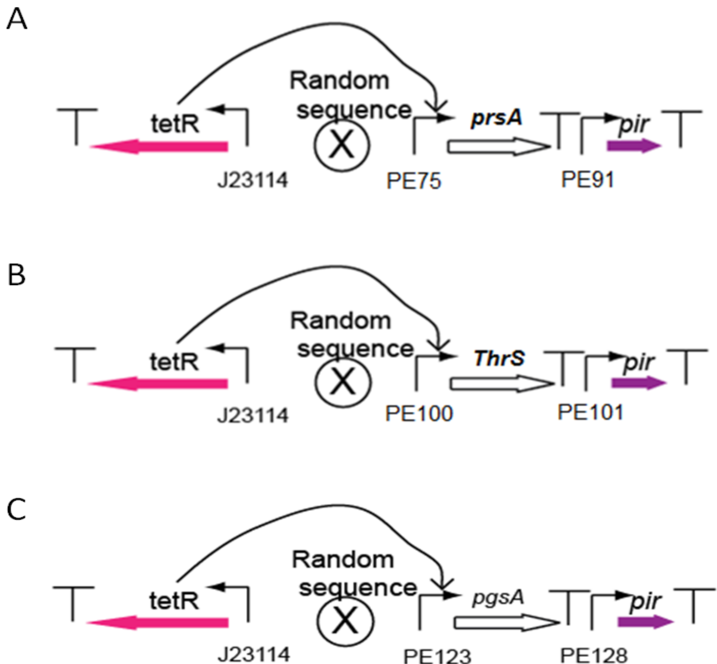

Figure 23: Constructs designed to create a conditionally replicating plasmid.
Each construct had the *tetR* gene cloned downstream of the J23114 constitutive promoter on one side and the *essG* and *pir* gene cloned downstream of weak or medium strength tetO-containing promoters on the other side separated by a random sequence. T represents the terminator. Three constructs were designed using three different essential genes: *prsA* gene (23A), *thrS* (23B) and *pgsA* (23C) gene. The *pgsA* construct was designed by Clara MacDonald, an honours student working under my supervision.

Table 8: Primers designed to create the *prsA* construct with the *prsA gene,* the *pir* gene and the tetO-containing promoters. Overlaps to adjacent regions are indicated in the lower case font and enzyme cut sites are underlined. (KpnI restriction enzyme site was added to the forward primer and SalI restriction enzyme site was added to the reverse primer).

	Expected size of constructs	Forward (5'-3')	Reverse (5'-3')
tetR gene with the J23114 constitutive promoter	~800bp	gtgagcgcgcgtaatacgactcactG GTACCTTAAGACCCACTTT CACATTTAAGT	cgcaccacgtgaacgattGGTA AACCCTGTGGCCTGT G
PE-75 promoter	~100bp	caatcgttcacgtggtgcgtCGCT CTAGAACTAGTTTAATT TTTTTAT	caggcatgcgtccattggtGGAT CCGTCACTATTAATA GGCT
prsA gene	~980bp	tattaatagtgacggatccACCAA TGGACGCATGCCT	tcccccctttttggttaactTTAG TGTTCGAACATGGCA GA
PE-91 promoter	~180 bp	agttaaccaaaaagggggggaTTTT ATCTCCCCTTTAATTTT TCCTCGCTCTAGA	ctcacatgcccttcccccGGAT CCTTTCGTAGTGAGA GTTA
pir gene	~920bp	ctcactacgaaaggatccGGGGG AAGGGCATGTGAGACT CAAGGTCATGATGGAC G	cactaaagggaacaaaagctgga GTCGACTCACCCCTTA GCTTTTTTGG

3.2.2. Creating the pSEVA-kan plasmid

The constructs designed were cloned into the MCS of the plasmid pSEVA514 between the T1 and the T0 terminators (Silva-Rocha *et al*, 2013, Figure 26). The PSEVA514 plasmid was chosen as this plasmid contains an R6K origin of replication which depends upon the trans-acting replication factor Pi, for plasmid replication to initiate (Kittleson *et al*, 2011). The tetracycline resistance marker in the pSEVA514 plasmid was replaced by the kanamycin (Km) marker from pSEVA256, as the *tetR* gene

was important for the construct to function properly. This pSEVA514-kan vector with the kanamycin (Km) marker was transformed into the *E. coli* DIAL strain 164H which contained the *pir* gene in its chromosome.

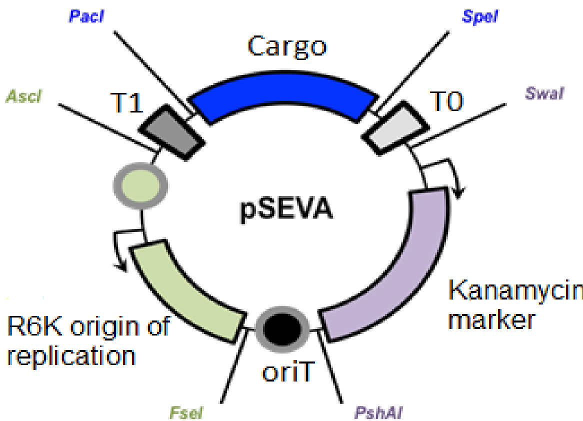

Figure 24: pSEVA514-kan plasmid with the R6K ori which depends upon the Pi protein, a product of the *pir* gene, to initiate replication (Adapted from Silva-Rocha *et al*, 2013)

3.2.3. Cloning the assembled constructs into the pSEVA-kan plasmid

The assembled constructs and the pSEVA514-kan plasmid were amplified using primers in Table 9. The PCR products were digested with KpnI and SalI restriction enzymes, ligated and transformed into electrocompetent *E. coli* DH10B cells. The cells were grown for 2 hours in LB broth with kanamycin and anhydrotetracycline (ATc) inducer and plated on LB-Km plates containing ATc (100ng/µl). The colonies obtained were screened for the presence of the assembled construct by PCR amplification and agarose gel electrophoresis.

Table 9: Primers designed to clone in the assembled *essG* constructs into pSEVA514-kan plasmid. Underline indicates the SalI and KpnI restriction sites added.

	Expected size of constructs	Forward (5'-3')	Reverse (5'-3')
pSEVA-514 kan plasmid	~5500bp	ataagtGTCGACCTGCAGGCATGCA	ataagtGGTACCGAGCTCGAATTCGC
Assembled *essG* constructs	3000bp-4000bp	GTGAGCGCGCGTAATACGACTCACTGGTACCTTAAGACCCACTTTCACATTTAAGT	CACTAAAGGGAACAAAAGCTGGAGTCGACTCACCCCTTAGCTTTTTTGG

3.2.4. Testing the conditionally replicating plasmids

The *E. coli* strains containing pSEVA plasmid with different assembled constructs were grown overnight in LB broth without kanamycin and ATc inducer. Appropriate dilutions were plated on LB with kanamycin to test the percentage of cells surviving in the absence of ATc inducer. The total number of colonies was determined by inoculating the colonies in LB without any selection of antibiotics and plating the culture on LB plates. The percentage of *E. coli* colonies surviving on LB-Km plates was calculated by dividing the number of colonies that appeared on LB-Km plates to the total number of colonies on LB plates without any antibiotics. **The pSEVA514-kan plasmid containing the *prsA essG* assembled construct (Figure 23A) was called pCP101**.

3.2.5. Preparing pREDi-pSEVA electrocompetent cells

The *prsA essG* knockout was created using the pREDi system. This chromosomal gene knockout method is mediated by a special plasmid called pREDi which contains an arabinose-inducible promoter. This promoter drives the expression of the lambda-red recombination proteins, which replaces the target genomic region with the marker-containing linear DNA cassette (Sung *et al*, 2011).

