In vitro Studies on the Sputum Phenotype of *Mycobacterium tuberculosis*

Su-Min Lee

Thesis submitted to The University of Leicester for the degree of Doctor of Philosophy

May 2012

Department of Infection, Immunity and Inflammation
College of Medicine, Biological Sciences and Psychology
University of Leicester
University Road, Leicester, LE1 7RH

Abstract

Two defining characteristics of *Mycobacterium tuberculosis* (*Mtb*) in sputum are gene expression and the production of lipid bodies (LBs). Previous analysis of the *Mtb* sputum transcriptome demonstrated that a population of persister-like bacilli predominate in sputum. LBs are intracellular structures consisting of triacylglycerol. Their formation is environmentally regulated, and occurs in response to a number of stresses, including conditions which induce a non-replicating persistence phenotype. Based upon these findings, *Mtb* in sputum may express a unique, transmission-adapted phenotype.

Comparison of the gene expression of *Mtb* in sputum against *Mtb* exposed to *in vitro* stimuli demonstrated that no obvious single growth condition fully replicated the sputum transcriptome. In contrast, *Mtb* cultures exposed to multiple stimuli, including phosphate-buffered saline or RPMI medium, nitric oxide, cholesterol, oleic acid and static incubation had gene expression that correlated significantly to that in sputum. However, the exact combination of stresses is yet to be defined.

It was demonstrated in these studies that the presence of LBs as a single factor does not influence the transmission adaptability of *Mtb* bacilli as tested here. An increased proportion of LB-positive cells did not confer increased survival to transmission stresses (desiccation and ultraviolet radiation). This was confirmed using two independent methods of LB induction (triacylglycerol synthase overexpression and nitric oxide exposure). Analysis of *Mtb* binding to macrophages also showed that an increased proportion of LB-positive cells did not result in increased bacterial binding. However, the *Mtb* multi-stimulus cultures showed increased macrophage binding as compared to control cultures, although this remains to be confirmed.

It was concluded that *Mtb* gene expression in sputum occurs secondary to multiple stimuli, and that LBs may represent an epiphenomenon of the *Mtb* sputum phenotype as far as transmission is concerned. If in fact *Mtb* in sputum are transmission-adapted, other factors must contribute to this phenotype.

Acknowledgements

I am extremely grateful for the support I have received from my family, colleagues and friends during these past years. First and foremost, I would like to thank my supervisor, Professor Mike Barer, for his support, advice, encouragement since I started working in the lab as an undergraduate in 2003. I am extremely thankful for his belief in my research in both allowing me to undertake this PhD and for also assisting in the funding of my degree.

I would like to thank the International Journal of Experimental Pathology for funding my PhD and to Professor Ian Lauder for putting my name forward for this bursary. I would also like to thank the University of Leicester for their assistance in funding through their award of the International Student Scholarship.

I could not have done this without the help of a number of people throughout the course of my lab work and writing up. I'd like to thank Dr. Natalie Garton for her help and knowledge throughout my PhD, Andrew Bell for his valuable assistance in microscopy, Cat 3 lab work, fluorescence/lipid body analysis and proofreading and Dr. Kelvin Lau for all his gene expression advice and scripting help. I also appreciate the proofreading and friendship of Adrian and Jessica Scott and Jessica Loraine.

Many thanks go to Dr. Galina Mukamolova for her general advice, Dr. Bernie Burke for his help and guidance with regards to macrophages and cell culture and to Helen Pearson, Hannah Brewin and Dr. Aras Kadioglu for their cell culture assistance and knowledge. And thank you to Sheila, Pam and Elizabeth for keeping the department running.

I would like to thank all my friends and colleagues in the department during my time here, including Emma Beards, Adam Berg, Eddy Cheah, Abdul Binjomah, Nino Iakobachvili, Helen Smith and the rest of Lab 136. Thank you to Ben Attwood, James Alix, Clare Shannon, Mark Taylor, Inga Kelpanides, Kevin Thomas, Mike Payne, Nick Roberts and all my other friends for their friendship and encouragement throughout my PhD. You all know who you are.

Finally, and most importantly, I'd like to thank Mom, Dad and Su-Lin for all your love and support my whole life.

List of Common Abbreviations

ADC	Albumin Dextrose Catalase
AFB	Acid-Fast Bacilli
Ag85	Antigen 85
AIDS	Acquired Immune Deficiency Syndrome
ATP	Adenosine Triphosphate
BCG	Bacille Calmette-Guérin
CD	Cluster of Differentiation (Cell Surface Protein)
cDNA	Complementary Deoxyribonucleic Acid
CFU	Colony Forming Unit
CoA	Coenzyme A
COPD	Chronic Obstructive Pulmonary Disease
CoV	Coefficient of Variation
CR	Complement Receptor Type (1-4)
Ct	Cycle Threshold
DAG	Diacylglycerol
DGAT	Diacylglycerol Acyltransferase
DNA	Deoxyribonucleic Acid
DosR	Dormany Survival Regulator
DOTs	Directly Observed Therapy – Short Course
dsDNA	Double-stranded DNA
ESAT	Early Secretory Antigenic Target
FAD	Flavin adenine dinucleotide
FAS	Fatty Acid Synthetase
FAS-1	Type 1 Fatty Acid Synthetase
FAS-2	Type 2 Fatty Acid Synthetase
FCS	Foetal Calf Serum
gDNA	Genomic Deoxyribonucleic Acid
GDP	Genome Directed Primer
H_2O_2	Hydrogen Peroxide
HIV	Human Immunodeficiency Virus
ICL	Isocitrate lyase
IFN-γ	Interferon-Gamma
IgG	Immunoglobulin G
iNOS	Inducible Nitric Oxide Synthase
kDa	Kilodalton
LB(s)	Lipid Body (Bodies)
LCFA	Long Chain Fatty Acid
M	Molar
mAGP	Mycolyl Arabinogalactan-Peptidoglycan complex
MDR-TB	Multi-drug Resistant tuberculosis
µg	Microgram
µl	Microlitre
mg	Milligram
MIC	Minimum Inhibitory Concentration
ml	Millilitre

MPN	Most Probable Number
MOI	Multiplicity of Infection
Mtb	*Mycobactrium tuberculosis*
NAD	Nicotinamide Adenine Dinucleotide
NK Cells	Natural Killer Cells
nm	Nanometre
noRT	'No Reverse Transcription'
NRP	Non-replicating Persistence
O_2	Oxygen
O_2-	Superoxide
OADC	Oleic Acid Albumin Dextrose Catalase
OD	Optical Density
PA	Phosphatidic Acid
PCR	Polymerase Chain Reaction
PBS	Phosphate-Buffered Saline
PDIM	Phthiocerol Dimycocerosate
PHA	Polyhydroxyalkanotes
PHB	Poly-β-hydroxybutyrate
PIM	Phosphatidylinositol Mannoside
PSR	Phosphate Starvation Response
qPCR	Quantitative Polymerase Chain Reaction
qRT-PCR	Quantitative Reverse Transcription Polymerase Chain Reaction
RNA	Ribonucleic Acid
Rpf(s)	Resuscitation Promoting Factor(s)
rRNA	Ribosomal Ribonucleic Acid
RNI	Reactive Nitrogen Intermediates
ROS	Reactive Oxygen Species
RPMI	Roswell Park Memorial Institute
RT-PCR	Reverse Transcription Polymerase Chain Reaction
SD	Standard Deviation
SDS	Sodium Docedyl Sulphate
SOD	Superoxide Dismutase
Sp-A	Surfactant Protein A
TA	Toxin-Antitoxin
TAG	Triacylglycerol
TDM	Trehalose 6′-Dimycolate
Tgs	Triacylglycerol Synthase
TMM	Trehalose 6′-Monomycolate
TNF	Tumour Necrosis Factor
UV	Ultraviolet
WE	Wax Ester
WHO	World Health Organization
WS	Wax Synthase
XDR-TB	Extremely Drug Resistant Tuberculosis

List of Figures

List of Tables

List of Equations

Table of Contents

Chapter 1

Introduction

1.1 General Introduction

Tuberculosis continues to be a major global health issue and is the leading cause of death from a curable infectious disease (Dye et al., 2009). Clinical disease is caused by infection with the bacterium *Mycobacterium tuberculosis* (*Mtb*) from the genus *Mycobacterium*, a group of bacteria characterised by their high lipid content (Wayne and Kubica, 1986). The cell envelope contains a number of unusual lipids that underpin the characteristic 'acid-fast' staining profile of mycobacteria and highlight the requirement of mycobacteria for long chain fatty acids (LCFA), required for synthesis of these cell wall lipids (Takayama et al., 2005).

Recent work has demonstrated that *Mtb* accumulates lipid, in the form of triacylglycerol, as intracellular droplets, or lipid bodies (LBs), within the cell cytoplasm (Garton et al., 2002). The production of LBs indicates the adoption of an altered cell physiology by *Mtb*. This may be significant in the host environment, where it has been suggested that *Mtb* uses fatty acids as opposed to carbohydrates *in vivo* (Honer zu Bentrup and Russell, 2001). These fatty acids may be stored in and utilised from triacylglycerol stored as LBs. It has since been proposed that the formation and utilisation of LBs is a dynamic process linked with the metabolic activity of *Mtb*. LB production and utilisation may occur as the cells enter and exit a state of non-replicating persistence, respectively (Sherratt, 2008).

Analysis of *Mtb* in expectorated sputum revealed a unique phenotype, not previously seen in *in vitro* studies (Garton et al., 2008). Although conventionally thought to contain aerobically replicating bacilli, *Mtb* in sputum were shown to express transcriptional signatures consistent with slowly replicating or non-replicating mycobacteria and the utilisation of lipids. The switch to a dormant-like state by *Mtb* leads to a long incubation period of the pathogen and is a feature thought to be expressed clinically as a latent, asymptomatic stage of disease (Ahmad, 2011). While in this state of dormancy or non-replicating persistence, *Mtb* bacilli display a number of traits beneficial to survival, including phenotypic tolerance to the cidal action of antibiotics and resistance to multiple forms of stresses. The stimuli underlying this sputum phenotype and the contribution of LBs to this proposed state of transmission adaptation is currently unknown and is focus of this thesis.

As the issues introduced here are relevant to the studies relating to *Mtb,* LBs, and sputum gene expression and transmission, these are reviewed in more detail in the following sections.

1.2 Mycobacteria and *Mycobacterium tuberculosis*

Members of the genus *Mycobacterium* are rod-shaped bacilli, ranging from 0.2-0.6 × 1.0-10μm in size and conventionally are characterised as aerobic, non-motile and slow-growing. Mycobacteria are close relatives of *Corynebacterium* and *Nocardia* and in common with them, possess deoxyribonucleic acid (DNA) with high levels of guanine and cytosine (G+C) (Wayne and Kubica, 1986). These members are further classified into the genus based upon their 'acid-fastness' and cell wall containing mycolic acids with chains 60-90 carbon atoms in length.

'Acid-fastness' refers to the ability of mycobacteria to resist decolouristion by acidified alcohol after being stained by a cationic dye (e.g. carbol fuchsin) (Wayne and Kubica, 1986). With this characteristic, mycobacteria are often referred to as 'acid-fast bacilli' (AFB). This property is associated with a high lipid content in the cell wall, of which mycolic acids and waxes form major constituents (Goodfellow and Minnikin, 1980). Acid-fast staining techniques are often used to identify mycobacteria in clinical specimens, and although other bacterial species share this staining characteristic, they stain less intensely than mycobacteria due to shorter mycolic acids.

The genus *Mycobacterium* contains over 120 species, naturally divided into fast- and slow-growing species. In practice, the 'slow-growers' are species of mycobacteria that require greater than 7 days incubation at optimal temperatures to produce colonies from highly diluted inocula. These slow-growing species include some of the oldest known pathogens, including *Mtb* and *Mycobacterium leprae*, the causative agents of tuberculosis and leprosy, respectively (Wayne and Kubica, 1986). The 'fast-growers' can form colonies in fewer than 7 days and are generally non-pathogenic, although they may establish infection in immunocompromised hosts (Howard and Byrd, 2000, Wayne and Kubica, 1986).

1.2.1 *Mycobacterium tuberculosis* complex

Tuberculosis is caused by members of the *Mtb* complex which consists of closely related organisms of the genus *Mycobacterium*, that share 99.9% similarity at the nucleotide level for some loci and identical 16S sequences (Brosch et al., 2002, Cole, 2002, Cousins et al., 2003, Sreevatsan et al., 1997). The genome of members of the *Mtb* complex differs from other bacteria in its high G+C content and the dedication of a large proportion of its genes to the coding of lipogenesis and lipolysis related enzymes (Cole et al., 1998).

The complex was initially thought to comprise six members, but two further members have been recently identified (Aranaz et al., 2003, Cousins et al., 2003). In addition to *Mtb*, which causes the vast majority of human tuberculosis cases, the complex includes *Mycobacterium africanum,* which is a major cause of human tuberculosis in Sub-Saharan Africa. *Mycobacterium bovis* is a member that infects a wide variety of animals and humans. The attenuated variant of *M. bovis* is known as Bacillus Calmette-Guérin (BCG); isolated after 230 serial passages of *M. bovis,* it has been used extensively as a live vaccine against tuberculosis (Cole, 2002). *Mycobacterium canetti* is member that can cause human tuberculosis, although it is rarely encountered (Cole, 2002). The three remaining members of the complex primarily infect non-human hosts. *Mycobacterium microti* is the tuberculosis agent in voles, and is avirulent in humans (Cole, 2002, Cosma et al., 2003). The two recently discovered members are *Mycobacterium caprae*, a recently reclassified strain infecting goats, and *Mycobacterium pinnipedii*, the causative agent of tuberculosis in seals (Aranaz et al., 2003, Cousins et al., 2003).

Bacteriologically, members of the *Mtb* complex are straight or slightly curved rods occurring singly or in threads and ranging from 0.3-0.6 × 1-4μm in size. Bacilli are strongly acid-fast, and under optimal growth conditions (37°C, pH 6.4-7.0), have a generation time of 14-15h (Wayne and Kubica, 1986). Growth tends to occur in a serpentine, cord-like mass, with bacteria showing a parallel orientation. On solid media, *Mtb* forms colonies that are rough, raised and thick, with a nodular or wrinkled surface and irregular margin. Colonies may become pigmented to an off-white, buff or yellow colour (Wayne and Kubica, 1986).

1.2.2 The *M. tuberculosis* Cell Wall

The unique cell wall structure, containing a large proportion of complex lipids, is arguably the most defining characteristic of *Mtb*. Mycolic acids, high molecular weight α-alkyl,β-hydroxy branched fatty acids 60-90 carbon atoms in length, represent key lipid components of the mycobacterial cell wall (Brennan, 2003). The *Mtb* cell envelope can be described as the cell membrane, and the inner section and outer section of the cell wall (Brennan, 2003). The inner section comprises a layer of peptidoglycan covalently attached to arabinogalactan, a polysaccharide, which attaches to mycolic acids. This is also known as the mycolyl arabinogalactan-peptidoglycan (mAGP) complex and forms the cell wall core. The outer section comprises of the outer leaflet, consisting mainly of polysaccharides, proteins, free fatty acids and lipids, as well as lipoarabinomannan (Brennan, 2003, Daffe and Draper, 1998). This organisation of the cell envelope is shown in Figure 1 and the components are discussed further below.

The plasma membrane and peptidoglycan layer of mycobacteria possesses similarities to those in other bacteria, being composed of a phospholipid bilayer (Minnikin et al., 2002). However, the *Mtb* cell membrane appears to be asymmetrical, with the outer membrane layer being thicker than the inner one. This asymmetry is thought to be due to the presence of phosphatidylinositol mannosides (PIMs), mannoside derivatives of phosphatidic acid that anchor lipoarabinomannan and lipomannan to the cell membrane (Minnikin et al., 2002).

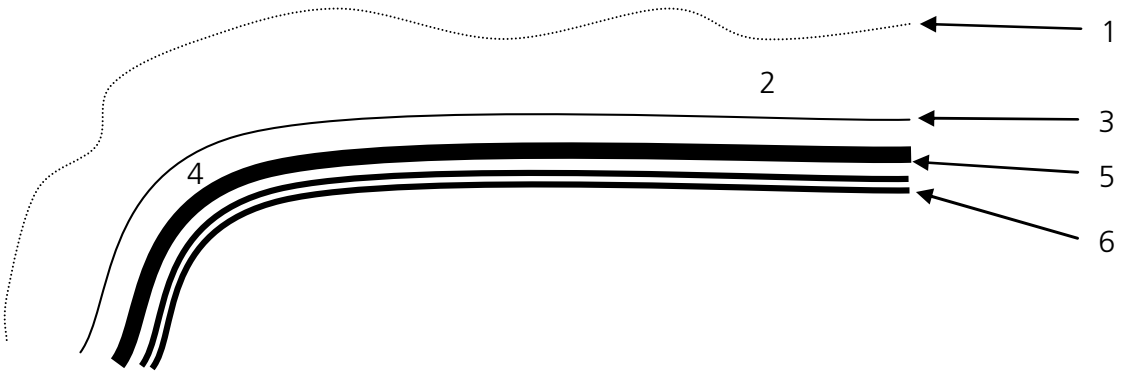


Figure 1: Organisation of *M. tuberculosis* Cell Envelope

Layers are not to scale. (1) Position of the superficial lipids of the outer leaflet; (2) Outer leaflet, consisting mainly of polysaccharide and protein; (3) Position of buried lipids; (4) Mycolic acid layer; (5) Peptidoglycan plus arabinogalactan; (6) Cell membrane (outer layer is thicker than the inner layer) (Adapted from Daffe and Draper, 1998).

Adjacent to the cell membrane is the mycobacterial peptidoglycan layer, which is similar to that found in other bacteria (Brennan and Nikaido, 1995). Bacterial peptidoglycan differs in that the muramic acid residue is N-acetylated to glycolic acid, while in mycobacteria, this residue is N-glycosylated (Brennan and Nikaido, 1995). Arabinogalactan, a polysaccharide comprising arabinose and galactan, is connected to the mycobacterial peptidoglycan via a phosphodiester link. In turn, mycolic acids are esterified to arabinogalactan through terminal arabinose residues to form the mAGP complex (Brennan and Nikaido, 1995).

Mycolic acids are composed of two structures: a meromycolate moiety (main chain) and an α-branch. The α-branch is consistent between all mycolic acids, while the main chain may differ and be used to differentiate between classes of mycolic acid, including α-, methoxy- and keto-mycolic acids (Takayama et al., 2005). The α-mycolic acids have two cyclopropane

rings but do not contain an oxygen functional group other than the β-hydroxy group. Conversely, the keto- and methoxy-mycolic acids are oxygenated and have oxygen-containing polar modifications in the distal part of the main chain. Non-polar modifications are located at both the distal and proximal end of the meromycolate chain (Takayama et al., 2005).

Within the cell wall, it has been proposed that the mycolic acids are attached perpendicular to the cell membrane in a juxtaposed arrangement (Minnikin et al., 2002). The packing arrangement is modulated by the functional groups of the mycolic acids and imparts a fluidity gradient to the cell wall, with the external cell wall section being more fluid than the tightly packed internal section (Brennan and Nikaido, 1995). The mycolic acid layer is covered by an outer leaflet structure comprised of extractable lipid, polysaccharide and protein (Chatterjee, 1997).

The cell wall also possesses waxes, which are esters of long chain alcohols and fatty acids; in mycobacteria, these consist of dimycocerosates of phthiocerols, or PDIM. Phthiocerols are long chain methoxyglycols with mycocerosic acids (multi-methyl branched fatty acids) (Minnikin et al., 2002). PDIMs interact with mycolic acids, form part of the outer section of the mycobacterial cell envelope and are thought to play a role in the regulation of immune response in infected tissue (Minnikin et al., 2002, Russell, 2007). Also dispersed amongst the mycolic acids are the lipopolysaccharides, lipomannan and lipoarabinomannan, which span the whole cell wall depth (Minnikin et al., 2002). Both LM and LAM are considered to be part of the free lipid component of the cell envelope, as they are extractable following exposure to various solvents. Lipoarabinomannan specifically, is an *Mtb* virulence factor and has immunomodulatory effects, including the induction of signalling events important in the host response to *Mtb* (Brennan and Nikaido, 1995, Schlesinger et al., 1994).

Finally, a range of extractable glycolipids are present in the outer section of the mycobacterial cell envelope. Trehalose, a non-reducing disaccharide comprised of two glucose units linked through an α,α,-1,1-glycosidic linkage, form the basic unit of these glycolipids (Elbein et al., 2003). Members of the *Mtb* complex produce characteristic glycolipids; for example, in W-Beijing *Mtb* strains, the production of a unique phenolic glycolipid is thought to underpin the increased virulence of these strains (Russell, 2007). Cord factor, also known as trehalose dimycolate (TDM), in particular, has been studied heavily. TDM is a glycolipid that causes cells to aggregate in cord-like formations *in vitro*, and provide a thick layer that protects *Mtb* from both chemicals and the host immune system (Takayama et al., 2005). Furthermore, TDM is also antigenic and plays a role in the host immune response to *Mtb* infection (Hunter et al., 2006).

The mycobacterial cell wall underpins a number of *Mtb* characteristics. As introduced, certain cell wall components play a role in the host immune response to *Mtb* infection. Furthermore, the presence of mycolic acids underlies the acid fast staining characteristic, while the low permeability of the cell wall may contribute to mycobacterial resistance to antibiotics (Goodfellow and Minnikin, 1980, Jarlier and Nikaido, 1994).

1.3 Tuberculosis

1.3.1 Clinical Disease

The majority of tuberculosis cases present clinically as pulmonary disease, resulting from primary progression from an initial infection or reactivation from latent infection (see 1.4.1). However, tuberculosis can display extremely varied disease expression and any organ of the body can be involved, as extra-pulmonary disease (Hopewell and Jasmer, 2005).

A persistent cough is the most common symptom of pulmonary tuberculosis; it may not be productive early on, but as the disease progresses and inflammation and necrosis occur, mucoid or purulent sputum is often produced (Hopewell and Jasmer, 2005). Haemoptysis (expectoration of blood or blood-stained sputum) is often a presenting symptom due to erosion of bronchial circulation (Hopewell and Jasmer, 2005, Quast and Browning, 2006). Further, patients may experience dyspnoea (painful breathing), pleuritic chest pain, anorexia and weight loss, night sweats, general malaise, fever and lassitude. These constitutional symptoms are not specific for tuberculosis, and many patients may be altogether asymptomatic, even in extensive disease (Cook and Zumla, 2002, Frieden et al., 2003, Hopewell and Jasmer, 2005, Quast and Browning, 2006).

Extra-pulmonary tuberculosis results from the dissemination of the bacilli via the bloodstream during the early course of infection (Brandli, 1998). Disseminated or miliary tuberculosis is more common in human immunodeficienty virus (HIV)-seropositive individuals and reflects an inability of the immune system to contain the infection. (Frieden et al., 2003, Hopewell and Jasmer, 2005). This failure following primary infection or re-infection can lead to proliferation and dissemination of the bacteria throughout the body; this usually involves the lungs, and multi-organ involvement is common (Hopewell and Jasmer, 2005). Due to inadequate immune response, a granulomatous reaction may not be seen, but instead, a diffuse pattern of lymphocytic infiltration and oedema are observed (Hopewell and Jasmer, 2005). As a consequence of the multi-system involvement, patients suffer from non-specific symptoms, including fever, weight loss, anorexia, weakness and wasting (Hopewell and

Jasmer, 2005). Mortality from miliary tuberculosis is high despite chemotherapy, and may be related to delays in diagnosis (Frieden et al., 2003). Among non-HIV infected individuals, lymphatic tuberculosis is the most common presentation of extrapulmonary disease, particularly in women and young children; it may present as a painless swelling of one or more lymph nodes (Frieden et al., 2003, Hopewell and Jasmer, 2005).

Other clinical tuberculosis manifestations include involvement of the central nervous system, the most serious consequence (Frieden et al., 2003). Involvement can include meningeal inflammation as well as space-occupying lesions (tuberculomas) of the brain. The symptoms can present clinically as chronic meningitis, with headache, fever and changed mental status, and are due to the presence of bacteria as well as the inflammatory host response (Frieden et al., 2003). Neurological manifestations include cranial-nerve palsies and motor, sensory and cerebellar defects, depending on where the tuberculomas are located. Meningeal tuberculosis is more common in children under the age of 5 and HIV-infected individuals, and requires prompt identification and treatment (Brandli, 1998, Frieden et al., 2003, Hopewell and Jasmer, 2005). Tuberculosis may also involve any bone or joint, although Pott's disease, i.e. spinal involvement, is the most common bone manifestation (Frieden et al., 2003). Genitourinary tuberculosis is uncommon and can involve the renal and male and female genital tracts; symptoms are difficult to distinguish from other genitourinary tract infections (Hopewell and Jasmer, 2005).

Diagnosis of tuberculosis is usually made on the basis of clinical examination, radiological features and bacteriologic examination (Cook and Zumla, 2002, Hopewell and Jasmer, 2005, Zumla et al., 1999). Physical findings in clinical examination generally are not helpful in defining the disease; chest signs are often absent or limited to fine or apical cracks. Advanced cases may lead to areas of dullness on percussion or localised wheezing (Hopewell and Jasmer, 2005). Pulmonary tuberculosis nearly always causes abnormalities on chest x-rays; primary disease is often seen as a middle or lower lung infiltrate, associated with ipsilateral hilar lymphadenopathy, while re-activation tuberculosis is usually associated with abnormalities in the upper lung lobes (Brandli, 1998, Hopewell and Jasmer, 2005). Cavitation, or destruction of the lung tissue, may be seen; subsequent healing may result in the development of a scar, loss of lung tissue and calcification (Brandli, 1998, Hopewell and Jasmer, 2005).

Bacteriologic evaluation can confirm the presence of active disease through the detection of AFB in sputum and the isolation of bacteria through culturing (Frieden et al., 2003, Hopewell and Jasmer, 2005). Culturing is performed on both solid and liquid media; the former allows examination of colony morphology and the identification of mixed cultures, while the latter

enables more rapid diagnosis (Frieden et al., 2003). In addition to these examinations, the rapid detection of *Mtb* in sputum can be achieved through PCR assays of DNA, with the sensitivity and specificity being as high as 95% and 98%, respectively for this technique (Frieden et al., 2003).

1.3.2 Treatment of Tuberculosis

Tuberculosis chemotherapy began in 1944 with the discovery of streptomycin by Schatz, Bugie and Waksman; however, mono-therapy using streptomycin resulted in the emergence of drug resistance and treatment failure (Mitchison, 2005). The first modern anti-tuberculosis drug to be discovered was isoniazid in 1952. Following the use of isoniazid monotherapy, the resultant emergence of isoniazid resistance led to its abandonment and future drugs were selected based upon their ability to prevent isoniazid resistance when used as part of a combination chemotherapy regimen. Pyrazinamide was introduced in 1961, followed by ethambutol in 1961 and rifampicin in 1967. These four drugs, along with streptomycin, are today considered the first-line anti-tuberculous drugs (Mitchison, 2005).

Current tuberculosis treatment involves a two-phase, multi-drug regimen. The treatment consists of an initial, intensive phase of four drugs, isoniazid, rifampicin, pyrazinamide and ethambutol, taken daily for two months, followed by a continuation phase consisting of rifampicin and isoniazid for a further four months, 3 times per week. This regimen has been shown to have good efficacy and low toxicity, and has made tuberculosis into a curable disease with reduced incidence in industrialised countries (du Toit et al., 2006, Zhang and Amzel, 2002).

However, despite the availability of an effective cure, tuberculosis remains a major global health problem. The complex and lengthy treatment regimens are expensive, particularly for developing nations, where 95% of tuberculosis cases are focused (Zhang and Amzel, 2002). Further, due to the multi-drug approach and length of chemotherapy, patients are vulnerable to side effects, poor compliance and slow improvement, resulting in poor cure rates, relapse and development of multi-drug resistant tuberculosis (MDR-TB). The problem of patient compliance has been combatted by the World Health Organization (WHO) introduced Directly Observed Therapy – Short Course (DOTs), which requires each dose of chemotherapy administered to a patient to be directly supervised by a health or social worker (McKinney, 2000). DOTs is extremely effective, with a 95% cure rate; however, it is difficult to implement and expensive, and as a result, only 1 in 5 people are treated in this manner

(McKinney, 2000, Zhang and Amzel, 2002). Steps must be taken to both increase the cases of DOTs and to develop drugs to reduce treatment time and compliance (McKinney, 2000).

The length of drug treatment is not a result of inefficacy of drug action (although this is increased due to drug resistance), but down to the physiology of bacteria *in vivo* (McKinney, 2000). The penetration of drugs into lesions and cavities in lungs has been shown to occur at levels beyond the minimum inhibitory concentration (MIC), so it seems likely that delayed drug activity *in vivo* is due to the presence of physiologically heterogeneous bacterial populations (McKinney, 2000). Mitchison (1979) proposed four distinct populations of bacteria: rapidly growing bacilli, slow growing or intermittently growing bacilli, non-growing/dormant bacilli and bacilli in an acidic environment. Rapidly growing bacteria are killed efficiently by isoniazid within the first two weeks of treatment, although the drug is ineffective against slow and intermittent growers (McKinney, 2000, Zhang and Amzel, 2002). The remainder of the treatment period is required to kill the residual population of bacteria; rifampicin is thought to eliminate intermittently growing bacteria with spurts of metabolism, while pyrazinamide kills bacteria in acidic environments and some dormant bacilli (McKinney, 2000, Mitchison, 2004). Thus, to reduce the treatment time, the physiology of *Mtb* must be fully understood, and effort must be made to eliminate this dormant, slow-growing and persistent population of cells.

1.3.3 Epidemiology

Mtb is a highly effective pathogen and today, tuberculosis remains the leading cause of death from a curable infectious disease (Dye et al., 2008). In recent data released by the WHO, there were approximately 8.8 million new tuberculosis cases and 1.4 million deaths in 2010, with a global incidence of 128 cases per 100,000 population worldwide (World Health Organization, 2011). In addition to active cases, it has been estimated that as many as one-third of the world's population is latently infected with *Mtb,* with up to 23% of these individuals expected to reactivate in their lifetime (Dye, 2006, Honer zu Bentrup and Russell, 2001).

The trend over the past decade has been an drop in the incidence of tuberculosis across 5 of the 6 country groups as classified by the WHO (Dye et al., 2009). A report by Dye et al. (2009) concluded that the reasons for the disease trends were strongly associated with biological, social and economic determinants of each region. National tuberculosis control programs have been vital in the diagnosis, treatment and subsequent reduction in transmission, although at this point in time, they have not had a major, detectable impact on

the incidence of tuberculosis (Dye et al., 2009). The exception to this decrease in disease incidence is Sub-Saharan Africa, where there was a rise in tuberculosis incidence, which may be partly attributed to the high incidence of HIV in this geographical region (Corbett et al., 2006, Dye et al., 2009). While *Mtb* infected individuals normally have a 10-20% lifetime risk of developing clinical disease, in HIV co-infected individuals, the annual risk alone can be greater than 10% (Corbett et al., 2006).

1.3.3.1 Multi-drug resistant tuberculosis

MDR-TB is an increasing global problem, with an estimated 650,000 cases worldwide in 2010 and threatens current tuberculosis control programmes (Caminero, 2010, World Health Organization, 2011). It is defined as resistance to both isoniazid and rifampicin, with or without resistance to other anti-tuberculosis drugs (Ormerod, 2005). Globally, MDR-TB is not a problem of the same magnitude as drug-susceptible tuberculosis, but MDR-TB has high prevalence in hot spots, in particular China, Russia and countries of the former Soviet Union (Aziz et al., 2006, Espinal, 2003, Caminero, 2010). While there is a molecular aspect to resistance, MDR-TB is often a man-made phenomenon and the geographical MDR-TB hot spots have a history of poor tuberculosis control and uncontrolled use of anti-tuberculosis drugs (Caminero, 2010).

MDR-TB is a particular problem because treatment is more complex, toxic, expensive and less successful than conventional treatment (Aziz et al., 2006). The loss of response to the frontline bactericidal drug, isoniazid, and the main sterilising drug, rifampicin, leads to patients being infectious for longer periods of time (Ormerod, 2005). The future of MDR-TB globally will depend on the efforts that countries are willing to make to control tuberculosis and will require prompt diagnosis, implementation of treatment for the full time course, prevention of transmission and careful introduction of second-line drugs (Espinal, 2003). However, the resistance to second-line drugs has been documented in extensively drug resistant tuberculosis (XDR-TB). XDR-TB is defined as resistance to isoniazid, rifampicin and at least three classes of second-line drugs (Gandhi et al., 2006). With the available drugs, XDR-TB patients are left with few treatment options, leaving a need for the development of new drugs (Gandhi et al., 2006).

1.4 Pathogenesis of Tuberculosis

The phagocytosis of *Mtb* bacilli by alveolar macrophages represents the first step in the pathogenesis of tuberculosis. The transmission of tuberculosis is reviewed below in Paragraph 1.5. The progression of the natural course of tuberculosis is shown in Figure 2. Following this phagocytosis, the infection is often arrested by the bactericidal activities of the activated macrophages, such as reactive nitrogen and oxygen intermediates, acidic pH and lysosomal enzymes (reviewed in Paragraph 1.4.3) (Parrish et al., 1998, Smith, 2003, Young and Duncan, 1995). However, *Mtb* possesses a number of survival mechanisms once inside the macrophage to avoid and resist killing by these bactericidal mechanisms (Smith, 2003). In a naïve host, the surviving bacteria undergo a period of rapid expansion in the absence of a specific, limiting immune response; infected macrophages then spread via the lymphatic system to the hilar lymph nodes (Honer zu Bentrup and Russell, 2001, Frieden et al., 2003). The macrophage therefore has a varying role during *Mtb* infection; it is the primary host cell during the defence against infection, while providing the key environment for bacterial replication and dissemination, as well as a possible niche for persistent infection (Cosma et al., 2003).

The state of disease which occurs during the course of tuberculosis infection is a product of both host and microbial factors (Hopewell and Jasmer, 2005, Russell, 2007). Active disease occurs when the immune system cannot adequately contain the replication and proliferation of bacilli and occurs most often in children under the age of 5, or in immunocompromised adults such as HIV-infected individuals (Frieden et al., 2003). In the majority of individuals, cell-mediated immunity develops 2-8 weeks following initial infection (Frieden et al., 2003). The growth of *Mtb* within the alveolar macrophages leads to a localised proinflammatory response through the release of chemokines and Tumour Necrosis Factor (TNF), recruiting mononuclear cells (i.e. monoctes, lymphocytes and neutrophils) from blood vessels to the site of infection (Cosma et al., 2003, Russell, 2007, Smith, 2003, Young and Duncan, 1995). The chemokines first recruit natural killer (NK) T cells to the site of infection, followed by $CD4^+$, $CD8^+$ and γδ T cells and B cells (Russell, 2007). Mycobacteria-specific $CD8^+$ T cells are cytotoxic for infected macrophages, directly killing intracellular bacilli via a granule-associated protein, granulysin (Tufariello et al., 2003). $CD4^+$ T cells on the other hand produce interferon-γ (IFN-γ), which activates macrophages, and other cytokines important in the control of infection, such as interleukin 2 and TNF alpha. The cells may also be necessary for the induction or function of $CD8^+$ T cell responses or the production of other important cytokines by macrophages or dendritic cells (Tufariello et al., 2003). The cytokines and chemokines released by both sets of T cells are controlled by IFN-γ and amplify cellular

recruitment (Russell, 2007). All these cells are the building blocks for the granuloma or 'tubercle', the classic lesion of tuberculosis (Russell, 2007).

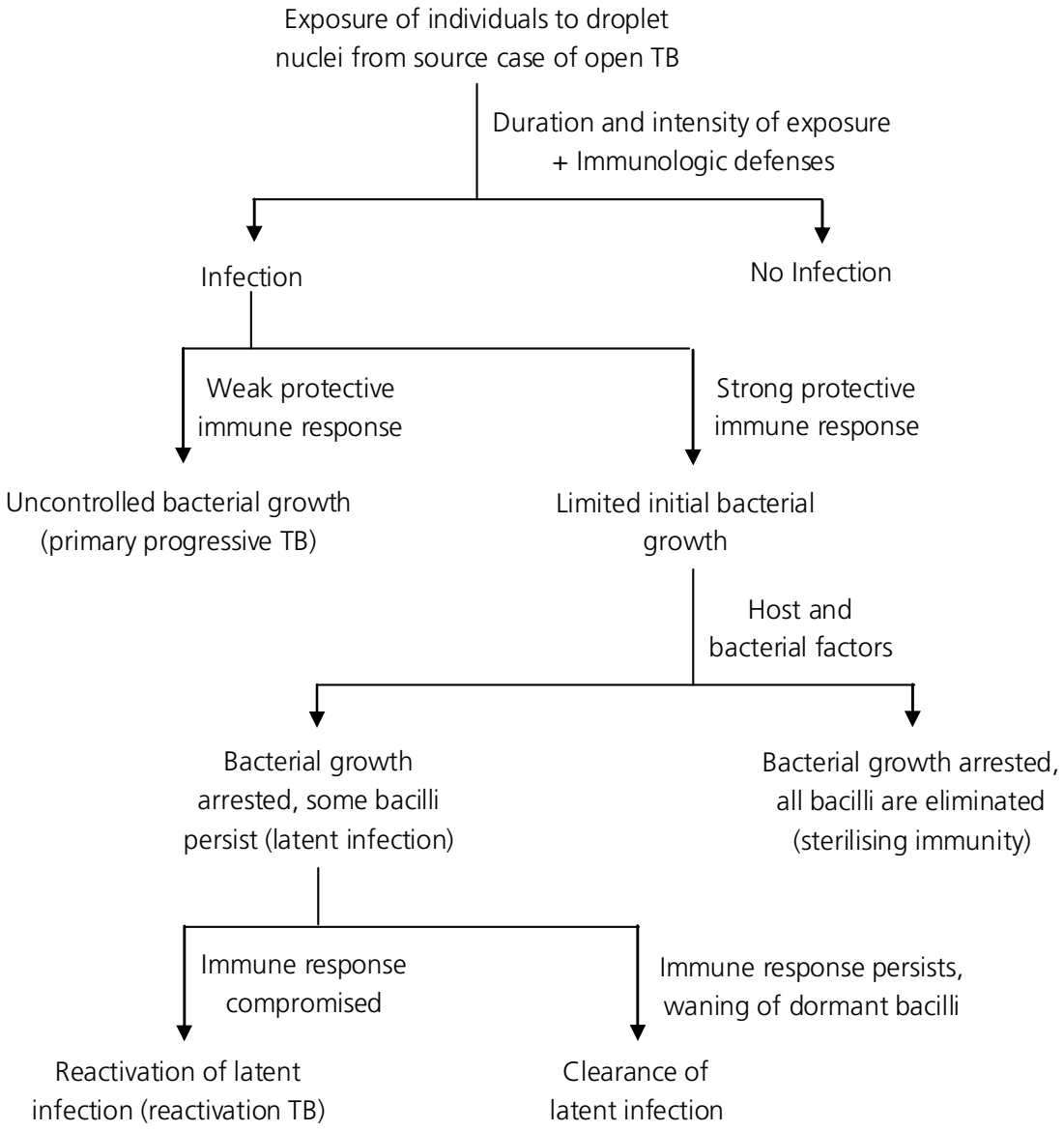


Figure 2: Progression of the Natural Course of Tuberculosis

The figure shows the natural course of events and outcomes in an immunocompetent patient following exposure to aerosolised droplet nuclei containing *Mtb* expectorated by a case of sputum smear-positive pulmonary tuberculosis (Figure adapted from Ahmad, 2011).

The granuloma forms around the foci of infection and this tissue response typifies a 'containment' phase of infection, by which the immune system limits the dissemination of infection (Russell, 2007). However, granuloma formation also benefits the pathogen, as T cells which can activate and subsequently destroy infected macrophages are restricted to the periphery of the structure (Honer zu Bentrup and Russell, 2001). The granuloma consists of a core of infected macrophages surrounded by foamy macrophages (highly lipid-laden macrophages) and other mononuclear phagocytes. A mantle of lymphocytes forms around these macrophages, in association with a fibrous layer of collagen and other extracellular matrix components, which delineates the structure from normal lung tissue (Honer zu Bentrup and Russell, 2001, Russell, 2007).

As the granuloma matures, it develops a fibrous sheath and the number of blood vessels penetrating it diminishes. The centre core consists of a cheese-like semi-solid structure that is low in oxygen and rich in lipids and proteins from dead cells (Honer zu Bentrup and Russell, 2001, Russell, 2007). As the centre of the granuloma loses is vascular appearance, it becomes necrotic. Mycobacteria are thought to be located within macrophages and necrotic tissue in the centre of granulomas, although several studies have found bacteria associated with macrophages in the peripheral leukocytic infiltrate outside the fibrotic capsule (Russell, 2007). The physiological state of bacilli within the granuloma is unknown and it is thought that persistent bacilli can survive in the granuloma for decades (Honer zu Bentrup and Russell, 2001). There is, however, a great deal of heterogeneity in granuloma structure and these can be in varying stages of evolution (Cosma et al., 2003).

During the 'containment' phase of infection, there is no overt sign of disease and the host cannot transmit the infection to others (Russell, 2007). The host displays a delayed type hypersensitivity response to mycobacterial antigens (Tufariello et al., 2003). The bacilli can survive within the host tissues for years and many patients will remain latently infected with *Mtb* for the rest of their lives with no clinical presentation of disease; it is unclear whether a host can ever completely eliminate a latent infection (Tufariello et al., 2003). Latent tuberculosis is further introduced in 1.4.1. In a latently infected individual, a change in immune status, often caused by old age, malnutrition, HIV co-infection or any condition that reduces the number, or impairs the function, of $CD4^+$ T cells, can lead to reactivation of disease. As numbers of $CD4^+$ T cells drop, the granuloma does not function as well, production of IFN-γ may decrease, and macrophages become less activated (Tufariello et al., 2003). The defects lead to a disorganised granuloma that is less able to control infection; the granuloma caseates, i.e. decays into a structure-less mass of cellular debris, ruptures and releases viable, infectious bacilli into the airways, resulting in a productive cough and transmission (Russell, 2001, Russell, 2007).

1.4.1 Latent Tuberculosis Infection

One of the reasons *Mtb* as a pathogen is successful is due to its ability to survive within the host for prolonged periods without being eradicated by the immune system. As previously discussed, *Mtb* infection is normally arrested by the immune system, and contained within the granuloma structure, without development of active disease. At this point, disease is considered latent, and clinically, a patient will have a positive tuberculin skin test (Boshoff and Barry, 2005). Latent disease may last for the lifetime of the patient, or the patient may develop reactivation disease. Once latently infected, a patient has a 2 to 23% chance of developing clinical disease in their lifetime (Parrish et al., 1998). Therefore, *Mtb* is able to reside in the host for decades, yet is able to maintain the ability to establish infection when suitable conditions arise.

The term latency only refers to the state of clinical disease; dormancy refers to the state of the bacillus, defined as 'a reversible state of low metabolic activity in a unit that maintains viability' (Kaprelyants and Kell, 1993). While little is still known about the physiology of *Mtb* during this stage of infection, it has been suggested that bacteria maintain a dormant or 'non-replicating, persistent' state during latency, brought about by the inhibitory growth conditions in the granuloma and macrophage (Boshoff and Barry, 2005, Parrish et al., 1998, Wayne and Sohaskey, 2001). An identifying feature of bacilli in sputum was the expression of slow growth signals, suggesting that these cells may have some relationship with *Mtb* dormant bacilli and disease latency. Furthermore, dormant cells to some extent show characteristics beneficial to survival, including increased tolerance to cell stresses such as the cidal action of antibiotics (Garton et al., 2008).

A variety of models have been used to study tuberculosis latency: *in vitro* through hypoxia (Wayne and Hayes, 1996), stationary phase (Voskuil, 2004) and nutrient depletion (Betts et al., 2002) or *in vivo* in the Cornell mouse model (Gomez and McKinney, 2004). In the Cornell model, mice are infected with a high dose of *Mtb,* treated with isoniazid and pyrazinamide; the mice enter a 'sterile state' with no culturable bacteria, but one-third reactivate spontaneously (Gomez and McKinney, 2004). However, while shedding light on disease latency, each *in vitro* model only examined a single host stress, and murine infection has been documented to differ from that in humans, e.g. limited granuloma formation.

Taking the results of all these models, Cardona (2009) proposed a dynamic model of latent tuberculosis infection. During infection, *Mtb* grows well inside phagosomes; however, as the macrophages become necrotic, some bacilli are released into the extracellular environment of the granuloma and stop replicating. This arrest may result from hypoxic and acid environments, as opposed to a mature immune response, in addition to bacterial enzymes

released from dead eukaryotic cells (Cardona, 2009). Hypoxia and low pH have both been shown cause such an arrest in *in vitro* studies (Gomes et al., 1999, Wayne and Sohaskey, 2001). Actively growing bacteria are killed as the immune response develops but non-replicating bacilli resist and survive (Cardona, 2009). Foamy macrophages also emerge from chronic inflammation, secondary to the phagocytosis of cellular debris rich in fatty acids and cholesterol (Caceres et al., 2009). These foamy macrophages also phagocytose extracellular, non-replicating bacilli; these bacilli do no replicate within these activated membranes, but are not killed because of their persistent state. The infected foamy macrophages may go on to drain towards the bronchial tree and begin endogenously re-infect the lung in a new location (Cardona, 2009). Higher oxygen tension in these new locations may allow bacilli to reactivate from this persistent state.

This proposed model describes latent infection and dormancy as a dynamic, active process and must still be confirmed experimentally. The contribution of other host stresses on mycobacterial persistence remains to be defined; for instance, nitric oxide has been shown to cause a reversible growth arrest in *Mtb* (Voskuil et al., 2003).

1.4.1.1 Location of Bacilli during Latent Infection

A number of studies have been undertaken to determine the location of *Mtb* during latent infection to define the conditions encountered by bacilli. Direct examination of infected lung tissue revealed lesion pathology heterogeneity within a single specimen (Boshoff and Barry, 2005). Open tuberculous cavities in patients with AFB smear-positive sputum yielded positive microscopy and cultures from bacteria isolated from these lesions. In contrast, closed lesions from smear-negative tuberculosis patients contained bacilli detectable by microscopy, but only 0.1% grew in culture; the studies were unable to determine if the bacilli were not viable or culturable (Salkin and Wayne, 1956). Other studies demonstrated that homogenized tissue from fibrous, caseous lesions from the lung apices were capable of causing infection in guinea pigs, while homogenates from encapsulated or calcified lesions rarely did (Reviewed by Gomez and McKinney, 2004). Furthermore, a proportion of homogenates of superficially normal lung tissue have been shown in studies to be capable of infecting guinea pigs (Gomez and McKinney, 2004).

Past studies have suggested that *Mtb* may lose its characteristic acid-fastness during infection. Khomenko (1987) showed that non-acid-fast bacilli isolated from tuberculosis lesions regained this staining property following *in vitro* liquid medium culturing. A more recent study compared lung tissue from patients with active tuberculosis and latent, non-progressive

tuberculosis (Ulrichs et al., 2005). The authors showed that a number of samples were negative for AFB microscopically, but positive on culture; this contrasted with the caseous lesions that had access to the bronchial system, which contained high numbers of AFB. Further immunohistological straining demonstrated that for both active and latent tuberculosis, bacilli were found in the cavity wall of the granuloma (latent disease) or cavity (active disease), peripheral infiltrations and within small granulomas (Ulrichs et al., 2005). In particular, *Mtb* was detected in antigen presenting cells at the periphery or in the necrotic centres of lesions examined. Russell et al. (2007) also demonstrated that *Mtb* is located within macrophages of the granuloma peripheral leukocyte infiltrate. Taken together, these studies suggest that *Mtb* is contained within the granuloma during latent infection. The fact that bacilli resides within macrophages suggests that the immune system is crucial in containing disease activity.

Current evidence describing the location of dormant *Mtb* in human infection is limited and conflicting as even healthy lung tissue can harbour *Mtb* bacilli. However, a common theme in the studies is the appearance of bacilli residing in the necrotic centres and macrophages of the granuloma.

1.4.2 DosR regulon

The DosR regulon may play a role in mycobacterial persistence and latent disease. Oxygen tension is a factor frequently associated with the establishment and maintenance of this latency (Wayne and Sohaskey, 2001). Based upon this relationship, Sherman et al. (2001) identified a set of 48 genes that were upregulated in response to short term oxygen depletion. The study exposed culture flasks to less than 1% oxygen saturation for 2 hours. Within this gene set, a two-component response regulator pair Rv3133c/Rv3132c was identified. Deletion of the regulator, Rv3133c, led to a reduced expression of the hypoxic response gene, α-crystallin, while deletion of the sensor kinase, Rv3132c led to no specific phenotype being observed (Sherman et al., 2001). Further work showed that the two-component regulatory system actually consists of two sensor histidine kinases, DosS (Rv3132c) and DosT (Rv2027c), in addition to DosR, the cognate response regulator (Roberts et al., 2004). DosS functions as a redox sensor while DosT acts as a hypoxia sensor.

Extending this work, Park et al. (2003a) demonstrated that disruption of Rv3133c, termed DosR for Dormancy Survival Regulator, led to a dramatically muted transcriptional response to hypoxia, relative to expression under normoxic conditions. Furthermore, 26 of the 27 genes most powerfully upregulated in response to hypoxia required the presence of DosR.

The authors also showed that DosR bound to copies of a motif upstream of the α-crystallin gene; this motif was located upstream of the majority of the hypoxically induced genes. These results established the role of DosR as the primary mediator of the *Mtb* hypoxic response (Park et al., 2003a).

These studies have demonstrated the subset of genes induced in response to this one condition, hypoxia. These 48 genes have been termed the 'dormancy regulon' because they are upregulated in *Mtb* in a dormancy inducing condition; however, it does not represent a definitive set of genes for all of these conditions. In two independent studies examining non-replicating persistence in the Wayne model, 42 and 36 genes out of the 48 DosR genes were induced in response to non-replicating persistence (Muttucumaru et al., 2004, Voskuil, 2004). Furthermore, 35 dosR genes were induced in a steady state chemostat culture exposed to hypoxia (Bacon et al., 2004), 27 were induced in stationary phase culture (Voskuil, 2004), and 31 were induced in a static culture allowed to settle for 30 minutes (Kendall et al., 2004). Hypoxia is considered to be the trigger for induction in these latter two studies.

In addition to hypoxia, the dosR regulon was expressed in *Mtb* in response to a non-toxic level of nitric oxide, which reversibly inhibited respiration and growth (Voskuil et al., 2003). The induction of the regulon was confirmed in a study treating *Mtb* with two different nitric oxide releasing agents (Ohno et al., 2003). Kendall et al. (2004) also showed induction of the dosR regulon on exposure to s-nitroglutathione (GSNO, a nitric oxide releasing agent), ethanol and hydrogen peroxide, but not heat or cold shock. This regulon induction makes the interpretation of the specific contribution of each condition complicated. *Mtb* genes may be regulated by a number of environmental signals simultaneously, and these results suggest that the signalling pathways in response to these conditions are overlapping (Boshoff and Barry, 2005, Ohno et al., 2003, Warner and Mizrahi, 2006).

There is limited evidence showing that the dosR regulon is required for *Mtb* survival during latent disease in humans. However, *in vitro* and animal studies have shed light as to its importance in infection. A recent study by Leistikow et al. (2010) showed that an *Mtb* mutant incapable of dosR regulon induction resulted in a 10,000-fold defect in anaerobic dormancy survival and that the dosR regulon is important in the resumption of growth from an anaerobic or nitric oxide induced nonrespiring state. Furthermore, members of the regulon are upregulated in activated murine macrophages (Schnappinger et al., 2003) and in infected mouse lung tissue (Voskuil et al., 2003). The importance of the dormancy regulon in animal infection however, is ambiguous. In guinea pigs, dosR is required for infection to progress to the granuloma stage and disruption of dosR attenuates *Mtb* (Malhotra et al., 2004). In mice, a similar disrupted mutant was shown to be more virulent than the wild type strain (Parish et

al., 2003a). This may reflect differences in the animal model as unlike the guinea pig, granuloma formation is limited in the mouse.

Interestingly, isocitrate lyase is not a member of the DosR regulon; it is upregulated in *Mtb* infected murine macrophages (Schnappinger et al., 2003), expressed on entry into non-replicating persistence in the Wayne model (Wayne and Sohaskey, 2001) and is required for chronic infection in mice (McKinney et al., 2000). While the dormancy regulon appears to be important in the adaptation to hypoxia and other stresses, it does not represent the full response to the range of conditions encountered by *Mtb in vivo*, be it the granuloma or activated macrophage. Additional genes identified by further studies may also play an important role and will further our understanding of *Mtb* responses. One of these studies by Rustad et al. (2008) demonstrated a core of 230 stably induced genes, termed the Enduring Hypoxic Response, that were found to be expressed for a much longer period during hypoxia than the dosR genes. These genes may be involved in the regulation and maintenance of bacteriostasis and that the DosR regulon may simply be involved in the preparation of *Mtb* for growth arrest by the sequestration of nutrients and the triggering of regulatory changes (Rustad et al., 2008).

1.4.3 Stresses or stimuli experienced in the host by *Mtb*

The transcriptome and phenotype of *Mtb* in sputum is likely to be produced through exposure to stimuli encountered within the host, but the exact stimulus or combination of stimuli that leads to this gene expression is unknown. Furthermore, many of these stimuli may be associated with *Mtb* bacilli entering a stage of non-replicating persistence leading to latent disease. Thus, to be able to fully understand the sputum transcriptome, a working knowledge of the possible stresses and stimuli encountered by *Mtb* throughout the course of infection is required.

A number of host stresses may be encountered by bacilli throughout the course of infection. These stresses conventionally include, but may not be limited to, damage to the mycobacterial cell wall including detergent stress, reactive nitrogen intermediates, reactive oxygen species, hypoxia, nutrient starvation, phosphate deprivation, acid stress, heat shock and shearing forces (Stallings and Glickman, 2010, Saviola, 2010). These stresses may be encountered from early infection through to the intraphagosomal environment (Russell, 2003).

The response of *Mtb* to each of these host-derived stresses is yet to be fully understood. Previous studies have made use of a number of *in vitro* dormancy models and *in vivo* animal

models which are not fully representative of the *in vivo* environment (Voskuil, 2004). For example, the set of genes differentially regulated in response to hypoxia is yet to be defined, with variations possibly due to the dissimilar experimental conditions (Rustad et al., 2008, Voskuil, 2004). Finally, there appears to be some overlap in the response to various stresses: the dosR regulon, for example, is induced under multiple conditions, including exposure to hypoxia and nitric oxide (Voskuil et al., 2003). The *in vivo* stresses listed above are reviewed below.

1.4.3.1 Damage to the Mycobacterial Cell Wall

Mycobacteria have a characteristic cell wall architecture, with it being thicker than many bacteria and possessing mycolic acids (Wayne and Kubica, 1986). The cell wall and membrane is the first defence against host stresses and *Mtb* is known to modify its cell wall architecture in response to *in vitro* stress experiments (Stallings and Glickman, 2010). These changes include cell wall thickening and modifications to the surface lipid and protein composition (Cunningham and Spreadbury, 1998, Stallings and Glickman, 2010).

It is thought that the main directed cell wall stress comes from antimicrobial peptides and proteins present in the airway surface fluid coating the respiratory epithelium and within the macrophage phagolysosome (Stallings and Glickman, 2010). A number of these proteins have been identified, such as β-defensins and granulysin, and their mode of action is thought to be through the alteration of membrane permeability (Stallings and Glickman, 2010). Pulmonary surfactant may also play a role in cell wall stress; the surfactant system consists of a complex array of proteins and lipids that reduce surface tension of the alveoli (Ferguson and Schlesinger, 2000). As a detergent, pulmonary surfactant may directly damage the cell wall. Intriguingly, surfactant may also play a role in innate immunity, through the modification of *Mtb*-macrophage interactions, although these mechanisms are complex and poorly understood (Ferguson and Schlesinger, 2000).

In vitro studies have made use of detergents or lysozyme to test mycobacterial survival to cell surface stress (Stallings and Glickman, 2010). In particular, sodium docedyl sulphate (SDS) is a detergent which causes cell wall damage on exposure. In *Mtb,* SDS stress leads to the upregulation of the stress-responsive sigma factor *sigE*, which activates a further 23 genes including 4 transcription factors (Manganelli et al., 2001). As certain SDS-activated genes are also upregulated during the growth of *Mtb* in macrophages, macrophages, like SDS, may cause damage the lipid components of the bacterial wall (Schnappinger et al., 2003).

It is interesting to note that components within the mycobacterial cell wall also contribute to the bacteria's virulence and pathogenesis, in addition to protecting the bacillus against attack (Makinoshima and Glickman, 2005, Smith, 2003). Changes to the outer surface of *Mtb* may affect the susceptibility of the bacillus to other host-derived stresses, and may even alter the actual host response (Rajni et al., 2011). Thus, while mycobacteria face cell stresses during infection, the changes in host response make the task of defining the bacterial response even more difficult.

1.4.3.2 Hypoxia

In vivo, the number of bacilli within a lesion correlates well with the degree of oxygenation, suggesting that oxygen may limit growth of *Mtb* during infection (Canetti, 1955a, Park et al., 2003a). Similarly, the most common site of reactivation of latent *Mtb* infection is within the oxygenated environment of the lung. Thus, while considered to be an obligate aerobe, *Mtb* is able to survive in hypoxic environments and a large volume of research has been directed at the impact of the gaseous environment on *Mtb*. The observation of hypoxic areas in *Mtb* infected primate and guinea pig granulomas supports the relevance of this research (Stallings and Glickman, 2010). The first *in vitro* model of *Mtb* hypoxia was the Wayne model and was described in Paragraph (Wayne and Hayes, 1996). The Wayne model demonstrated the response of *Mtb* to the gradual withdrawal of oxygen and survival to anaerobiosis. Bacteria grown in a stirred culture with defined headspace experienced progressive oxygen depletion, eventually leading to hypoxia and growth arrest. The hypoxic bacteria are able to enter two sequential states of non-replicating persistence, NRP1 at 72 hours and NRP2 at 240 hours (Wayne and Sohaskey, 2001). The Wayne model and subsequent hypoxia experiments have been used to understand the physiologic adaptation to hypoxia.

The early hypoxic response is controlled by the dosR-dosS/dosT two component system. It and the Enduring Hypoxic Response were discussed in Paragraph 1.4.2.

1.4.3.3 Reactive Nitrogen Intermediates

Once *Mtb* bacilli are internalised, the immune system may employ the production of reactive nitrogen intermediates (RNI) and reactive oxygen species (ROS) as a defence mechanism. Nitric oxide and other NO-derived RNI, produced by inducible nitric oxide synthase (iNOS) of activated macrophages, have been demonstrated to be a major host defence mechanism against acute and persistent *Mtb* infection in the mouse model of infection, although this may

differ in human disease (Flynn and Chan, 2001). Mice lacking this iNOS suffer from increased susceptibility to *Mtb* and a host of other bacterial infections, suggesting that nitric oxide and RNI form a major *in vivo* stress for AFB (MacMicking et al., 1995).

Phenotypically, exposure to nitric oxide inhibits aerobic respiration and reversibly slows replication of *Mtb* (Voskuil et al., 2003). This corresponds to the reported effects of hypoxia, and could indicate a common mechanism in the generation of this phenotype; to some degree this involves the 48-gene dosR regulon (Voskuil et al., 2003), upregulated under both conditions. This phenotype and gene expression studies suggest a similar response by *Mtb* to both nitric oxide and low O_2 exposure. As nitric oxide is released by activated macrophages and acts as a eukaryotic signalling molecule (Martin et al., 2000), *Mtb* may use it, in addition to hypoxia, as an environmental signal allowing the bacillus to adapt its metabolism to host-derived stresses.

1.4.3.4 Reactive Oxygen Species

In addition to RNI stress, mycobacteria may be exposed to reactive oxygen species (ROS) through innate immunity. The human innate immune response to infection involves the production of superoxide (O_2^-) via nicotinamide adenine dinucleotide phosphate (NADPH) oxidase ($gp91^{phox}$ and $gp47^{phox}$) proteins in phagocytic cells and can lead to damage of all components of the cell, including proteins, lipids and DNA (Stallings and Glickman, 2010). ROS may play a role in the control of replication in early infection, as mice unable to produce ROS are partially inhibited in their ability to control *Mtb* growth in an aerosol infection model (Voskuil et al., 2011). However, the deficiency of NADPH oxidase in macrophages has no effect on mycobacterial growth, suggesting that *Mtb* has mechanisms to resist killing by ROS (Stallings and Glickman, 2010, Voskuil et al., 2011).

Mtb uses a variety of mechanisms to control damage by O_2^- including the direct scavenging of the reactive species, by degrading O_2^- to water and molecular oxygen by the action of superoxide dismutase (SOD), catalase (KatG) and peroxidase (as a complex with peroxynitrite reductase) (Voskuil et al., 2011). Crillo et al. (2009) have also identified a mycobacterial gene locus (*mel2*) that contains luciferase-like genes involved protection against oxidative stress; luciferases are thought to scavenge hydrogen peroxide (H_2O_2) by converting it to water, light and an oxidised intermediate.

More recently, Voskuil et al. (2011) looked at the transcriptional response to both ROS and RNI. *Mtb* was found to be highly resistant to low levels of H_2O_2 and was only killed by high levels of oxidative stress and at lower levels, oxidative stress was not found to arrest bacterial

growth. There was significant overlap in the transcriptional response to both H_2O_2 and nitric oxide, but did not include the dosR regulated genes (Voskuil et al., 2011). Of note, the expression of certain oxidative stress defence genes was constitutive, suggesting that bacilli are prepared for exposure to ROS (Voskuil et al., 2011). These studies suggest that while *Mtb* encounters oxidative stress *in vivo,* bacilli have evolved appropriate mechanisms to counter this oxidative stress. Some of these protective mechanisms may be shared with and further complicates the *Mtb* response to RNI.

1.4.3.5 Nutrient Starvation and Phosphate Deprivation

Mtb is likely to inhabit a nutrient limited environment during infection. Bacilli isolated from lung lesions have displayed altered morphology and staining properties (loss of acid-fastness) compared to bacilli grown *in vitro*; cultures starved in distilled water displayed these same properties but regained their acid-fastness when restored to a rich growth medium (Betts et al., 2002). Further evidence pointing to *in vivo* nutrient starvation involved the use of auxotrophic *Mtb* strains. When starved of nutrients, bacteria must synthesise their own essential molecules; however, these strains, unable to produce these molecules, were shown to be attenuated in mouse models of infection (Stallings and Glickman, 2010).

In many bacteria, amino acid biosynthesis genes are upregulated during nutrient starvation, via the stringent response (Stallings and Glickman, 2010). The stringent response is a stress response that occurs in bacteria, as well as plant chloroplasts, in reaction to a variety of stress conditions. The response is a global regulatory mechanism, where the transcription of stable RNA is inhibited through the production of the hyperphosphorylated guanine nucleotides, ppGpp and pppGpp ((p)ppGpp) (Stallings et al., 2009). The production of (p)ppGpp is catalysed by the enzyme RelA (Primm et al., 2000). Studies have shown *Mtb* ΔrelA mutants had poor long-term survival during both starvation and hypoxic conditions; these mutants had similar virulence to wild-type strains during early infection in mice, but poor long-term persistence (Primm et al., 2000). The effect of relA knockout on long-term survival suggests that the stringent response is necessary for persistence; the exact stresses involved in the upregulation of the response in *Mtb*, nutrient deprivation or otherwise remain to be determined.

A specific subset of nutrient starvation stress is the limitation of phosphate *in* vivo. Phosphate is essential in various biomolecules, including membrane lipids, complex carbohydrates and nucleic acid and as a result, the acquisition and metabolism of phosphate from the environment is required for the survival of bacteria (Schaechter et al., 2006). Phosphate is

acquired through several transport systems; Rengarajan et al. (2005) showed specific phosphate transport genes *pstA1, pstC1, pstS3* and *phoT* to be important for the survival of *Mtb* in macrophages.

It has been suggested that *Mtb* bacilli residing in macrophage phagosomes encounter this inorganic phosphate limitation (Rengarajan et al., 2005). The depletion of phosphate in the environment has been shown to restrict the growth of *Mtb* in mouse and guinea pig lungs in a dose-dependent manner (Rifat et al., 2009). Furthermore, Rifat et al. (2009) showed that a number of genes are differentially expressed during phosphate deprivation, which they termed the phosphate starvation response (PSR). The PSR response included the upregulation of the *senX3-regX3* regulatory system, which is required for the induction of a number of phosphate transport genes; deletion or disruption of either *senX3* or *regX3* led to the attenuation of persistent infection in mice (Parish et al., 2003b, Rifat et al., 2009, Stallings and Glickman, 2010). Finally, as the PSR featured several genes similarly regulated under hypoxia and nutrient starvation, for example *sigE* and *relA* (Rifat et al., 2009). The results suggest that *Mtb* may face a phosphate deprivation, in addition to a nutrient starvation, *in vivo* and this may play a role in mycobacterial persistence.

1.4.3.6 Acid Stress

A key host defence following the internalisation of *Mtb* by macrophages is the fusion of phagosomes containing AFB with lysosomes, to destroy the bacteria through exposure to acid and hydrolase enzymes. This low pH damages DNA, proteins and lipids and disrupts biochemical reactions (Russell, 2003). In the unactivated macrophage, *Mtb* is able to block phagosome maturation and lysosome fusion (Russell, 2001). However, when the macrophage is activated with IFN-γ, the block on maturation is overcome, and phagosomes are able to fully mature and fuse with lysosomes (Russell, 2001).

Wild-type *Mtb* strains are quite resistant to acid stress in low pH media; further, certain *Mtb* mutants that allow for macrophage phagolysosome fusion and acidification are not attenuated in unactivated macrophages. These results indicate that *Mtb* has other mechanisms protective against acid stress (Stallings and Glickman, 2010). Specifically, genes allowing bacilli to maintain intra-bacterial pH despite an acidic environment (such as the proteases Rv3671c and Rv2224c) are crucial for survival (Stallings and Glickman, 2010). In addition to specific genes, genome wide changes may occur; studies have demonstrated the induction of 81 genes in response to acidic conditions, including a number of genes involved in lipid

metabolism (including *icl*), suggesting that *Mtb* begins to express genes required for long term survival upon exposure to acidic conditions (Fisher et al., 2002).

1.4.3.7 Heat Shock

As infection progresses and is combated by the host immune system, a fever develops as a consequence; this pyrexia may subject *Mtb* to a heat stress (Saviola, 2010). Heat shock subjected on proteins within the bacterial cytoplasm and cell wall leads to unfolding and denaturation. The heat shock response in *Mtb* is a complex adaptive system that allows cells to survive increases in ambient temperature and may be induced by a range of stress conditions in addition to heat (Stewart et al., 2002).

Two major regulons of the heat shock response in *Mtb* are the Hsp70 (DnaK) regulon and Hsp60 (GroE) response. These two regulons are repressed by HspR and HrcA, respectively, both control a total of 40 genes combined (Stewart et al., 2002). However, these genes and their encoded proteins may only constitute part of the overall adaptive response, as genes encoding sigma factors, sigB, sigE and sigH, are also induced following exposure to heat stress; the upregulation of sigB suggests overlap between the heat shock and general stress responses (Manganelli et al., 1999).

The majority of heat shock proteins act as molecular chaperones, to ensure proper folding, translocation and assembly of protein structures (Saviola, 2010). Partial disruption of the heat shock response through disruption of *hspR* in *Mtb* has an impact on virulence, with bacilli having impaired ability to establish chronic infection (Stewart et al., 2002). The *Mtb* Δ*sigE* mutant was shown to be defective in its ability to grow inside both human and murine unactivated macrophages and more sensitive to destruction by activated murine macrophages (Manganelli et al., 2001). Further, the *Mtb* Δ*sigH* mutant is more sensitive than the wild type strain to heat shock and oxidative stresses (Wang et al., 2011). These results indicate that the heat shock response plays an important role in the virulence of *Mtb,* although it remains to be determined if the specific stress of heat shock is the major cause of induction *in vivo*.

1.4.3.8 Shearing Forces

A poorly understood response is that of *Mtb* to shearing forces. Shear stress is defined as force per unit area and practically may be considered as pressure, cell surface frictional shear and tensile forces to counter and externally applied force (Braddock et al., 1998). These shearing forces may occur both during inspiration and expiration. The cough, a defence mechanism

used to clear excessive secretions and foreign materials from the airways, places the greatest force on bacilli (Donahue, 2002). Typically, a cough begins with an inspiration of air followed by an expiratory effort. This expiration is preceded by a closure of the glottis which allows greater pressure and higher air flow rates to be generated. During this stage, a narrowing of the central airway promotes increased shearing forces from the high velocity airflow (Donahue, 2002). Throughout the process of coughing or expectoration of sputum, *Mtb* bacilli are exposed to these shearing forces.

Only one study has been dedicated to the study of shearing forces and *Mtb* (Hall-Stoodley et al., 2006). The force was used to demonstrate the binding of *Mtb* to innate immunity molecules such as human fibronectin and surfactant protein (Hall-Stoodley et al., 2006). However, response of *Mtb* to the stress was not assessed. As gene expression changes in response to shearing forces occur in eukaryotic host cells in fluid systems, it seems reasonable to assume that transcriptional changes also occur in *Mtb* bacilli in response to this stress and requires further investigation (Braddock et al., 1998).

1.4.4 Genes in Dormant *M. tuberculosis*

Dormancy and the subsequent reactivation of *Mtb* are dynamic processes that require a number of genes and their products. A variety of genes and molecular mechanisms have been implicated in, or may be associated with *Mtb* persistence and as discussed, the sputum transcriptome featured the induction of a number of these genes, in addition to slowly or non-replicating signals. Some genes thought to be associated with mycobacterial dormancy have already been reviewed in this introduction, including *dosR*, *relA* (part of the bacterial stringent response) and *sigH, sigE, hsp70* and *hspR* (upregulated in response to heat shock). Genes associated with lipid storage and utilisation (including *tgs1* and *icl1*) are discussed below in Paragraphs 1.7.1.1 and 1.7.2, respectively. Other genes implicated with dormancy that are not introduced elsewhere are discussed here.

As expected, a number of genes, belonging to the dosR regulon are associated with mycobacterial dormancy. The most dramatic transcriptional activation belonging to the dosR regulon during hypoxia is the induction of *hspX* (*acr*), which encodes for a 16kDa α-crystallin-like heat shock protein (Sherman et al., 2001). The hspX protein acts as a molecular chaperone, preventing thermally induced aggregation of other proteins and may also play a role in the stabilisation of cellular structures as it has been demonstrated to localise to the cell wall (Hu et al., 2006). However, the protein may be redundant in its capacity as a chaperone, as deletion of *hspX* has no impact on the survival of *Mtb* to heat shock (Hu et al., 2006).

Deletion of *hspX* in *Mtb* leads to increased growth in both resting and activated macrophages and in BALB/c mice, although the lung damage cause by the Δ*hspX* strain was similar to the wild type strain (Hu et al., 2006). Conversely, overexpression of the gene leads to a decrease in the logarithmic growth rate of *Mtb* (Yuan et al., 1996). However, Yuan et al. (1998) in an independent study showed that *hspX* deletion led to attenuation of *Mtb* in macrophages. The conflicting evidence leaves many questions with regards to the role of α-crystallin in *Mtb,* but suggests that the *hspX* gene is associated with slow growth.

NarK2, also a member of the dosR regulon, encodes a nitrite transporter, is a member of the *narK2-narX* operon. In *Mtb,* it has been suggested that on adaptation to hypoxic conditions, the reduction of nitrate, through the action of nitrate reductase, serves as an energy, as opposed to nitrogen, source. Under these conditions, nitrate serves as a final electron acceptor in the proton motive force (Sohaskey and Wayne, 2003). Intriguingly, the *narK2-narX* operon does not encode a functional nitrate reductase; deletion of *narX* had no effect on the reduction of nitrate in *Mtb* (Sohaskey and Wayne, 2003). In fact, the non-dosR regulated *narGHJI* operon is the only nitrate reductase that is active in *Mtb*; a *narG M. bovis* BCG deletion mutant has been shown to be attenuated in severe combined immune deficiency (SCID) mice, as compared to the wild type strain (Sohaskey and Wayne, 2003).

As discussed, β-oxidation of fatty acids may provide an important energy source for *Mtb* during infection. However, this metabolism of fatty acids produces the reducing equivalents NADH and $FADH^2$, which must be re-oxidized for the cycle to continue (Boshoff and Barry, 2005). It is thought that nitrate is present in the phagosome of activated macrophages and granulomas as a degradation product of nitric oxide and RNI (Boshoff and Barry, 2005). *Mtb* may reduce this nitrate via the *narGHJI* nitrate reductase, restoring the redox balance during growth on fatty acids (Boshoff and Barry, 2005). The nitrite-efflux system encoded by *narK2* may then be used to remove the resulting nitrite waste. A study by Shi et al. (2005) lends support to this theory, as expression of *narK2* was upregulated during late state infection of the mouse lung, suggesting that nitrate reduction is increased during *Mtb* infection. However, the use of alternative electron acceptors has not yet been demonstrated in *Mtb* (Boshoff and Barry, 2005).

Sigma factors act as transcriptional regulators by reversibly associating with the prokaryotic RNA polymerase and are responsible for promoter regulation (Manganelli et al., 1999). They allow organisms to adjust their transcription rapidly and several are induced upon exposure of *Mtb* to environmental stresses (Wang et al., 2011). SigE and sigH were described in previous sections. In addition to these two factors, SigF, encoded by the *sigF* gene, is thought to play a role in stress gene regulation as it is strongly induce on exposure to nitrogen limitation, cold

shock, hypoxia, ROS, nutrient starvation and entry into stationary phase (Wang et al., 2011). Studies on *Mtb* strains with disrupted or deleted *sigF* genes showed that they were less virulent than the wild type in mice, based upon time of death, unable to reach the same bacterial load in murine lungs as the wild type and formed fewer lesions (Wang et al., 2011). These studies indicate that SigF appears to be important during late stage disease. This is supported by Gomez et al. (1997).

Along with dosS/dosT-dosR, other two-component regulatory systems may be involved with persistence. *MprAB* encodes a two-component system implicated in virulence during the persistence stage of infection; *mprA* binds to motifs upstream of *sigB* and *sigE in vitro*, to regulation the expression of both genes, suggesting a role in the general stress response. Furthermore, inactivation of *mprA* enhances the survival of *Mtb* in unactivated macrophages but impairs its persistence in murine lungs and spleens (Pang and Howard, 2007). It has been noted that disruption of *mprA* affects the growth of *Mtb* at multiple stages during the infective process; intriguingly however, during exponential growth, this disruption upregulates stress related genes of the dosR regulon. The *mprAB* interactions are clearly complex; while it appears clear that the genes are associated with the stress response in *Mtb*, the roles of *mprAB* in persistence and gene regulation in exponential growth need further study.

While many genes are implicated in persistent *Mtb,* a number of questions still remain. Firstly, it must be established that the *in vitro* conditions thought to induce dormancy truly are accurate. Functional studies need to be performed to elicit the role of some of implicated genes. Finally, the regulatory networks and interaction between these dormancy genes in underpinning persistence must still be established.

1.5 Transmission of Tuberculosis

Tuberculosis is spread via the respiratory system, by airborne droplet nuclei, which are particles of 1-5μm in diameter that contain *Mtb*. A study by Fennelly et al. (2004) demonstrated this variability in the size distribution of tuberculous aerosols, finding some particles to be greater than 7μm in diameter. The majority of the particles were found to be in a respirable range, generally defined in studies as including droplets from 1-10μm in diameter. The particles containing *Mtb* settle very slowly and can remain airborne for minutes to hours after expectoration, due in part to their small size (Beggs et al., 2003, Frieden et al., 2003). A person with pulmonary tuberculosis, coughing frequently, can introduce thousands of these droplet nuclei into the air. Furthermore, Riley and O'Grady (1961) demonstrated that droplet nuclei could also be produced during sneezing and speaking. It is therefore not surprising that tuberculosis transmission often occurs in poorly ventilated and overcrowded areas, such as prisons or low-income housing areas (Beggs et al., 2003). Aerosol transmission is further reviewed in Chapter 5.