To create an *essG* knock-out strain, the first step was to make electrocompetent cells of *E. coli* containing the pREDi and pCP101 plasmid. For this, electrocompetent pREDi cells were prepared as described in Sung *et al*, 2011. The pCP101 (conditionally replicating plasmid) was transformed into the pREDi electrocompetent cells. The strains containing both plasmids were selected by plating the cultures on LB agar plates containing ampicillin, Km and ATc, and incubating the plates at 30°C. PREDi-pCP101 electrocompetent cells were prepared as described by Sung *et al*, 2011.

3.2.6. Creating knockout strains of *E. coli* using pREDi system

The chloramphenicol (Cm) antibiotic marker and 500bp flanking regions surrounding the *E. coli prsA* essential gene were amplified using primers in Table 9 and assembled together using yeast mediated recombination. Around 400-600ng of this linear DNA cassette was transformed into pREDi-pCP101 electrocompetent cells and electroporated at 2.5kV, 25µF and 200Ω. The cells were grown in LB for 1 hour at 30°C, plated on LB plates with Km and Cm and incubated at 30°C for 2 days. ATc was added to the media throughout this experiment because the viability of the cell was dependent upon the presence of the inducer. The putative *prsA* knock-out strains were determined by PCR amplifying the Cm cassette in the *E. coli* chromosome and sequencing the cassette.

Table 9: Primers designed to amplify the Cm marker and the *E. coli prsA* flanking regions to create the linear chloramphenicol cassette. Overlaps to adjacent regions are indicated in the lower case font.

	Expected size of constructs	Forward (5'-3')	Reverse (5'-3')
Cm marker	~925bp	gcagcgggtctttgagccgggttcgaCTTATTCAGGCGTAGCAAC	tggacgcatgcctgaggttcttctcACCGAATAAATACCTGTGACG
prsA gene left flanking region	~500bp	cgacggccagtgagcgcgcgtaatacgactcactggatccGACCAAAGCGTGCC	cgcctggttgctacgcctgaataagTCGAACCCGGCTCAAAGAC
prsA gene right flanking region	~500bp	atcttccgtcacaggtatttattcggtGAGAAGAACCTCAGGCATGC	attaaccctcactaaagggaacaaaagctggaggatccAGTGGTATCTGGTGGCGC

3.2.7. Testing the system using *E. coli prsA* essential gene

To test this system, the *prsA* essential gene from *E. coli* was PCR amplified and cloned into the SalI and the BamHI sites of the pBBR-MCS5 gentamicin (Gm) resistant plasmid (Kovach *et al*, 1995) using primers shown in Table 10. This plasmid was transformed into the knockout strain and the Km and ATc inducer were withdrawn to dilute out the pCP101 plasmid containing the *prsA* essential gene. The culture was grown overnight at 37°C and plated on LB-Gm plates to select for the pBBR-MCS5 plasmid containing the WT *prsA essG*.

Table 10: Primers designed to clone the *E. coli* WT *prsA* gene into the pBBR-MCS5 plasmid. Underline indicates the restriction sites added. BamHI restriction enzyme site was added to the forward primer and SalI restriction enzyme site was added to the reverse primer.

	Expected size of constructs	Forward (5'-3')	Reverse (5'-3')
pBBR-MCS5 plasmid	~4800bp	AATCAAGGATCCACT AGTTCTAGAGC	AATCAAGTC GAC CTCGAGGGGGGGCC
E. coli WT *prsA* gene	~980bp	AATCCAGGATCCTCTTT GAGCCGGGTTCGA	AATCCAGTCGACCGG TGA CAACGTCACCTT

3.2.8. Cloning the *Salmonella prsA* into the pBBR-MCS5 plasmid and transforming into the knockout strain

As the *Salmonella prsA* essential gene had 91% sequence identity with the *E. coli prsA* essential gene and there was no need for codon harmonization due to the similar G+C content, we decided to test if the *S. typhimurium* SL1344 *prsA gene* could support growth of the *E. coli* strain in the absence of the WT *E. coli prsA* gene. To achieve this, Dr. Francis Nano cloned the *prsA Salmonella* gene into the MCS of the pBBR-MCS-5 plasmid (Kovach *et al*, 1995) and transformed this plasmid into the knockout strain. Km and Atc inducer were withdrawn to dilute the pCP101 plasmid harbouring the *E. coli essG*. The cells were grown overnight at 37°C and plated on LB-Gm plates. The plates were incubated at 37°C and the colonies that appeared on the plates were tested for the presence of the *Salmonella prsA essG* and the absence of pCP101 plasmid by replica plating the colonies on LB plates containing Gm and Km antibiotic markers.

3.3. Results

3.3.1. *EssG* constructs assembled using yeast mediated recombination

The main objective of this research was to design a method to screen and select for TS essential genes generated by EP-PCR. As our aim was to design a conditionally replicating plasmid which replicated in the presence of the inducer ATc and diluted out in the absence of ATc, we designed three constructs that differed in the *essG*, the RBS of the *pir* gene and the strength of the *tetR*-controlled promoters (Figure 23). The constructs were assembled together using yeast mediated recombination as described previously. The positive clones were determined by isolating the pRS416 plasmid DNA from yeast colonies that grew on URA-DO plates, amplifying the assembled construct using primers specific to the constructs designed (Table 8) and analyzing the PCR products by agarose gel electrophoresis. Correct size PCR products (~3kb and ~4kb) were obtained suggesting putative positive clones (Figure 25). These clones were further confirmed by sequencing.

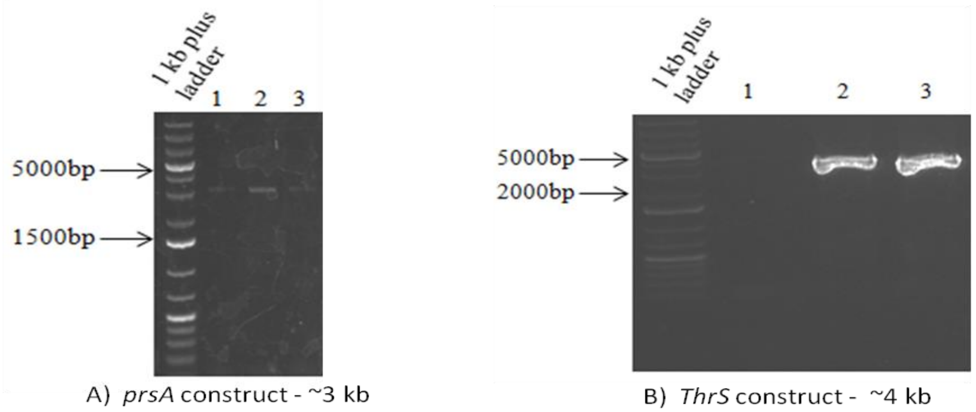

Figure 25: Screening putative clones containing correctly assembled constructs for directed evolution. A) Three positive clones (~3kb in size) were obtained for the *prsA* construct in which the *prsA essG* was assembled together with the *pir* gene, the PE75 and PE91 *tetR* controlled promoters and the WT *tetR* gene (Figure 23A) B) Two positive clones (~4kb in size) were obtained for the *thrS* construct in which the *thrS essG* was assembled together with the *pir* gene, the PE100 and PE101 *tetR* controlled promoters and the *tetR* gene with the J23114 promoter (Figure 23B).

3.3.2. pSEVA514-kan plasmid with *prsA* construct dilutes out efficiently

To determine which of the three plasmids containing the assembled constructs dilute out efficiently in the absence of ATc inducer, the plasmids were transformed into *E. coli* DH10B cells as explained in the methods section. The cells were grown for 18 hours in the absence of ATc inducer and plated on kanamycin and ATc to test if the plasmid was lost in the absence of inducer. The pCP101 plasmid with the PE-75 and PE-91 promoter combinations and the *prsA* essential gene worked best. This plasmid diluted out successfully in 18 hours and was used for further studies. The plasmid containing the *thrS* and *pgsA* construct did not dilute out efficiently in the absence of inducer (Table 11).

Table 11: Percentage of *E. coli* colonies with the construct containing plasmid which survive on LB-kan plates upon withdrawing kanamycin and ATc inducer.

Conditionally replicating plasmids cloned into *E. coli*	**Percentage of *E. coli* colonies surviving when grown in LB for 18 hours and plated on kan**
pSEVA514-kan with *prsA* construct (PE75 and PE91 tetO-promoters)- pCP101	0.05%
pSEVA514-kan with *thrS* construct (PE75 and PE91 tetO-promoters)	99%
pSEVA514-kan with *pgsA* construct (PE75 and PE91 tetO-promoters)	99%

3.3.3. *prsA* essential gene knockout created

The *prsA essG* knockout was created using the pREDi system as described in the methods section. Approximately 25 colonies were obtained on LB-Cm-ATc plates upon transforming the linear Cm cassette into the *E. coli* strain containing the pCP101 plasmid, suggesting putative *prsA essG* knockouts. These putative knockouts were screened using primers that would give an amplified product if Cm was present on the *E. coli* chromosome (Figure 26A). The putative positive clones determined by this screen were

tested for the absence of WT *E. coli prsA* using specific primers which would amplify the *E. coli prsA* gene in the chromosome (Figure 26B). Two putative positive clones #24 and #25 were confirmed by sequencing. The #25 *prsA* knockout strain was used for further experiments.

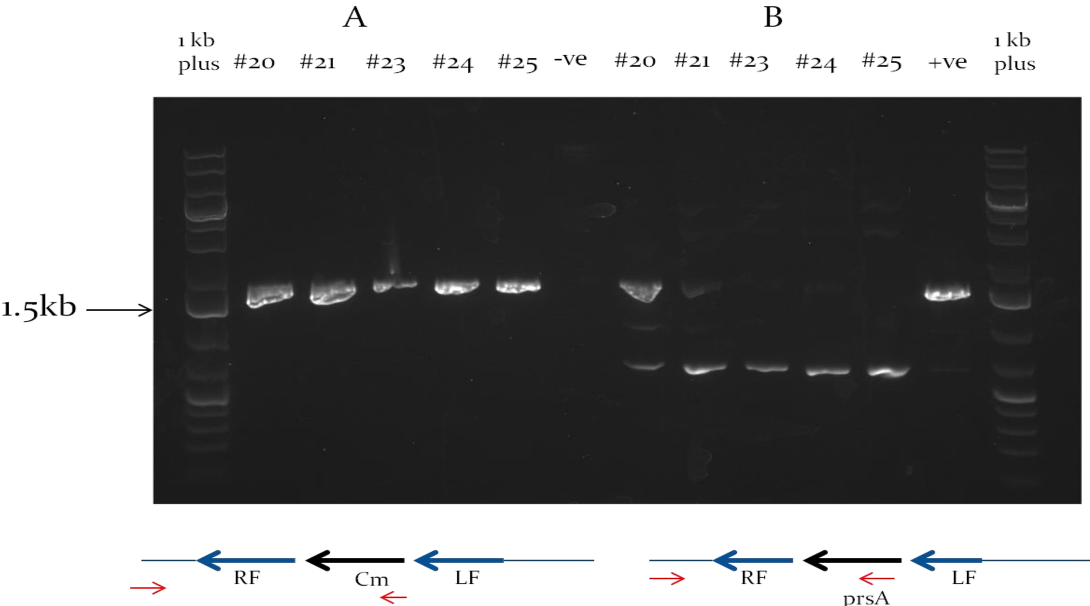

Figure 26: Screening for *prsA essG* gene knockouts.
Clones numbered #20, #21 and #23 to #25 were screened for the presence of the chloramphenicol (Cm) marker inserted into the *E. coli* genome. The primers used for screening are represented using red arrows in the figures below the gel. A) Screening putative clones for the presence of Cm marker in the *E. coli* chromosome using primers amplifying the Cm gene at one end and the region outside *prsA* right flanking region at the other end. B) Screening putative clones for the deletion of the native *prsA* gene in the *E. coli* chromosome using primers amplifying the *prsA* gene at one end and the region outside the *prsA* right flanking region at the other end. The putative positive clones give a desired band of 1.6kb with primers which amplify the Cm gene and no PCR product upon amplification using *prsA* gene specific primers (see primers Table 9 and Table 10). The #24 and #25 *prsA* knockouts were further confirmed by sequencing.

3.3.4. *Salmonella* WT *prsA* supports growth of the *E. coli* strain

The *Salmonella prsA* shares 91% nucleotide identity with the *E. coli prsA* and has a GC content of 54.5%, very close to the *E. coli prsA* gene (GC content-53%). As codon harmonization might be necessary before testing the *prsA M.TB* essential gene in the *E. coli* knockout strain, we tested whether the *Salmonella prsA* gene could support the growth of the knockout *E. coli* strain. To achieve this, WT *Salmonella prsA* was cloned into the pBBR MCS-5 Gm resistant plasmid and the colonies containing this plasmid were selected by plating on LB-Gm plates, as described in the methods. The colonies that grew on LB-Gm plates had lost the pCP101 plasmid, as confirmed by replica plating on LB-Km plates. This suggests that *Salmonella* WT *prsA* supports growth of the *E. coli* strain in the absence of the WT *E. coli prsA* gene.

3.4. Discussion and future directions

Our research group pioneered the use of psychrophilic essential genes as a tool for engineering TS bacterial pathogens. While we have found great potential in using psychrophilic essential genes, we have also found that they have limitations. Thus, we sought a complementary approach in creating stable TS strains by using the directed evolution of essential genes as a way of engineering a TS phenotype into pathogens. One advantage of using the directed evolution approach is that the resultant TS essential genes engineered by directed evolution will have few nucleotide changes when compared to those of the wild type and thus would be very close to the natural genes. As a result, there will be less selective pressure on these genes to revert to a temperature-resistant form.

3.4.1. Successful screening and selection system designed

We have designed a successful screening and selection approach which is universal. Our approach to make a gene TS involves subjecting the gene to EP-PCR and transforming this clone bank of mutants into an *E. coli* strain in which the native essential gene has been deleted and this gene is supplied on a conditionally replicating plasmid (figure 22). This screening and selection system was designed in *E. coli* which is a preferred model organism to carry out genetic manipulations.

The R6K origin of replication depends upon the trans-acting replication factor Pi to initiate plasmid replication (Shafferman *et al*, 1982). The Anderson lab at the University of California, Berkeley demonstrated that the R6K ori plasmid copy number can be evaluated by optimizing the level of expression of the pi protein (Kittleson *et al*, 2011). We therefore cloned the *pir* gene containing construct into the pSEVA plasmid with the R6K origin of replication. We used a combination of weak or medium strength *tetR*-controlled promoters and ribosome binding sites to control the pi protein expression. We

also used a weak *tetR*-controlled promoter to control the expression of the essential gene product. We wanted a combination that would give us a low plasmid copy number and produce just enough *essG* to maintain cell viability. The WT *tetR* gene was cloned in front of the J23114 constitutive promoter. Both the *tetR* and *essG* modules were assembled with the *pir* gene (Figure 23).

When ATc inducer was added, the repression of the *tetR* gene was relieved. This allowed the expression of both the *essG* and the *pir* gene. In the absence of ATc, the essential gene was no longer expressed and the *pir* gene was turned off, which resulted in the plasmid diluting out. During this time, the viability of the cell was supported by the wild-type *essG* which was supplied on a broad host range pBBR-MCS5 low-copy plasmid. Among the different promoter and ribosome binding site combinations used, the PE-75 and pE-91 promoter combination worked the best resulting in the plasmid diluting out efficiently.