Following inhalation, the infectious droplet nuclei lodge in the alveoli of the distal airways (Frieden et al., 2003). *Mtb* is then taken up by alveolar macrophages, which initiates in a cascade of events that can result in either successful containment of infection or the progression to active disease. These events were summarised previously (1.4).

A number of factors may affect the probability of a bacillus establishing infection and are shown in Figure 3. There are therefore many risk factors for the transmission of tuberculosis that determine the infectivity of the source (e.g. expulsive force of cough), effectiveness of the host defences (e.g. co-existing illness) and properties of the bacillus (e.g. virulence factors, genetic mutations, etc.) (Dharmadhikari and Nardell, 2009). The variation in infectiousness is highlighted by a study performed in the 1960s exposing guinea pigs to air from a tuberculosis hospital ward. Patient infectiousness was analysed through the exposure of guinea pigs to air extracted from hospital wards; results demonstrated that 4% of the tuberculosis patients on a particular ward produced greater than 73% of the infections guinea pigs (Sultan et al., 1960). More recent molecular epidemiology studies in New York City and San Francisco have also shown such variation, suggesting that approximately 10% of patients are considered to be highly infectious and thus, transmit the majority of disease (Hamburg and Frieden, 1994).

The type and site of disease appears to partially decide the infectivity of a source. It has been argued that cavitation of lung lesions, harbouring up to 10^8-10^9 organisms, is the event responsible for the propagation of *Mtb*. Predictably, patients may be more or less infectious based upon the site of disease; an example of this was evidenced by Riley et al. (1961) who observed that a laryngeal tuberculosis patient infected 15 guinea pigs within 3 days on an

experimental ward, far more than even pulmonary tuberculosis patients. Transmission of extra-pulmonary disease such as Pott's disease (spine) or kidney tuberculosis, is even rarer, as these sites do not communicate with the external environment (Dharmadhikari and Nardell, 2009).

Cough frequency also appears to be important in its association with transmission. Loudon and Spohn (1969) found an association between overnight cough frequency and an increased transmission of tuberculosis among household contacts. A more recent study by Fennelly et al. (2004) generated data consistent with this finding; increased cough frequency led to higher tuberculous aerosol production. However, the presence of culturable organisms in cough generated aerosols decreases rapidly after the first few weeks of anti-tuberculosis treatment, even though sputum smears may still be positive (Fennelly et al., 2004).

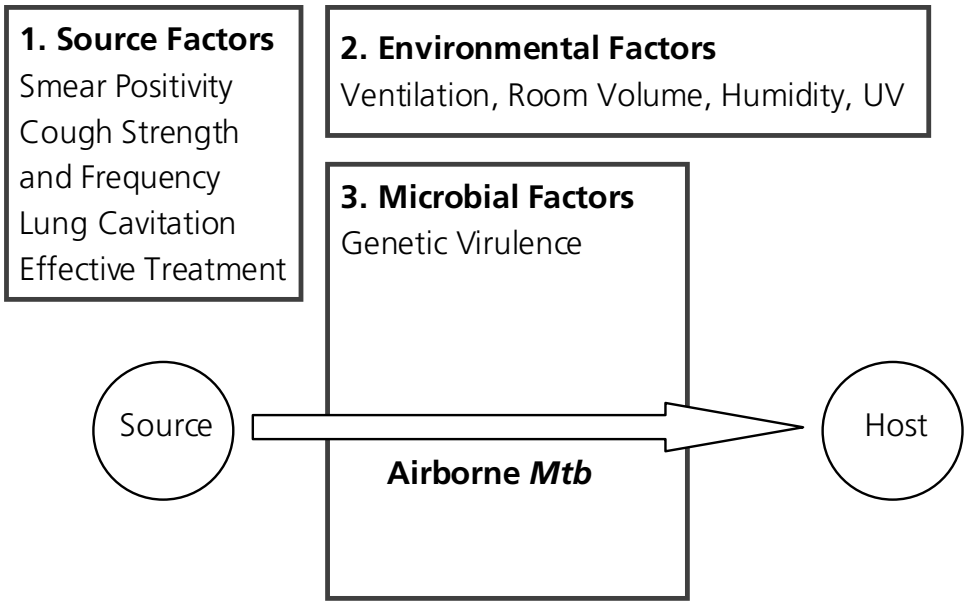


Figure 3: Tuberculosis Transmission Factors

The figure shows the factors affecting the transmission of tuberculosis including source factors, environmental factors and microbial factors (Figure adapted from Dharmadhikari and Nardell, 2009).

Environmental factors directly impinge on the survival of *Mtb* as the bacilli travel between hosts. A number of stresses may be involved, including desiccation, oxygen, ozone,

temperature, humidity and radiation (Cox, 1989). Desiccation and ultraviolet light radiation are discussed in Chapter 5.

The dose required for infection is unknown, although classic studies in guinea pigs by Riley et al. have demonstrated that 1-10 bacilli delivered as a non-dispersed aerosol can produce pulmonary tuberculosis. Autopsies performed on infected guinea pigs often showed the formation of solitary lesions, secondary to the inhalation of a single infectious droplet nucleus (Reviewed in Lever et al., 2000). As a result, these studies suggest that there is a high risk of breathing a large infectious dose of *Mtb* immediately after release. Individuals in close contact with source cases are more likely to receive these greater infectious doses.

Numerous factors contribute to the effectiveness of the transmission of tuberculosis. Taken together, studies suggests that close and prolonged contact is required for the transmission of the disease, and the highest risk of inhaling *Mtb* occurs immediately following aerosolisation (Hickman et al., 1995, Lever et al., 2000). Transmission rates increase further in cavitating disease in patients with respiratory disease (Dharmadhikari and Nardell, 2009).

1.5.1.1 Host Factors affecting Vulnerability to Infection

Post-mortem and guinea pig studies have demonstrated that a single inhaled droplet nucleus may initiate infection in certain individuals, underscoring the large variability in human susceptibility, possibly due to differences in immune function, genetics and co-morbidity (Dharmadhikari and Nardell, 2009). Once *Mtb* enters the human lung, interactions between the bacilli and host immune system and the influence of co-morbid medical conditions determine whether infection and progression to clinical disease are established. Innate immunity involving macrophages acts as the first line of defense, followed by cell-mediated immunity and delayed-type hypersensitivity reactions; any defect in these systems could make the host more vulnerable to infection, enhancing transmission (Dharmadhikari and Nardell, 2009). Epidemiological studies have also suggested enhanced risk of *Mtb* infection associated with ethnic or racial background (Stead et al., 1990).

The risk factors increasing host susceptibility to the reactivation of latent *Mtb* infection compared to initial infection may be altogether different. For example, studies have shown that exposure to silica dust and silicosis lead to an increased rate of tuberculosis reactivation; silica may both enhance *Mtb* growth and alter host immunological pathways (Dharmadhikari and Nardell, 2009). However, in recent times, the most important co-morbid factor increasing host vulnerability to infection is HIV or acquired immune deficiency syndrome (AIDS) co-infection (Dye et al., 2008). Worldwide, approximately one-third of the 40 million people

infected with HIV/AIDS are co-infected with *Mtb* and the risk of these individuals developing clinical disease can be as high as 10% per year (Gandhi et al., 2006). The basis behind this increased risk may be due to the host depletion in CD4 T-cells, thought to play a key role in the containment of latent infection (Dharmadhikari and Nardell, 2009). While important in latent infection, it is unclear if HIV infection increases overall risk of primary infection.

1.5.1.2 Entry of M. tuberculosis into Host Cells

Mtb, being a facultative intracellular pathogen, has developed numerous mechanisms for entering human macrophages. For this reason, it differs from most obligate extracellular pathogens which have evolved mechanisms to avoid phagocytosis and the host immune system.

Complement proteins, present in serum, opsonise bacteria for phagocytosis by the C3b or C3bi complement receptors on macrophages (Aderem and Underhill, 1999). Several receptors that participate in the phagocytosis of complement-opsonised particles, including complement receptor type 1 (CR1), type 3 (CR3) and type 4 (CR4), are expressed on macrophages. These receptors occur in two distinct structural forms (Ernst, 1998). CR1 is a single chain trans-membrane protein, which binds C3b and C4b, but not C3bi, and is involved mainly in particle binding. CR3 and CR4 are heterodimeric proteins which bind C3bi and are responsible for particle internalisation (Aderem and Underhill, 1999, Ernst, 1998). *Mtb* can also activate the alternative pathway of complement activation, resulting in opsonisation with C3b and C3bi complement components. Bacteria that are opsonised with these ligands can bind to CR1, CR3 and CR4 and are internalised into phagosomes (Ernst, 1998).

Uniquely, *Mtb* has an additional mechanism for acquiring opsonic C3 peptides. Bacilli can recruit the complement fragment C2a to form a C3 convertase and generate active C3b. This mechanism occurs in the absence of components of the alternative or classical pathways (Schorey et al., 1997). In addition to acquiring C3 peptides by two mechanisms, *Mtb* can also bind to CR3 via two distinct domains. Opsonised *Mtb* can bind to CR3 at the C3bi binding domain, but unopsonised bacilli can use endogenous capsular polysaccharides to interact with a β-glucan binding site (Ernst, 1998). *Mtb* can therefore utilise multiple mechanisms to exploit complement receptors and gain entry into the macrophage.

Human monocyte-derived macrophages may also bind and internalise virulent *Mtb* via mannose receptors. Mannose receptors are monomeric transmembrane proteins, expressed on mature macrophages, but not on fresh monocytes. It is unclear how many ligands for

mannose receptors *Mtb* possesses, but the main ligand is lipoarabinomannan (LAM) (Schlesinger et al., 1994). LAM is abundant in the mycobacterial cell wall, peripherally exposed and contains terminal mannose residues that interact with the receptors. As the expression of mannose receptors is downregulated by interferon-gamma, it is likely that the mannose receptors play an important role in *Mtb* phagocytosis early in infection (Ernst, 1998).

Surfactant protein A (Sp-A) may also play a role in mycobacterial uptake, although the interactions are largely undefined. Sp-A enhances macrophage binding and uptake of *Mtb*; Sp-A may either act as an opsonin, or may modulate a receptor involved in the binding of bacilli; in fact, mannose receptors may account for some of the binding activity (Downing et al., 1995). However, the mechanisms involved in Sp-A and mycobacteria interactions demand further study.

Finally, other receptors possibly involved in macrophage binding include CD14 (cluster of differentiation 14), scavenger receptors and Fcγ receptors. CD14 is a membrane protein that is a high affinity receptor for lipopolysaccharides of gram-negative bacteria (Ernst, 1998). They can also bind LAM, but it is unclear if they can mediate phagocytosis without another membrane protein. Scavenger receptors bind polyanionic macromolecules, such as lipopolysaccharides of gram-negative bacteria. Again, it has been shown that they can bind *Mtb*, but it is as yet unknown if they can activate internalisation of the bacilli. Fcγ receptors may play a role in the internalisation of Immunoglobulin G (IgG) coated mycobacteria, in individuals with circulating antibodies to *Mtb* (Ernst, 1998).

Mtb has numerous and diverse ligands on its cell surface and can engage multiple receptors of different types. However, the question still remains: why has a pathogen that can survive and replicate extracellularly, developed multiple mechanisms to gain entry to the intracellular environment of cells that are meant to kill it? The sheer number of mechanisms that lead to phagocytosis of the bacillus suggest that *Mtb* has found the intracellular environment of macrophages to be especially advantageous. One possibility is that the macrophage promotes dissemination of *Mtb* by assisting transport of bacilli across epithelial layers. Studies have suggested that *Mtb* gene expression is altered on phagocytosis into macrophages; these genes may play roles in the survival, growth and spread of bacteria (Ernst, 1998). Another theory is that the bacilli are modified within the macrophage, with specific changes to surface proteins, carbohydrates or lipids. These modifications may be advantageous, and may aid bacterial dissemination to distant sites (Ernst, 1998). Considerable additional work is required to understand these complex interactions between *Mtb* and the macrophage.

1.6 *M. tuberculosis* in Sputum

A prominent feature of clinical tuberculosis is the expectoration of sputum. Expectoration is defined as 'the act of coughing up and spitting out the material produced in the respiratory tract' (Farzan, 1990). Sputum provides an easily obtainable sample of the *in vivo* bacterial population as it leaves the body. In this sense, bacilli in smear-positive tuberculous sputum may express characteristics that are required for or enhance transmission. However, until recently, studies on the biological characteristics of *Mtb* in sputum have been limited.

Work by Garton et al. (2008) has shed some light on the characteristics of *Mtb* in sputum. All sputum samples examined showed that a percentage of bacilli that accumulated LBs (further discussed in Paragraph 1.7). In addition to the presence of LBs in all samples examined, examination of the sputum transcriptome revealed gene expression significantly different to conventional *in vitro* aerobic growth conditions. Further, the transcriptome signatures suggest that the bacillary population is dominated by slowly or non-replicating bacilli (Garton et al., 2008). The gene expression in sputum is outlined in detail below (1.6.1).

Outside of the bacilli, little is known about the composition of tuberculous sputum; the research that has been performed on sputum has largely been in relation to diseases such as cystic fibrosis and chronic obstructive pulmonary disease (COPD) (Voynow and Rubin, 2009). Sputum consists of a number of components, may be heterogeneous and its composition can vary greatly from patient to patient and sample to sample. Cellular components can include eosinophils, neutrophils, macrophages, lymphocytes, metachromatic cells and bronchial epithelial cells (Belda et al., 2000). Its key molecular components include mucous glycoproteins, mucins, immunoglobulins (Jeffery and Li, 1997). Furthermore, blood and saliva are often noted as a proportion of expectorated sputum samples, and it is reasonable to assume that other natural respiratory secretions and microbiota may comprise a percentage of expectorated sputum samples.

The exact source from which tubercle bacilli in AFB positive sputum are derived is unknown. Tuberculosis cavities form when the lesion involves the wall of an airway and the structure of a tuberculosis cavity (from the inside out) consists of the necrotic zone lining the margins of the cavity, the cavitary capsule and the atelectatic zone around the cavity (Canetti, 1955b). As the capsule and atelectatic zone contain few bacteria, Canetti (1955b) demonstrated that the bacteria originate from the necrotic margin of these caseated lesions within the lung. Based upon the microarray data from Garton et al. (2008), dormant bacilli also represent a potential source of *Mtb* in sputum, as the transcriptome showed the presence of slowly or non-replicating transcriptional signatures. Potential locations of latent bacilli are reviewed in Paragraph 1.4.1.1.

If bacteria are derived from these lesions, one could extrapolate that other components are expelled in addition to the bacilli. The tuberculous lesions/granulomas comprise a number of components. The pathogenesis of tuberculosis is reviewed later in this chapter, but to summarise, after *Mtb* infection, the infection site becomes organised into a granuloma, consisting of infected macrophages surrounded by macrophages, monocytes and multi-nucleated giant cells (Russell, 2007). Following this, the macrophages begin to die and form caseum, the granuloma becomes increasingly necrotic, the lesion liquefies and can rupture into the lung airway, facilitating transmission (Russell, 2007). The caseum is complex in composition, and may possess a number of lipid species including cholesterol esters, cholesterol and triacylglycerol (Kim et al., 2010). These dynamic tuberculosis lesions may contribute, in addition to *Mtb* bacilli, a variety of living and necrotic eukaryotic cells and lipid-rich material present in sputum. However, the composition of tuberculosis sputum and source of bacilli *in vivo* remain to be determined experimentally.

1.6.1 Gene Expression in Sputum

It has long been the prevailing view that *Mtb* bacilli in sputum are undergoing active aerobic replication. However, the microarray study by Garton et al. (2008) of the transcriptome of tuberculous sputum from 4 independent patients dispelled this theory, revealing that gene expression of the AFB in sputum differs greatly from this aerobic growth theory. Comparison of the transcriptome of *Mtb* recovered from human sputum to these aerobic cultures demonstrated gene expression unique to this previously characterised *in vitro* growth condition (Garton et al., 2008). Specifically, compared to aerobic growth, 182 genes were significantly induced in sputum, while 334 genes were significantly repressed. It was notable that nearly twice as many *Mtb* genes were downregulated than upregulated in sputum.

As *Mtb* in the Wayne model had previously been shown to accumulate LBs in similar levels to those found in sputum, the transcriptome was compared to two stages of non-replicating persistence (NRP1 and NRP2) prepared using the Wayne model and a mixed RNA preparation consisting of aerobic and NRP2 cells mixed in the proportion of 70:30 (w/w total RNA), respectively (Garton et al., 2008). This data is displayed in Figure 4 (Garton et al., 2008).

No single or obvious combination of defined *in vitro* growth conditions corresponded to the transcriptional signature from sputum. There was however, some overlap in the pattern of expression between sputum, NRP2 and other published studies. *Mtb* in sputum had a number of similarly regulated genes when compared to both slow growth and NRP Stage 2 (NRP2

clustered genes are highlighted in Figure 4, Boxes 1 and 2) (Garton et al., 2008). Also, when compared to a previously established murine macrophage infection model (339 induced and 111 repressed genes on infection), *Mtb* in sputum showed similar expression distributions (Schnappinger et al., 2003).

In addition to the genome wide expression changes, changes in the transcription of functional groups and regulons were notable. Of the upregulated genes, members of the dosR regulon were most prominently activated (Garton et al., 2008). The dosR regulon was discussed in detail in Paragraph 1.4.2. In the sputum transcriptome, 11 of the original 48 dosR regulon genes were identified to be upregulated greater than 2.5-fold. These genes included *tgs1*, *hspX* and *narK2*, genes implicated with dormancy. Interestingly, of the 11 genes, Rv3133c, the *dosR* gene, was not shown to be upregulated in sputum.

In addition to dosR, a noticeable feature is the upregulation of a number of lipid catabolism and cholesterol utilisation genes in the sputum transcriptome. Most notably, isocitrate lyase, *icl1*, was induced 3.61-fold in sputum. The upregulation in particular demonstrates a shift to the utilisation of lipids, as opposed to carbohydrates, as a carbon and energy source. I*cl1* is the key glyoxylate shunt enzyme, essential for the metabolism of fatty acid and virulence in *Mtb,* and is discussed in more detail below (McKinney et al., 2000, Munoz-Elias and McKinney, 2005). With regards to cholesterol utilisation, three genes from the putative *kstR* regulon in particular were upregulated greater than 2.5-fold in sputum; the regulon consists of 74 genes in *Mtb* including *mce4* and *igr* and is reputed to control for the expression of genes used for the utilization of diverse lipids, particularly cholesterol, as energy sources (Kendall et al., 2007, Kendall et al., 2010). Furthermore, the regulation of 64 genes that may be involved in cholesterol metabolism was, as a group, significantly upregulated in sputum as compared to aerobic growth (Cole et al., 1998, Garton et al., 2008).

Of the repressed genes, the downregulation of aerobic respiration and ribosomal genes was prominent. Regulation of 21 aerobic respiration genes and 45 ribosomal genes was shown to be decreased in both sputum and NRP2 as compared to aerobic growth (Garton et al., 2008). In addition, the reduction in the efficiency of the aerobic respiration chain is confirmed by the downregulation of *nuoB* (type-I NADH dehydrogenase), *qcrC* (cytochrome bc_1 complex) and *ctaD* (aa_3-type cytochrome *c* oxidase) (Garton et al., 2008, Shi et al., 2005). These three genes, along with *atpA* (adenosine triphosphate (ATP) synthase alpha chain) and *atpD* (ATP synthase beta chain), coincided with the repressed genes previously observed in bacillary stasis of the chronic murine infection model (Shi et al., 2005).

Altogether, the most prominent feature of the sputum transcriptome was the marked difference in gene expression as compared to aerobic growth. These changes included dosR

activation, lipid utilisation, slow growth and repression of aerobic respiration and ribosomal genes.

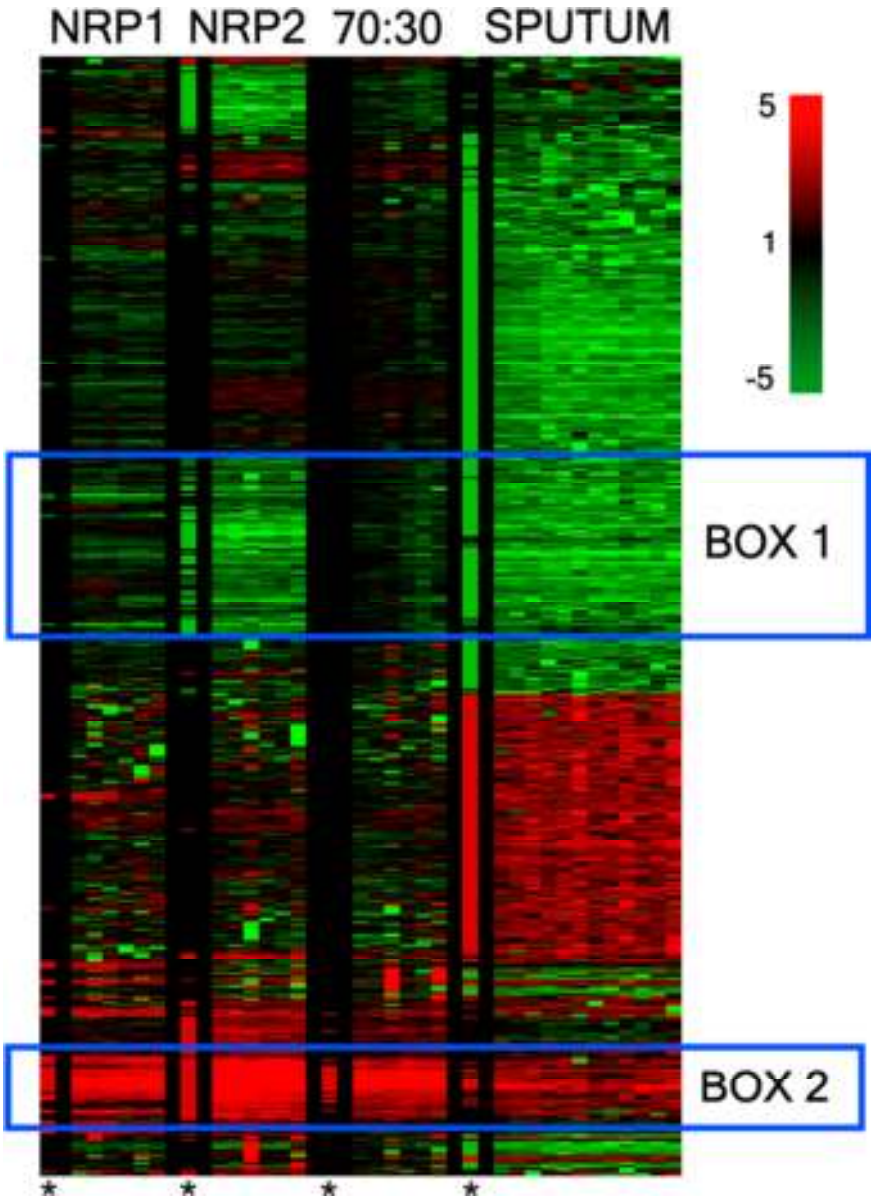


Figure 4: Heat Map of Genes Differentially Regulated in Sputum and Non-Replicating Persistence versus Aerobic Culture

The figure displays a heat plot showing genes regulated differentially against *Mtb* H37Rv in NRP1, NRP2, 70:30 aerobic growth/NRP2 and by infecting *Mtb* in sputum (in columns, from left to right, respectively) (Garton et al., 2008). Upregulated and downregulated genes are marked in the colours red and green, respectively. Asterisked columns signify conditions in which genes were identified to be significantly differentially expressed as compared to aerobic growth. Boxes 1 and 2 highlight genes similarly clustered to NRP Stage 2. Box 2 represents the dosR regulon.

1.7 Lipid Bodies in Mycobacteria

LBs consist of intracellular, spherical, lipid-filled inclusions, first reported by Burdon in 1946 and more recently recognised in AFB in sputum from a patient suffering from tuberculosis in The Gambia (Garton et al., 2002). However, as acid-fast staining does not discriminate *Mtb* from other acid-fast mycobacteria, and as one-third of tuberculosis cases in The Gambia are caused by *M. africanum*, a further study was conducted to examine the occurrence of LBs in AFB in sputum (de Jong et al., 2007, Garton et al., 2008). As mentioned, LBs were found to be a universal feature of AFB in sputum samples from both Leicester and The Gambia (Garton et al., 2008). Figure 5 demonstrates LBs in AFB-positive sputum.

At the time of this report, it was concurrently discovered that tgs1, a triacylglycerol synthase catalysing the final step of triacylglycerol production (discussed below; 1.7.1.2) is expressed in mycobacteria as they enter *in vitro* non-replicating persistence and are exposed to a variety of stresses (Daniel et al., 2004, Sirakova et al., 2006). Based on this finding and as triacylglycerol had previously been shown to be the major component of LBs, subsequent work by Garton et al. found that when placed under similar *in vitro* conditions of non-replicating persistence, *Mtb* accumulated LBs to levels comparable to those found in sputum (Garton et al., 2002, Garton et al., 2008). Furthermore, it was demonstrated that *Mtb* produced LBs in response to stresses including two models of hypoxia, nitric oxide exposure, heat and cold shock, acid pH and alkaline pH (Sherratt, 2008).

As discussed in 1.6.1, analysis of the sputum transcriptome by Garton et al. (2008) showed upregulation of both *dosR* and *tgs1*, suggesting a potential link between LB formation and bacilli in a state of slow or non-replication. Induction of *tgs1* and other *tgs* genes and subsequent triacylglycerol accumulation has also been directly demonstrated in response to slow withdrawal of oxygen, nitric oxide exposure and acidic and static growth conditions (Daniel et al., 2004, Sirakova et al., 2006). Also, as a member of the dosR regulon, *tgs1* is linked to the adaptation to non-replicating persistence in a hypoxic shift-down model (Park et al., 2003a) and in response to nitric oxide exposure (Voskuil et al., 2003). In addition, Schnappinger et al. (2003) have shown a modification in *Mtb* gene transcription upon infection in murine macrophages; *dosR* and *tgs1* were both upregulated on infection with expression increasing further in response to activation of these macrophages.

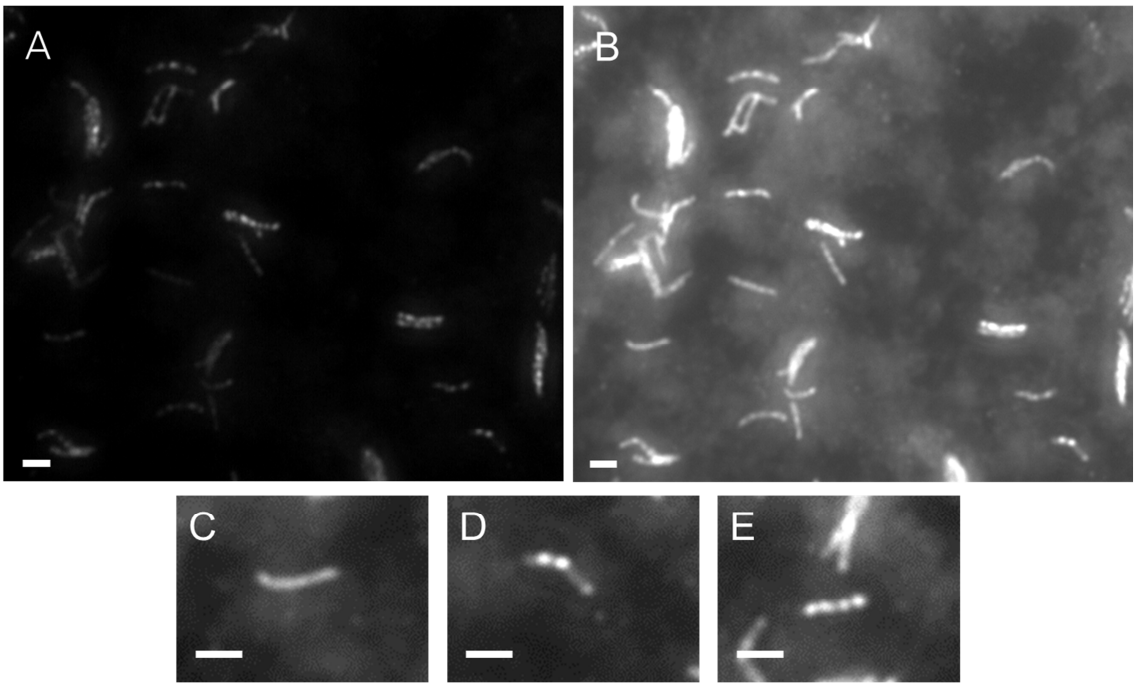


Figure 5: Lipid Bodies in Tuberculosis Sputum Samples

The figure shows images of Auramine-Nile red stained AFB positive sputum smears. (A) Auramine fluorescence image; (B) Corresponding Nile red fluorescence image. The auramine filter is a long-pass filter that attenuates shorter wavelengths and transmits longer wavelengths; thus, some red light (LBs) is visible in the auramine image; (C) AFB with no LBs; (D) AFB with 2 LBs (E) AFB with 4 LBs. Scale bar is 2µm.

The evidence above suggests that *tgs1* is an important 'dormancy' related gene, and that *Mtb* may form LBs in response to conditions that activate the DosR regulon. However, the association is not so straightforward. The metabolic states produced by both hypoxia and nitric oxide exposure are still being defined. While both conditions are capable of inducing dosR and result in overlapping gene expression patterns, their effects on redox processes are expected to be different and may produce two distinct metabolic states (Boshoff and Barry, 2005). One such characteristic of these states may be growth arrest. Using hypoxia and nitric oxide as examples, both stresses have been demonstrated to reversibly arrest growth of *Mtb* (Voskuil et al., 2003, Wayne and Hayes, 1996). As a result, while LBs and *dosR* expression might be related, LB formation may actually occur as a result of the metabolic state that the bacillus is in and requires further investigation.

A link between LB formation and metabolic activity was further evidenced when sputum cultures were examined for a so-called 'time to positivity' (Garton et al., 2008). Sputum samples graded 3+ smear positive on microscopic examination were cultured in a BACTEC

960 system for diagnostic purposes. The percentage of LB-positive cells correlated significantly to the time it took for *Mtb* growth to be detected, lending support to the hypothesis that LBs are related to the metabolic state in mycobacteria, particularly the percentage of non-replicating persistent AFB (Garton et al., 2008).

1.7.1 Biochemical Basis of Lipid Bodies

A number of processes may underlie the production of LBs in mycobacteria. These factors, including a general overview of the storage of lipid in prokaryotes, triacylglycerol synthesis and the formation of LBs in mycobacteria, are introduced and reviewed in the following sections.

1.7.1.1 Lipid Storage in Prokaryotes

Lipids, along with their functions within membranes of bacteria, can act as storage molecules in prokaryotes. Almost all prokaryotes accumulate one type of lipid, unless this storage would not confer the bacteria an advantage, like when in a nutrient rich environment, for example (Waltermann and Steinbuchel, 2005). Most commonly, this storage molecule is poly-β-hydroxybutyrate (PHB) or other polyhydroxyalkanotes (PHA) and these both PHB and PHA accumulate within the bacterial cytoplasm in the form of inclusion bodies (Waltermann and Steinbuchel, 2005).

In contrast to PHB and PHA, triacylglycerol is commonly used as a storage lipid in eukaryotes and plant seeds, and along with wax esters, is considered to be unusual prokaryotic storage molecules (Alvarez and Steinbuchel, 2002). Triacylglycerols are fatty acid triesters of glycerol and their biosynthesis has been shown to be a common feature of the actinomyces group, including *Mycobacterium, Streptomyces, Nocardia* and *Rhodococcus* and the Gram-negative genus, *Acinetobacter* (Alvarez and Steinbuchel, 2002). This triacylglycerol is stored in the form of LBs within the bacterial cytoplasm and may account for up to 70% of the dry weight of a cell (Alvarez and Steinbuchel, 2002).

A number of functional roles have been suggested for triacylglycerol LBs. From research in *Rhodococcus opacus* and *Streptomyces lividans,* these triacylglycerol LBs are thought to act predominantly as energy and carbon storage bodies (Waltermann and Steinbuchel, 2005). Other functional roles have also been suggested for triacylglycerol LBs. This includes the use of LBs as deposits for toxic or surplus fatty acids from phospholipid biosynthesis or to maintain of a bacterial hydration through the oxidation of fatty acid hydrocarbon chains

(Waltermann and Steinbuchel, 2005). However, the literature regarding the mobilisation and degradation of stored triacylglycerol is limited (Waltermann and Steinbuchel, 2005).

In addition to triacylglycerol, esters of long chain alcohols and fatty acids, or wax esters (WE), are also used as lipid storage molecules in bacteria (Waltermann and Steinbuchel, 2005). This occurs mainly in the genus *Acinetobacter* but has been described in *Moraxella, Micrococcus*, *Fundibacter, Corynebacterium, Mtb* and *Nocardia* (Waltermann and Steinbuchel, 2005). WE have been noted to be stored as either spherical or rectangular inclusion bodies (Waltermann and Steinbuchel, 2005).

Studies on *Acinetobacter calcoaceticus* have led to the identification of a novel wax synthase/ diacylglycerol acyltransferase (WS/DGAT) which has given some insight into the synthesis of these lipids in prokaryotes (Kalscheuer and Steinbuchel, 2003). The WS/DGAT catalyses the final step of both triacylglycerol and WE biosynthesis, where an acyl group from acyl-CoA is esterified to diacylglycerol or a long chain alcohol (Kalscheuer and Steinbuchel, 2003). Homologues of this gene were identified in *Mtb* and were subsequently studied by Daniel et al. (2004); these homologues and the biosynthetic pathway for triacylglycerol are presented in the following section.

1.7.1.2 Triacylglycerol Synthesis and Lipid Body Formation in Mycobacteria

As described previously, triacylglycerol is stored in mycobacteria in the form of LBs; the formation of LBs appears to be environmentally linked, based on previous work in the group (Garton et al., 2002, Sherratt, 2008). In *Mycobacterium smegmatis,* LBs were formed following cultivation in low nitrogen broth and utilised during carbon starvation. In *Mtb,* LBs have been shown to be synthesised under conditions of hypoxia, nitric oxide exposure, acidic conditions, heat and cold shock (Sherratt, 2008). Thus, Barer and Garton (2009) have proposed a role for the LB structure as a possible storage unit for LCFAs as triacylglycerol, buffering against fatty acid toxicity and providing a source of fatty acids for the biosynthesis of cellular components. Mycobacteria have an extraordinary requirement for LCFAs. While the cell is actively growing, large amounts of LCFAs are required to synthesise new cell envelope components; however, exogenous fatty acid sources, while advantageous are also potentially bactericidal, due to their detergent like actions (Kondo and Kanai, 1972). The model is outlined in Figure 6. Within the model for LBs is a proposed a model for the metabolic pathway of triacylglycerol formation in mycobacteria (Barer and Garton, 2009). Initially, LCFAs may enter the cell via a fatty acid transport system or may passively diffuse into the

cell. Mycobacteria may also synthesise LCFA *de novo* or through the elongation of short chain or medium chain fatty acids.

The mycobacteria possess two types of fatty acid synthetase (FAS): type 1 (FAS-1), found in eukaryotes and certain prokaryotes and type 2 (FAS-2), found in bacteria and plants (Takayama et al., 2005). The mycobacterial FAS-1 catalyses the *de novo* synthesis of short chain fatty acyl primers and is able to produce two classes of fatty acid: fatty acids with 16-18 carbon chains, used in phospholipid synthesis and fatty acids with 24-26 carbon chains, used in mycolic acid and multimethyl branched fatty acid synthesis (Bhatt et al., 2007). FAS-1 also possesses the ability to elongate fatty acids obtained from the environment to be used in mycolic acid synthesis (Schweizer and Hofmann, 2004). Mycobacterial FAS-2, on the other hand, is not capable of *de novo* fatty acid synthesis. Instead, the FAS-2 enzymes serve to elongate medium chain length fatty acids produced by FAS-1 (Takayama et al., 2005).