3.4.2. Future directions to screen TS essential genes

We will mutate the *prsA Salmonella* gene by using Mutazyme II polymerase (Rasila *et al*, 2009) or Taq polymerase and clone it into the MCS of the pBBR-MCS5 plasmid. We will transform this plasmid into the *E. coli prsA* knock-out strain containing the conditionally replicating plasmid and withdraw the inducer. The surviving cells will contain a functional copy of the *essG* on the pBBR plasmid. We will screen for functional TS mutants by replica plating at high and low temperatures. Once we have standardized the experiments using *Salmonella prsA*, we will engineer TS *M.TB* essential genes using the same approach. The TS mutants obtained will be codon optimized for *M.TB* and substituted into the *M.TB* chromosome using allelic gene replacement.

Potential pitfalls may be that the TS essential genes function poorly, are inactivated at different temperatures or do not function at all. To solve this problem, we will test different TS variants of the same essential genes or use completely different essential genes. Moreover, as we are working with TS essential genes derived from directed evolution of the native *M.TB* gene, the problem of poor functionality is less likely compared to using Arctic essential genes.

3.4.3. This screening and selection approach is universal

We have demonstrated that the native *prsA* essential gene from gram-negative *Salmonella* can support the growth of the knock-out *E. coli* strain in the absence of *E. coli prsA* gene. Colleagues in my lab have attempted replacing the *prsA* essential gene with essential genes from other organisms like *Mycobacterium tuberculosis* and achieved success. As *Salmonella* and *Mycobacteria* are widely divergent organisms with different GC compositions, this technique may be used to screen and select TS essential genes from a plethora of microbes. This approach will allow us to select from a large number of TS variants with desired inactivation temperatures.

Conclusions

The objective of this research was to develop a stable TS *M.TB* strain which can be used as a potential vaccine candidate to reduce the global health burden caused by TB. As a step towards achieving this goal, we have engineered a stable TS *M. smegmatis* strain with a desired restrictive temperature of 37°C. We have demonstrated the functioning of the psychrophilic *ligA* gene in *M. smegmatis* and have obtained primary *M.TB* merodiploids, which will be selected and screened further in their ability to impart a TS phenotype. Considering that the amino acid identity of the *ligA* gene in *M. smegmatis* and *M.TB* strain is 80%, there is a good possibility that the *ligAPh* gene will impart TS to *M.TB*. We have also designed a novel approach to screen and select for TS essential genes by directed evolution. This approach could be used to engineer TS strains of different pathogens like *S. typhi, F. tularensis* and *M. tuberculosis* that are cleared by an immune response that includes an important T-cell mediated component.

Bibliography

Aharoni, A., Gaidukov, L., Khersonsky, O., McQ Gould, S., Roodveldt, C., & Tawfik, D. S. (2005). The "evolvability" of promiscuous protein functions. *Nature Genetics*, *37*(1), 73–76. http://doi.org/10.1038/ng1482

Arbues, A., Aguilo, J. I., Gonzalo-Asensio, J., Marinova, D., Uranga, S., Puentes, E., … Martin, C. (2013). Construction, characterization and preclinical evaluation of MTBVAC, the first live-attenuated *M. tuberculosis*-based vaccine to enter clinical trials. *Vaccine*, *31*(42), 4867–4873. http://doi.org/10.1016/j.vaccine.2013.07.051

Baba, T., Ara, T., Hasegawa, M., Takai, Y., Okumura, Y., Baba, M., … Mori, H. (2006). Construction of Escherichia coli K-12 in-frame, single-gene knockout mutants: the Keio collection. *Molecular Systems Biology*, *2*, 2006.0008. http://doi.org/10.1038/msb4100050

Balasubramanian, V., Pavelka, M. S., Bardarov, S. S., Martin, J., Weisbrod, T. R., Mcadam, R. A., … Jacobs, W. R. (1996). Allelic exchange in *Mycobacterium tuberculosis* with long linear recombination substrates. *Journal of Bacteriology*, *178*(1), 273–279. http://jb.asm.org/content/178/1/273.long

Barkan, D., Stallings, C. L., & Glickman, M. S. (2011). An improved counterselectable marker system for mycobacterial recombination using galK and 2-deoxy-galactose. *Gene*, *470*(1-2), 31–36. http://doi.org/10.1016/j.gene.2010.09.005

Baxby, D. (1977). The origins of vaccinia virus. *The Journal of Infectious Diseases*, *136*(3), 453–455. http://www.ncbi.nlm.nih.gov/pubmed/198484

Behar, S. M., & Ph, D. (2011). The next generation: tuberculosis vaccines that elicit protective CD8+ T cells. *Expert Review of Vaccines*, *6*(3), 441–456. http://doi.org/10.1586/14760584.6.3.441.

Behr, M. (2002). BCG--different strains, different vaccines? *The Lancet Infectious Diseases*, *2*(February), 86–92. doi:10.1016/S1473-3099(02)00182-2

Belshe, R. B., Edwards, K. M., Vesikari, T., Black, S. V, Walker, R. E., Hultquist, M., … Connor, E. M. (2007). Live attenuated versus inactivated influenza vaccine in infants and young children. *The New England Journal of Medicine*, *356*, 685–696. http://doi.org/10.1056/NEJMoa065368

Blander, J. M. (2007). Signalling and phagocytosis in the orchestration of host defence. *Cellular Microbiology*, *9*, 290–299. http://doi.org/10.1111/j.1462-5822.2006.00864

Bloom, J. D., & Arnold, F. H. (2009). In the light of directed evolution: pathways of adaptive protein evolution. *Proceedings of the National Academy of Sciences of the United States of America*, *106 Suppl* , 9995–10000. http://doi.org/10.1073/pnas.0901522106

Bloom, J. D., Meyer, M. M., Meinhold, P., Otey, C. R., MacMillan, D., & Arnold, F. H. (2005). Evolving strategies for enzyme engineering. *Current Opinion in Structural Biology*, *15*, 447–452. http://doi.org/10.1016/j.sbi.2005.06.004

Brändli, O. (1998). The clinical presentation of Tuberculosis. *Respiration*, *65*, 97–105. http://www.ncbi.nlm.nih.gov/pubmed/9580919

Calmette A. (1931). Preventive vaccination against tuberculosis with BCG. *Proceedings of the Royal Society of Medicine*, *24*, 85–94. http://www.ncbi.nlm.nih.gov/pmc/articles/PMC2182232/

Cann, A. J., Stanway, G., Hughes, P. J., Minor, P. D., Evans, D. M. A, Schild, G. C., & Almond, J. W. (1984). Reversion to neurovirulence of the live-attenuated sabin type 3 oral pollovirus vaccine. *Nucleic Acids Research*, *12*(20), 7787–7792. http://doi.org/10.1093/nar/12.20.7787

Chan, J., Xing, Y., Magliozzo, R. S., & Bloom, B. R. (1992). Killing of virulent *Mycobacterium tuberculosis* by reactive nitrogen intermediates produced by activated murine macrophages. *The Journal of Experimental Medicine*, *175*(April), 1111–1122. http://doi.org/10.1084/jem.175.4.1111

Clem, A. S. (2011). Fundamentals of vaccine immunology. *Journal of Global Infectious Diseases*. http://doi.org/10.4103/0974-777X.77299

Daniel, T. M. (2006). The history of tuberculosis. *Respiratory Medicine*, *100*, 1862–1870. http://doi.org/10.1016/j.rmed.2006.08.006

Day, C. L., Tameris, M., Mansoor, N., Van Rooyen, M., De Kock, M., Geldenhuys, H., … Hanekom, W. A. (2013). Induction and regulation of T-cell immunity by the novel tuberculosis vaccine M72/AS01 in South African adults. *American Journal of Respiratory and Critical Care Medicine*, *188*(4), 492–502. http://doi.org/10.1164/rccm.201208-1385OC

Dubes, G. R., & Chapin, M. (1956). Cold-Adapted Genetic Variants of Polio Viruses. *Science*, *124*(3222), 586–588. http://www.sciencemag.org/content/124/3222/586.2.extract

Duplantis, B. N., Bosio, C. M., & Nano, F. E. (2011). Temperature-sensitive bacterial pathogens generated by the substitution of essential genes from cold-loving bacteria: Potential use as live vaccines. *Journal of Molecular Medicine*, *89*, 437–444. http://doi.org/10.1007/s00109-010-0721-3

Duplantis, B. N., Osusky, M., Schmerk, C. L., Ross, D. R., Bosio, C. M., & Nano, F. E. (2010). Essential genes from Arctic bacteria used to construct stable, temperature-sensitive bacterial vaccines. *Proceedings of the National Academy of Sciences of the United States of America*, *107*, 13456–13460. http://doi.org/10.1073/pnas.1004119107

Ehrt, S., & Schnappinger, D. (2009). Mycobacterial survival strategies in the phagosome: Defence against host stresses. *Cellular Microbiology*, *11*(May), 1170–1178. http://doi.org/10.1111/j.1462-5822.2009.01335.

Etienne, G., Laval, F., Villeneuve, C., Dinadayala, P., Abouwarda, A., Zerbib, D., … Daffé, M. (2005). The cell envelope structure and properties of *Mycobacterium smegmatis* mc2155: Is there a clue for the unique transformability of the strain? *Microbiology*, *151*, 2075–2086. http://doi.org/10.1099/mic.0.27869-0

Eum, S. Y., Kong, J. H., Hong, M. S., Lee, Y. J., Kim, J. H., Hwang, S. H., … Barry, C. E. (2010). Neutrophils are the predominant infected phagocytic cells in the airways of patients with active pulmonary TB. *Chest*, *137*, 122–128. http://doi.org/10.1378/chest.09-0903

Flynn, J. L., & Chan, J. (2001). Immunology of tuberculosis. *Annual Review of Immunology*, *19*, 93–129. http://www.ncbi.nlm.nih.gov/pubmed/11244032

Gerdes, S. Y., Scholle, M. D., Campbell, J. W., Bala, G., Ravasz, E., Daugherty, M. D., … Oltvai, Z. N. (2003). Experimental Determination and System Level Analysis of Essential Genes in Escherichia coli MG1655. *J Bacteriol.*, *185*(19), 5673–5684. http://doi.org/10.1128/JB.185.19.5673

Giver, L., & Arnold, F. H. (1998). Combinatorial protein design by in vitro recombination. *Current Opinion in Chemical Biology*, *2*, 335–338. http://doi.org/10.1016/S1367-5931(98)80006-9

Glass, J. I., Assad-Garcia, N., Alperovich, N., Yooseph, S., Lewis, M. R., Maruf, M., … Venter, J. C. (2006). Essential genes of a minimal bacterium. *Proceedings of the National Academy of Sciences of the United States of America*, *103*(2), 425–430. http://doi.org/10.1073/pnas.0510013103

Grode, L., Ganoza, C. A., Brohm, C., Weiner, J., Eisele, B., & Kaufmann, S. H. E. (2013). Safety and immunogenicity of the recombinant BCG vaccine VPM1002 in a phase 1 open-label randomized clinical trial. *Vaccine*, *31*(9), 1340–1348. http://doi.org/10.1016/j.vaccine.2012.12.053

Hesseling, A C., Rabie, H., Marais, B. J., Manders, M., Lips, M., Schaaf, H. S., … Beyers, N. (2006). Bacille Calmette-Guerin Vaccine – Induced Disease in HIV-Infected and HIV-Uninfected Children. *Clinical Infectious Diseases*, *42*, 548–58. http://www.ncbi.nlm.nih.gov/pubmed/16421800

Hossain, M. M., & Norazmi, M.-N. (2013). Pattern recognition receptors and cytokines in Mycobacterium tuberculosis infection--the double-edged sword? *BioMed Research International*, *2013*, 179174. http://doi.org/10.1155/2013/179174

Husson, R. N., James, B. E., & Young, R. a. (1990). Gene Replacement and Expression of Foreign DNA in Mycobacteria. *Journal of Bacteriology*, *172*(2), 519–524. http://www.ncbi.nlm.nih.gov/pmc/articles/PMC208472/

Ji, Y., Zhang, B., Van, S. F., Horn, Warren, P., Woodnutt, G., … Rosenberg, M. (2001). Identification of critical staphylococcal genes using conditional phenotypes generated by antisense RNA. *Science (New York, N.Y.)*, *293*(5538), 2266–2269. http://doi.org/10.1126/science.1063566

Jr, M. S. P., & Jr, W. R. J. (1999). Comparison of the Construction of Unmarked Deletion Mutations in *Mycobacterium smegmatis* , *Mycobacterium bovis Bacillus Calmette-Guérin* and *Mycobacterium tuberculosis* H37Rv by Allelic Exchange. *Journal of Bacteriology*, *181*(16), 4780–4789. http://www.ncbi.nlm.nih.gov/pubmed/10438745

Kalpana, G. V, Bloom, B. R., & Jacobs, W. R. (1991). Insertional mutagenesis and illegitimate recombination in mycobacteria. *Proceedings of the National Academy of Sciences of the United States of America*, *88*(June), 5433–5437. http://doi.org/10.1073/pnas.88.12.5433

Kapopoulou, A., Lew, J. M., & Cole, S. T. (2011). The MycoBrowser portal: A comprehensive and manually annotated resource for mycobacterial genomes. *Tuberculosis*, *91*(1), 8–13. http://doi.org/10.1016/j.tube.2010.09.006

Kearse, M., Moir, R., Wilson, A., Stones-Havas, S., Cheung, M., Sturrock, S., … Drummond, A. (2012). Geneious Basic: An integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics*, *28*(12), 1647–1649. http://doi.org/10.1093/bioinformatics/bts199

Kittleson, J. T., Cheung, S., & Anderson, J. C. (2011). Rapid optimization of gene dosage in E. coli using DIAL strains. *Journal of Biological Engineering*, *5*(1), 10. http://doi.org/10.1186/1754-1611-5-10

Kleinnijenhuis, J., Oosting, M., Joosten, L. a B., Netea, M. G., & Van Crevel, R. (2011). Innate immune recognition of *Mycobacterium tuberculosis*. *Clinical & Developmental Immunology*, *2011*, 405310. http://doi.org/10.1155/2011/405310

Klingler, K., Tchou-Wong, K. M., Brändli, O., Aston, C., Kim, R., Chi, C., & Rom, W. N. (1997). Effects of mycobacteria on regulation of apoptosis in mononuclear phagocytes. *Infection and Immunity*, *65*(12), 5272–5278. http://iai.asm.org/content/65/12/5272

Korbel, D. S., Schneider, B. E., & Schaible, U. E. (2008). Innate immunity in tuberculosis: myths and truth. *Microbes and Infection*, *10*(9), 995–1004. http://doi.org/10.1016/j.micinf.2008.07.039

Korycka-Machala, M., Rychta, E., Brzostek, A., Sayer, H. R., Rumijowska-Galewicz, A., Bowater, R. P., & Dziadek, J. (2007). Evaluation of NAD+-dependent DNA ligase of mycobacteria as a potential target for antibiotics. *Antimicrobial Agents and Chemotherapy*, *51*(8), 2888–2897. http://doi.org/10.1128/AAC.00254-07

Korzeniewska-Kosela, M. (1994). Tuberculosis in Young Adults and the Elderly. A Prospective Comparison Study. *Chest Journal*, *106*, 28. http://doi.org/10.1378/chest.106.1.28

Kovach, M. E., Elzer, P. H., Hill, D. S., Robertson, G. T., Farris, M. a, Roop, R. M., & Peterson, K. M. (1995). Four new derivatives of the broad-host-range cloning vector pBBR1MCS, carrying different antibiotic-resistance cassettes. *Gene*, *166*, 175–176. http://doi.org/10.1016/0378-1119(95)00584-1

Lee, J., Hartman, M., & Kornfeld, H. (2009). Macrophage apoptosis in tuberculosis. *Yonsei Medical Journal*, *50*(1), 1–11. http://doi.org/10.3349/ymj.2009.50.1.1

Lew, J. M., Kapopoulou, A., Jones, L. M., & Cole, S. T. (2011). TubercuList - 10 years after. *Tuberculosis*, *91*(1), 1–7. http://doi.org/10.1016/j.tube.2010.09.008

Liu, J., Tran, V., Leung, A. S., Alexander, D. C., & Zhu, B. (2009). BCG vaccines: their mechanisms of attenuation and impact on safety and protective efficacy. *Human Vaccines*, *5*(January 2015), 70–78. http://doi.org/10.4161/hv.5.2.7210

Maassab, H. F., & Bryant, M. L. (1999). The development of live attenuated cold-adapted influenza virus vaccine for humans. *Reviews in Medical Virology*, *9*, 237–244. http://doi.org/10.1002/(SICI)1099-1654(199910/12)9:4.