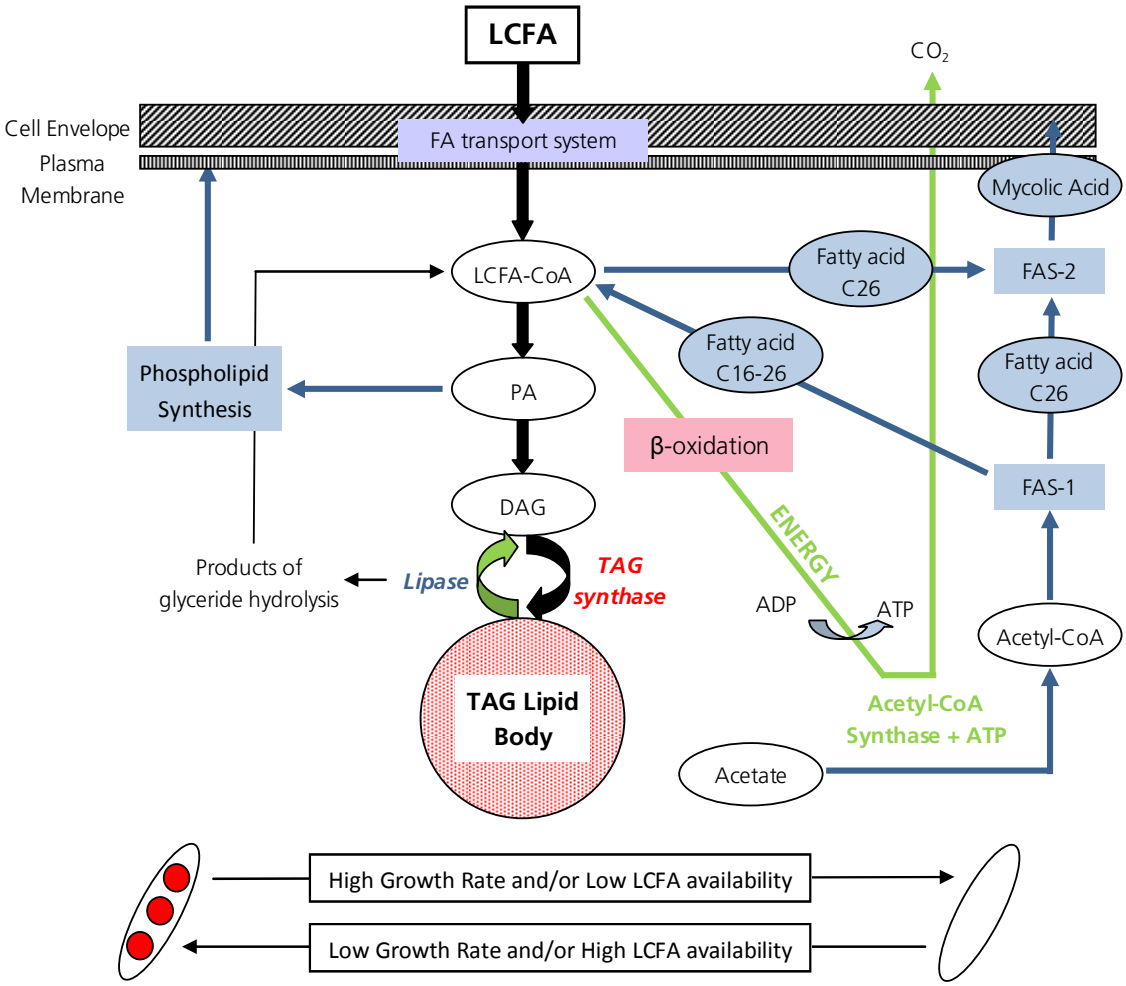
Acquired LCFA are activated to LCFA-CoA by coenzyme-A (acetyl-CoA), an adenosine triphosphate (ATP) dependent reaction catalysed by a fatty acid CoA synthase (Gurr and James, 1975). These activated fatty acids may be directly used in the synthesis of complex cell wall lipids or phospholipids during periods of growth or replication. When bacterial growth slows or stops completely in response to unfavourable environmental conditions or nutrient starvation, the biosynthesis of phospholipids and cell wall lipids is likely to arrest and surplus LCFA are channelled into the production of triacylglycerol. This triacylglycerol may be drawn upon in future as a carbon or LCFA source (Waltermann and Steinbuchel, 2005). At first, triacylglycerol may be located within the cell wall until saturation, at which point LBs are formed in the cytoplasm. The synthesis of triacylglycerol may also occur secondary to the detoxification of the exogenous cellular environment through the removal of environmental LCFA (Kondo and Kanai, 1972). The synthesis of triacylglycerol is outlined in Figure 7.

The final step of triacylglycerol biosynthesis is the addition of a fatty acid chain to the precursor, diacylglycerol, catalysed by a diacylglycerol acyltransferase (DGAT) enzyme. A dual-function WS/DGAT was first identified from *A. calcoaceticus* and fifteen mycobacterial genes have since been identified through their shared homology with this WS/DGAT (Daniel et al., 2004, Kalscheuer and Steinbuchel, 2003). When cloned and expressed in *Escherichia coli (E.coli)*, the gene products showed DGAT activity, with Rv3130c displaying the highest activity. The gene products had very low wax ester synthase activity, and were therefore renamed triacylglycerol synthases (Tgs); the most active of these genes, Rv3130c, was designated as *tgs1* (Daniel et al., 2004). *Tgs1* is a member of the dosR regulon and is upregulated under stress conditions leading to triacylglyerol accumulation in *Mtb* (Sirakova et

al., 2006). The formation of LBs in *Mtb* has since been demonstrated in response to all of the mentioned stress conditions (Sherratt, 2008).

Once stored as triacylglycerol, LCFA may be liberated through the action of a lipase, to be used in the synthesis of cell wall lipids or phospholipids, or as an energy source in the absence of a carbohydrate carbon source. The *Mtb* genome has been annotated with 21 putative lipase genes, although none of them showed triacylglycerol hydrolysis activity (Cole et al., 1998). However, three further lipase genes have subsequently been annotated through their homology with an *Mtb* strain W17 putative lipase (Deb et al., 2006). One of these genes, LIPY, was shown to have hydrolase activity and a *Mtb* ΔLIPY mutant was shown to have greatly impaired triacylglycerol hydrolysis under conditions of starvation (Deb et al., 2006).

The studies on *tgs1* and LIPY outlined here lend support to the proposed model of LB formation in *Mtb*. Under conditions of stress, which lead to periods of slow growth, excess LCFA is funnelled into triacylglycerol and LB synthesis. This triacylglycerol can be hydrolysed back to LCFA to replenish cell reserves.



Caption Overleaf.

Figure 6: Proposed Model for Lipid Bodies in the Regulation of the Long Chain Fatty Acid-CoA Pool in Mycobacteria

LBs may play a role in the regulation of long chain fatty acid (LCFA) flow through the mycobacterial cell (Barer and Garton, 2009). LCFA is stored as triacylglycerol (TAG) (Figure 7), which may be mobilised when required through the action of lipases; subsequent β-oxidation of fatty acids yields energy as ATP (Gurr and James, 1975). If conditions permit the bacilli to grow or replicate, these LCFA may be activated by coenzyme A (CoA), may be elongated via FAS-2 and used for the synthesis of mycolic acids or other LCFA derivatives. Alternatively, LCFA may be used for phospholipid turnover during growth via a biosynthetic intermediate, Phosphatidic acid (PA). For the *de novo* synthesis of short-chain fatty acids via the FAS-1, acetyl-CoA acts as the base molecule; fatty acids are formed through condensation of acetyl-CoA and malonyl-CoA (Takayama et al., 2005).

Finally, LBs may be formed based upon the environmental balance of available LCFA and cell growth conditions. During rapid cell growth, LCFA flux is also rapid due to cell wall lipid and phospholipid turnover. Conversely, if cell growth slows or is arrested, biosynthesis of cell wall lipids and phospholipids is downregulated and LCFA may be shuttled into TAG synthesis for storage. TAG may also be present in the cell wall, until saturation, at which point it is again stored as LBs. Excess environmental LCFA may also be diverted into TAG LB storage, to minimise intra- and extra-cellular toxicity.

TAG synthesis is shown using black arrows, other lipid biosynthetic pathways are shown in blue and lipid degradation is shown in green. Enzymes are italicised. Metabolic pools are enclosed within ovals. (Figure adapted from Sherratt, 2008, Barer and Garton, 2009)

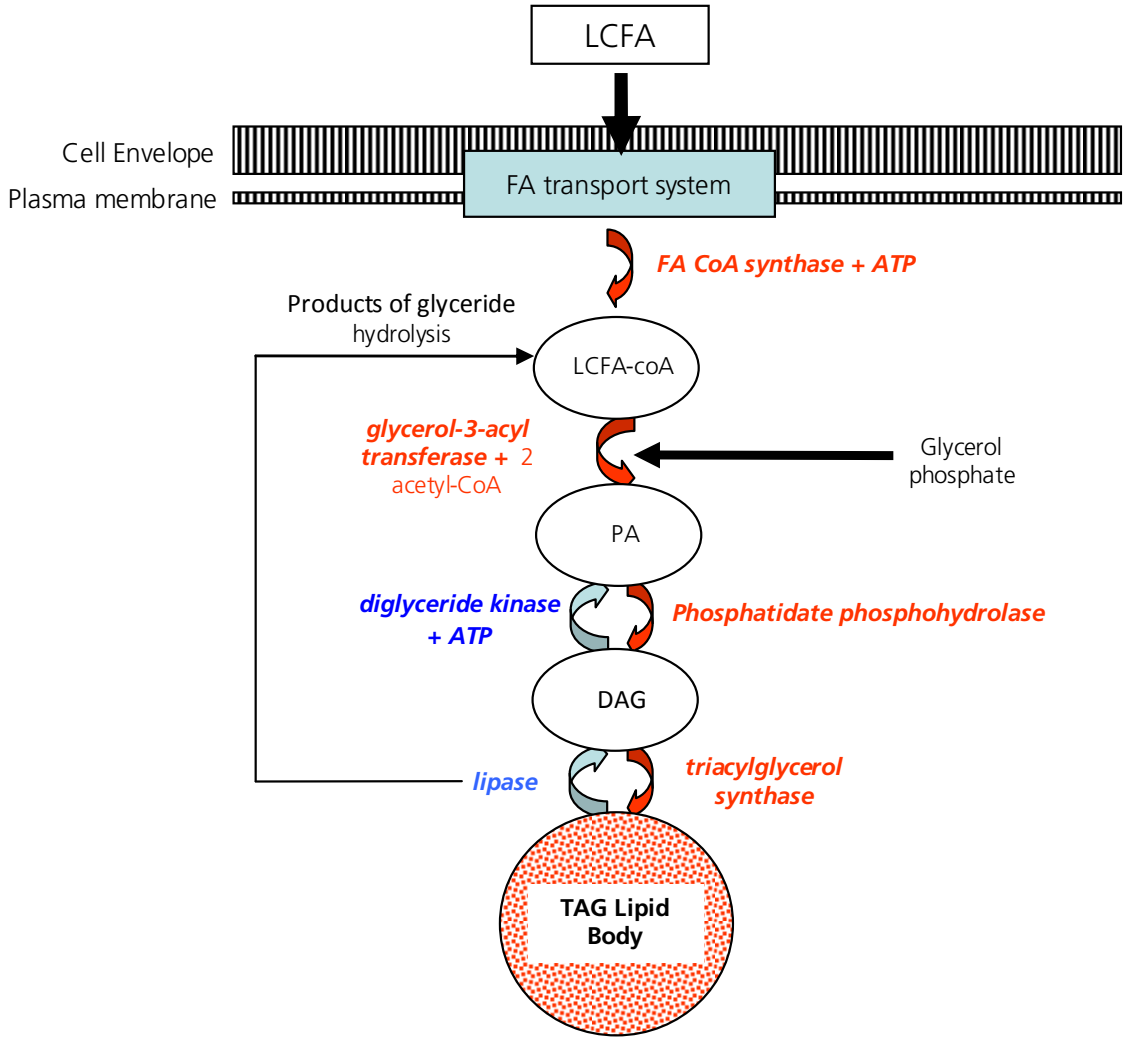


Figure 7: Proposed Model for Triacylglycerol Formation in Mycobacteria

The figure shows a proposed model for the synthesis of triacylglycerol (TAG) in mycobacteria. Fatty acids (FA) are first acquired by the cell; long chain fatty acids (LCFA) are subsequently activated by coenzyme-A (acetyl-CoA), an adenosine triphosphate (ATP) dependent reaction catalysed by fatty acid CoA synthase (Gurr and James, 1975). Glycerol phosphate is esterified by LCFA-CoA to form lysophosphatidic acid., which is acylated further to yield phosphatidic acid (PA) (Gurr and James, 1975). Diacylglycerol (DAG) is then formed through the dephosphoration of phosphatidic acid or the acylation of an existing monoglyceride (Gurr and James, 1975). Finally, LCFA-CoA availability exceeds cell usage, it may be esterified to DAG to form TAG. The reactions may be reversed through the action of a lipase to break down TAG and DAG to replenish the LCFA-CoA pool. The complete reaction is represented in simplified form in the figure.

The enzymes are highlighted in italics. Anabolic reactions are shown in red and catabolic reactions are shown in blue. (Figure adapted from Sherratt, 2008, Barer and Garton, 2009)

1.7.2 Utilisation of Fatty Acids and Cholesterol in *M. tuberculosis*

Fatty acid metabolism plays an important role in the general metabolism of *Mtb* as there is increasing evidence to suggest that the tubercle bacillus uses fatty acids as a carbon and energy source *in vivo,* largely because of the variety and quantity of lipids available within mammalian cells and the granuloma (Cole et al., 1998, Wheeler and Blanchard, 2005). The importance of fatty acids is displayed in the approximately 250 distinct enzymes involved in fatty acid metabolism in *Mtb* compared with only 50 in *E. coli*; many of the genes encoding components of fatty acid oxidation systems are replicated in the genome (Cole et al., 1998).

Segal and Bloch (1956) described the preferential metabolic response of *Mtb* isolated from lungs of infected mice on media where fatty acids were the major carbon source. McKinney et al. (2000) demonstrated that isocitrate lyase 1 (ICL1), an enzyme involved in fatty acid metabolism, is essential for the late stage *Mtb* persistence in the mouse model of infection. Further work revealed that *Mtb* possesses two *icl* homologues: *icl1* and *icl2*. Deletion of *icl1* resulted in reduced bacterial loads in chronic infection while deletion of *icl2* had little effect on bacterial growth. Bacteria deprived of both genes however, were unable to grow in mice and were rapidly cleared from both the lungs and spleen (Munoz-Elias and McKinney, 2005). Further evidence supporting the use of fatty acids *in vivo* was provided by Schnappinger et al. (2003), who showed *icl1,* which encodes for ICL1, to be upregulated during infection of macrophages.

While *in vivo* grown mycobacteria have been suggested to be largely lipolytic, rather than lipogenic, the source and type of fatty acid used by *Mtb* during infection is unknown (Munoz-Elias and McKinney, 2006). The mature granuloma is rich in lipids derived from dead cells, as well as bacteria, and *Mtb* may acquire this source of fatty acids (Honer zu Bentrup and Russell, 2001). Other sources of fatty acids may include lung surfactant, which is rich in fatty acids and can be internalised by macrophages, or the utilisation of macrophage triacylglycerol stores following phagocytosis (Munoz-Elias and McKinney, 2006). Interestingly, *Mtb* may metabolise fatty acids liberated through the action of a lipase from internally stored triacylglycerol, as described in 1.7.1.2 (Daniel et al., 2004).

Two pathways are specifically required for the metabolism of fatty acids: the catabolic β-oxidation cycle that degrades fatty acids to acetyl-CoA units, or, if glycolytic substrates are absent, the anaplerotic glyoxylate cycle (Munoz-Elias and McKinney, 2005).

The main pathway for the metabolism of fatty acids in bacteria and eukaryotes is the β-oxidation cycle. β-oxidation yields acetyl-CoA from fatty acids through successive rounds of degradation, which is used in the Krebs cycle (Munoz-Elias and McKinney, 2005). The first

step of β-oxidation is the introduction of a *trans* double bond to the hydrocarbon chain of an activated fatty acid; this is catalyzed by an acyl-CoA dehydrogenase with flavin adenine dinucleotide (FAD) as a co-factor (Gurr and James, 1975); *Mtb* possesses 36 homologues of the *noc* gene, which encodes for acyl-CoA dehydrogenase (Cole et al., 1998). Water is then introduced across this *trans* double bond of the unsaturated acyl-CoA to form β-hydroxyacyl-CoA.

The dehydrogenation step follows next, where β-hydroxyacyl-CoA is converted into a β-keto fatty acid by the enzyme β-hydroxyacyl-CoA-dehydrogenase. Finally, thiolytic cleavage of the keto acid occurs, where the –SH group of the CoA displaces an acetyl-CoA moiety. The result is a loss of two carbon atoms from the length of the molecule. This whole process is repeated until the carbon chain has been reduced to acetyl-CoA groups. Each cycle generates energy in the form of one molecule each of $FADH_2$, NADH and acetyl-CoA (Gurr and James, 1975).

In the event that acetyl-CoA cannot enter the Krebs cycle, it can be directed into the glyoxylate shunt pathway. This pathway is the first step leading to the movement of carbon into gluconeogenesis, which is the only mechanism by which the bacilli can retain carbon from fatty acids (Russell, 2001). The gating enzyme to the pathway, ICL, converts isocitrate into succinate and glyoxylate. This is followed by the addition of acetyl-CoA to glyoxylate to form malate (catalysed by malate synthase) (Russell, 2001). The glyoxylate shunt pathway therefore circumvents the two decarboxylation steps of the Krebs cycle to allow the incorporation of two carbon molecules, such as acetate obtained from β-oxidation into the cycle (Wall et al., 2005). This pathway, unique to prokaryotes, lower eukaryotes and plants, allows for the synthesis of carbohydrates from fatty acids, while supplying the Krebs cycle with intermediates (Honer zu Bentrup and Russell, 2001, Russell, 2001). The glyoxylate shunt pathway is shown in Figure 8.

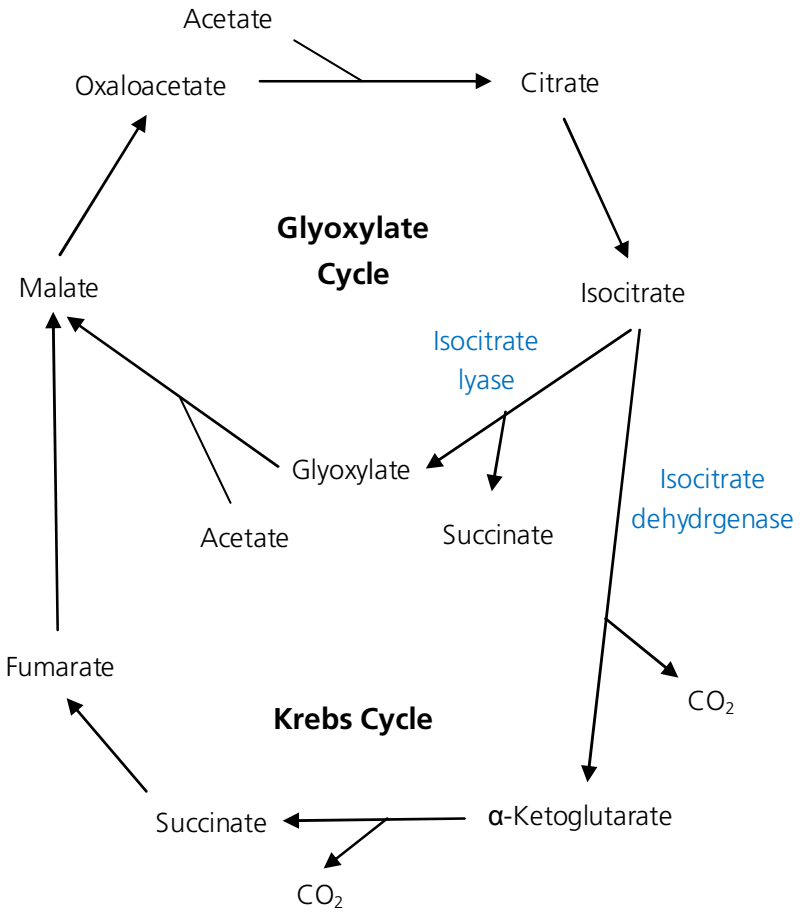


Figure 8: Glyoxylate Shunt Pathway

The pathway is required for the synthesis of carbohydrates and supply of Krebs intermediates when fatty acid is the carbon source (Russell, 2001). The β-oxidation of fatty acids generates acetate which is incorporated into the cycle, bypassing two decarboxylation steps in the Krebs cycle. The gating enzyme, isocitrate lyase, converts isocitrate into glyoxylate and succinate and contrasts to the Krebs cycle gating enzyme, isocitrate dehydrogenase, which produces α-ketoglutarate (Russell, 2001).

More recently, cholesterol has been identified as a lipid important in *Mtb* infection. *Mtb* has the ability to both accumulate and utilise cholesterol, even when grown in media containing a carbon source other than cholesterol (Brzostek et al., 2009). Pandey and Sassetti (2008) showed that *Mtb* can use cholesterol as a carbon source; carbon from the sterol ring is converted to carbon dioxide, suggesting this was used for energy generation via the Krebs cycle.

The molecular basis behind cholesterol utilisation in *Mtb* is only now being elucidated. From this aspect, it appears that two gene loci are particularly important: *mce4* and *igr*. *Mce4*

comprises 6 genes and has been predicted to encode a multi-subunit transport system, of which cholesterol is a substrate, while the *igr* locus also has 6 genes, encoding proteins related to lipid metabolism (Miner et al., 2009). On deletion of *mce4, Mtb* could not import cholesterol, and thus had a marked growth defect on media containing cholesterol as the only carbon source. Furthermore, *Mtb* was unable to convert cholesterol to CO_2 and could not incorporate cholesterol into the cell membrane (Pandey and Sassetti, 2008). This subsequently led to the attenuation of the mutant in activated macrophages and attenuation in the mouse model of infection during late time-points (Pandey and Sassetti, 2008). The authors concluded that *Mtb* requires cholesterol for the establishment of chronic infection. Concurrent studies also identified the *igr* operon as its deletion (Rv3540-5c) led to an attenuated mutant in macrophages (Chang et al., 2009). For this reason, it was termed *igr* for 'intracellular growth'. The operon was noted to be upregulated in macrophages and required for survival in mouse spleens. The annotated genes suggest a role in lipid metabolism, but studies had been unable to identify the lipid substrate utilised (Miner et al., 2009).

Interestingly, concurrent studies identified a gene cluster for its homology to one in *Rhodococcus* sp. strain RHA1, which encodes for cholesterol catabolism (Van der Geize et al., 2007). This regulon contains genes involved in both cholesterol and lipid metabolism and under the control of kstR (encoded by *kstR,* Rv3574) and kstR2 (encoded by *kstR2,* Rv3557c), protein regulators which fall into the TetR family that act as a repressors (Kendall et al., 2007, Kendall et al., 2010). The cluster contained both *igr* and *mce4* and indicated that *igr* may be associated with cholesterol metabolism. Subsequent experiments showed that an *igr* deletion mutant was unable to grow in the presence of cholesterol and could not be rescued through the addition of other carbon sources (Chang et al., 2009). A Δ*igr*Δ*mce4* double knockout mutant exhibited partial growth in the presence of cholesterol, was able to infect and kill macrophages and grow in murine lungs and spleen (Chang et al., 2009). These results demonstrate that *igr* is associated with cholesterol metabolism, as deletion leaves *Mtb* unable to degrade cholesterol, leading to a build-up of toxic metabolites and cell death; the additional deletion of *mce4* blocks cholesterol uptake, rescuing the cell (Chang et al., 2009).

Cholesterol appears to be available during infection and may act as a nutrient during the latter stages of infection. However, recent evidence has shown that it is not an essential source of nutrition during infection (Yang et al., 2011). The mutation of Rv1106c (*hsd*), a gene required by *Mtb* for growth on cholesterol as a sole carbon source, had no impact on *Mtb* growth in macrophages or guinea pigs, showing that *Mtb* does not rely on cholesterol as its only energy source *in vivo* (Yang et al., 2011). How this impacts human infection remains to be seen. Thus, while these studies indicate a role for cholesterol in *Mtb* infection, how and when it is utilised remains unclear.

1.8 Aims and Objectives of this Study

The overall aims of this study were to examine *in vitro,* two of the defining characteristics of *Mtb* in sputum, the transcriptome and accumulation of lipid bodies and begin to study the potential roles of these features in the transmission of tuberculosis.

The specific objectives were to:

- Establish a well validated polymerase chain reaction based assay capable of the recognising whether the gene transcriptional pattern in sputum had been replicated *in vitro*.
- Define candidate growth conditions or stimuli that may underlie the gene expression of *Mtb* in sputum.
- Attempt to determine the impact of lipid bodies on the transmission related properties of *Mtb*, including survival to desiccation and UV light and the primary process of *Mtb* binding to macrophages.

Chapter 2

Materials and Methods

2.1 Bacterial Strains and Plasmids

Table 1: Bacterial Strains used in this Study

Strain	Description	Source
M. tuberculosis H37Rv	Virulent laboratory strain	Laboratory Stocks
M. tuberculosis H37Rv pSMT3-*tgs1*	Laboratory strain with pSMT3 plasmid overexpressing *tgs1*	(Garton, 2011)
M. tuberculosis H37Rv pSMT3	Laboratory strain with pSMT3 plasmid	(Garton, 2011)

2.2 Laboratory Reagents and Culture Media

2.2.1 Chemicals and Media

All chemicals were obtained from Thermo Fisher Scientific (Loughborough, Leicestershire, UK) or Sigma-Aldrich Company Limited (Poole, Dorset, UK), unless otherwise stated. Growth media was obtained from Becton Dickinson Biosciences (Oxford, UK) or Sigma-Aldrich Company Limited.

2.2.2 Growth Media

Middlebrook 7H9 Broth

Middlebrook broth was prepared by dissolving 4.7g of broth powder in 900ml double-distilled water containing 2.5g glycerol. The solution was sterilised at 121°C for 17 minutes. Prior to use, the broth was supplemented with Albumin-dextrose-catalase (ADC) supplement at a concentration of 10% (v/v) and 10% (w/v) Tween-80 at a concentration of 5% (v/v).

Middlebrook 7H10 Agar

Middlebrook 7H10 agar was prepared by dissolving 19g of agar powder in 900ml double-distilled water containing 6.25g glycerol. The agar was boiled for 30 minutes to fully dissolve the powder and sterilised at 121°C for 17 minutes. Prior to use, the agar was supplemented with Oleic acid-albumin-dextrose-catalase (OADC) at a concentration of 10% (v/v).

Albumin-Dextrose-Catalase (ADC) Supplement

ADC was prepared by dissolving the following components in 150ml of double-distilled water:

Bovine Serum Albumin fraction V	7.50g
D-Glucose	3.00g
Sodium Chloride	1.28g
Catalase	6.0mg

The solution was centrifuged at 6371 × g for 30 minutes in a Beckman Coulter (High Wycombe, UK) Avanti J-E refrigerated superspeed centrifuge to remove undissolved solids. The supplement was filter-sterilised (0.2µm filter, Nalgene, Hereford, UK) and stored at 4°C.

Oleic Acid-Albumin-Dextrose-Catalase (OADC) Supplement

OADC was prepared as per ADC supplement, with the addition of oleic acid solution (1% w/v) in 0.2M sodium hydroxide. Prior to centrifugation, the solution was sonicated (Decon FS 100, Hove, UK) for 30 minutes to emulsify the oleic acid.

Phosphate-buffered Saline (PBS)

Phosphate-buffered saline (PBS) was prepared using PBS tablets (Sigma-Aldrich). One PBS tablet was dissolved in 200ml double-distilled water to give a solution with 0.01M phosphate buffer, 0.002M potassium chloride and 0.137M sodium chloride at pH7.4. The solution was autoclaved at 121°C for 20 minutes.

Oleic acid solution

Oleic acid solution was prepared by dissolving 100µl of oleic acid in 100ml 5% (w/v) bovine serum albumin, to give a stock concentration of 3.35mM. The solution was sonicated for 1 hour to allow the oleic acid to enter solution before being filter sterilised (0.2µm filter). The solution was stored at 4°C. Prior to use, the solution was melted at 50°C. When liquefied, the oleic acid solution was cooled to 37°C and added to growth medium at 10% (v/v).

Preparation of media with cholesterol

The preparation of media supplemented with cholesterol was adapted from Pandey et al. (2008). Cholesterol (10mg) was dissolved in 1ml (1:1 w/v) of tyloxapol/ethanol solution at 80°C for 15 minutes. The 1ml of cholesterol/tyloxapol/ethanol mix was added to 5ml of the respective media (7H9, PBS or RPMI-1640) at 80°C for a further 5 minutes. The combined mixture was added to 94ml of the respective media at 50°C to reach a final volume of 100ml. The complete solution was cooled to room temperature. Prior to use, the solution was pre-warmed to 37°C.

Complete RPMI-1640 supplemented with Foetal Calf Serum

RPMI-1640 media was purchased (Gibco) pre-supplemented with 1mM sodium pyruvate, 2mM L-glutamine, 4500mg/L glucose, 10mM HEPES and 1500mg/L sodium bicarbonate. Prior to use, the media was supplemented with 10% (v/v) heat-inactivated Foetal Calf Serum (FCS), 1% (v/v) Penicillin-Streptomycin supplement (Sigma) and 1% (v/v) ITS supplement (Sigma). Penicillin-Streptomycin supplement contained 10000 units penicillin and 10mg streptomycin per ml. ITS supplement contained 1.0mg/ml recombinant human insulin, 0.55mg/ml human transferrin and 0.5µg/ml sodium selenite. The addition of β-mercaptoethanol was excluded from the media. For experiments which made use of *M. tuberculosis* (*Mtb*), Penicillin/Streptomycin was excluded. Supplemented media were stored at 4°C and used within 1 month. Supplemented media was pre-warmed to 37°C prior to use.

Sautons Broth

Sautons broth was prepared with the following reagents dissolved in 900ml of distilled water:

Potassium dihydrogen orthophosphate	0.5g
Magnesium sulphate	0.5g
L-Asparagine	4.0g
Glycerol	75.0g
Ferric ammonium citrate	50.0mg
Citric Acid	2.0g
1% (w/v) zinc sulphate	0.1ml
Tween-80	0.5g

The pH was adjusted to 7.0 and the solution was made up to a final volume of 1 litre. The solution was sterilised at 121°C for 17 minutes and stored at room temperature.

Youmans Broth

Youmans broth was prepared from the following reagents in 900ml distilled water:

Potassium dihydrogen orthophosphate	5.0g
Magnesium sulphate	0.5g
L-Asparagine	5.0g
Glycerol	40.0g
Ferric ammonium citrate	50.0mg
Sodium citrate	2.5g

The pH was adjusted to 7.0 and the solution was made up to a final volume of 1 litre. The solution was sterilised at 121°C for 17 minutes and stored at room temperature.

Artificial Saliva

Artificial saliva composition was obtained and adapted from Shellis et al. (1978). The constituents listed in Table 2 were dissolved in 1 litre of double distilled water.

Table 2: Adapted Composition of Artificial Saliva (Shellis, 1978)

Constituent	Concentration (mg/L)	Constituent	Concentration (mg/L)
Ammonium chloride	233	Albumin	25
Calcium chloride, dihydrate	210	Urea	173
Magnesium chloride, hexahydrate	43	Uric acid	10.5
Potassium chloride	1162	Creatinine	0.1
Potassium di-hydrogen orthophosphate	354	Choline	13
Potassium thiocyanate	222	Amino Acid Mix	41
Sodium citrate	13	Vitamin Mix	0.8
Sodium hydrogen carbonate	535	Alpha Amylase	3×10^5 units
di-Sodium hydrogen orthophosphate	375		

Glycoprotein was excluded from the artificial saliva mixture as the addition of glycoprotein in Shellis et al. (1978) was primarily to increase saliva viscosity. The Amino Acid and Vitamin Mixes, described by Shellis et al. (1978) were substituted with RPMI 1640 Amino Acids Solution (Sigma-Aldrich) and RPMI 1640 Vitamins Solution (Sigma-Aldrich), respectively. The complete artificial saliva solution was filter sterilised and stored at 4°C. The solution was pre-warmed to 37°C prior to use.

2.3 Cultivation of Bacteria

2.3.1 Stock Cultures

Frozen stocks of *Mtb* were prepared by mixing 500µl of exponentially growing culture (doubling every 24 hours, assessed through optical density readings) with an equal volume of 50% (v/v) glycerol solution (in double distilled water and filter-sterilised). The 1ml stock cultures were stored in 1.5ml cryovials (NUNC, Thermo Fisher) at -80°C.

2.3.2 Cultivation of *M. tuberculosis*

2.3.2.1 Standard Culture Techniques

Mtb is an ACDP Category 3 hazardous pathogen. All work was performed within a Class 1 or Class 2 microbiological safety cabinet, contained within a Category 3 Containment laboratory suite and in accordance with the suite Code of Practice. All liquid and solid cultures were double-bagged and the external surfaces sterilised with 70% (v/v) ethanol when removed from the hood and when under incubation.

Starter cultures of *Mtb* were prepared by adding a thawed stock aliquot of *Mtb* to 5ml of 7H9/ADC. The culture was incubated at 37°C with shaking at 100rpm. After 7-10 days, the culture was used to inoculate four 10ml aliquots of 7H9 in 50ml screw-top centrifuge tubes and these cultures were further incubated for 3-5 days. The cultures were used to inoculate 100-200ml of media (7H9 or Sautons) at a calculated OD_{580} of 0.05 in roller bottles (Greiner, Stonehouse, UK) or 30-50ml of media in screw-top, plastic Erlenmeyer flasks. The roller bottles were incubated at 37°C in a rolling incubator for 3-5 days before use. The Erlenmeyer flasks were incubated at 37°C in a shaking incubator (at 100rpm) for 3-5 days prior to use. Samples were taken from *Mtb* cultures 24h prior to experiments being performed and were Auramine O and gram stained to check for contamination. Cultures were discarded after 5 passes.

2.3.2.2 Enumeration of Colony-Forming Units

Colony-forming unit (CFU) counting was carried out by the drop plate method, a modified version of the Miles and Misra (surface viable count) method. Serial decimal (ten-fold) dilutions of the cell suspension were made in aliquots of 7H9 broth with Tween (without ADC supplement). Plates of Middlebrook 7H10 agar were divided into numbered sectors, and three 20µl drops from each dilution were plated out in the appropriate sector on duplicate plates. The plates were sealed with laboratory sealing film (Nescofilm, Osaka, Japan), inverted and statically incubated in double bags at 37°C. Colonies were counted after 2-3 weeks using a dissection microscope at ×4 magnification. Plates were counted weekly until the colony counts stabilised.

Final counts of 10-100 colonies (averaged over the six replicate spots) were used for the final concentration of CFUs, using the formula:

$$\frac{CFU}{ml} = C \times 50 \times D$$

C = Average colony count per 20µl spot

D = Dilution factor

2.3.2.3 Measuring Optical Density

The optical density/absorbance of an *Mtb* culture was measured at a wavelength 580nm using a Jenway 6300 spectrophotometer. 1ml volumes of culture were transferred into 1.5ml cuvettes and were sealed with autoclave tape and laboratory sealing film. Thick cultures (OD_{580} greater than 1.0) were diluted 1:10 prior to measurement. Cultures were measured against a blank of the relevant culture media (7H9, Sautons or RPMI).

2.3.2.4 Nitric Oxide Exposure

Solutions of the nitric oxide donor, Spermine NONOate (SPER/NO, Axxora, Nottingham, UK) and its control compound, Spermine tetrahydrochloride (SPER.4HCL) were prepared to a concentration of 10mM in sterile distilled water. The two solutions were added to two separate exponential cultures of *Mtb* to give a final concentration of 100µM. SPER/NO was powder was handled under a flow of nitrogen gas due to its oxygen sensitive nature. The

solution was prepared immediately prior to use, and discarded after 30 minutes. The solution of SPER.4HCL was stored at -20°C.

All nitric oxide experiments were carried out in culture grown in 7H9 broth, using a final concentration of 100μM of SPER/NO or SPER.4HCL to match conditions by Daniel et al. (2004) and Sherratt (2008). Experiments were performed after 4 hours of nitric oxide exposure, yielding the highest levels of lipid bodies.

2.3.3 Generation of Lipid Bodies/Exposure to Stimuli

Mtb was subjected to nitric oxide as the main lipid body producing exposures. Other stimuli used included stationary phase incubation, cholesterol or oleic acid. A roller bottle or Erlenmeyer flask of *Mtb* culture was prepared as described above. The culture was incubated at 37°C for 4-5 days and growth was monitored by measurement of optical density.

An exponentially growing culture of OD_{580} of 0.4-0.6 was exposed to either of the two exposures or to the control condition for the appropriate time. Following exposure, aliquots were removed and frozen for lipid body assessment. All samples from the same experiment were batch thawed, immobilised and stained with Nile red or LipidTOX stain (see section 2.4). The rest of the lipid body positive cultures were used for the experiment in question.

2.4 Staining and sample preparation for fluorescence microscopy

2.4.1 Preparation of fluorescent probes, staining solutions and reagents

2.4.1.1 Nile Red

A stock of Nile Red (Molecular Probes, Invitrogen) was prepared at 0.5mg/ml in ethanol. The stock solution was filtered with a syringe filter (2μm, Acrodisc) to remove any crystals in solution. The stock bottle was wrapped in aluminium foil and stored at -20°C. The stock was further diluted in ethanol to make a 10μg/ml solution for staining cells directly. The working solution was stored at room temperature in a sealed amber bottle.

2.4.1.2 LipidTOX Red

A stock of LipidTOX™ Red neutral lipid stain (Invitrogen) was stored at -20°C in the dark. Prior to use, the stock was defrosted and diluted 1:50 (v/v) in PBS.

2.4.1.3 Auramine O

Auramine O solution was prepared as below:

Solution 1: 0.1g of Auramine O powder was dissolved in 10ml 95% ethanol.
Solution 2: 3.0g of phenol crystals were dissolved in 87ml distilled water.

The two solutions were mixed and stored in a tightly stoppered amber bottle, away from heat and light. The solution was stored at room temperature for up to 3 months.

2.4.1.4 Acid Alcohol

Acid alcohol was prepared by adding 0.5ml concentrated hydrochloric acid to 100ml of 70% ethanol.

2.4.1.5 Potassium Permanganate

A 0.5% (w/v) aqueous solution of potassium permanganate was prepared by dissolving 5.0g of potassium permanganate in 1L of distilled water. The solution was stored at room temperature in a tightly sealed bottle.

2.4.2 Labelling and Staining Protocols

2.4.2.1 Auramine O Staining of M. tuberculosis

Mtb was immobilized onto coverslips as described in 2.4.3.1. The smears were flooded with Auramine O solution for 15 minutes, rinsed with water and decolourised by flooding with acid alcohol for 15 minutes. The slides were rinsed with water and covered in 0.5% (w/v) potassium permanganate solution for 1 minute to reduce non-specific background staining and rinsed thoroughly with water. The slides were blotted and allowed to dry. The smears were mounted in PBS and glass coverslips were applied and sealed with clear nail varnish.

2.4.2.2 Labelling of Mycobacterial Cells with Nile Red

For *Mtb* in pure culture, bacilli were immobilised to coverslips prior to staining, as described in 2.4.3.1. Coverslips were placed in foil wrapped Petri dishes and were washed three times in PBS. The coverslips were covered in 10µg/ml Nile red solution in ethanol for 10 minutes and washed in three further changes of PBS. Coverslips were dried prior to mounting.

2.4.2.3 Labelling of Mycobacterial Cells with LipidTOX

For *Mtb* in pure culture, bacilli were immobilised to coverslips prior to staining, as described in 2.4.3.1. Coverslips were overlaid in the diluted LipidTOX™ stain. The coverslips were covered and incubated in the dark at 37°C for 30 minutes. Following staining, the coverslips were rinsed with water and dried, prior to mounting.

2.4.3 Preparation of slides and coverslips for microscopy

2.4.3.1 Immobilisation of bacteria onto glass coverslips

Mtb bacilli were immobilised on 19mm glass coverslips using a universal tube silicone chamber system developed by Walker. The universal chamber system was constructed as in Figure 9 and 50µl volumes of cell suspension were dispensed into each of the wells. The universal chamber was sealed with laboratory sealing film and centrifuged at 1000 × g for 10 minutes (Mistral 3000L centrifuge). The sealed system avoided the need for fixation of the bacilli prior to immobilisation on coverslips. After centrifugation, the supernatant was removed from each of the wells prior to the block being dismantled. Coverslips were subjected to formaldehyde fixation (see below) prior to removal from the containment suite. The coverslips were dried, mounted on a glass slide in 10% glycerol (v/v) in PBS and sealed with clear nail varnish.

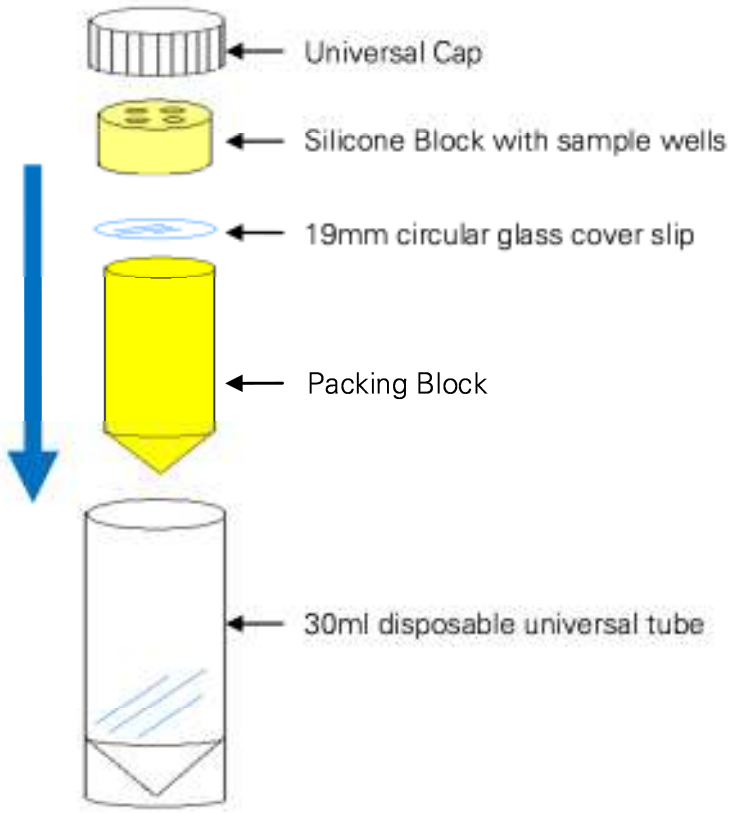


Figure 9: Assembly of Universal Tube Silicone Chamber System for the Immobilisation of *M. tuberculosis* onto coverslips

The silicon chamber system was assembled as demonstrated by the blue arrow. *Mtb* culture was dispensed into one of the four holes located on the silicone block. Image adapted from (Sherratt, 2008).

2.4.3.2 Formaldehyde treatment of coverslips and slides

Slides and coverslips of immobilised *Mtb* bacilli were treated with formaldehyde to allow for their safe removal from the containment suite. The slides/coverslips were placed in a glass slide box. The open slide box was placed within a larger plastic box, alongside a tissue soaked in 23% (w/v) formaldehyde. The plastic box was sealed, and the slides/coverslips were fixed overnight before being removed from the microbiological safety cabinet.

2.4.4 Recording of Fluorescence Images

Stained *Mtb* coverslips were mounted in PBS for microscopy. Smears were viewed using a Nikon Diphot 300 inverted microscope with a 100W mercury light source. Cells were visualised using a 100X magnification oil immersion lens using immersion oil (Citifluor). Images were recorded using a 12/10bit, high speed Peltier-cooled CCD camera (FDI, Photonic Science, East Sussex, UK) using Image–Pro Plus (Media Cybernetics, U.S.A) software. A 11001V2 Blue (excitation: 470±40nm; emission: >515nm, Chroma Technology Corporation) and a G-2A (excitation: 510-560nm; emission: 590±10nm, Nikon) filter sets were used for epifluorescence microscopy.

During the course of this study, the microscope was updated and images were subsequently recorded using a 12bit, high speed CCD camera (Retiga Exi Fast 1394, QImaging, Surrey, BC Canada) using InVivo v3.2.2 build 48 (Media Cybernetics, U.S.A) software. A Chroma 62001 Filter set was used (DAPI excitation: 350±25nm; emission: >445-471nm; FITC, excitation: 480±7.5nm; emission: >513-546nm; TRITC, excitation: 4597±12.5nm; emission: >597-655nm.

2.4.5 Acquisition of Fluorescence Intensity from Fluorescence Images

Fluorescence intensity data was acquired from fluorescence images using the method developed by Andrew Bell at the University of Leicester (Bell, 2010). Images were blinded using ImageJ (National Institutes of Health, Besthda, Maryland). For each phase image, regions of interest were identified by ImageJ through thresholding; the region of interest was applied to the corresponding fluorescent image and the fluorescence was measured (Bell, 2010).

The R Project Software Environment for Statistical Computing (R Development Core Team, GNU General Public License and The University of Auckland, New Zealand) was used to filter out non-cellular fluorescence information; this was performed through the identification of single cells using a mathematical equation (Bell, 2010). Fluorescence intensity was presented as a factor of area (cell size), i.e. grayscale units per unit area (μm^2) (Bell, 2010).

2.5 Ribonucleic Acid (RNA) Manipulation Techniques

2.5.1 RNA Precautions

RNA manipulations were carried out using precautions to minimise the degradation of RNA by ribonuclease (RNase) contamination. Gloves were worn for all manipulations and were changed frequently. New sterile disposable plastic ware and RNase-free filter pipette tips were used, and all equipment and work areas were treated with RNase-Zap (Ambion) to remove surface RNase decontamination. Dedicated RNase free reagents stocks were used and kept separate from normal laboratory reagents. RNA was stored at -80°C and manipulations were carried out on ice to reduce RNase activity.

2.5.2 Preparation of Guanidine Thiocyanate (GTC) Solution

Guanidine thiocyanate (GTC) solution was prepared by mixing the following reagents with 200ml double-distilled water:

Guanidine thiocyanate	295.4g
Sodium N-lauroyl sarcosine	2.5g
1M sodium citrate (pH 7.0)	12.5ml
Tween-80	5ml

The solution was incubated overnight at 37°C statically to allow the GTC to fully dissolve. The solution was made up to a final volume of 500ml with double-distilled water. Immediately prior to use, 7µl/ml β-mercaptoethanol was added to the GTC solution. GTC solution was stored in a sealed, glass bottle, away from heat and light.

2.5.3 RNA extraction from *M. tuberculosis* H37Rv

RNA was extracted from 5ml of aliquots of *Mtb* culture. GTC solution was added at a ratio of four parts to one part culture volume to stabilise the RNA. The cells were pelleted at 2000 × g for 30 minutes and the supernatant was removed. The cell pellet was resuspended in 200µl GTC solution and is transferred to a 1.5ml screw-top micro-centrifuge tube. The suspension was pelleted at 8000 × g for 5 minutes and the supernatant was removed using fine-tipped filter pipette tips. The pellet was overlaid with 1ml of TriReagent/Trizol (Sigma) and the samples were stored at -80°C. Trizol works by maintaining RNA integrity while at the same time disrupting and breaking down cells and cell components.

The samples were defrosted and ceramic beads (Lysing Matrix A, Q-Biogene, UK) were added to the Trizol and cells. The bacteria were lysed using a reciprocal shaker (Ribolyser, FastPrep FP120, ThermoSavant) set at 6.5 m/s (vertically) for 45 seconds. The samples were allowed to cool at room temperature for 10 minutes before the addition of 200µl of chloroform. The samples were vortex-mixed and centrifuged (Heraeus FRESCO 17 centrifuge, Thermo Electron Corporation) at 16000 × g for 3 minutes to separate the aqueous and organic phases. The aqueous phase containing RNA was transferred to a new tube; care was taken to avoid carry-over of interphase material. The aqueous phase was washed and re-extracted with the addition of an equal volume of chloroform. The tubes were centrifuged to separate the phases, the aqueous phase was transferred to a new tube and an equal volume of isopropanol was added. Glycoblue co-precipitant was added to each tube to a final concentration of 150µg/ml and the tubes were mixed by inversion. The RNA was precipitated overnight at -20°C.

Following precipitation, the tubes were centrifuged at 16000 × g for 10 minutes at 4°C to pellet the RNA and glycoblue co-precipitant (blue pellets). The supernatant was removed, and the pellets were washed once with 70% (v/v) ethanol and once with 95% (v/v) ethanol. The ethanol was removed and the pellet was allowed to air dry before resuspension in 100µl of RNase-free water. Purification of RNA samples is documented in Chapter 3.

2.5.4 RNA quantification by spectroscopy

The concentration of RNA was estimated by spectroscopy (Nanodrop – Thermo Scientific). A 1µl sample of each RNA extract was placed on the Nanodrop and the absorbance at 260nm was measured. RNase-free water was used as a blank.The concentration of RNA was calculated using the following equation (A_{260}=1 is equivalent to 40µg/µl):

$$C = \frac{\mathrm{A} \times 40 \times \mathrm{D}}{1000}$$

C = Concentration (µg/µl)

A = Absorbance at 260nm

D = Dilution factor

1000 = Correction (converts ml to µl)

2.5.5 Reverse-Transcription Polymerase Chain Reaction (RT-PCR) using Genome-Directed Primers

Extracted RNA was reverse transcribed to produce complementary DNA (cDNA). Reverse-transcription was performed using SuperScript II Reverse Transcriptase (Invitrogen) and Mycobacterial genome-directed primers (mtGDPs), which were designed to amplify all known open reading frames of the *Mtb* H37Rv genome. The mtGDPs, consisting of 37 heptamers and octamers, were used to prime for first-strand cDNA synthesis. The following reagents were added to a RNAse-free PCR tube on ice:

RNA	x µl*
dNTPs (10mM)	1.5µl
mtGDPs (10pmol/µl)	1.5µl
H_2O	up to 18µl

*The volume of RNA suspension used in the reaction varied on a per sample basis. If available, 0.5µg of RNA was used per reaction. However, if the concentration was lower, half the available RNA was used for the reaction and the volume made up with water. The concentration of extracted RNA was divided by 0.5 to give the volume in µl to be added to the reaction.

The mixture was heated to 65°C for 5 minutes and then snap-cooled on ice to disrupt the secondary structure of the RNA. The following reagents were added to the reaction:

5X Superscript II reaction buffer (Invitrogen)	6µl
DTT 0.1M	3µl
RNasin Ribonuclease Inhibitor(Promega) (20units/µl)	1.5µl

The mixture was incubated at 25°C for 2 minutes to allow for the primers to anneal and then placed directly on ice. Finally, 1.5µl (300units) of SuperScript II reverse transcriptase was added to each reaction, prior to the final incubation steps. A duplicate reaction without the addition of reverse transcriptase was run in parallel to the test sample. This acted at a "no-RT" control, to monitor the presence of genomic DNA (gDNA) contamination. The tubes were incubated at 25°C for a further 10min, then at 42°C for 50min to allow for reverse transcription of the RNA to occur. The reaction was inactivated by incubation at 70°C for 15min.

Following reverse transcription, samples were diluted 1:4 (v/v) with water. The samples were stored at -20°C.

2.6 Statistical Analysis

Calculations and standard deviations were calculated using Excel 2010 (Microsoft Corp.). Complex statistical analyses were performed using Prism 5 (GraphPad Software, Inc.) statistical software. Heat map and cluster analysis was performed using the R Project Software Environment for Statistical Computing.

Chapter 3

Development of Quantitative Real-Time Reverse Transcription – Polymerase Chain Reaction Assays for mRNA Quantification

3.1 Introduction

Quantitative, real-time reverse transcription - polymerase chain reaction (qRT-PCR) is a powerful method for the amplification, detection and quantification of trace amounts of messenger RNA (mRNA) (Heid et al., 1996). The method allows for accurate expression profiling of selected genes. However, the procedure is complex and problems may be associated with its true sensitivity, reproducibility and specificity (Pfaffl et al., 2002).

Two quantification strategies are generally used: absolute quantification and relative quantification/expression. In absolute quantification, the qPCR signal is related to input mRNA copy number through the use of a calibration curve; calibration curves may comprise a variety of samples, including DNA, RNA or plasmids (Pfaffl, 2006, Pfaffl et al., 2002, Pfaffl and Hageleit, 2001). Relative quantification examines the expression ratio of a target gene compared to a reference gene; it demonstrates the changes of mRNA levels of a gene across multiple samples, relative to levels of another gene transcript. (Pfaffl, 2006, Livak and Schmittgen, 2001). For the purposes of this study, the term 'expression' indicates the transcription of a messenger RNA (mRNA) product, but not necessarily a functional protein or gene product.

Absolute quantification is useful in situations where determination of absolute transcript copy number is necessary (Livak and Schmittgen, 2001). This may include the comparison of expression data between different days and laboratories, as a calibration curve is generally non-changing (Pfaffl et al., 2002). However, the reliability of absolute quantification requires identical amplification efficiencies for the sample RNA and the product used in the calibration curve (Pfaffl, 2006). In other cases, it may be unnecessary to quantify absolute copy numbers and relative expression change is adequate.

Relative quantification does not require a calibration curve and any transcript can be used as a reference (Pfaffl, 2006); the units of relative expression cancel out, and these relative quantities can therefore be compared across multiple experiments (Pfaffl, 2006, Livak and Schmittgen, 2001). Absolute quantification can also be calculated to show relative expression, but requires extra and unnecessary steps to do so.

Both absolute and relative quantification methods have merits and disadvantages and their selection and use depends on the requirements of the experiment being conducted.

3.1.1 Normalisation

To normalise is 'to bring or return to a normal or standard condition or state' (Oxford Dictionaries, 2010). In terms of gene expression data in PCR assays, the normalisation of data allows for comparison between experiments, by controlling for variations in starting genetic material and cell transcriptional activity; it reduces the effect of experimental error by standardising variations in cycling performance and enzymatic efficiencies during qRT-PCR (Pfaffl, 2006).

The material from different biological samples can vary in mass, cell number, RNA integrity and quantity or experimental treatment (Wong and Medrano, 2005). However, the calculation of relative expression must be made on a constant measure of RNA or amount of cells (Pfaffl et al., 2002). Ideally, mRNA levels would be standardised to the total cell number or total amount of extracted RNA or gDNA; unfortunately, due to uncontrollable variations in the production of accurate cell counts, and extraction and purification of RNA or DNA, this type of normalisation is often impractical (Wong and Medrano, 2005). Thus, to avoid these problems, the results of real-time PCR results may be normalised against a selected reference gene (Pfaffl et al., 2004, Suzuki et al., 2000). This normalisation can be incorporated into the relative quantification mathematical models – reviewed in 3.1.2).

An ideal reference gene should be expressed constitutively, regardless of changes to experimental conditions (Wong and Medrano, 2005). Therefore, the choice of reference gene is critical. Any variation in the expression of the reference gene may obscure genuine expression changes and produce calculated artefacts (Bustin et al., 2005).

Previously, genes thought to have stable expression, known as housekeeping genes, have been used as controls in gene expression assays (Wong and Medrano, 2005). Many housekeeping genes however, are affected by certain experimental treatments or biological processes, and no single gene is constitutively expressed in all cells and under all experimental conditions (Suzuki et al., 2000, Thellin et al., 1999). Therefore, when using a single housekeeping gene, the validation of its constitutive expression under the experimental test condition is important.

In recent years, there has been the increasing application of software tools for the evaluation of reference gene expression for accurate normalisation (Pfaffl et al., 2004, Vandesompele et al., 2002). These tools can be used to determine the most stable set of reference genes (from a tested set) and calculate a gene expression reference index. The use of a reference index provides an alternate normalisation strategy to the single housekeeping gene (Vandesompele et al., 2002).

Ribosomal RNA (rRNA) has also been used as a reference gene for normalisation (Wong and Medrano, 2005). However, there are problems associated with the use of rRNA to normalise mRNA gene measurements. While the polymerases for rRNA and mRNA are the same in prokaryotes, changes in polymerase activity may not affect both types of RNA equally (Wong and Medrano, 2005).

Thus, the optimum normalisation method, using single or multiple housekeeping genes, or ribosomal RNA should be examined further.

3.1.2 Mathematical Models

The basic equation describing PCR amplification is shown in Equation 1 (Rutledge and Cote, 2003); N_C represents the number of amplicon/gene molecules or copy number, N_0 is the number of initial target/gene molecules or copy number, E is the fractional amplification efficiency and C is the number of PCR cycles. Assuming an efficiency of 1 (100%), each PCR cycle produces a doubling in the number of molecules/copy numbers (Rutledge and Cote, 2003). Further, after any number of cycles, the quantity of N_C is dependent on N_0; by rearranging the equation, N_0 may be calculated if N_C is quantified and amplification efficiency is known.

Equation 1: PCR Amplification

$$N_C = N_0 \cdot (E+1)^C$$

$$N_0 = \frac{N_C}{(E+1)^C}$$

For comparisons between reactions, the calculation of N_0 is simplified by examining each reaction at a point when it contains a set amount of amplified DNA. A qRT-PCR reaction produces a fluorescence curve as the gene product is amplified; a typical fluorescence curve is displayed below in Figure 10. To examine a point when samples contain a set amount of amplified DNA, a fluorescence threshold is chosen, from which a fractional threshold cycle (Ct) value is calculated; this Ct is defined as a theoretical point at which the fluorescence threshold is reached by each amplification reaction (Livak and Schmittgen, 2001, Rutledge and Cote, 2003).

Using the fluorescence threshold method, N_C becomes a constant, as displayed in Equation 2; N_t is the number of molecules/copies at the fluorescent threshold, and Ct is threshold cycle number.

Equation 2: Calculation of Starting Number of Molecules using Fluorescence Threshold

$$N_0 = \frac{N_t}{(E+1)^{Ct}}$$

Absolute quantification may be performed by comparing samples to a standard curve constructed from the amplification and quantification of known dilutions of DNA or RNA (Wong and Medrano, 2005). The standard curve is derived by taking the logarithm of Equation 2, as displayed in the line one of Equation 3. The equation can be rearranged (Equation 3, line two) into the general structure of a line: y = mx + b.

Equation 3: Absolute Quantification (Standard Curve)

$$Log(N_0) = Log(N_t) - Log[(E+1)^{Ct}]$$
$$Log(N_0) = Log(N_t) - Log(E+1) \cdot Ct$$
$$Log(N_0) = -Log(E+1) \cdot Ct + Log(N_t)$$

In absolute quantification, N_t is constant, and amplification efficiency is assumed to be consistent in all individual amplification reactions. With these constants, a plot of $Log(N_0)$ versus Ct yields a line where:

Slope = -log(E + 1),
$E = 10^{-Slope} - 1$
Intercept = $log(N_t)$ (Rutledge and Cote, 2003)

This is displayed visually in Figure 10.

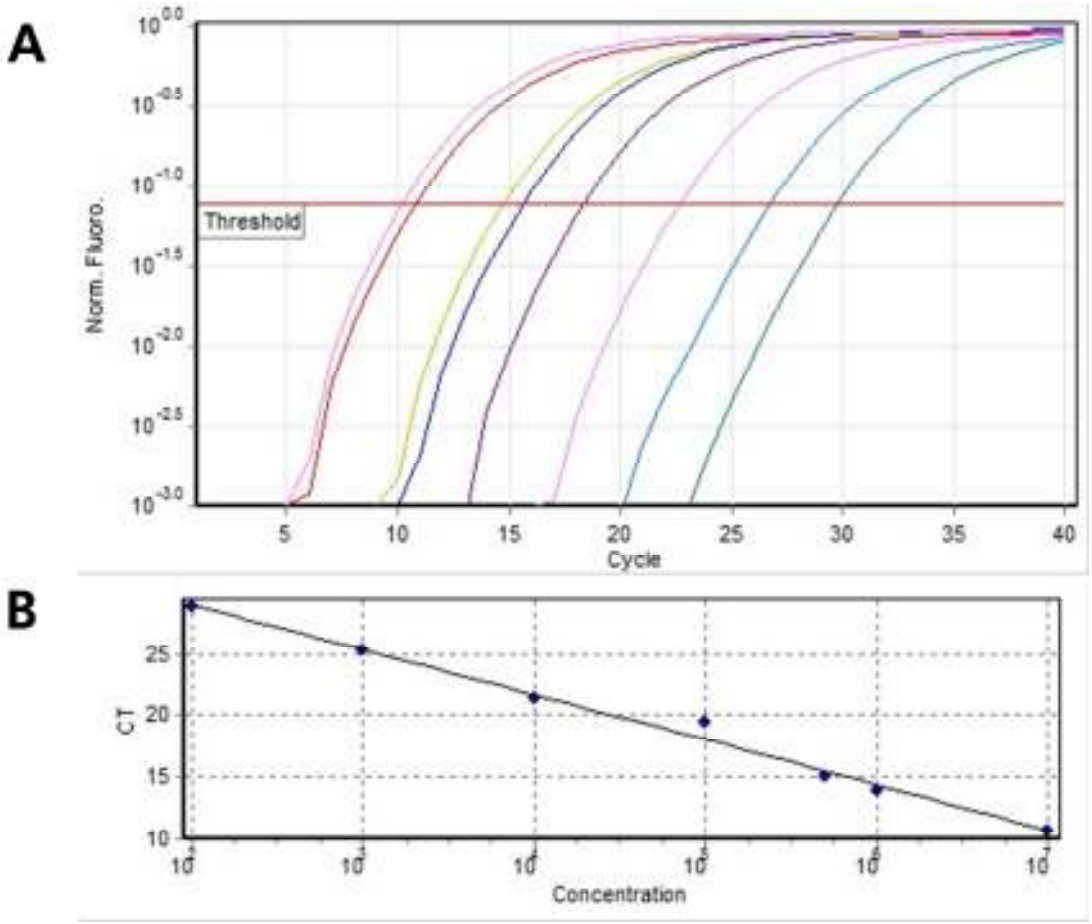


Figure 10: Fluorescence output of a typical amplification run (A) and a constructed standard curve (B)

Figure 10A displays a fluorescence output of a typical amplification run, with a separate reaction conducted for each concentration of DNA standards; seven dilutions of DNA concentration were included. The plot of Log Fluorescence versus Cycle number displays the amplification runs graphically. The threshold line is the reference point at which the threshold cycle is taken. Figure 10B is the standard curve generated by graphing the threshold cycle versus the log of the DNA concentration. The standard curves are only valid if all the individual PCR reactions have similar amplification efficiency and if this efficiency remains constant over the number of cycles it takes to reach the fluorescence threshold (Rutledge and Cote, 2003).

A variety of mathematical models have also been developed to calculate relative expression ratios (R) of single samples (Pfaffl, 2006). The relative expression of a chosen gene is calculated in relation to another, as discussed above; this can be achieved using the comparative Ct, or 'delta delta Ct' (ΔΔCt) method (see Equation 16 in Appendix 2 for equation derivation) (Pfaffl, 2006, Livak and Schmittgen, 2001). The method does not include efficiency correction, and thus, it is assumed that an optimal doubling of DNA occurs during each real-time PCR cycle (Livak and Schmittgen, 2001). Variations in efficiency lead to the generation of errors whilst using this method.

The relative expression may also be calculated for single samples with correction for less than optimal amplification efficiencies (Souaze et al., 1996). The calculation is displayed as Equation 4. The equation is based upon Equation 15 of the 'delta delta Ct' derivation, but incorporates the actual amplification efficiencies, as opposed to assuming perfect efficiency. The model incorporates both gene quantification and normalisation in a single calculation.

Equation 4: Pfaffl Model - Efficiency Corrected Relative Expression Calculation Models, based on One Sample

$$Ratio = \frac{(E_{target})^{\Delta Ct_{target}\,(control\,-\,sample)}}{(E_{reference})^{\Delta Ct_{reference}\,(control\,-\,sample)}}$$

Further, the efficiency adjusted relative expression calculation model may be modified based on multiple samples, as displayed in Equation 5, by using the mean delta-Ct of the target and reference for the control and sample (Pfaffl et al., 2004). The equation can be further modified by comparing the gene of interest to multiple reference genes, by creating a reference index, which consists of three or more reference genes (Pfaffl et al., 2004, Vandesompele et al., 2002).

Equation 5: Efficiency Corrected Relative Expression Calculation Models, based on Multiple Samples

$$Ratio = \frac{(E_{target})^{\Delta Ct_{target}\,(MEAN\ control\,-\,MEAN\ sample)}}{(E_{reference})^{\Delta Ct_{reference}\,(MEAN\ control\,-\,MEAN\ sample)}}$$

In each of the relative expression models, the target gene is normalised to the expression of one or multiple reference genes. It must therefore be confirmed that the reference or housekeeping gene used has stable expression and is not influenced by any applied treatment.

Whilst the mathematical models of qRT-PCR have been significantly reviewed in the literature, the best model for the studies presented in this thesis must be analysed and verified.

3.1.3 Reaction Efficiency

Amplification reactions (including replicates) often result in significantly different amplification curves, due to both sample-to-sample and run-to-run variations (Pfaffl, 2006). These variations result from variations in reaction efficiency during each run. Efficiency has a major effect on amplification kinetics and is influenced by all PCR reaction reagents (e.g. primers, DNA polymerase, etc.) (Pfaffl, 2006). The maintenance of constant reaction efficiencies between samples is desirable for the reliable comparison between samples and the minimisation of calculated errors; this is crucial for the analysis of the relationship between test and reference samples, the basis of all relative quantification (Pfaffl, 2006). Therefore, mathematical models that correct for variations in efficiency (Equation 4 and Equation 5) are recommended.

For efficiency to be corrected, it must first be calculated on the basis of fluorescence increase in the exponential phase of the PCR (Pfaffl, 2006). It may be calculated directly by examining a standard curve for the fluorescence increase, or indirectly, by fitting fluorescence amplification data in the form of the fluorescence curve to a mathematical model, such as the sigmoidal or logistic models (Pfaffl, 2006). The direct methods have the advantage of being independent of background fluorescence. Background fluorescence often interferes with the indirect data fit into mathematical models, and this is the case with SYBR green assays; SYBR green binds indiscriminately to the small groove of the double-stranded DNA (dsDNA) helix and thus show low and constant background fluorescence (Lutfalla and Uze, 2006, Pfaffl, 2006).

Efficiency is often calculated directly by the 'Dilution method' using data from a standard curve (as in Figure 10) using Equation 6 (Wong and Medrano, 2005).

Equation 6: Dilution Method of Efficiency Calculation

$$E = 10^{-1/slope}$$

For example, an image of a typical standard curve is shown in Figure 10. If the slope = -3.5, the efficiency is calculated as in Equation 7.

Equation 7: Example of Dilution Method of Efficiency Calculation

$$slope = -3.5$$
$$\therefore E = 10^{-1/-3.5} = 1.93$$

Therefore, an efficiency of 1.93 calculated, demonstrating that after each reaction cycle, there is 1.93 times the number of molecules, as compared to the previous cycle. This efficiency is may be expressed alternatively as fraction or a percentage, e.g. 0.93, or 93%.

Alternatively, efficiency can be calculated from the fluorescence increase in the exponential phase of a log scale plot of fluorescence history. As this is a direct method, the analysis is independent of background fluorescence. This efficiency calculation is most reliably performed using a software application such as LinRegPCR (Figure 11) (Ramakers et al., 2003, Ruijter et al., 2009).

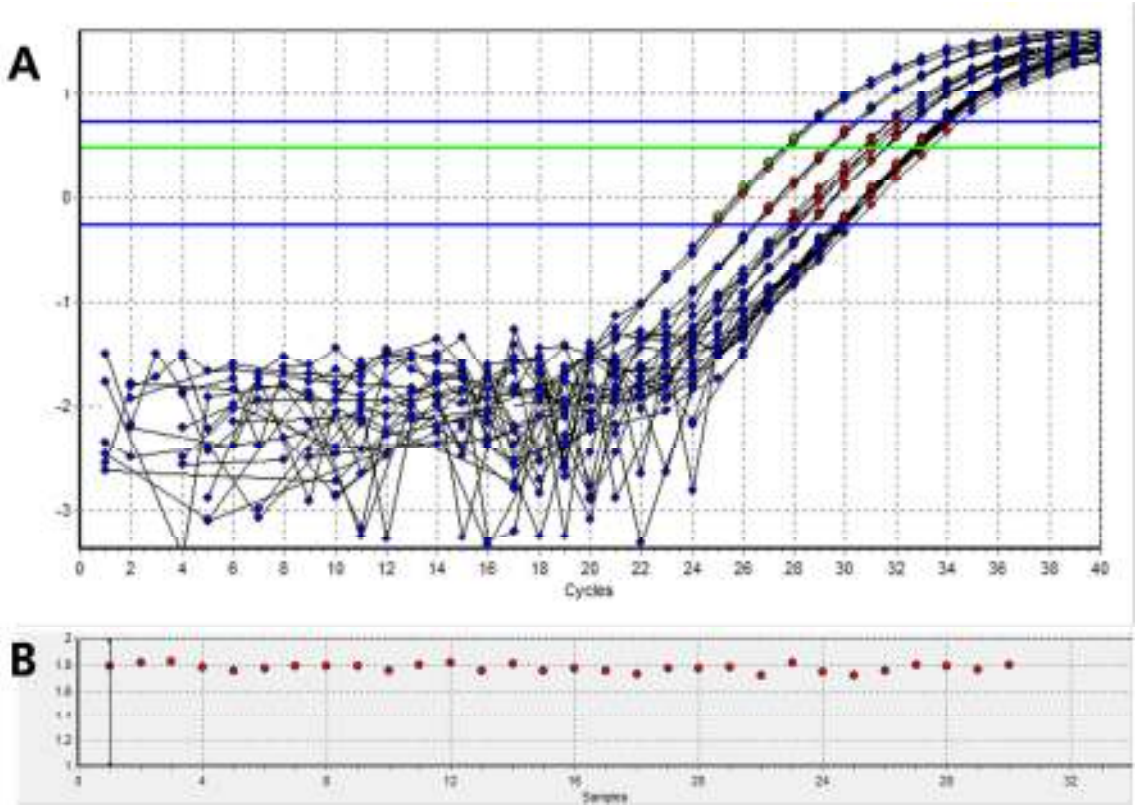


Figure 11: Example Analysis Screen from LinRegPCR

LinRegPCR produces a linear regression plot that is drawn from fluorescence data points (data is acquired after each PCR reaction cycle). The slope of this regression line is calculated between an automatically selected 'Window of Linearity' (WoL), represented by the blue lines in Graph (A), and remains the same for all sample reactions in a given PCR run. This resulting slope of the regression line represents the PCR efficiency and is displayed for each sample reaction in Graph (B). In addition, LinRegPCR automatically calculates baseline (i.e. background) fluorescence (Ramakers et al., 2003, Ruijter et al., 2009).

3.1.4 Genomic DNA Contamination

While RT-PCR is an extremely powerful tool for the analysis of gene expression in bacteria, the majority of techniques used for total RNA isolation yield RNA samples that contain large amounts of DNA contamination (Ambion Incorporated, 2007b, Matthews et al., 2002). Unfortunately, PCR assays are unable to discriminate between reverse-transcribed complementary DNA (cDNA) and gDNA and therefore, the gDNA contamination must be minimised to prevent inaccuracies in obtained results (Tondeur et al., 2004).

A simple method for the detection of gDNA is the inclusion of a "no-RT" control during reverse-transcription steps (Ambion Incorporated, 2007b). RNA is prepared and processed as for a reverse transcription reaction, but the reverse transcriptase is left out of the overall reaction. Thus, the RNA is not reverse transcribed and cDNA is not produced, leaving only the contaminant gDNA.

Conventionally, DNase I treatment can be used to digest gDNA from RNA extracts (Ambion Incorporated, 2007c). However, opinion within the scientific community is split as to whether DNase I treatment completely solves the problem of DNA contamination. It has been suggested that inactivation steps of DNase I can compromise the performance of RT-PCR reactions (Ambion Incorporated, 2007b). Matthews et al. (2002) concluded that DNase I is insufficient for removing DNA from RNA samples and can degrade RNA.

Previous results in this lab have demonstrated varying levels of DNA contamination in RNA samples from acid-fast bacilli positive sputum depending on the mRNA target (Lee, 2007). Examination of *tgs1* and *icl1* transcripts revealed 13.5% and 1.3% gDNA contamination respectively (Lee, 2007). This variation is likely to be dependent on both variations in efficacy of DNase I, in addition to the gene expression level.

While DNase I was perceived to cleave DNA non-specifically, it has been recognised to show some sequence preference. This pattern recognition is often limited to sequences of three or four nucleotides and DNase I often prefers purine-pyrimidine sequences (Ambion Incorporated, 2007a, Campbell and Jackson, 1980). Further, while DNase I has been shown to generally cleave dsDNA, variations in the metal cation activators present in solution can lead to variations in the kind of cleavage (i.e. single-stranded nicks as opposed to double-stranded cuts) (Campbell and Jackson, 1980). DNase I is also sensitive to the structure of the minor groove; any distortions in the groove can significantly affect its ability to cleave DNA (Ambion Incorporated, 2007a).

To establish the most effective DNA digestion methods, they will be studied further, specifically using mycobacterial RNA extracts and the mycobacterial genes to be analysed in future.

3.1.5 Aims

The development of quantitative real-time RT-PCR assays for the quantification of gene transcripts in *M. tuberculosis* (*Mtb*) is discussed in this chapter, including:

- Comparison and selection of absolute or relative quantification.
- Analysis of real-time qPCR reaction efficiencies.
- Selection of internal control genes and calculation of a reference index.
- Analysis and selection of genomic DNA removal protocols.
- Validation of qPCR primers and cycling conditions.

3.2 Materials and Methods

3.2.1 Exposure of *M. tuberculosis* to Phosphate-buffered saline

Phosphate-buffered saline (PBS) was prepared as described in 2.2.2. The PBS was pre-warmed to 37°C. The 30ml *Mtb* H37Rv culture was transferred to a 50ml centrifuge tube and centrifuged for 10 minutes at 3200 x g. The supernatant was removed and the bacterial pellet was resuspended in an equal volume PBS.

3.2.2 Exposure of *M. tuberculosis* to RPMI-1640 media

Complete RPMI-1640 media was prepared as described in 2.2.2. The complete RPMI-1640 media was pre-incubated to 37°C. The 30ml *Mtb* H37Rv culture was transferred to a 50ml centrifuge tube and centrifuged for 10 minutes at 3200 x g. The supernatant was removed and the bacterial pellet was resuspended in an equal volume RPMI-1640 media.

3.2.3 Preparation of SDS solution

5% (w/v) sodium docedyl sulphate (SDS) solution was prepared by dissolving 5g of SDS powder in 100ml of distilled water. The solution was heated to 37°C for 30 minutes to fully dissolve the SDS powder. The solution was filter sterilised (0.2µm filter) and stored at 4°C.

3.2.4 DNase Treatment Protocols

3.2.4.1 RNeasy Mini Column Cleanup of RNA

Extracted mycobacterial RNA was processed through an RNeasy Mini Column (Qiagen) cleanup according to the manufacturer's instructions. The RNeasy Mini Column was based on a high salt buffer system and a selectively-binding silica membrane; RNA was bound to the membrane, and other, non-binding contaminants were washed away.

Buffer RLT (350µl) was added to each 100µl sample of RNA and mixed well. To this mixture, 250µl of absolute ethanol was added and mixed thoroughly by pipetting. Each sample (700µl) was transferred to an RNeasy Mini spin column (placed in a 2ml centrifuge tube) and centrifuged at 9600 x g for 15 seconds. The flow through was discarded. Buffer RPE (500µl) was applied to each spin column. The columns were centrifuged for 15 seconds at 9600 x g

to wash the spin column membrane and the flow through was discarded. The columns were washed in the same manner with 500 µl Buffer RPE again, but with a centrifugation step of 2 minutes. The columns were transferred to a new collection tube, and centrifuged for a further 1 minute to remove excess buffer. The columns were transferred to fresh 1.5ml collection tubes and 50µl RNase free water was applied to each column to elute the bound RNA. The tubes were incubated at room temperature for 5 minutes prior to centrifugation for 1 minute at 9600 x g.

3.2.4.2 RNase-free On-column DNase Digestion

The RQ1 On-column DNase (Qiagen) digestion was applied as a supplemental treatment to the RNeasy Mini Column cleanup. The treatment was performed after the centrifugation of the 700µl sample (containing RNA, Buffer RLT and absolute ethanol) onto the spin column. Following this step, 350µl RW1 buffer was applied to each sample. The tubes were centrifuged for 15 seconds at 9600 x g and the flow through was discarded. The on-column digestion was performed by carefully adding a mixture of 10µl DNase I and 70µl RD1 buffer to each column membrane. This was incubated at room temperature for 15 minutes, following which, the samples were washed with the application of 350µl of Buffer RW1. The tubes were centrifuged for 15 seconds at 9600 x g. This was followed by the first Buffer RPE wash step in the RNeasy Mini Column cleanup protocol (3.2.4.1).