Maassab, H. F., & DeBorde, D. C. (1985). Development and characterization of cold-adapted viruses for use as live virus vaccines. *Vaccine*, *3*(December), 355–369. http://doi.org/10.1016/0264-410X(85)90124-0

Maglione, P. J., & Chan, J. (2009). How B cells Shape the Immune Response against *Mycobacterium tuberculosis*. *European Journal of Immunol.*, *39*(3), 676–686. http://doi.org/10.1002/eji.200839148.

Mahairas, G. G., Sabo, P. J., Hickey, M. J., Singh, D. C., & Stover, C. K. (1996). Molecular analysis of genetic differences between *Mycobacterium bovis* BCG and virulent *M. bovis*. *Journal of Bacteriology*, *178*(5), 1274–1282. http://www.ncbi.nlm.nih.gov/pubmed/8631702

Mansoor, N., Scriba, T. J., Kock, M. De, Tameris, M., Keyser, A., Little, F., … Hanekom, W. a. (2010). Infant HIV-1 Infection Severely Impairs the Bacille Calmette- Guerin Vaccine-Induced Immune Response. *J Infect Dis*, *199*(7), 1–17. http://doi.org/10.1086/597304.

Markham, J. F., Morrow, C. J., Scott, P. C., & Whithear, K. G. (2015). Safety of a temperature-sensitive clone of *Mycoplasma synoviae* as a live vaccine. *Avian Diseases*, *42*(4), 677–681. http://www.ncbi.nlm.nih.gov/pubmed/9876835

Mccullum, E. O., Williams, B. A R., Zhang, J., & Chaput, J. C. (2010). In Vitro Mutagenesis Protocols. *Methods in Molecular Biology*, *634*(3), 103–109. http://doi.org/10.1007/978-1-60761-652-8

McWhinnie, R. L., & Nano, F. E. (2014). Synthetic promoters functional in *Francisella novicida* and *Escherichia coli*. *Applied and Environmental Microbiology*, *80*(1), 226–234. http://doi.org/10.1128/AEM.02793-13

Mihret, A. (2012). The role of dendritic cells in *Mycobacterium tuberculosis* infection. *Virulence*, *3*(7), 654–9. http://doi.org/10.4161/viru.22586

Miyazaki, K., Wintrode, P. L., Grayling, R. A, Rubingh, D. N., & Arnold, F. H. (2000). Directed evolution study of temperature adaptation in a psychrophilic enzyme. *Journal of Molecular Biology*, *297*, 1015–1026. http://doi.org/10.1006/jmbi.2000.3612

Montagnani, C., Chiappini, E., Galli, L., & de Martino, M. (2014). Vaccine against tuberculosis: what's new? *BMC Infectious Diseases*, *14*(Suppl 1), S2. http://doi.org/10.1186/1471-2334-14-S1-S2

Moura, G., Pinheiro, M., Silva, R., Miranda, I., Afreixo, V., Dias, G., … Santos, M. a S. (2005). Comparative context analysis of codon pairs on an ORFeome scale. *Genome Biology*, *6*(3), R28. http://doi.org/10.1186/gb-2005-6-3-r28

Murray R. Patrick, Rosenthal S. Ken, Micheal A. Pfaller. 2005. *Mycobacterium tuberculosis*. *Medical Microbiology- Fifth Edition.* Philadelphia, PA: Elsevier

Noirot, P., & Kolodner, R. D. (1998). DNA strand invasion promoted by *Escherichia coli* RecT protein. *Journal of Biological Chemistry*, *273*(20), 12274–12280. http://doi.org/10.1074/jbc.273.20.12274

O'Garra, A., Redford, P. S., McNab, F. W., Bloom, C. I., Wilkinson, R. J., & Berry, M. P. R. (2013). *The immune response in tuberculosis*. *Annual Review of Immunology* (Vol. 31). http://doi.org/10.1146/annurev-immunol-032712-095939

Ottenhoff, T. H. M., & Kaufmann, S. H. E. (2012). Vaccines against tuberculosis: Where are we and where do we need to go? *PLoS Pathogens*, *8*(5). http://doi.org/10.1371/journal.ppat.1002607

Parikh, A., Kumar, D., Chawla, Y., Kurthkoti, K., Khan, S., Varshney, U., & Nandicoori, K. V. (2013). Development of a new generation of vectors for gene expression, gene replacement, and protein-protein interaction studies in mycobacteria. *Applied and Environmental Microbiology*, *79*, 1718–1729. http://doi.org/10.1128/AEM.03695-12

Parish, T., Gordhan, B. G., McAdam, R. A, Duncan, K., Mizrahi, V., & Stoker, N. G. (1999). Production of mutants in amino acid biosynthesis genes of Mycobacterium tuberculosis by homologous recombination. *Microbiology (Reading, England)*, *145 (Pt 1*(1999), 3497–503. http://www.ncbi.nlm.nih.gov/pubmed/10627047

Raja, A. (2004). Immunology of tuberculosis. *Indian Journal of Medical Research*, *120*(October), 213–232. http://doi.org/10.4084/MJHID.2014.027

Rasila, T. S., Pajunen, M. I., & Savilahti, H. (2009). Critical evaluation of random mutagenesis by error-prone polymerase chain reaction protocols, *Escherichia coli* mutator strain, and hydroxylamine treatment. *Analytical Biochemistry*, *388*(1), 71–80. http://doi.org/10.1016/j.ab.2009.02.008

Redford, P. S., Murray, P. J., & O'Garra, A. (2011). The role of IL-10 in immune regulation during *M. tuberculosis* infection. *Mucosal Immunology*, *4*(3), 261–270. http://doi.org/10.1038/mi.2011.7

Romano, M., & Huygen, K. (2012). An update on vaccines for tuberculosis – there is more to it than just waning of BCG efficacy with time. *Expert Opinion on Biological Therapy*, 1–10. http://doi.org/10.1517/14712598.2012.721768

Sakamoto, K. (2012). The Pathology of *Mycobacterium tuberculosis* Infection. *Veterinary Pathology*, *49*(3), 423–439. http://doi.org/10.1177/0300985811429313

Shafferman, A., Kolter, R., Stalker, D., & Helinski, D. R. (1982). Plasmid R6K DNA replication. *Journal of Molecular Biology*, *161*, 57–76. http://doi.org/10.1016/0022-2836(82)90278-9

Silva-Rocha, R., Martínez-García, E., Calles, B., Chavarría, M., Arce-Rodríguez, A., De Las Heras, A., … De Lorenzo, V. (2013). The Standard European Vector Architecture (SEVA): A coherent platform for the analysis and deployment of complex prokaryotic phenotypes. *Nucleic Acids Research*, *41*(November 2012), 666–675. http://doi.org/10.1093/nar/gks1119