3.2.4.3 Turbo DNase Treatment

Turbo DNase (Ambion) treatment was applied to RNA samples adjusted to a volume of 50µl. To each sample, 0.1 volume (5µl) 10X Turbo DNase buffer and 1µl Turbo DNase were added. The sample was mixed gently and incubated at 37°C for 30 minutes. Another 1µl Turbo DNase was added to the reaction and the sample was incubated at 37°C for another 30 minutes. Following incubation, the sample was either added directly to an RNeasy Mini Cleanup Column or the DNase was inactivated. For DNase inactivation, 10µl of DNase inactivation reagent was added to each sample and mixed thoroughly. The mixture was incubated at room temperature (at least 22-26°C) for 5 minutes and mixed occasionally to re-disperse the inactivation reagent. The sample was centrifuged at 9600 x g for 3 minutes to pellet the DNase inactivation reagent, and the RNA was transferred to a fresh tube.

3.2.4.4 RQ1 DNase Treatment

RQ1 RNase-free DNase (Promega) treatment was applied to RNA samples adjusted to a volume of 42µl. To each sample, 5µl (0.1 volume) RQ1 RNase-free DNase 10X Reaction buffer, 1.25µl RNasin (Promega) RNase inhibitor and 2µl RQ1 DNase were added. The sample was mixed gently and incubated at 37°C for 30 minutes.

Following incubation, the sample was run through an RNeasy Mini Column to remove any trace of DNase and to further purify the RNA.

3.2.5 Quantification of RNA using NanoDrop™ 1000 Spectrophotometer

RNA was quantified using the NanoDrop™ 1000 Spectrophotometer at a wavelength of 260nm. 1µl of undiluted RNA suspension was placed on the NanoDrop™ pedestal. Measurement and calculation was automated by the spectrophotometer software. RNA suspension was measured against a blank of RNase-free water.

3.2.6 Primer Design

The selection of candidate reference/housekeeping genes was based on genes previously used in *Mtb* studies. The primer sets were based on gene sequences obtained from the TubercuList online database of the sequenced *Mtb* H37Rv genome (Institut Pasteur, 2004).

Primers were designed using the Lasergene (DNAStar) suite of DNA analysis software. EditSeq (DNAStar) was used for analysis of sequences. Primer Select (DNAStar) was used for selection of primer sequences from EditSeq, using specified conditions. Primer Select was also used to check for self-dimer, pair-dimer and hairpin formations.

Following primer design, the primer pairs were tested using *In silico* PCR amplification (University of the Basque Country, 2010) to determine that the correct band size was amplified and that the primers specifically amplified only one band within the entire *Mtb* genome.

3.2.7 Quantitative Real-Time Polymerase Chain Reaction using SYBR Green

Real-time Quantitative Polymerase Chain Reaction (qPCR) was performed on the cDNA produced during the RT-PCR step. Absolute™ SYBR QPCR Green Master-Mix (ABgene) was

used for the reactions. The Master-Mix was prepared according to manufacturer's instructions and stored in the dark at -20°C prior to use. The mix contained SYBR Green, reaction buffer, dNTPs, $MgCl_2$ and a hot-start DNA polymerase.

Strips of 4 0.1ml Rotor-gene (Corbett Research/Qiagen) tubes were used for the reactions. The following were added to each 0.1ml strip tube:

Template cDNA	1µl
Forward Primer	1µl
Reverse Primer	1µl
2X SYBR Green Master-Mix	12.5µl
H_2O	9.5µl
Total	**25µl**

Primer concentrations varied between each of the individual assays, to reduce primer-dimer formation.

Cycling conditions (including annealing temperatures) were common for all the reactions. The second acquisition temperature varied with differing primer sets (see Appendix 1). The qPCRs were run in a Rotor-gene 6000 machine (Corbett Research). The following cycling conditions were used for each reaction:

Temperature Hold	56°C	2 minutes
Temperature Hold	95°C	15 minutes
Cycling (40X)	95°C	30 seconds
	59°C	30 seconds
	72°C	20 seconds
	82-87°C	20 seconds
Melting	50-99°C	1°C per 5 seconds

Fluorescence acquisition was performed at the end of the 72°C and 82-87°C cycling steps. The latter acquisition temperature for each primer set is shown in Appendix 1. The higher acquisition temperature was included to help eliminate errors caused by primer-dimer formation. Reverse-transcribed cDNA sample reactions were run in triplicate and no-RT control reactions were run in duplicate.

For absolute quantification, a standard curve of gDNA was included with each run. A concentration of purified *Mtb* CDC1551 gDNA was obtained from the University of

Colorado. The DNA concentration was converted into copies per microlitre by using the Avogadro number (6.023×10^{23}) and the molecular weight of each genome copy (average number of bases in the genome multiplied by the average molecular weight of a pair of nucleic acids [660 Da]). 10 fold dilutions of this stock were kept in aliquots at -20°C and used as DNA standards. The DNA standards included in each run were: 10^7, 10^6, 5×10^5, 10^5, 10^4, 10^3 and 10^2 genome copy numbers per microlitre, respectively. The standard curve was constructed by plotting the threshold cycle (Ct) corresponding to each standard vs. the value of their corresponding number of genome copies (genome copies/µl). This process was automated by the Rotor-gene 6000 (Corbett) software.

3.2.8 Software Analysis

LinRegPCR software (Ramakers et al., 2003, Ruijter et al., 2009) and Rotor-gene 6000 Software Series 1.7 (Corbett Research) were used to analyse the efficiencies of PCRs through the direct analysis of the fluorescence application curve and analysis of a dilution series respectively. LinRegPCR software was also used to assign Ct values to each sample for relative quantification.

GeNorm software was used for the analysis and selection of housekeeping genes. A gene expression stability measure (*M*) was calculated by the program to determine the expression stability of control genes on the basis of non-normalised expression levels (Vandesompele et al., 2002). The measure was based on the average pairwise variation of one gene compared to all studied reference genes (Penning et al., 2007). Any variation in the expression ratios between two genes (over different growth conditions) was reflective of the fact that one (or both) of the genes were not constantly expressed. Genes with the lowest *M* values were deemed to have the most stable gene expression (Vandesompele et al., 2002).

Pairwise comparisons of genes or normalising factors were calculated using GeNorm (shown in Table 3). For the calculation of the gene stability measure, *M*, the pairwise comparison for a candidate housekeeping gene was calculated against each of the other candidate housekeeping genes for all the samples. The *M* value was the standard deviation of all the pairwise comparisons and was repeated for each candidate housekeeping gene (Vandesompele et al., 2002).

The *M* value stability ranking was used to calculate a normalisation factor. The normalisation factor was based upon the average of a selected number of control genes; the geometric mean was used for averaging of control genes, as opposed to arithmetic mean, as the former was

better for controlling for outliers and abundance differences between different genes (Vandesompele et al., 2002). The calculation of the geometric mean is shown in Equation 8.

Table 3: Pairwise Comparison of Genes/Normalising Factors, using GeNorm, for the Calculation of *M* values and Determination of Optimum Number of Reference Genes

	Gene A	Gene B	Pairwise Comparison
Sample 1	a_1	b_1	$Log_2(a_1/b_1)$
Sample 2	a_2	b_2	$Log_2(a_2/b_2)$
Sample 3	a_3	b_3	$Log_2(a_3/b_3)$
…	…	…	…
Sample n	a_n	b_n	$Log_2(a_n/b_n)$
			Standard Deviation of Pairwise Comparison = *V/M*

Equation 8: Geometric Mean

$$Geometric\ Mean = \sqrt[n]{a \cdot b \cdot c \cdot \ldots \cdot n}$$

GeNorm was further used to determine the number of reference genes needed to obtain an adequate normalisation factor (Vandesompele et al., 2002). The calculations in Table 3 were also used for this purpose. Normalisation factors were calculated based upon the *M* value stability ranking, starting with the most stable genes. The pairwise variation $V_{n,n+1}$ was calculated between two sequential normalisation factors (NF_n and NF_{n+1}), for all samples per growth medium, with n = 2 to 9. A pairwise variation *V*, of less than 0.15 demonstrated that the addition of an extra housekeeping gene had no significant effect on the normalisation factor (Vandesompele et al., 2002).

In addition to GeNorm, Normfinder software was used in the analysis and selection of housekeeping genes (Andersen et al., 2004). The stability of candidate genes was automatically generated by the software based upon a mathematical model of variation in gene expression (Andersen et al., 2004).

Microsoft Excel was used for calculation of relative gene expressions, standard deviations and co-efficiencies of variation. An example of the Excel spreadsheets used to calculate quantities for real time RT-PCR data is shown in Figure 12. All other statistics were performed using GraphPad Prism 5 (GraphPad Software) statistical software.

A

$$Raw\ Q = Mean\ E^{(Min\ Ct - Sample\ Ct)}$$ e.g. $D8$8 ^ (C6 - C2)

	A	B	C	D	E	F
1	Sample	Gene	Ct	Efficiency	Raw Quantitiy	Mean Raw Quantity
2	Exponential Phase	hspX	20.12	1.905	0.0146	**0.0151**
3	Exponential Phase	hspX	19.98	1.875	0.0177	
4	Exponential Phase	hspX	20.15	1.932	0.0130	
5	Stationary Phase	hspX	13.8	1.900	0.8572	**0.9323**
6	Stationary Phase	hspX	13.56	1.882	1.0000	
7	Stationary Phase	hspX	13.65	1.998	0.9396	
8			**Mean E**	**1.915**		
9						
10						
11	Sample	Gene	Ct	Efficiency	Raw Quantitiy	Mean Raw Quantity
12	Exponential Phase	thyA	25.22	1.855	0.7810	**0.7307**
13	Exponential Phase	thyA	25.34	1.824	0.7316	
14	Exponential Phase	thyA	25.46	1.829	0.6795	
15	Stationary Phase	thyA	24.88	1.833	0.9643	**0.9700**
16	Stationary Phase	thyA	24.82	1.860	1.0000	
17	Stationary Phase	thyA	24.91	1.861	0.9456	
18			**Mean E**	**1.844**		
19						

Min Ct for Gene

Efficiency values imported from LinRegPCR

B

	A	B	C	D	E	F	G	H	I
1	**Mean Raw Quantities**								
2		**Housekeeping**				**Target**			**Normalisation**
3	**Sample**	thyA	dfrA	polA	aroA	hspX	icl1		**Factor**
4	Exponential Phase	0.7307	0.8623	0.8122	0.6591	0.0051	0.1260		0.7621
5	Stationary Phase	0.9700	0.7206	0.7122	0.8532	0.9323	0.8104		0.8073
6									
7									

$$NF = \sqrt[4]{Q_{thyA} \times Q_{dfrA} \times Q_{polA} \times Q_{aroA}}$$

for each sample

C

$$Normalised\ Q = \frac{Raw\ Quantity}{NF}$$

J	K	L	M	N	O
	Normalised Quantities			**Fold Changes**	
	Target Genes			**Stationary/Exponential**	
	hspX	icl1		Gene	Fold Change
	0.0067	0.1653		hspX	172.57
	1.1549	1.0039		icl1	6.07

$$Fold\ Change = \frac{Test\ Normalised\ Q}{Control\ Normalised\ Q}\ \ e.g.\ \frac{Stationary\ Q}{Exponential\ Q}$$

Caption Overleaf

Figure 12: Example of Excel spreadsheets used to calculate relative quantities from qRT-PCR data

The figure shows example Excel (Microsoft) spreadsheets used to calculate relative quantities from qRT-PCR data using the Pfaffl relative quantification method. Spreadsheet (A) shows the calculation of raw quantities from the mean amplification efficiency and Ct values for both housekeeping genes (e.g. *thyA*) and target genes (e.g. *hspX*). The triplicate amplification reaction quantities are averaged. Spreadsheet (B) shows the calculation of the normalisation factor (i.e. the geometric mean of the raw quantities of *thyA, dfrA, polA* and *aroA*. Spreadsheet (C) demonstrates the normalisation of the target genes (e.g. *hspX* and *icl1*) by dividing the raw quantities by the normalisation factor. The fold changes are calculated by dividing the test sample normalised quantities (e.g. stationary phase culture) by the control sample normalised quantities (e.g. exponential phase culture).

3.3 Results

3.3.1 Comparison of Direct Efficiency Calculations

Variations in amplification reactions can result in significantly different fluorescence amplification curves (Pfaffl, 2006). Even small changes in efficiency can generate false expression ratios and therefore, efficiency evaluation is essential for all qRT-PCR experiments. The amplification efficiencies of both gDNA and cDNA were examined to identify the presence of any efficiency/amplification variation. The most appropriate efficiency calculation method was also determined by testing the two direct efficiency calculations: the dilution (standard curve) method and the 'fluorescence increase in the exponential phase' of PCR method.

RNA was extracted from exponentially growing *Mtb* H37Rv using a Trizol extraction method (2.5.3) and was treated with a RNeasy Mini Cleanup Column (3.2.4.1) and RNase-free On-column DNase (3.2.4.2). The RNA was reverse-transcribed in a RT-PCR reaction (2.5.5). The resultant cDNA was used in qRT-PCR reactions, using the primer sets for the 16S, *sigA*, *tgs1*, *hspX*, *ctaD* and *qcrC* genes.

Efficiencies were calculated directly using the 'fluorescence increase in exponential phase' method and were automated using the program LinRegPCR. The efficiencies calculated using the Dilution method were automated by the Rotor-gene 6000 software.

Table 4 shows the results for the direct efficiency calculations for 6 different primer sets (16S, *sigA, tgs1, hspX, ctaD* and *qcrC*) and calculated using fluorescence curve fit for gDNA and cDNA and using the dilution method. The table results were based upon the average efficiencies of one qRT-PCR run; each of the 6 gDNA dilutions (10^7, 10^6, 10^5, 10^4, 10^3 and 10^2) were run in triplicate (18 PCR reactions total) and 3 biological cDNA samples were run in triplicate (9 PCR reactions total). The no reverse transcriptase controls were included in each reaction run to exclude gDNA contamination. The average efficiencies plus or minus the standard deviations were calculated and included for the gDNA, cDNA and Dilution methods.

An unpaired Student's t-test was used to compare the means of cDNA efficiencies calculated by the two direct efficiency calculation methods: 'fluorescence increase during exponential phase' and the dilution method. The calculated means differed significantly for 16S ($p=0.0055$) and *tgs1* ($p<0.0001$). The means did not differ significantly for the other 4 primer sets: *sigA* ($p=0.4971$), *hspX* ($p=0.9105$), *ctaD* ($p=0.5932$) and *qcrC* ($p=0.1634$).

Table 4: Direct efficiency calculations for various samples and genes over two runs

Sample	# of Samples	Average Efficiencies					
		16S	***sigA***	***tgs1***	***hspX***	***ctaD***	***qcrC***
gDNA	18	1.79 ± 0.053	1.76 ± 0.065	1.88 ± 0.035	1.84 ± 0.035	1.90 ± 0.040	1.89 ± 0.052
cDNA	9	1.86 ± 0.064	1.74 ± 0.046	1.91 ± 0.013	1.83 ± 0.019	1.90 ± 0.054	1.91 ± 0.036
Dilution Method	3	1.70 ± 0.030	1.72 ± 0.035	1.79 ± 0.035	1.83 ± 0.13	1.92 ± 0.042	1.94 ± 0.014

The results were analysed using an unpaired Student's t-test to analyse any statistical variation in the mean between calculated efficiencies for gDNA and cDNA via 'fluorescence increase during exponential phase'. The values for 16S were significantly different (p=0.025), while the means were not significantly different for all the other samples: sigA (p=0.43), tgs1 (p=0.086), hspX (p=0.66), ctaD (p=0.81) and qcrC (p=0.63). This is displayed visually in Figure 13.

To determine the need for efficiency correction in the quantification calculations, the mean efficiencies of the six primer sets were compared using a one-way ANOVA. The mean gDNA efficiencies for the 6 primer sets were significantly different ($p<0.0001$). A similar p-value was calculated ($p<0.0001$) when comparing the means of the cDNA efficiencies across the 6 primer sets.

Significant differences were shown between the mean efficiencies of primer sets, suggesting that efficiency correction was mandatory for the correct calculation of expression change and reduction of error. While the differences in mean efficiency for gDNA and cDNA were only significant for one primer set (16S), the difference in means still varied up to 3% in the other primer sets. Further, the mean efficiencies calculated using the two direct calculation methods for 16S and *tgs1* differed significantly. It was deemed inappropriate to use gDNA in standard curves to calculate the efficiency of cDNA reactions and the 'fluorescence increase' method was used for all future efficiency calculations.

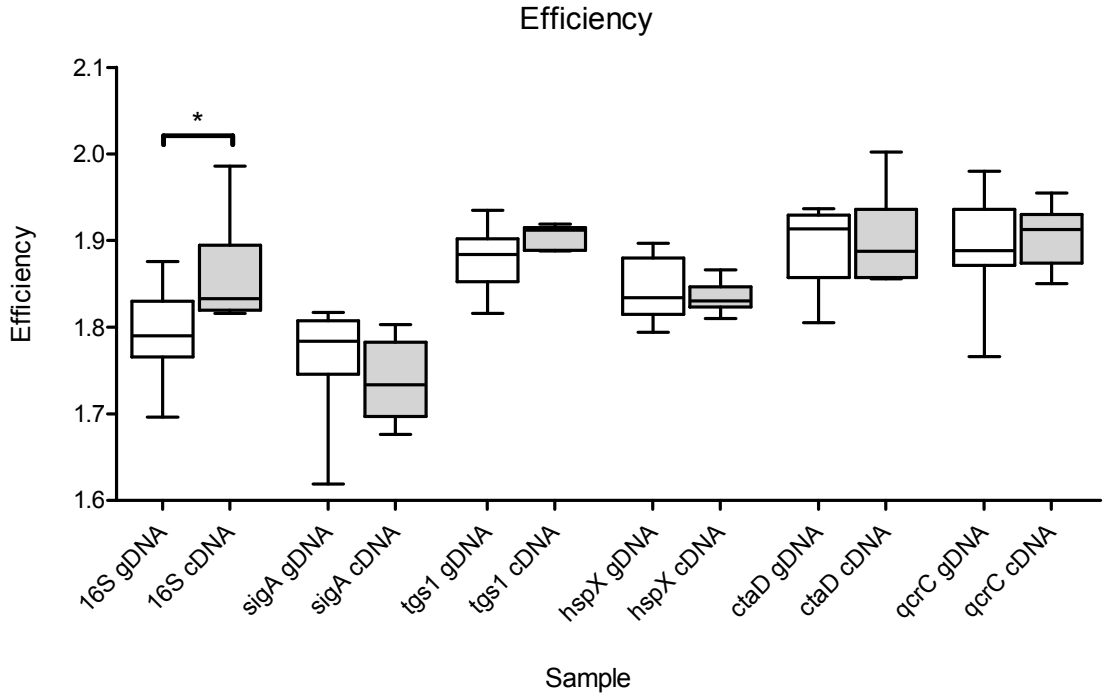


Figure 13: Box Plot showing efficiencies for gDNA and cDNA calculated from fluorescence increase

This box plot demonstrates the average efficiencies for genomic DNA and cDNA (calculated from fluorescence increase using LinRegPCR) for each of the 6 primer sets. The box represents the lower and upper quartiles with means; whiskers represent the range of data. The mean efficiencies calculated from the dilutional method are not displayed. A star indicates a significant unpaired t-test ($p<0.05$).

3.3.2 Comparison of Absolute and Relative Quantification

Absolute and relative quantification methods provide two alternative methods for quantifying the gene expression/mRNA transcription in biological samples. Absolute quantification has previously been used in the laboratory to calculate gene expression with a gDNA standard curve. Relative quantification differs by calculating the expression of a gene in relation to another; a standard curve is not required. As gDNA and cDNA were shown in 3.3.1 to vary in amplification efficiency, the use of a standard curve for quantification may not be desirable. Therefore, the calculated expression results of both absolute and relative quantification methods were compared to determine the most appropriate method for this study.

RNA was extracted as above from exponentially growing *Mtb* H37Rv and *Mtb* H37Rv pSMT3-*tgs1*. *Mtb* H37Rv pSMT3-*tgs1* was a bacterial strain developed by Natalie Garton (2011) to overexpress the *tgs1* gene in comparison to the H37Rv laboratory strain and is further discussed in Appendix 3. The expression of two genes expected to have high and low expression were compared: *tgs1* (high expression) and *hspX* (low expression). Quantitative PCRs were performed in triplicate.

For the standard curve, the variation of absolute quantification values can be difficult to assess precisely, due to the exponential scale of Ct values. Therefore, the standard deviations in Ct were used to estimate the variation expressed as plus/minus percentage of molecules based on Equation 9 (Rutledge and Cote, 2003).

Equation 9: Calculation of Variation in Absolute Quantification Standard Curves

$$\pm \,\% \, Molecules = [(E + 1)^{SD} - 1] \times 100\%$$

SD was the standard deviation in Ct generated from replicate amplifications and E was the efficiency determined via the 'fluorescence increase' method (expressed as a percentage).

Table 5 shows the variation in triplicate absolute quantification standard curves, for each of 3 PCR runs. Overall, the average standard deviation of Ct values for each run ranged from 0.32 to 0.39. This corresponded to an estimated variation in molecules that ranged from ±20.84 to ± 26.26%, using amplification efficiencies taken from the 'fluorescence increase' method.

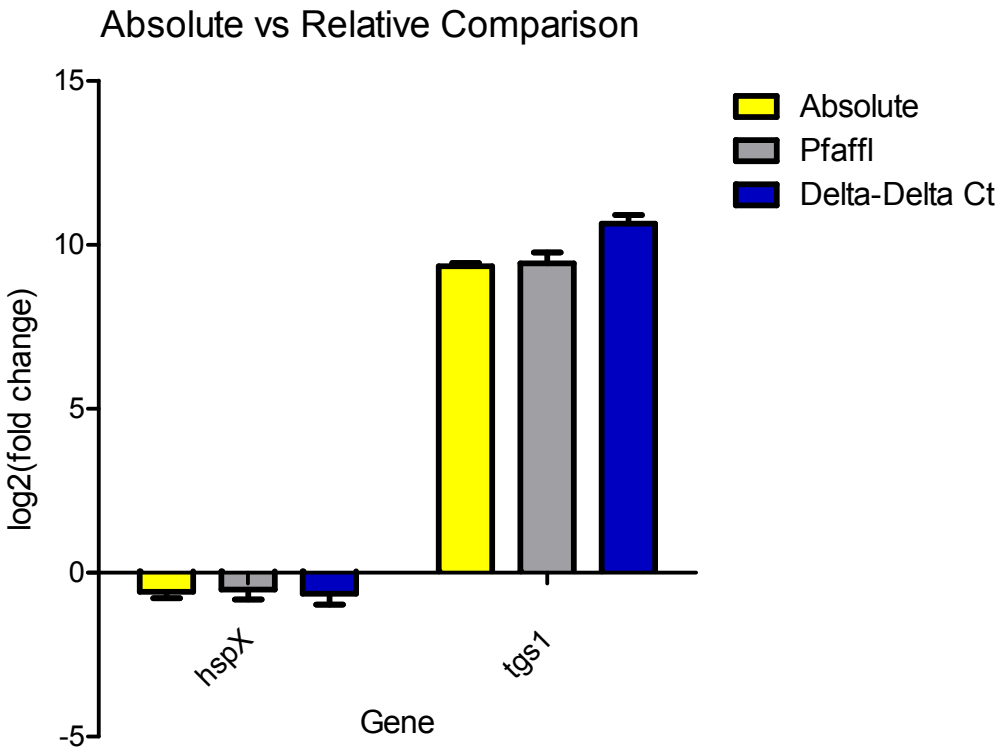
Table 5: Variation in Absolute Quantification Standard Curves

	Run 1 (Ct)	Run 2 (Ct)	Run 3 (Ct)
10^7	10.30 ± 0.11	9.02 ± 0.44	10.41 ± 0.31
10^6	13.76 ± 0.36	12.61 ± 0.28	14.26 ± 0.12
10^5	17.99 ± 0.34	16.15 ± 0.42	18.33 ± 0.54
10^4	21.61 ± 0.65	20.69 ± 0.39	21.88 ± 0.41
10^3	25.49 ± 0.59	24.80 ± 0.43	26.36 ± 0.35
10^2	29.16 ± 0.30	27.73 ± 0.29	29.60 ± 0.18
Average SD	± 0.39	± 0.38	± 0.32
% Mol	± 26.26	± 25.54	± 20.84

The expression fold changes were calculated for each individual PCR, using the Absolute quantification, delta-delta Ct (relative) and Pfaffl (relative) methods. Figure 14 displays the averaged fold changes for *Mtb* pSMT3-*tsg1* against *Mtb* H37Rv for the genes *hspX/sigA* and *tgs1/sigA*. *SigA* was used for normalisation, as this experiment was performed prior to the assessment of candidate housekeeping genes (3.3.4).

A one-way analysis of variance (ANOVA) was used to compare the Pfaffl, delta-delta Ct and absolute quantification methods. The values for the three methods differed significantly for *tgs1/sigA* (p=0.0013) but not for *hspX/sigA* (p=0.8715). Paired t-tests were performed to identify the statistically different quantification method. The Pfaffl method did not differ significantly from the Absolute quantification method (p=0.6429) whilst the ΔΔCt method was significantly different from the Absolute values (p=0.0058).

Standard deviation between PCR runs was lower for the Absolute Quantification method as compared to the Relative Expression methods. However, an F-test to compare variances showed no statistical difference between the absolute quantification and relative expression results.



	tgs1/sigA	*hspX/sigA*
M. tuberculosis pSMT3-tgs1 vs *M. tuberculosis* H37Rv		
Absolute	9.34 ± 0.10	-0.58 ± 0.20
PfaffI	9.43 ± 0.33	-0.52 ± 0.30
Delta-Delta Ct	10.64 ± 0.27	-0.64 ± 0.34

Figure 14: Calculated Fold Changes for each Gene using Absolute and Relative Quantification

The averaged fold changes for *M. tuberculosis* pSMT3-*tsg1* against *M. tuberculosis* H37Rv calculated using the Absolute Quantification, Delta-Delta Ct (relative quantification) and PfaffI (relative quantification) methods are displayed above. Two genes were selected and are displayed: *tgs1* and *hspX*; both genes were normalised using *sigA*. Scale is $\log_2$; 1 log is equivalent to a 2 fold change.

3.3.3 Genomic DNA Removal

The extraction of RNA from biological samples often results in gDNA contamination. The removal of this contamination is mandatory to prevent any bias in the quantification results, as quantitative PCR assays are unable to differentiate between cDNA and gDNA (Tondeur et al., 2004). As previous DNase protocols yielded higher than expected gDNA levels (Lee, 2007), two new DNase protocols, RNase-free On-column DNase (Qiagen) and Turbo DNase (Ambion), and a combination of the two protocols were tested for their efficiency in removing gDNA from contaminated RNA extracts.

RNA was extracted using a Trizol extraction method (2.5.3**)** from exponential phase *Mtb* H37Rv to a crude extract phase. The RNA was further processed in duplicate via seven methods, shown in Figure 15.

Each of the 7 tested samples was reverse-transcribed with genome-directed primers into cDNA. The reverse transcribed cDNA was used in SYBR green qRT-PCR reactions as described in 3.2.7. Both test samples and noRT controls were run in triplicate for this experiment.

Figure 16 shows the remaining gDNA copy numbers for three genes (*sigA, tgs1* and *hspX*) for each of the gDNA cleanup methods; methods 1 and 2 were excluded due to lack of a DNase step in the methods.

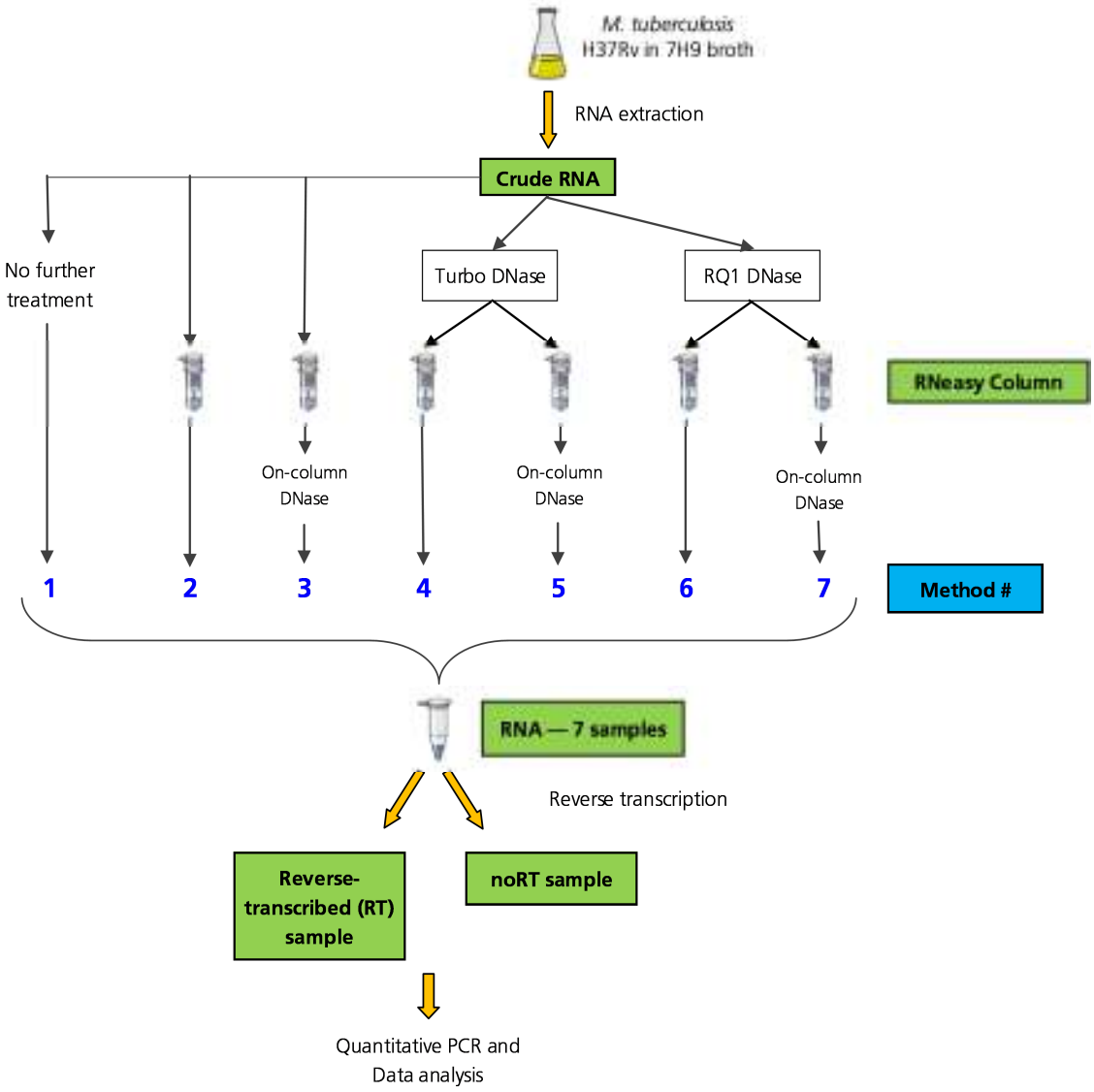


Figure 15: Schematic of experimental design for the investigation of DNase treatment options.

The schematic shows the experimental design for the investigation of DNase treatment options via 7 treatment methods from RNA extraction to Quantitative PCR and data analysis.

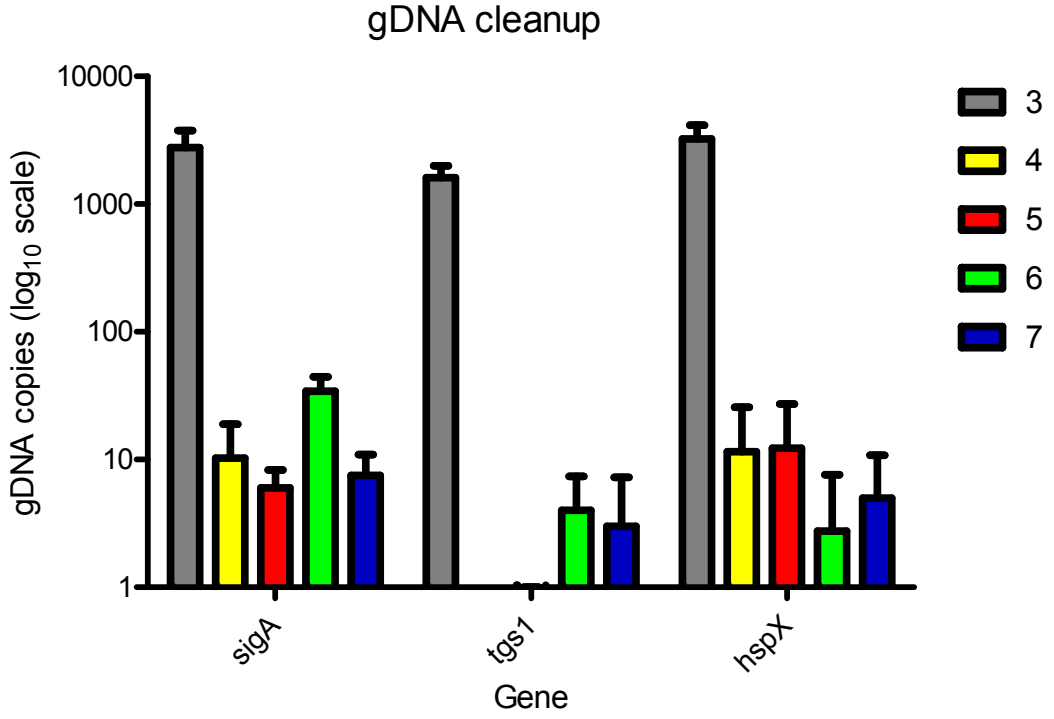


Figure 16: Comparison of Genomic DNA Removal Protocols

The bar graph shows the remaining genomic DNA copy numbers for three genes (*sigA, tgs1* and *hspX*), for each of the gDNA cleanup methods, averaged over 2 experimental (DNase treatment) replicates and 3 PCR replicates; methods 1 and 2 were excluded due to lack of a DNase step in the methods. The methods were as follows:

3. RNeasy Mini Column with RNase-free On-column DNase treatment
4. Turbo DNase and RNeasy Mini Column
5. Turbo DNAase and RNeasy Mini Column with RNase-free On-column DNase treatment
6. RQ1 DNase and RNeasy Mini Column
7. RQ1 DNase and RNeasy Mini Column with RNase-free On-column DNase treatment

The error bars demonstrate the standard deviation across the replicates. The table shows the numerical values (in gDNA copy numbers) for each of the gDNA cleanup methods.

The use of multiple DNase methods (on-column plus off-column methods) in the cleanup of RNA generally appeared to be the most efficient in removing gDNA contamination. The two processes, 5 and 7 as described above, proved to be the most and equally effective. These were differentiated by examining the return in RNA gene copy numbers (Table 6).

Table 6: RNA return for gDNA treatments (5 and 7)

Process	RNA copy numbers	
	sigA	*tgs1*
5	$2.56 \times 10^4 \pm 1.88 \times 10^3$	$3.63 \times 10^2 \pm 3.71 \times 10^1$
7	$2.01 \times 10^4 \pm 2.13 \times 10^3$	$3.05 \times 10^2 \pm 1.56 \times 10^2$

An unpaired t-test used to analyse the RNA return for gDNA treatments 5 and 7 demonstrated that the difference was not-statistically significant (*sigA* p=0.2888, *tgs1* p=0.4909). Method #5 was selected and used for all further steps based upon its ease of use and as the method requires fewer manual handling steps than Method #7.

3.3.4 Selection of Internal Housekeeping Genes

Any variation in the expression of a reference gene produces calculated artefacts and hides genuine gene expression changes (Bustin et al., 2005). It is therefore vital to select a housekeeping gene that is expressed constitutively. Traditional housekeeping genes have been used without prior validation of their merit (Vandesompele et al., 2009). In previous studies, sigA mRNA transcripts had been used for the normalisation of samples (Lee, 2007). To confirm or exclude the use of *sigA* and to select the most stable candidate, a selection of 11 candidate housekeeping genes were examined using reference gene validation software.

3.3.4.1 Selection and Primer Design

Further housekeeping genes (mRNA) were analysed along with *sigA* to select the genes with the most stable expression. Housekeeping genes were chosen from a literature search of *Mtb* housekeeping genes used in other studies. The ribosomal RNA gene, 16S, was also included in this study. Table 7 displays the housekeeping genes chosen and their functions.

The housekeeping genes were chosen on the basis of previous use as a housekeeping gene in another study, or the submission of a gene to the Oswaldo Cruz Institute Database as a possible housekeeping gene (Oswaldo Cruz Institute, 2010). A gene was excluded from the study if shown to be differentially expressed in previous studies. Co-expressed genes were also excluded from the study. The number of candidate housekeeping genes was limited to a total of 11 genes.

Primers for each gene were designed to have a specific annealing temperature of approximately 59°C to match the annealing conditions of previously designed primers. Amplicon size was designed to be between 80 and 220bp in length if possible. A list of primers for each selected housekeeping gene is displayed in Appendix 1.

The primers were tested to verify that they amplified a correctly sized band, by running a SYBR green qRT-PCR assay with a common annealing temperature of 59°C. The products were imaged using agarose gel electrophoresis as shown in Figure 17.

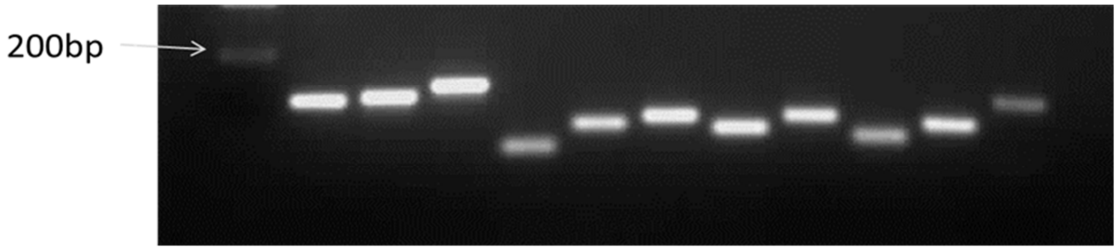


Figure 17: Gel electrophoresis of qRT-PCR products using housekeeping gene primer sets

PCR was used to amplify specific genes in *M. tuberculosis* CDC1551 genomic DNA obtained from Colorado State University. The gel electrophoresis displays these amplified gene bands, in order from left to right: pre-prepared DNA ladder, *gyrB* (178bp), *lysA* (185bp), *polA* (214bp), *fbpB* (76bp), *ftsZ* (120bp), *sigA* (160bp), *dfrA* (111bp), *thyA* (139bp), *aroA* (98bp), *mtrA* (124bp) and *16S* (173bp).