Snapper, S. B., Melton, R. E., Mustafa, S., Kieser, T., & Jacobs, W. R. (1990). Isolation and characterization of efficient plasmid transformation mutants of Mycobacterium smegmatis. *Molecular Microbiology*, *4*(11), 1911–1919. http://doi.org/10.1111/j.1365-2958.1990.tb02040.x

Steingart, K. R., Henry, M., Ng, V., Hopewell, P. C., Ramsay, A., Cunningham, J., … Pai, M. (2006). Fluorescence versus conventional sputum smear microscopy for tuberculosis: a systematic review. *Lancet Infectious Diseases*, *6*(September), 570–581. http://doi.org/10.1016/S1473-3099(06)70578-3

Sturgill-Koszycki, S., Schlesinger, P. H., Chakraborty, P., Haddix, P. L., Collins, H. L., Fok, A K., … Russell, D. G. (1994). Lack of acidification in Mycobacterium phagosomes produced by exclusion of the vesicular proton-ATPase. *Science (New York, N.Y.)*, *263*, 678–681. http://doi.org/10.1126/science.8303277

Sun, R., Skeiky, Y. A W., Izzo, A., Dheenadhayalan, V., Imam, Z., Penn, E., … Sadoff, J. C. (2009). Novel recombinant BCG expressing perfringolysin O and the over-expression of key immunodominant antigens; pre-clinical characterization, safety and protection against challenge with *Mycobacterium tuberculosis*. *Vaccine*, *27*, 4412–4423. http://doi.org/10.1016/j.vaccine.2009.05.048

Sung, Bon Hyun, Jun Hyoung Lee, and Sun Chang Kim. 2011. “Scarless chromosomal gene knockout methods.” *Strain Engineering-Methods in Molecular Biology* 765: 27–42

Tameris, M. D., Hatherill, M., Landry, B. S., Scriba, T. J., Snowden, M. A., Lockhart, S., … McShane, H. (2013). Safety and efficacy of MVA85A, a new tuberculosis vaccine, in infants previously vaccinated with BCG: A randomised, placebo-controlled phase 2b trial. *The Lancet*, *381*(9871), 1021–1028. http://doi.org/10.1016/S0140-6736(13)60177-4

Van Kessel, J. C., & Hatfull, G. F. (2007). Recombineering in *Mycobacterium tuberculosis*. *Nature Methods*, *4*(December), 147–152. http://doi.org/10.1038/nmeth996

Van Kessel, J. C., & Hatfull, G. F. (2008). Recombineering mycobacteria and their phages. *Nature Reviews. Microbiology*, *6*(11), 851–857. http://doi.org/10.1038/nrmicro2014

Verver, S., Warren, R. M., Beyers, N., Richardson, M., Van Der Spuy, G. D., Borgdorff, M. W., … Van Helden, P. D. (2005). Rate of reinfection tuberculosis after successful treatment is higher than rate of new tuberculosis. *American Journal of Respiratory and Critical Care Medicine*, *171*, 1430–1435. http://doi.org/10.1164/rccm.200409-1200OC

Weerdenburg, E. M., Peters, P. J., & van der Wel, N. N. (2010). How do mycobacteria activate CD8+ T cells? *Trends in Microbiology*, *18*(December), 1–10. http://doi.org/10.1016/j.tim.2009.10.004

Weiner, J., & Kaufmann, S. H. E. (2014). Recent advances towards tuberculosis control: Vaccines and biomarkers. *Journal of Internal Medicine*, *275*, 467–480. http://doi.org/10.1111/joim.12212

White, M. D., Bosio, C. M., Duplantis, B. N., & Nano, F. E. (2011). Human body temperature and new approaches to constructing temperature-sensitive bacterial vaccines. *Cellular and Molecular Life Sciences*, *68*, 3019–3031. http://doi.org/10.1007/s00018-011-0734-2

WHO Global tuberculosis report 2013. Geneva: WHO Press; 2013.

Wilkinson, A., Day, J., & Bowater, R. (2001). Bacterial DNA ligases. *Molecular Microbiology*, *40*, 1241–1248.

Yang, X. Y., Chen, Q. F., Li, Y. P., & Wu, S. M. (2011). *Mycobacterium vaccae* as adjuvant therapy to anti-tuberculosis chemotherapy in never-treated tuberculosis patients: A meta-analysis. *PLoS ONE*, *6*(9), 0–7. http://doi.org/10.1371/journal.pone.0023826

Young, D., Stark, J., & Kirschner, D. (2008). Systems biology of persistent infection: tuberculosis as a case study. *Nature Reviews. Microbiology*, *6*(july), 520–528. http://doi.org/10.1038/nrmicro1919

Zhao, H., Giver, L., Shao, Z., Affholter, J. A., & Arnold, F. H. (1998). Molecular evolution by staggered extension process (StEP) in vitro recombination. *Nature Biotechnology*, *16*, 258–261. http://doi.org/10.1038/nbt0898-773

Appendix

Table A1: Essential genes selected for this study and their functions.

Gene Name	Enzyme encoded by the gene	Function
ligA	NAD+ dependent DNA ligase	DNA replication, recombination and repair
prsA	Ribose-phosphate pyrophosphokinase	Nucleotide biosynthesis
thrS	Threonyl t-RNA synthetase	Role in protein translation
pgsA1	CDP-diacylglycerol-glycerol-3-phosphate 3-phosphatidyltransferase	Glycerophospholipid metabolism

Table A2: Primers designed to sequence the *ligAPh* gene in the TS *M. smegmatis* strain.

Primer 1	TGACGCCCTTGGGATTTGCC
Primer 2	ATCATCTCGCCGCGTGCCT
Primer 3	GGCGACCACAGCACCGCA
Primer 4	GCTGACATGGCCGAGGATCG

Table A3: Primers designed to amplify the hygromycin resistant cassette and clone immediately downstream of the *ligA* gene, in the *M.TB* left flank or 200bp into the *M.TB* right flank. Underline indicates the restriction sites added (BamHI restriction site was added to the forward and reverse primer).

Constructs to be amplified	Expected size of the products	Forward (5'-3')	Reverse (5'-3')
hygR gene with promoter	~1500bp	AGCTGGGGATCCTCAC GCATGCGGTGGTAC	AGCTGGGGATCC TTAACGGTGAAG GGTCAGGC
ligA-M.TB construct (to clone *hygR* gene 200bp into the right flanking regions)	~8500bp	AGCTGGGGATCCACCA TGGTCGCGGCTCCAT	AGCTGGGGATCC GGTATCCATGTCG CGCCA
ligA-M.TB construct (to clone the *hygR* gene 200bp into the left flanking regions)	~8500bp	AGCTGGGGATCCCCGC AGCAGCGACCTCTT	AGCTGGGGATCC CGGATCAAGGTA CCACCGC
ligA-M.TB construct (to clone the *hygR* gene downstream of the *ligA* gene)	~8500bp	AGCTGGGGATCCGAG CGTAACGCCACTGC	AGCTGGGGATCC GGCGAACCTTGGT TGCCA

Table A4: Primers designed to clone the *galK* gene into the PST-KO vector. Underline indicates the restriction sites added (BspHI restriction site was added to the forward primer and NheI restriction site was added to the reverse primer).