Table 7: Housekeeping genes for *M. tuberculosis*

Gene	Gene Name	Product Function	Reference
Rv0005	*gyrB*	DNA gyrase/topoisomerase (type II) (subunit B) – negatively supercoils closed circular dsDNA in an ATP-dependent manner	(Oswaldo Cruz Institute, 2010, Unniraman et al., 2002)
Rv1293	*lysA*	Probable diaminopimelate decarboxylase – involved in the last step of the biosynthesis of lysine	(Gokulan et al., 2003, Oswaldo Cruz Institute, 2010)
Rv1629	*polA*	Probable DNA polymerase I – exhibits DNA polymerase activity and 3′-5′ and 5′-3′ exonuclease activity	(Institut Pasteur, 2004, Huberts and Mizrahi, 1995)
Rv1886c	*fbpB (85B)*	Secreted antigen 85-B (Antigen 85 Complex B) – possesses mycolyltransferase activity required for the biosynthesis of trehalose dimycolate	(Oswaldo Cruz Institute, 2010, Hellyer et al., 1999)
Rv2150c	*ftsZ*	Cell division protein – essential for cell division; the intracellular concentration of the protein FtsZ is thought to be critical for septum formation in mycobacteria	(Delogu et al., 2006)
Rv2703	*sigA*	RNA polymerase sigma factor A – an initiation factor that promotes attachment of RNA polymerase to specific initiation sites	(Garton et al., 2008)
Rv2763c	*dfrA (folA)*	Dihydrofolate reductase – essential step for *de novo* glycine and purine synthesis	(Oswaldo Cruz Institute, 2010, Argyrou et al., 2006)
Rv2764a	*thyA*	Probable thymidylate synthase – involved in deoxyribonucleotide biosynthesis	(Institut Pasteur, 2004, Rengarajan et al., 2004)
Rv3227	*aroA*	5-enolpyruvylshikimate-3-phosphate synthetase – involved in the biosynthesis of chorismate within the biosynthesis of aromatic amino acids	(Parish and Stoker, 2002)
Rv3246c	*mtrA*	Two component, sensory transduction, transcriptional regulatory protein – starts transcriptional activity	(Fol et al., 2006, Oswaldo Cruz Institute, 2010)
16S	*rrs*	Ribosomal RNA – 16S – component of small prokaryotic ribosomal subunit (30S)/interacts with 23S subunit	(Garton et al., 2008)

3.3.4.2 Expression levels of candidate reference genes

The cDNA from multiple growth conditions were used to obtain a view into the differential expression of the housekeeping genes between these conditions. *Mtb* H37Rv was prepared under five different growth conditions, with three biological replicates of each:

- Exponential phase culture in 7H9 broth
- Stationary phase culture in 7H9 broth (culture at constant OD for three consecutive days)
- Exponential phase culture in 7H9 broth with SDS treatment (SDS added to a final concentration of 0.05% (w/v))
- Exponential phase culture in RPMI-1640 medium (as described in 3.2.2)
- Exponential phase culture in phosphate-buffered saline (as described in 3.2.1)

RNA was extracted from each of the cultures using a Trizol extraction method (2.5.3). The crude extract was processed to remove gDNA using process 5 in 3.3.3. The clean RNA was reverse-transcribed using genome-directed primers (GDPs). The cDNA from each sample was included in a qRT-PCR reaction. Triplicate cDNA reactions were performed for each run, and all samples were run simultaneously for each primer set.

Figure 18 visually displays the variation in Ct values. The median Ct values fall between 6.96 (16S) and 26.33 (*polA*). The gene pattern therefore shows that 16S rRNA is expressed in much greater quantities as compared to the other 10 mRNA genes. The set of 10 Ct values for the mRNA housekeeping genes are significantly different (ANOVA, $p<0.0001$). However, the 10 mRNA genes are expressed at more similar level compared to 16S.

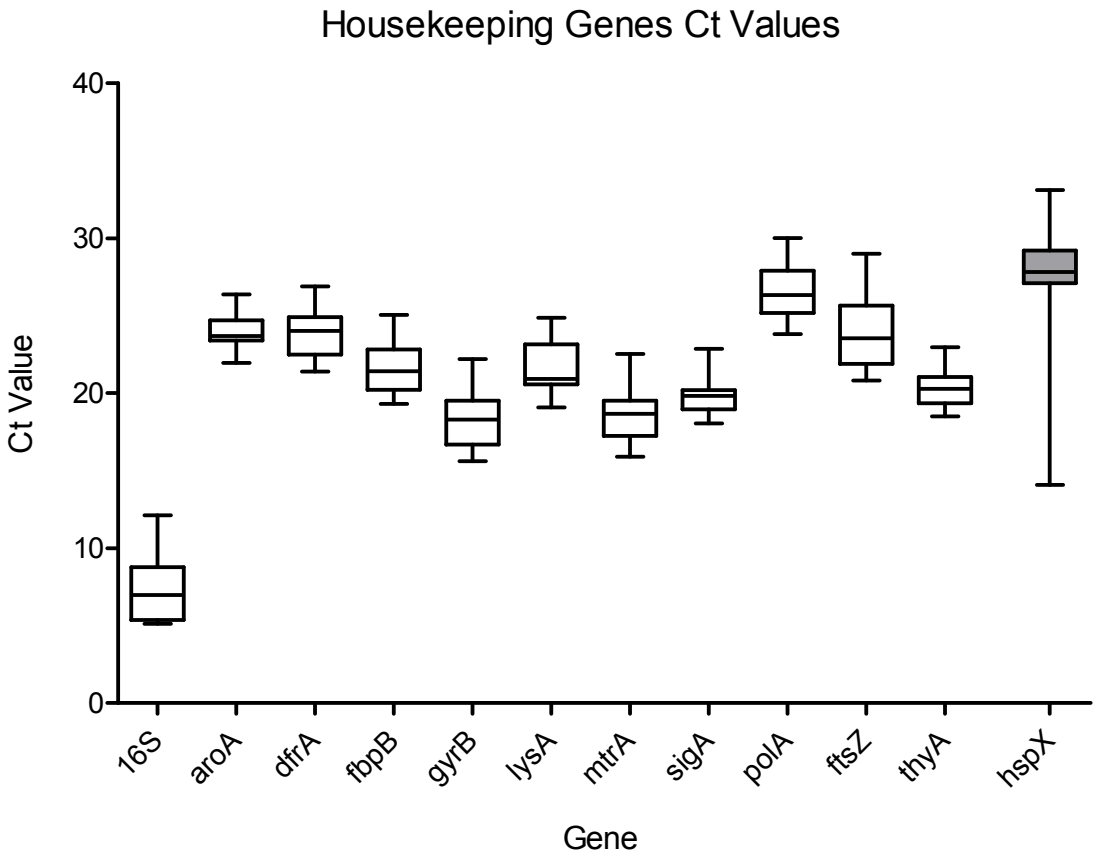


Figure 18: Expression levels of candidate housekeeping genes

The box and whisker plot shows the expression of each of the 11 candidate housekeeping genes for each of the 5 experimental growth conditions. The box represents the lower and upper quartiles with means; whiskers represent the range of data. The observed range of Ct values are distributed over a fairly large range (medians falling between 6.96 and 26.33) representing highly expressed genes (mean Ct values < 30); the graph does not take into account variations in starting genetic material. The differentially expressed target gene *hspX* is also included for comparison and is highlighted in grey; note the large range of data as compared to the 11 candidate housekeeping genes.

3.3.4.3 Expression stability of reference genes under different growth conditions

The geNorm program was used to calculate the gene expression stability measure *M* for each of the 14 genes; *M*. was defined as 'the average pairwise variation between a particular gene and all other control genes' (Vandesompele et al., 2002). The lowest *M* value characterized the gene with the most stable expression. After each series of calculations, the gene with the highest *M* value (least stable expression) was eliminated; the *M* values were recalculated and a ranking of genes according to their *M* values was generated. The *M* value for each of the candidate genes is shown in Figure 19.

Pairwise variation was calculated by geNorm to determine the optimal number of reference genes for normalisation. Two genes (n=2) were used for the calculation of this RT-PCR normalisation factor, and more housekeeping genes were included in a stepwise manner, until adding a (n+1)th gene had no significant effect to the new normalisation factor (Vandesompele et al., 2002). A cut-off value of 0.15 was used; below this, the addition of an additional reference gene was not required (Vandesompele et al., 2002).

Figure 20 shows the pairwise variation for each of the five growth conditions. The optimum number of genes in the normalisation factor was determined to be 4 genes. For each of the five growth conditions, the maximum number of housekeeping genes required was 4 (for Stationary Phase culture and PBS exposure). The other three growth conditions (exponential phase, RPMI exposure and SDS exposure) required fewer genes. Therefore, the inclusion of a fifth housekeeping gene in the normalisation factor was not required.

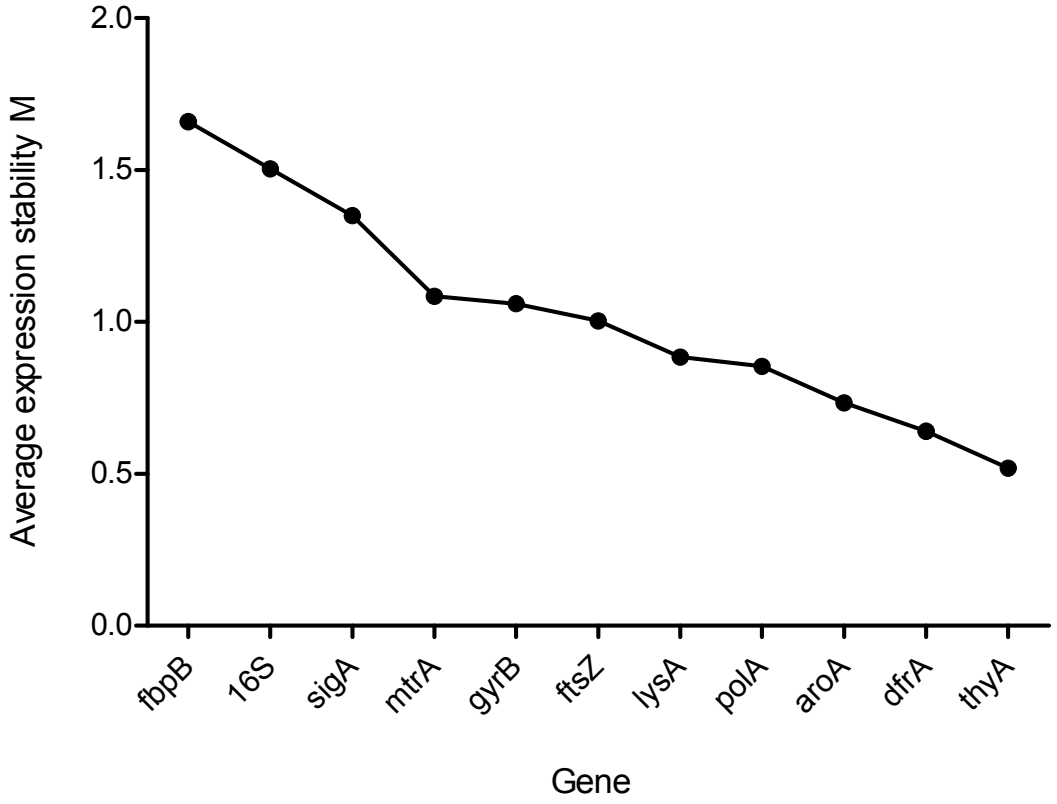
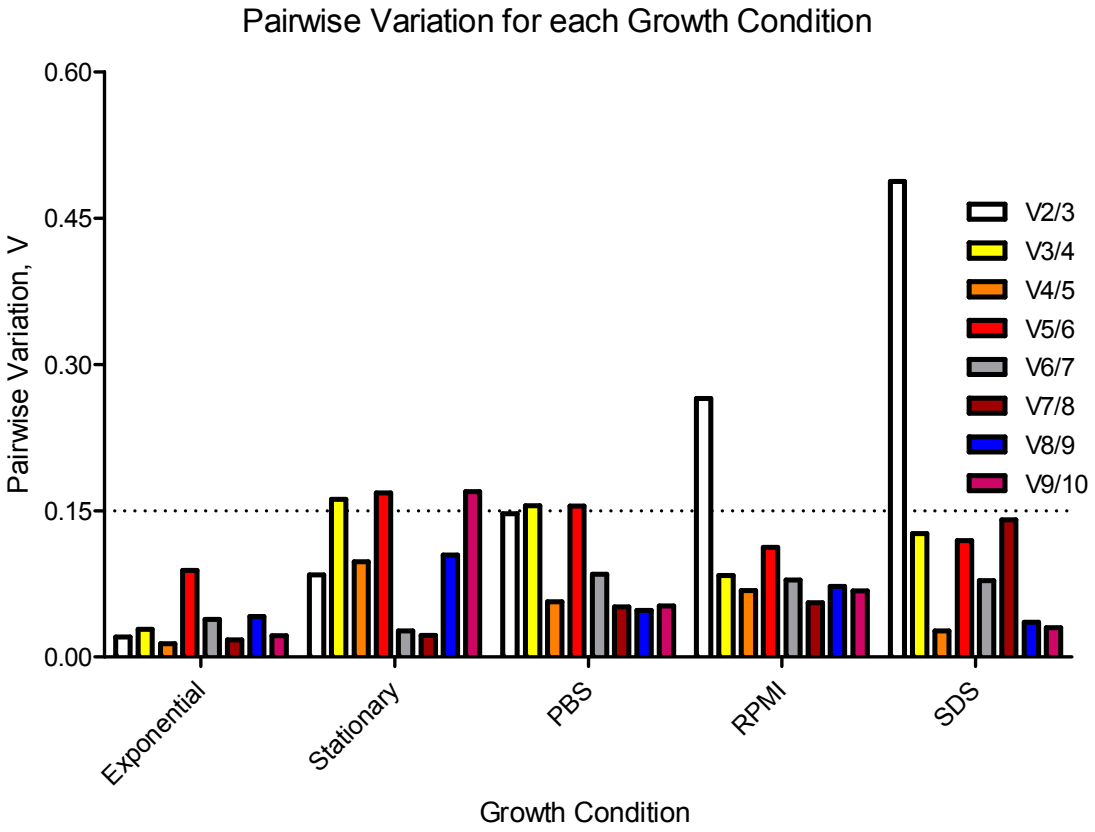


Figure 19: Selection of the most suitable reference genes for normalisation of *M. tuberculosis* expression samples

The figure shows the average expression *M* values of the 11 candidate genes. The lower the Average Expression Stability Value (M), the more stable the gene is. Therefore, genes to the left are less stable, and genes to the right are more stable. Of the 12 candidate genes, the three most stable genes are *thyA, dfrA* and *aroA,* while the three least stable genes are *fbpB,* 16S and *sigA*.



Growth Condition	Pair	Variation (<0.15)
Exponential Phase	V2/3	0.021
Stationary Phase	V4/5	0.098
PBS Exposure	V4/5	0.057
RPMI Exposure	V3/4	0.084
SDS Exposure	V3/4	0.127

Figure 20: Calculated Pairwise Variations for five growth conditions

The variable V defines the pairwise variation between two sequential normalisation factors. For example, V4/5 shows the variation of the normalisation factor calculated with 4 genes in relation to that calculated with five genes. A pairwise variation below 0.15 indicates that the addition of an extra gene has no significant contribution to the normalisation factor and does not need to be included in calculations. The table shows the first pair with variation below 0.15 for each of the 5 growth conditions and associated variation value.

Normfinder software analysis was also used to compare and/or verify the geNorm results. The order and stability values of the 11 genes as determined by Normfinder and the geNorm order are listed in Table 8.

Table 8: Normfinder Housekeeping Gene Results

Number	Normfinder Order		geNorm Order
	Gene	Stability Value	
1	*thyA*	0.116	*thyA*
2	*polA*	0.116	*dfrA*
3	*dfrA*	0.133	*aroA*
4	*aroA*	0.140	*polA*
5	*ftsZ*	0.145	*lysA*
6	*lysA*	0.146	*ftsZ*
7	*gyrB*	0.173	*gyrB*
8	*mtrA*	0.196	*mtrA*
9	*sigA*	0.205	*sigA*
10	16S	0.216	16S
11	*fbpB*	0.233	*fbpB*

The Normfinder order was similar to the geNorm order, with select differences. *ThyA* was the most stable gene as determined by both Normfinder and geNorm. The genes ranked 2-4 (*dfrA, aroA* and polA) were the same for both programs, but differed in their order. The genes *lysA* and *ftsZ* were ranked 5 and 6 in the geNorm order, but reversed in the Normfinder order. Finally, the ranking of the 5 least stable genes was identical as determined by both programs.

As the ranking of the candidate housekeeping genes differed for Normfinder as compared to geNorm, geNorm was used to examine the optimum number of genes in the normalisation factor using the Normfinder ranking. The optimum number of genes was again determined to be 4 genes (Results displayed in Appendix 4 – Supplementary Data). As for the geNorm ranking order, the maximum number of genes required in the normalisation factor was 4 (RPMI and SDS exposure). The other three growth conditions required fewer genes. Thus, for both ranking orders (Normfinder and geNorm), 4 genes were required for the normalisation factor, and used the same 4 genes (*thyA, dfrA, aroA* and *polA*).

3.3.4.4 Comparison of reference genes for normalisation

The four reference genes (*thyA, dfrA, aroA* and *polA*) were combined into a single 'normalising factor'. This normalising factor (NF4) was calculated using the geometric mean of the relative gene quantities as shown in Equation 10.

Equation 10: Normalising Factor Calculation using *thyA, dfrA, aroA* and *polA*.

$$Normalising\ Factor = \sqrt[4]{thyA \cdot dfrA \cdot aroA \cdot polA}$$

The use of NF4 was compared to normalisation using the conventional *sigA* gene. The effect of normalisation on gene expression was compared for both *sigA* and NF4. This comparison is displayed in Figure 21. Graph A shows a scatter plot comparing the fold changes of genes normalised to either *sigA* or NF4. The fold changes calculated using *sigA* and NF4 correlated significantly ($p < 0.0001$*, $R^2=0.7522$, Pearson's correlation).

Graphs B and C demonstrate the difference in the range of calculated fold changes for candidate housekeeping genes and 20 selected target genes (normalised to either *sigA* or the normalising factor). The target genes analysed In Graph C are listed and discussed in Chapter 4. The *sigA* and NF4 fold changes were significantly different for both the housekeeping genes ($p < 0.0001$*, paired t-test) and target genes ($p < 0.0001$*, paired t-test).

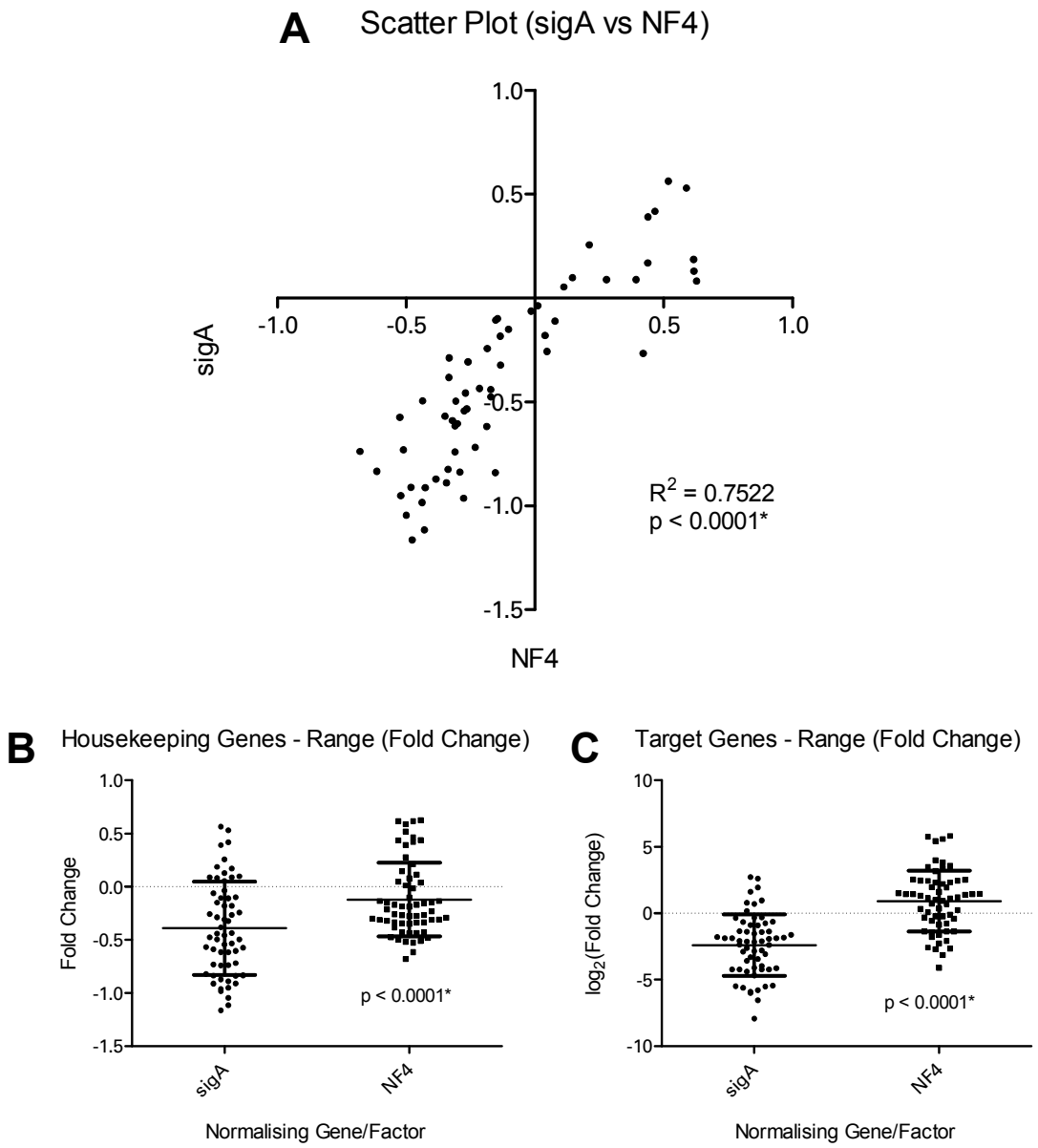


Figure 21: Comparison of Normalisation using *sigA* and NF4

Figure 21A is a scatter plot comparing fold changes normalised using *sigA* and NF4. The genes used for this scatter plot were the candidate housekeeping genes (not including *sigA, aroA, dfrA, thyA and polA*). Fold changes were calculated against exponential phase. The fold changes correlate significantly ($p < 0.0001$*, R^2=0.7522, Pearson's correlation).

Figure 21B and C show calculated fold changes for housekeeping and target genes for 4 growth conditions (PBS/RPMI/SDS/Stationary) calculated against exponential growth, respectively. Graph B shows the candidate housekeeping genes (not including *sigA, aroA, dfrA, thyA* and *polA* normalised to *sigA* and NF4. Graph C shows the fold changes for 20 target genes (see Chapter 4). The *sigA* and NF4 fold changes were significantly different for both the housekeeping genes ($p < 0.0001$*, paired t-test) and target genes ($p < 0.0001$*, paired t-test).

3.4 Discussion

3.4.1 Comparison of Direct Efficiency Calculations

A comparison of the direct efficiency calculations ('Fluorescence increase in exponential phase' vs. Dilution method) was made for the qRT-PCR assays used in this study. Furthermore, amplification efficiencies for the pure gDNA used in standard curves were compared to reverse-transcribed cDNA.

It can be seen in the results that the efficiencies varied significantly when examining gDNA and cDNA samples (for 16S), and also when comparing the 'fluorescence increase in exponential phase' vs. Dilution method of determining efficiency (for 16S and *tgs1*). While this variation was significant for only a couple genes, the variation was often greater than 2% within the same sample. Overall, the standard deviation in efficiencies was generally less than 5% for all the genomic and complementary DNA samples, analysed by fluorescence increase and dilution. Statistically, when the efficiencies across the 6 gene primer sets were examined, a one-way ANOVA revealed significant differences between gDNA and cDNA amplification efficiencies ($p<0.0001$), and also between the efficiency calculation methods ($p<0.0001$).

The dilution method makes use of a gDNA standard curve; however, the variations between gDNA and cDNA amplification efficiencies suggest that it may be inappropriate to apply efficiencies calculated using this method on cDNA assays. While cDNA may be used as an alternative to gDNA in the standard curve, efficiency would need to be calculated on a sample by sample basis, wasting valuable template (Wong and Medrano, 2005).

Examining fluorescence increase may be a better estimator of the true efficiency as only data from the exponential phase of amplification was analysed (Pfaffl, 2006). The efficiency of reactions generally vary from being stable and high in the early exponential phase of amplification, and then decline down to zero, due to the decay of polymerase activity and depletion of PCR reaction components (Wong and Medrano, 2005). The dilution method of efficiency calculation does not account for this loss of efficiency.

Ramakers et al. (2003) previously scrutinised the use of standards for estimating PCR efficiencies. The authors analysed standard curves based upon the serial dilution of a single cDNA sample; when a single sample was used, contamination could occur from a variety of reagents encountered during the RNA extraction and reverse transcription process, e.g. salt, phenol, chloroform or ethanol, and result in reduced amplification efficiency (Ramakers et al., 2003). However, dilution of the sample also resulted in the dilution of the contaminant and a reduction in its effect; therefore, with each dilution, the PCR efficiency increased. In an

example of this dilution effect, the authors showed that in contaminated cDNA samples, there was little relation between the efficiencies of the individual PCR reactions and the standard curve derived (dilution) efficiency (Ramakers et al., 2003).

While a contaminated sample was less applicable to the standard curves constructed in this study (as purified gDNA is used), the same principles applied to the extracted experimental cDNA samples used. This contamination may lead to minor differences in efficiency between individual cDNA samples. These minor variations in efficiency may be non-significant, but still generate differences in efficiency of greater than 2% between gDNA and cDNA samples. Even this small difference can still bias results greatly as studies have demonstrated that a difference in efficiency of only 3% (0.03) between the target and reference gene can create a difference of 242% (if $E_{target} > E_{Ref}$) after 30 cycles (Pfaffl, 2006). As the difference in efficiency increased, this false difference increased further, e.g. $\Delta E=0.05$ (432% calculated difference) and $\Delta E=0.10$ (1744% calculated difference) (Pfaffl, 2006).

The statistically different efficiency variations between gDNA and cDNA and between primer sets suggested that the use of direct fluorescence efficiency calculations was more appropriate than the Dilution method of efficiency calculation. Furthermore, due to these variations, efficiency correction appeared to be important in qRT-PCR assays, and is discussed further in 3.4.2.

3.4.2 Comparison of Absolute and Relative Quantification

Two methods of gene expression quantification, absolute quantification and relative quanfitication, were analysed. Replicate qRT-PCR assays were performed and the results from both methods were compared.

Relative expression was based on the expression of two genes, *tgs1* and *hspX,* normalised to the *sigA* mRNA housekeeping gene and expressed as a ratio. These normalised ratios from two *Mtb* strains were compared: H37Rv and H37Rv pSMT3-*tgs1*. This gave the fold change, for the gene of interest, between these two strains. Alternatively, absolute quantification, gave an absolute value of transcript copy numbers based upon a standard curve. As a result of this, absolute and relative quantification results were not immediately comparable, as they referred to different expression variables (Pfaffl, 2006).

However, if the results obtained by absolute quantification for *tgs1* and *hspX* were also normalised to *sigA* or another housekeeping gene, the unit 'copy numbers' would cancel out, and the absolute quantification result would be expressed as a ratio of the two genes. If these

gene ratios were compared between two *Mtb* strains, a similar fold change would be generated. In this manner, the results would be comparable to those obtained by the relative quantification method.

The three quantification methods were analysed using a one-way ANOVA and showed a significant difference for the *tgs1/sigA* fold change (p=0.0013) but not for the *hspX/sigA* fold changes (p=0.8715). The paired t-test used to compare the individual quantification methods revealed the ΔΔCt relative quantification method to be significantly different from absolute quantification.

The basis of absolute quantification includes the assumption that efficiencies are constant for all reaction samples. This has been shown not to be the case (see results 3.3.2 and discussion 3.4.1), especially when the samples vary considerably. For relative quantification, whilst the ΔΔCt method does not take into account efficiency, the Pfaffl model for relative quantification does include an efficiency correction. The *tgs1/sigA* results demonstrated this calculated difference in fold changes if efficiency is not corrected. For this reason, either absolute quantification or the Pfaffl model was deemed more appropriate than ΔΔCt for this comparison.

Due to the nature of absolute quantification, the method adds an extra level of variation through the use of a gDNA standard curve. To obtain high precision and reliability, the external standard curve must be highly defined and reproducible (Pfaffl and Hageleit, 2001). However, the replicate standard curves shown in Table 5 showed high estimated variation in number of molecules ranging from ±20.84 to ±26.26%. Furthermore, amplification efficiencies derived from standard curve gDNA and target cDNA can vary greatly as discussed previously (3.4.1). Therefore, the standard curve is a significant source of variation in gene expression calculations.

The Pfaffl method of relative expression corrects for differential efficiencies in samples, as the efficiency of each reaction and primer set is calculated individually. Thus, results derived from this method may be closer to the true cell situation, as opposed to results from absolute quantification (which uses a standard curve) or the ΔΔCt method of relative quantification (which does not correct for efficiency variations).

Due to these differences, the use of the Pfaffl method was deemed most appropriate for the quantification of gene expression in this study. The *in vitro* studies to be conducted will look at the gene expression ratio between two sets of genes. Relative quantification methods directly calculate the ratio of the target gene versus the reference gene, which circumvents the need for the standard curve used in absolute quantification. These conclusions are consistent

with those of Chini et al. (2007) who found the Pfaffl method to be most appropriate for quantification of gene expression in *Staphylococcus aureus*.

3.4.3 Genomic DNA Removal

The adequate removal of genomic DNA from RNA samples is vital to prevent bias in the results. As discussed, DNase I commonly shows some sequence preference in its digestion of DNA, generally cleaving only dsDNA and being affected by the structure of the minor groove (Ambion Incorporated, 2007a). DNase I often has a pattern of recognition of DNA base sequences limited to three or four nucleotides, and commonly prefers purine-pyrimidine sequences (Ambion Incorporated, 2007a, Campbell and Jackson, 1980). Moreover, depending on the metal activator cations present in solution, the kind of cleavage (i.e. double-stranded cuts or single-stranded nicks) of the DNA substrate catalysed by DNase I can vary (Campbell and Jackson, 1980).

The results presented compared various gDNA removal protocols. With the previous best available DNA method (RNeasy Mini Column/On-column DNase) there was still a very small proportion of gDNA that remained post-treatment.

Method #5 (RNeasy Mini Column/On-column DNase/Turbo Off-column DNase) and Method #7 (RNeasy Mini Column/On-column DNase/RQ1 DNase) demonstrated the best removal of contaminating gDNA from RNA samples and also good RNA return. Method #5 did show higher RNA return than Method #7, but the difference was not statistically significant when analysed by an unpaired t-test.

While both methods were acceptable for gDNA removal, Method #5 was selected based upon ease of use as the method required fewer manual handling steps. This method was used to clean and purify all future extracted RNA samples.

3.4.4 Selection of Internal Control Genes

As reviewed earlier, the establishment of a set of suitable reference 'housekeeping' genes is an essential prerequisite for the analysis of gene expression using RT-PCR and allows for the normalisation of differences in the amount of amplified RNA template (Pfaffl, 2006, Wong and Medrano, 2005, Radonic et al., 2004). These differences can be generated by: (i) differential amounts of starting RNA, (ii) variations in the quality of starting material (e.g. protein contamination, degraded samples, and (iii) variations in RNA preparation and

synthesis of cDNA (Radonic et al., 2004). The reference gene is exposed to the same preparation steps as the target genes (Radonic et al., 2004).

An adequate reference gene must display several criteria. It must show constitutive, stable expression in samples, not regulated or influenced by experimental procedures (Penning et al., 2007, Ohl et al., 2005, Radonic et al., 2004, Andersen et al., 2004). Further, the range of expression of the housekeeping gene should be similar to that of the target gene in the samples analysed (Ohl et al., 2005).

Traditional housekeeping genes are often used without extensive verification of their validity (Vandesompele et al., 2009). In *Mtb* studies *sigA* (mRNA), 16S (rRNA) and 23S (rRNA) are often used for normalisation purposes (Daniel et al., 2004, Garton et al., 2008, Haile et al., 2002). Furthermore, microarray analysis is now commonly used and makes use of global normalisation. Without validated housekeeping genes, the task of comparing quantitative RT-PCR results and microarray analyses is more difficult.

Therefore, to determine the most stable housekeeping gene or genes, 11 candidate reference genes were validated using geNorm and Normfinder algorithms, under five different *Mtb* growth conditions.

Candidate housekeeping genes were selected from a search of the relevant literature; genes used previously as housekeeping genes and genes submitted as possible housekeeping genes were included (Oswaldo Cruz Institute, 2010). The most commonly used housekeeping genes are *sigA* (sigma factor A) and the rRNA genes (16S and 23S). Thus, *sigA* and 16S were included and a further 9 genes were also selected and analysed for their suitability as housekeeping genes. Genes were selected from the literature or the Oswaldo Cruz Housekeeping Gene Database (Oswaldo Cruz Institute, 2010). The final list of 11 genes was shown in Table 7.

A number of genes were excluded due to conflicting gene expression data in the literature. The gene *recA* was suggested as a possible housekeeping gene (Oswaldo Cruz Institute, 2010), but was shown to have variable expression. Movahedzadeh et al. (1997) demonstrated *recA* expression to be inducible by DNA damage. Further, *rpoB* and *rpoC* were suggested as possible housekeeping genes (Oswaldo Cruz Institute, 2010). *RpoB* and *rpoC* encode DNA-directed RNA polymerase subunits (Miller et al., 1994); however, while their expression may be constant in an *in vitro* environment, recent microarray studies have demonstrated that expression of the two genes is highly upregulated in sputum (Garton et al., 2008). As further qRT-PCR work involves attempts to replicate this sputum transcriptome, the use of these two genes was deemed inappropriate.

Other candidate housekeeping genes include *groEL2*, *dnaJ* and *dnaK* (chaperone proteins). *GroEL2* prevents misfolding and promotes proper refolding of polypeptides and also modulates mycolic acid synthesis was suggested as a possible housekeeping gene (Pang and Howard, 2007); however, the gene has been shown to express a heat-shock protein and expression of the gene varied several-fold with temperature (Shinnick et al., 1988). *DnaJ* and *dnaK* are chaperone genes regulated on a single operon. Both genes play roles in *Mtb* infection and expression increases during this process (Stewart et al., 2004).

Co-expressed genes were excluded from the study, as such genes can falsely lower the calculated stability values. Co-expressed genes will have similar expression values, regardless of the stress imposed on the cells. For this reason, only one ribosomal subunit gene was included: 16S was selected over 23S due to its use in previous studies. Furthermore, *gyrA* and *gyrB* were both candidate housekeeping genes, but are located on a single operon (Unniraman et al., 2002). *GyrB* was selected for this study, while *gyrA* was excluded.

A number of algorithms and software programs incorporating several strategies are available for the analysis of housekeeping gene stability (Vandesompele et al., 2009). Commonly used algorithms/software include geNorm (Vandesompele et al., 2002), Normfinder (Andersen et al., 2004) and Bestkeeper (Pfaffl et al., 2004). GeNorm is heavily cited in the literature and the ranking method has been independently validated (Gabrielsson et al., 2005, Vandesompele et al., 2009). Normfinder and Bestkeeper are also well cited; however, Bestkeeper is limited to a maximum of 10 candidate genes (Vandesompele et al., 2009). Therefore, for the purposes of this study, geNorm and Normfinder were used for the selection of housekeeping genes.

The geNorm software tool works by analysing the expression stability (*M*) of one gene based on the average pairwise variation between all the studied genes. Over five test growth conditions (exponential phase, stationary phase, PBS-, RPMI- and SDS-exposure) this analysis determined the three candidate housekeeping genes with the most stable expression to be *thyA, dfrA* and *aroA*. S*igA* and 16S, two of the most commonly used housekeeping genes in *Mtb* gene expression studies were determined to actually be two of the least-stable genes of the 11 candidate genes tested.

16S rRNA is a component of the 30S subunit of prokaryotic ribosomes and has been used as a housekeeping gene in previous studies in the literature (Daniel et al., 2004, Madigan et al., 1997). The use of rRNA as reference genes for mRNA studies has proven to be controversial. Its use has been suggested by some authors, as variations in mRNA are weak in comparison to rRNA, and therefore have little impact in the modification of the total RNA level (Thellin et al., 1999). However, regulation of rRNA versus mRNA may differ, depending on the stimulus

involved (Thorrez et al., 2008). Further, a study by Tricarico et al. (2002) reported varying ratios of rRNA to mRNA. As shown in Figure 18, 16S bacterial rRNA appears in vastly greater abundance compared to some mRNA transcripts; due to this large ratio, it may be difficult to accurately measure both 16S rRNA and a rare transcript in the same RNA/cDNA dilution. Ideally, the reference/housekeeping genes should also have similar expression levels as the genes being targeted.

The low stability value for *sigA* was not unexpected. Despite the use of *sigA* as a single housekeeping gene in previous studies, a study by Manganelli et al. (1999) showed that *sigA* may be differentially expressed under different growth conditions. While *sigA* was constitutively expressed over a variety of growth conditions, including oxidative, acid, detergent, cold and heat stresses, three conditions (incubation in water, growth into stationary phase and static incubation) led to downregulation of *sigA* expression (Manganelli et al., 1999). This differential expression, in addition to the low stability ranking, demonstrates that the use of *sigA* as a single housekeeping gene is inappropriate.