Construct to be amplified	Expected size of the products	Forward (5'-3')	Reverse (5'-3')
PST-KO plasmid	~4200bp	ataagtTCATGAGCGTTTT TATTTGTTAACTGTT	aatctgGCTAGCACGCAA AAGAAAATGCCG
galK gene with promoter	~1400bp	ataagtTCATGAAGACCCC AGGCTTGACAC	aatctgGCTAGCGCAGGA GTTTCGTTCAGC

A5. Sequence of the codon harmonized *ligA* gene from the psychrophile *Pseudoalteromonas haloplaktis (TAC 125)*

ATGGCCAGCAGCATCTCGGAGCAGATCAACCATCTGCGTTCGACGCTCGAACAGCACTCG
TACAACTACTACGTGCTCGATACGCCCTCGATCCCGGATGCCGAATACGACCGTCTGCTG
CAGCAGCTCAGCGCGCTGGAGACGCAGCACCCGGAACTGATCACGGCCGACTCGCCGACG
CAGAAAGTGGGCGGTGCCGCGCTGTCGAAATTCGAGCAGGTGGCGCACCAGGTGCCGATG
CTGTCGCTCGATAACGCGTTCAGCGAGGATGAGTTCATCGCGTTCAACCGCCGTATCAAA
GAGCGTCTGATGTCGACCGAAGAACTGACGTTCTGTTGTGAACCGAAACTCGATGGCCTC
GCGGTGTCGATCATCTACCGTGATGGCGTGCTGGTGCAGGCCGCCACGCGCGGTGATGGG
CTCACCGGCGAGAACGTGACGCAGAACGTGAAAACCATCCGTAACGTGCCGCTGAAACTG
CGCGGTAGCGATTACCCGGCCGAACTCGAGGTGCGCGGCGAGGTGTTCATGGATAACGCA
GGCTTCGAGAAGTTCAACATCGAGGCCGAAAAACGTGGTGAAAAAGTGTTCGTGAACCCG
CGCAACGCCGCCGCCGGTAGCCTGCGCCAGCTCGACTCGAAAATCACCGCCAAACGCCCG
CTGATGTTCTACGCCTACAGCACCGGTCTGGTGGCCGACGGTAGCATCGCCGAGGATCAT
TACCAGCAGCTCGAAAAACTGACCGATTGGGGGCTGCCGCTGTGCCCGGAGACCAAACTG
GTGGAGGGCCCGCAGGCCGCGCTGGCCTACTACACCGACATCCTCACCCGCCGTGGCGAA
CTCAAATACGAGATCGATGGCGTGGTGATCAAAATCAACCAGAAAGCGCTGCAGGAGCGT
CTCGGCTTCGTGGCACGCGCGCCGCGTTGGGCGATCGCCTACAAGTTCCCGGCGCAGGAG
GAGATCACCAAACTGCTCGATGTGGAGTTCCAGGTGGGGCGTACCGGCGCGATCACGCCG
GTGGCACGCCTCGAACCGGTGTTCGTGGGTGGTGTGACCGTGTCGAACGCCACGCTGCAC
AACGGCGATGAGATCGCACGCCTCGGCGTGAAAGTGGGCGACACGGTGATCATCCGCCGT
GCCGGGGACGTGATCCCGCAGATCACGCAGGTGGTGCTCGAGCGCCGCCCGGATGATGCA
CGCGATATCGAGTTCCCGGTGACGTGCCCGATCTGTGACTCCCATGTGGAAAAAGTGGAG
GGTGAGGCGGTGGCGCGTTGTACCGGTGGTCTGGTGTGCCCGGCGCAGCGTAAACAGGCG
ATCAAACACTTCGCGTCGCGCAAAGCGCTCGATATCGACGGCCTCGGCGATAAAATCGTG
GATCAGCTCGTCGACCGCGAACTGATCAAAACCCCGGCCGATCTGTTCATCCTCAAGCAG
GGCCATTTCGAATCGCTCGAGCGTATGGGGCCGAAGTCGGCCAAAAACCTGGTGACCGCG
CTGCAGGACGCCAAAGCCACCACGCTGGCCAAGTTCCTCTACTCGCTCGGTATCCGTGAG
GCCGGTGAGGCCACCACGCAGAACCTGGCCAACCATTTCCTGACGCTCGAGAACGTGATC
AACGCAAGCATCGATTCGCTGACGCAGGTGTCGGATGTGGGCGAGATCGTGGCCACGCAT
GTGCGTAGCTTCTTCGCCGAACAGCATAACCTCGATGTGGTGAACGCGCTGGTGGAGCAG
GGTATCAACTGGCCGGAACTGACGCCGCCGTCGGCGCAGGAGCAGCCGCTCGCAGGCCTG
GTGTACGTGCTGACCGGTACCCTCAACACCCTCAACCGTAACGACGCCAAAGCGCGTCTG
CAGCAGCTCGGTGCCAAAGTGTCGGGTTCGGTGTCGGCCAAAACCGATGCGCTGGTGGCA
GGCGAGAAGGCAGGCTCGAAACTGACCAAGGCGCAGGACCTCGGTATCGATGTGCTGACC
GAGGAGGATCTGATCAACCTGCTGGAGCAGCATAACGGCTGA

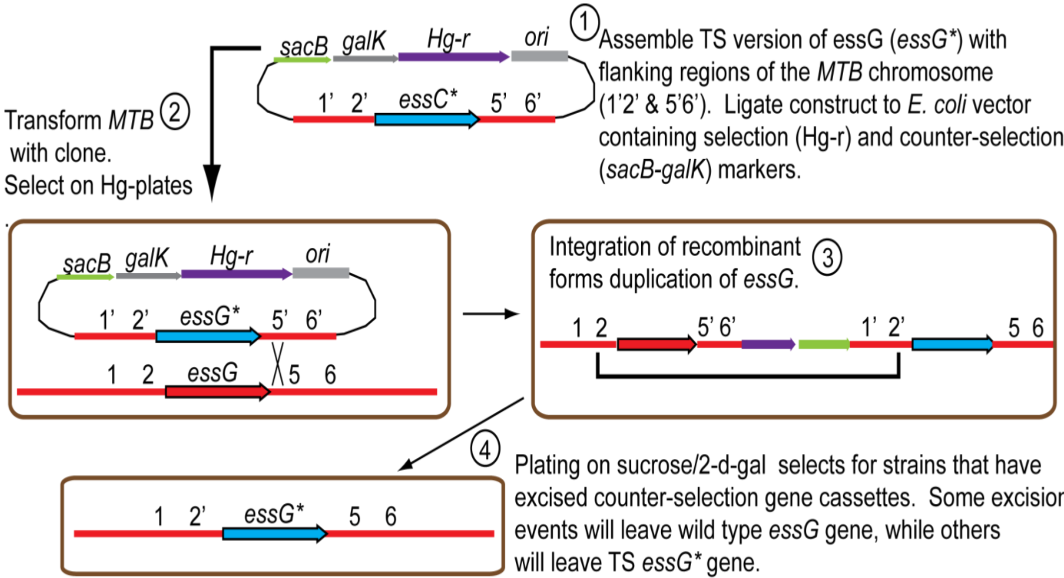

Figure A6: Allelic gene replacement of the TS essential genes created by directed evolution into *M.TB* using suicide vector (Figure credits: Dr. Francis E. Nano).

A7. Buffer recipes:

TL buffer - 100 mM Lithium Acetate (LiAc) in TE

TLP buffer- 45% PEG in 100mM LiAc/TE

Table A8: List of strains and vectors used in this study

Strains and vectors	**Characteristics**	**References**
M. tuberculosis Erdman/ H37Rv	WT *M. tuberculosis* strain	ATCC 35801/ ATCC 25618
M. *smegmatis* mc2155	WT *M. smegmatis* strain	ATCC 700084, Snapper et al, 1990
p2NIL vector	Mycobacterial suicide vector containing kanamycin resistant marker and SacB counterselectable marker	Korycka-Machala et al, 2007
*pSTKO-GalK vector	Mycobacterial suicide vector wih hsp60 promoter driving the expression of SacB and GalK counterselectable marker	Parikh et al, 2013; Barkan et al, 2010
pBBR-MCS5 vector	Broad host range Gentamycin resistance vector	Kovach et al, 1995
*pCP101 vector	Vector with PSEVA514 backbone and the *prsA-pir* gene construct.	Silva-Rocha et al, 2012

*Vectors made by Crystal T. Pinto