The Ct values for all 11 candidate genes confirmed that 16S was highly expressed in comparison to mRNA; 16S mean expression was over 500,000 fold greater compared with selected mRNA genes. The Ct values of the mRNA candidate housekeeping genes showed that while the mean values of these genes varied significantly, the expression level of these genes was more similar to the mRNA target genes than 16S rRNA.

Normfinder software uses its own unique approach to the estimation of expression variation and examines intra- and inter-sample variations (Andersen et al., 2004). The Normfinder results confirmed the 4 most stable housekeeping genes as determined by geNorm, albeit in a slightly different order. However, the resulting normalising factor is identical. The 5 least stable genes determined by Normfinder are identical the geNorm results; the software confirms the less stable expression of the commonly used *sigA* and 16S.

GeNorm and Normfinder use alternative methods for determining reference gene stability and variation. Therefore, the reproducibility and confirmation of the 4 most stable housekeeping genes using both geNorm and Normfinder methods gives high confidence regarding the stability of the selected reference genes. The variability in order of the top and middle ranking genes (e.g. *polA/aroA/dfrA* and *ftsZ/lysA*) may be due only to very slight differences in expression stability within these groups of genes. Minimal differences within the calculation algorithms may cause these differences in ranking order.

As discussed in Vandesompele et al. (2002), mRNA expression data is more accurate and reliable when normalised using the geometric mean of multiple internal reference genes, as

opposed to a single housekeeping gene. Bearing this in mind, geNorm analysis further demonstrated that, for the current study design, four reference genes (*thyA, polA, dfrA* and *aroA*) were required for normalisation for the analysis of gene expression in *Mtb*.

For each individual growth condition, exponential phase culture required 2 reference genes, RPMI and SDS required 3, and stationary phase culture and PBS needed 4 reference genes. Thus, the addition of a fifth gene was not required. The difference in normalisation requirements for each of the samples underlies the importance of the selection of a sufficient set of reference genes across all experimental treatments. Although the Normfinder gene list order differed from the geNorm order, analysis confirmed that the same 4 genes were required for the normalising factor.

The effect of normalisation using *sigA* and NF4 was demonstrated in 3.3.4.4. A scatter plot comparing fold changes using the two normalisation methods revealed significant correlation ($p < 0.0001$). While the two methods were significantly related, there was variation between the two methods with an R^2 value of 0.7522. The effect of normalisation (*sigA* and NF4) on fold changes was also compared for the candidate housekeeping genes (not including *aroA, polA, dfrA, thyA* and *sigA*) and 20 selected target genes (see Chapter 4); the expression of stationary phase culture, PBS-, RPMI- and SDS-exposure was calculated against exponential phase growth. The housekeeping gene expression in particular was expected to cluster near zero (no up- or down-regulation).

For both the housekeeping genes and target genes, the means of the calculated fold changes using *sigA* and NF4 are significantly different ($p < 0.0001$, paired t-test). The overall range of *sigA* and NF4 normalised gene expression for both the housekeeping and target gene sets was similar. However, the calculated *sigA* fold changes for both sets of genes appeared to be down-shifted as compared to the NF4 fold changes. Particularly for the housekeeping genes, the mean was expected to be at or close to zero; this was true for the NF4 normalisation, but not for *sigA*. Thus, due to the down-shifted fold changes, when *sigA* is used as a single housekeeping gene, small inductions of target genes may not be adequately represented.

The results suggest that the use of normalisation factor NF4 is more appropriate for this study than *sigA*. Firstly, s*igA* gene expression is known to vary under a variety of growth conditions (Manganelli et al., 1999). This was confirmed in this study by the low ranking of the *sigA* geNorm *M* stability value. Further examination suggested that the *sigA* results were down-shifted compared to the corresponding NF4 results; this down-shift was particularly notable in the housekeeping gene fold changes. Overall, for this study, it appeared that the use of NF4 was appropriate; genes were normalised to NF4 for all further experiments. The use of *sigA* and 16S as housekeeping genes is discussed further in Chapter 6.

3.5 Conclusions

In this chapter, a number of qRT-PCR techniques were examined to develop the most appropriate, precise and reproducible PCR assays for further gene expression studies. The studies fulfilled the main aims to evaluate and discuss absolute and relative quantification methods, efficiency calculations, gDNA removal methods and the selection of internal reference genes.

The specific findings in this chapter are as follows:

- The use of gDNA vs. cDNA or alternative efficiency calculation methods (dilution method vs. direct fluorescence analysis) for PCR assays can result in significantly different assay efficiencies.
- Fold changes calculated using the Pfaffl relative quantification method were not significantly different from absolute quantification, and were used in further studies as the method obviated the need of an external standard curve.
- A normalising factor (NF4) comprising the geometric mean of 4 housekeeping genes (*thyA, dfrA, aroA* and *polA*) was shown to provide the best normalisation of expression values. Commonly used *Mtb* housekeeping genes, *sigA* and 16S, had low stability as compared to other candidates.

Chapter 4

Gene Expression of *M. tuberculosis* in sputum and *in vitro*

4.1 Introduction

As introduced in Chapter 1, tubercle bacilli expectorated in sputum have generally been thought to originate from rapid and extensive growth at the edges of liquefied, granulomatous lung lesions (Garton et al., 2008). Sputum itself provides an easily obtainable sample of the *in vivo* bacterial population that must be eliminated by chemotherapy or otherwise may go on to transmit infection. While *in vitro* and *in vivo* (animal) gene expression studies aim to mimic a single *in vivo* stress, this population may display properties of bacilli that have been exposed to the multitude of stresses in the *in vivo* environment including but are not limited to: hypoxic, nutrient limiting, nitrosative, oxidative and acidic conditions (McKinney and Gomez, 2003, Rohde et al., 2007).

Further, as the bacteria are being expelled from the patient in question, this population may contain bacilli that are involved in the transmission of disease between hosts and must be targeted by antibiotic therapy. If they are in fact involved in transmission, the *M. tuberculosis* (*Mtb*) bacilli found in smear-positive sputum may express properties conducive to their virulence and survival within their new host.

The publication of the *Mtb* sputum transcriptome by Garton et al. (2008) revealed that the bacilli were in a state distinct from any previously established *in vitro* growth condition. Currently, the understanding of the *in vivo* factors involved in producing such a phenotype is limited. Thus, the replication of the sputum transcriptome *in vitro* may help to further the understanding of the nature of *Mtb* bacilli in sputum, as well as the stimuli responsible for this unique gene expression and the resulting bacterial phenotype.

4.1.1 Tuberculous Sputum

The current knowledge with regards to the composition of tuberculous sputum and source of bacilli *in vivo* was reviewed in Chapter 1. Despite the research into the composition of tuberculous lesions and sputum composition in other respiratory diseases, relatively little is known about tuberculous sputum. In addition to the composition of the sputum, other sources of contention include the state of tubercle bacilli in sputum and the timescale required for *Mtb* to enter a period of dormancy or non-replicating persistence. Both these issues are important if one is to fully replicate the gene expression of sputum *Mtb* in an *in vitro* setting.

The tubercle bacilli in sputum were originally thought to originate from rapid and extensive aerobic growth (Canetti, 1955b). However, the transcriptome of *Mtb* in sputum revealed the presence of slowly or non-replicating bacilli (Garton et al., 2008). The presence of these

bacilli fits into a hypothesis postulated by Mitchison, who suggested that there were several populations of *Mtb* bacilli *in vivo* (Mitchison, 2004). Mitchison (2004) proposed that the populations consisted of actively dividing bacilli, bacilli metabolising at a slower rate (or occasionally dividing), bacilli in an acidic environment, and non-growing/dormant bacilli or 'persistors'. It now seems that these non-growing/dormant bacilli comprise much of the bacillary population in sputum.

An important issue with regards to dormant bacilli is the timescale that is required for *Mtb* to enter this non-replicating persistence and clinical latency. Previous studies have demonstrated that growth arrest may occur after short exposures to stresses including nitric oxide, hypoxia and *in vivo* levels of vitamin C (Taneja et al., 2010, Voskuil et al., 2003, Wayne and Sohaskey, 2001). However, other *in vitro* models, e.g. progressive nutrient depletion, have required longer time course experiments to induce persistent organisms (Hampshire et al., 2004). Hence, the timeframe that is required for bacteria to enter persistence phase is still unknown and an issue of contention. In this study, it is postulated that the that bacilli become persistent over a time period consisting of hours, as opposed to days or months, due to the advent of host-derived stresses on *Mtb in vivo*.

4.1.2 Sputum Transcriptome

The *Mtb* sputum transcriptome was reviewed in Chapter 1 and was unique to established, *in vitro* growth conditions (Garton et al., 2008). For the purposes of this study, 20 genes were selected from the sputum transcriptome for further study. The genes were selected to reflect the trends seen in the sputum transcriptome and also included specific genes of interest. Of the 20 genes, 10 were upregulated and 10 were downregulated in sputum. The 20 genes and their selection are introduced here and further analysed and discussed in the Results (4.2.2) and Discussion (4.3.1), respectively.

4.1.2.1 Upregulated Genes

Out of the 182 genes that were significantly induced in sputum compared to aerobic growth, the 10 genes selected for use in this study are highlighted here. The upregulated genes selected and their products are displayed in Table 9 and fall into a number of functional classification groups.

Table 9: 10 Upregulated Genes and Products

Gene	Gene Number	Product
hspX	Rv2031c	Heat shock protein, hspX (alpha crystallin homolog)
Rv3551	Rv3551	Co-enzyme A transferase alpha subunit
tgs1	Rv3130c	Triacylglycerol synthase (diacylglycerol acyltransferase)
icl1	Rv0467	Isocitrate lyase (glyoxylate shunt enzyme)
narK2	Rv1737c	Probable Nitrate/Nitrite transporter (excretion of nitrite), narK2
virS	Rv3082c	Virulence-regulating transcriptional regulator, virS
mce3C	Rv1968	Mammalian cell entry family protein (possible cell invasion protein)
ltp2	Rv3540c	Probable lipid transfer protein or Keto Acyl-CoA Thiolase
Rv3180c	Rv3180c	Conserved hypothetical; possible toxin (part of toxin-antitoxin operon with Rv3181c)
ppsA	Rv2931	Phenolpthiocerol Synthesis Type-I Polyketide Synthase

Firstly, a number of lipid and cholesterol metabolism genes were found to be induced in sputum. The genes that may be involved in cholesterol metabolism (as identified by Cole et al. (1998)), as a group, were significantly upregulated as compared to aerobic growth. Individually, a number of genes that may be involved in β-oxidation (from the *fadB, echA, fadE* and *fadA* families) were upregulated >2.5 fold; furthermore, *icl1*, a key lipid metabolism was also more highly expressed compared to aerobic growth.

Of the lipid and cholesterol metabolism genes, ***icl1***, ***ltp2*** and **Rv3551** were selected as 3 of the target genes. *Icl1* was reviewed in Chapter 1. *Ltp2* (Rv3540c) is a member of the putative *kstR* regulon that controls cholesterol utilisation; *ltp2* is a probable ketoacyl-coA thiolase, involved in β-oxidation (Kendall et al., 2007). Furthermore, the gene has been identified as a member of the igr (Rv3545c-Rv3540c) operon, found to be essential for cholesterol metabolism, intracellular growth and virulence (Chang et al., 2009).

Rv3551 was selected out of interest as it is predicted to encode for a subunit of a CoA-transferase; it may be involved in the metabolism of cholesterol, as it has been identified to be a member of a cholesterol catabolic gene cluster (Van der Geize et al., 2007). Further, it may have significance in the virulence of *Mtb*, as it has been identified by Rengarajan et al. (2005) as essential for the adaptation and survival of the bacillus in macrophages.

Members of the DosR regulon of genes (reviewed in Chapter 1) were found to be upregulated in the microarray data. ***Tgs1*** (triacylglycerol synthase 1), ***hspX*** (also known as *acr*, encoding an alpha-crystallin-like protein) and ***narK2*** (a nitrate/nitrite transporter) were selected from the 48-gene regulon. All three genes were reviewed in detail in Chapter 1.

Rv3180c was selected as it represents a possible toxin, part of a toxin-antitoxin (TA) operon with Rv3181c (Ramage et al., 2009). The regulation of TA systems are based on the differential stability of the stable toxin and unstable antitoxin and the systems are unusually abundant in *Mtb*, although the reasons for this are unknown (Gerdes et al., 2005). It has been suggested that TA systems may function to stabilise the *Mtb* chromosome or protect adjacent regions of the chromosome from deletion (Ramage et al., 2009). They may work as stress response genes and there is evidence that some TA systems may be induced following exposure to stress conditions, allowing the cell to respond to the adverse conditions (Provvedi et al., 2009).

Ramage et al. (2009) identified 88 putative TA system candidates; expression of each of the candidates (10 were untested) in *M. smegmatis* demonstrated that 30 encoded for a functional toxin and paired antitoxin. Four confirmed toxins (as part of a TA system) were downregulated in sputum. However, Rv3180c was the only confirmed or untested (putative) toxin to be upregulated in sputum (Garton et al., 2008). The sheer number of TA systems in the *Mtb* genome suggests their importance and thus, Rv3180c was selected for further study.

Mce3C is a member of the mammalian cell entry (*mce*) genes and was found to be upregulated in sputum. Arruda et al. (1993) were the first to describe a DNA fragment from *Mtb* that conferred the ability to enter mammalian cells on *E. coli* cells and named the gene *mcel* for 'mammalian cell entry locus'. Following the release of the complete *Mtb* genome sequence, four *mce* operons were identified and designated *mce1* to *mce4* (Cole et al., 1998). *Mce3* was shown to be expressed in infection animal tissue; when disrupted, it has been shown by two independent studies to lead to the attenuation of *Mtb* in mice (Gioffre et al., 2005, Kumar et al., 2003, Senaratne et al., 2008).

VirS encodes for a transcriptional regulatory protein that is homologous to bacterial virulence regulating proteins in Shigella, Yersinia and various *E. coli* strains (Gupta et al., 1999). The virS protein is essential for the induction of the *mymA* operon in *Mtb,* which is upregulated at acidic pH and within macrophages; analysis of the operon suggested that the proteins encoded are involved in the modification of fatty acids required for the cell envelope (Singh et al., 2003).

PpsA encodes for a polyketide synthase, involved in the biosynthesis of ptthiocerol dimycocerosate (PDIM), a component of the cell wall (1.2.2). While the function of PDIM remains to be determined, it has been suggested that it has a role in countering the early immune response of the host (McKinney, 2000). A study by Kirksey et al. (2011) lends support to this theory, as PDIM has been shown to be required for PDIM-deficient variants of *Mtb* were more susceptible to IFNγ mediated immunity in mice. However, the importance of PDIM may be greater, as it may also be required for the virulence and resisting IFNγ-independent immunity also (Murry et al., 2009). The upregulation of *ppsA* in sputum may help the bacillus adapt to transmission, and was thus included in the set of 10 upregulated genes to be studied further.

4.1.2.2 Downregulated Genes

As for the upregulated genes, 10 of the 334 significantly repressed genes in sputum (compared to aerobic growth) are highlighted here; the genes and their products are shown in Table 10.

As genes involved in aerobic respiration and ribosomal functions were found to be particularly repressed in the microarray data, genes from these functional groups comprise the majority of the 10 downregulated genes selected. Five of the selected downregulated genes were involved in aerobic respiration, encoding the products as shown in Table 10: ***ctaD***, ***qcrC***, ***nuoB***, ***nuoL*** and ***atpD***. In particular, *ctaD, qcrC, nuoB* and *atpD* have been shown to be repressed during bacillary stasis in a chronic model (Shi et al., 2005).

RpsL was selected for the further analysis of ribosomal gene expression in this study. *RpsL* is a ribosomal gene, which encodes protein S12, part of the 30S ribosomal protein, and has been demonstrated to be highly expressed in other bacterial systems (Kenney and Churchward, 1996). In sputum, the gene was shown to be downregulated along with other ribosomal genes, suggesting activation of the stringent response (Stallings et al., 2009).

Rv1103c is a putative antitoxin, as part of a TA system with Rv1102c (Ramage et al., 2009). TA systems were discussed above. As Rv1103c was the only antitoxin to be downregulated in sputum, it was selected for further analysis; intriguingly, the paired toxin, Rv1102c, was not upregulated greater than 2.5 fold in sputum (Garton et al., 2008). No confirmed antitoxins were upregulated in sputum.

Table 10: 10 Downregulated Genes and Products

Gene	Gene Number	Product
Rv1103c	Rv1103c	Conserved hypothetical; possible mazE3, antitoxin, part of toxin-antitoxin operon with Rv1102c
atpD	Rv1310	Probable ATP synthase beta chain
Rv2141c	Rv2141c	Conserved hypothetical; possible protease
qcrC	Rv2194	Probable Ubiquinol-cytochrome C reductase, QcrC (cytochrome C subunit)
mmaA2	Rv0644c	Methoxy-mycolic acid synthase 2, mmaA2
ctaD	Rv3043c	Probable cytochrome C oxidase polypeptide I, ctaD
nuoL	Rv3156	Probable NADH dehydrogenase I (chain L), nuoL
mce1A	Rv0169	Mammalian cell entry family protein (possible cell invasion protein)
nuoB	Rv3146	Probable NADH dehydrogenase I (chain B), nuoB
rpsL	Rv0682	Probable 30S ribosomal protein S12, rpsL

Mce1A is a member of the *mce* family of genes; as described above (4.1.2.1), when discovered (named *mcel*), it individually conferred on *E. coli* the ability to enter non-phagocytic cells (Arruda et al., 1993). *Mce1A* however, is one of 12 *mce1* operon genes; unexpectedly, unlike the other *mce* operons, Shimono et al. (2003) demonstrated that the disruption of the complete *mce1* operon led to a hypervirulent mutant. The mutant was unable to persistently infect mouse lungs, but instead continued to replicate and killed the mouse (Shimono et al., 2003). In sputum, 6 of the 12 *mce1* genes are downregulated; however, at this time, the specific product or combination of *mce1* products that mediates the hypervirulence remains unknown.

As bacilli in sputum have a unique phenotype, there may be differences in the cell wall composition. ***MmaA2*** is downregulated in sputum and encodes for a methoxy-mycolic acid synthase, involved in the modification of mycolic acids. While *mmaA2* has not been extensively studied, it is a member of a locus of 4 genes. The closely related *mmaA4* is involved in virulence of *Mtb;* mutation of this gene reduces the production of methoxy-mycolates, but leads to greater resistance to oxidative stress (Smith, 2003).

Rv2141c was selected as it was the only gene out of 21 identified to be downregulated on nitric oxide exposure to be similarly repressed in sputum (Ohno et al., 2003). This was of particular interest, as many of the genes upregulated upon exposure to nitric oxide are not

specific for this stimulus and may be induced in response to other stresses. The gene itself encodes for a conserved hypothetical protein of unknown function.

4.1.3 Growth Conditions

In this study, one of the aims was to replicate the sputum transcriptome in an *in vitro* setting. As Garton et al. (2008) have shown that the gene expression of bacilli in sputum differs from aerobic, exponential growth, the use of conventional 7H9 or Sautons media as a base medium would be inappropriate for this study. While it was unlikely that a single growth medium would be adequate to fully replicate the *Mtb* gene expression in sputum, a number of preliminary conditions were selected to test.

The growth media were selected for a number of reasons, as described below. However, one requirement was that the conditions needed to be compatible for subsequent macrophage infection; in addition to the examination of gene expression, a further aim was to study the effect of sputum-like conditions on macrophage binding (Chapter 5).

As the sputum gene expression suggested that cells were slow growing or non-replicating/persistent, the expression of stationary phase *Mtb* culture was of interest and cultures were grown to stationary phase without the addition of other external stimuli.

RPMI-1640 was also of particular interest; the medium is commonly used for the culture of eukaryotic cells, e.g. THP-1 acute monocytic leukaemia cell line (Tsuchiya et al., 1980). As human cell lines are capable of growing in such a medium, RPMI-1640 was selected as it may provide a growth environment more representative of the *in vivo* environment than conventional *Mtb* culture medium, e.g. 7H9 or Sautons broth.

As reviewed in Chapter 1, the sputum transcriptome features the downregulation of aerobic respiration and ribosomal function genes. This repression of ribosomal function suggests activation of the bacterial stringent response; which can occur in response to a number of stresses, including nutrient deprivation (Stallings et al., 2009). As a number of the downregulated energy metabolism and ribosomal function genes in sputum coincided with genes shown to be repressed in response to nutrient starvation, PBS was selected as a growth medium as it has been used in past studies in the examination of this stress (Betts et al., 2002).

As introduced in Paragraph 1.4.3.1, mycobacteria may also face attacks to the cell surface. Studies using the detergent SDS to induce cell wall damage have identified genes upregulated in response to this damage; a proportion of these genes were also upregulated in *Mtb* growth in macrophages, suggesting similar cell wall damage occurs *in vivo* (Schnappinger et al.,

2003). Therefore, as SDS has been used in past studies and may simulate the cell wall damage in macrophages or secondary to the detergent effects of surfactant, exposure to SDS was selected as one of the initial growth conditions to be tested (Pang et al., 2007). While SDS exposure may replicate *in vivo* cell wall stress, SDS has also been used to lyse macrophages during infection experiments; thus, SDS was selected to examine how its use may alter the gene expression results.

The gene expression in artificial saliva was the last of the five growth media. *Mtb* bacilli may be exposed to saliva during the process expectoration as evidenced by the fact that there is often a large proportion of saliva in sputum samples (Simpson et al., 2004). Artificial saliva was used as it was easier to obtain and standardise as compared to saliva obtained from multiple patients.

4.1.4 Aims

In this chapter, the use of quantitative real-time RT-PCR to analyse the gene expression of *Mtb* exposed to various growth conditions and combinations of growth conditions is addressed. This was conducted with the aim of identifying stimuli responsible for the sputum transcriptome and replicating the gene expression of *Mtb* in sputum in an *in vitro* setting. Work discussed in this chapter includes:

- Selection of upregulated and downregulated genes to be analysed.
- Validation of qPCR primers and cycling conditions.
- Analysis of gene expression of *Mtb* cultures exposed to selected growth conditions and/or stimuli.
- Comparison of gene expression to sputum microarray data.
- Analysis of the production of lipid bodies as compared to *Mtb* in sputum.

4.2 Results

4.2.1 Experimental Strategy

The experimental strategy used to address the research objectives in this chapter is outlined in Figure 22.

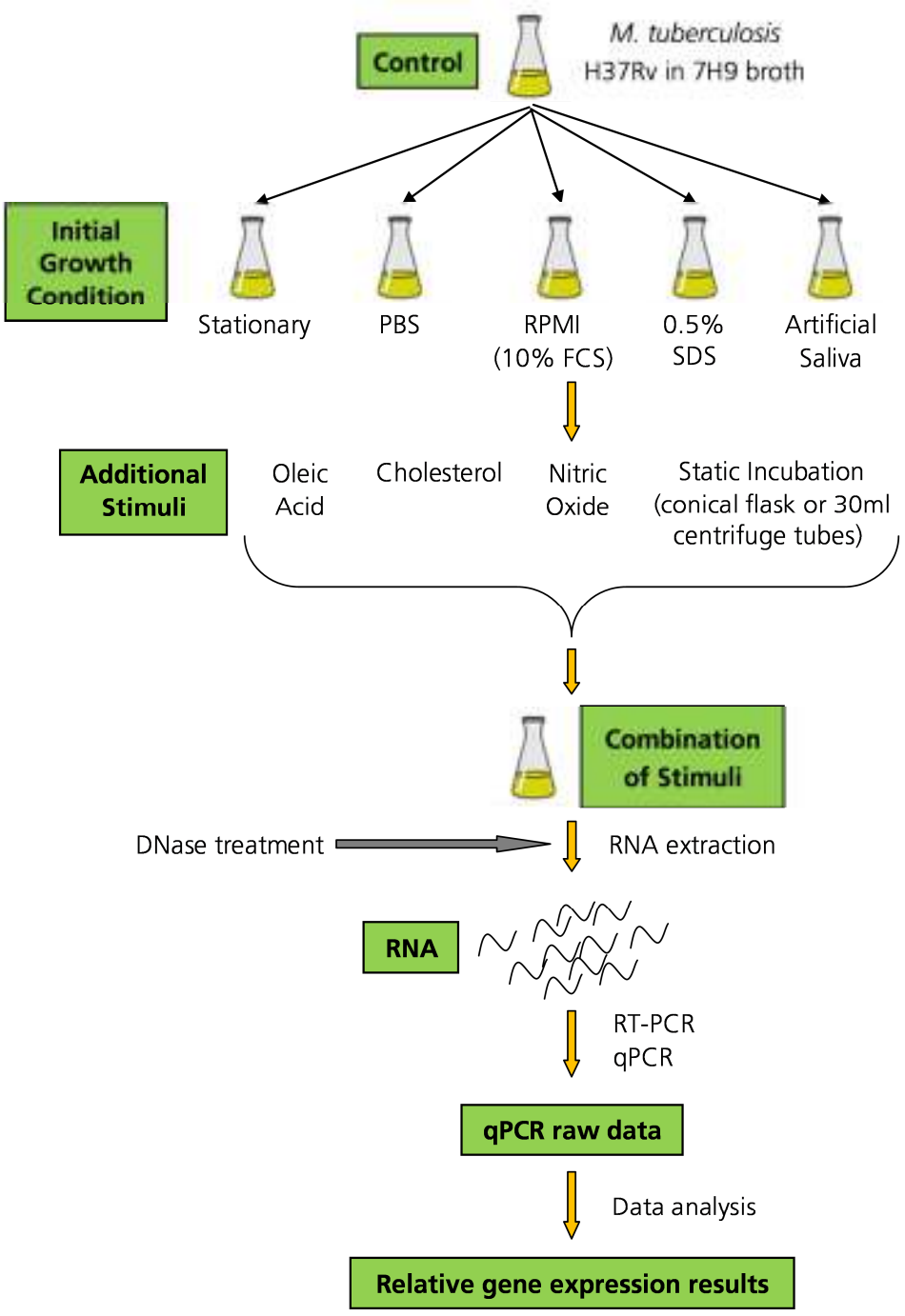


Figure 22: Pictorial Representation of Gene Expression Studies

The flowchart demonstrates the *Mtb* gene expression studies starting from *Mtb* in 7H9 broth through to Quantitative PCR and Data analysis.

The experiments were performed in a step-wise fashion. The gene expression of the initial five growth conditions (plus exponential phase control) was examined first, in biological triplicates. The addition of other stimuli (nitric oxide (NO), cholesterol, static incubation, oleic acid) was performed subsequently in duplicate. All PCRs were performed in triplicate.

The gene expression analysis presented below made use of the work described in Chapter 3, including relative quantification (Pfaffl) and the normalisation factor. Other preliminary work was not presented in the following results. Preliminary RNA extractions were performed on RPMI-1640 (also referred to as RPMI) medium to confirm the quality of RNA extracted from this medium. RNA was extracted and purified as described in Chapter 3. The RNA integrity was confirmed through imaging of intact 16S and 23S bands following gel electrophoresis.

Furthermore, RNA extractions and qRT-PCRs were performed following 1h, 4h, 8h and 24h of RPMI exposure to compare expressions changes. Gene expression changes were visible in all four time-points, but were minimal after 1h. Clumping was noted at the 4h, 8h and 24h time-points, but was worst at 8h and 24h. Thus, the 4h exposure was selected for further experiments. This exposure time was the appropriate balance for the induction of gene expression changes, minimization of clumping and for work within the C3 laboratory suite code of conduct.

4.2.2 Selection of Upregulated and Downregulated Genes

Microarray analysis of *Mtb* is useful for monitoring of expression changes for a large number of genes. However, the technique becomes expensive, complex and impractical when a large number of growth conditions are used, particularly for preliminary or screening experiments (Schena, 2003). Furthermore, due to the more limited dynamic range of fluorescent microarrays, key microarray results must be validated using more sensitive quantitative RT-PCR techniques (Groen, 2001). However, gene expression studies using quantitative RT-PCR to measure genome wide expression changes, while having been employed in other non-tuberculosis studies for up to 200 genes (Dolganov et al., 2001), would have similar problems as microarrays, with the technique becoming expensive and complex for the large number of growth conditions used in preliminary/screening experiments.

Thus, for this study, ten upregulated (>2.5 fold upregulation) and ten downregulated (>2.5 fold downregulation) were selected from the sputum microarray gene expression to be analysed through quantitative RT-PCR analysis (Garton et al., 2008). The genes were selected to be representative of the microarray gene expression trends, as introduced in Chapter 1; furthermore, genes of particular interest were also included. The 10 selected upregulated

genes and their products (or predicted products) were listed in Table 9 and introduced in 4.1.2.1. The 10 downregulated genes were listed in Table 10 and were introduced in 4.1.2.2.

Figure 23 shows the genes expression levels of the 20 selected genes in sputum. The microarray data was globally normalised to the 50th percentile (median) of all genes detected to be present on the array (Garton et al., 2008, Park et al., 2003b). The microarray data could not be normalised to sigA or the normalisation factor (NF4, geometric mean of *thyA, dfrA, polA* and *aroA*). As expected, none of the 4 normalising factor housekeeping genes (*thyA, dfrA, polA* and aroA) or sigA were upregulated or downregulated significantly; thus, expression data for these genes was unavailable (Garton et al., 2008).

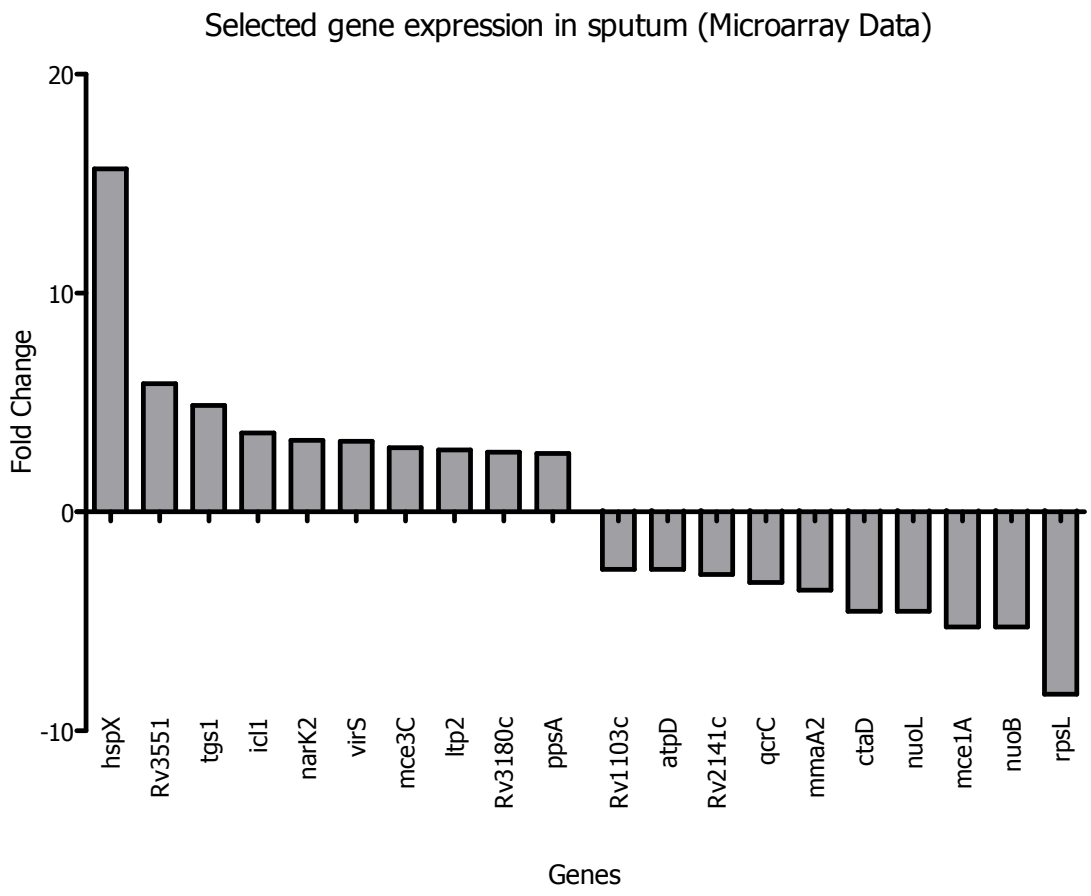


Figure 23: Selected Gene Expression of *M. tuberculosis* in sputum (Microarray Data)

The bar graph displays the microarray gene expression data of the 20 selected genes in sputum (Garton et al., 2008). The x-axis displays the genes and the y-axis displays the fold change of each respective gene. The figure is arranged with highest upregulation to lowest downregulation from left to right. The first 10 genes are upregulated and the second 10 genes are downregulated.

Within this arrangement, the upregulated and downregulated genes are displayed in descending order based upon fold change, i.e. highest upregulation to lowest downregulation. The numerical data for the bar graph is available in Appendix 4.

4.2.3 Validation of Quantitative PCR Assays

The validation of cycling conditions and primer sets for each transcript prior to their use in experiments is essential to confirm correct and efficient amplification of selected gene. The amplification was first confirmed using the *In silico* PCR amplification simulation program (University of the Basque Country, 2010). PCR assays were simulated using each primer set against the *Mtb* H37Rv genome, with up to 2 mismatches, to ensure that only a single gene product of the correct size was produced in each PCR assay.

Following the simulation, a qPCR assay for each of the primer sets using CDC1551 gDNA was conducted using the method described in (3.2.7). The cycling conditions used for all the target qPCRs were described in 3.2.7, and the annealing temperatures in Appendix 1. The amplified cDNA from each reaction and the corresponding noRT control was visualised by gel electrophoresis (2% (w/v) agarose) to confirm that a single band was amplified for each reaction and that no gDNA was amplified in the noRT control. A further negative control was included for each primer set. The gel electrophoresis for amplified gDNA using the 20 primer sets is shown in Figure 24. The primers used are listed in Appendix 1.

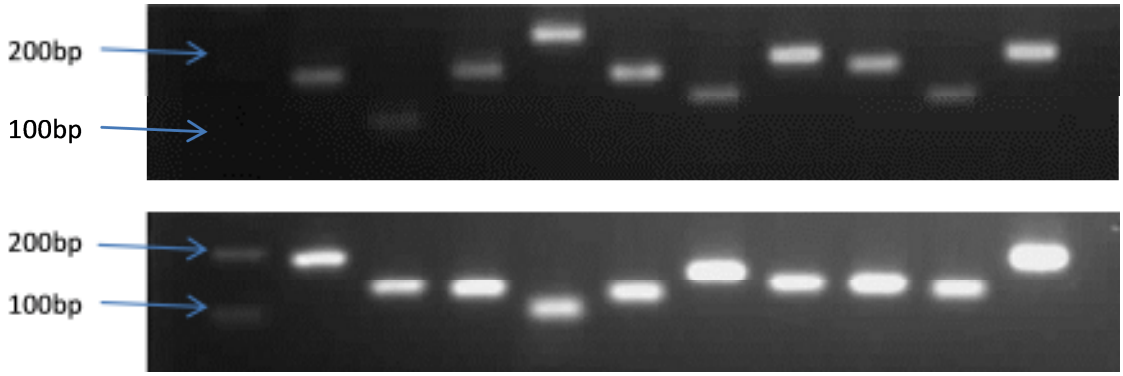


Figure 24: Amplified PCR reactions using each primer set, as visualised by gel electrophoresis

The figure shows the amplified PCR reactions as visualised by gel electrophoresis. The specific genes were amplified from *Mtb* CDC1551 genomic DNA obtained from Colorado State University.
The top row shows the upregulated set of genes (from left to right): pre-prepared DNA ladder, *icl1* (181bp), *narK2* (143bp), *mce3C* (187bp), *hspX* (235bp), *ppsA* (176bp), *virS* (144bp), *tgs1* (197bp), Rv3180c (179bp), *ltp2* (102bp) and Rv3551 (190bp).
The bottom row shows the downregulated set of genes (from left to right): pre-prepared DNA ladder, *mce1A* (186bp), *mmaA2* (133bp), *rpsL* (125bp), Rv1103c (91bp), *atpD* (112bp), Rv2141c (106bp), *qcrC* (130bp), *ctaD* (128bp), *nuoB* (118bp) and *nuoL* (167bp).

The gel electrophoresis for the amplified PCR reactions confirmed the amplification of a single band of the correct size for each of the 20 primer sets. The average efficiencies for all 20 PCR reactions were greater than the designated 80% cut-off limit. Further analysis of the SYBR green fluorescence melt curve confirmed a single product peak for each primer set and excluded the production of primer-dimers.

The average efficiencies of the PCR reactions were also analysed using LinRegPCR and displayed in Table 11. An efficiency value of 80% (1.80) was used as the cut-off value. The final primer sets used for the 20 target genes all had efficiencies above this cut-off.

Table 11: Average Efficiencies for Each PCR Reaction

Gene Name	Average Efficiency	Gene Name	Average Efficiency
hspX	1.835	*Rv1103c*	1.887
Rv3551	1.832	*atpD*	1.924
tgs1	1.863	Rv2141c	1.950
icl1	1.872	*qcrC*	1.916
narK2	1.952	*mmaA2*	1.895
virS	1.927	*ctaD*	1.878
mce3C	1.935	*nuoL*	1.925
ltp2	1.902	*mce1A*	1.915
Rv3180c	1.938	*nuoB*	1.915
ppsA	1.926	*rpsL*	1.953

Finally, the intra-run variation was assessed by calculating the coefficient of variation (CoV) across the 3 PCR technical replicates, for each primer set. The intra-run variation was calculated over 6 growth conditions and 3 biological replicates, giving a total of 18 CoV values. The average CoV (plus/minus standard deviation) for each primer set is shown in Table 12. The variation between the biological replicate PCR runs could not be calculated, as the biological replicates differed in total amount of cDNA. The average CoV was below 15% for all the primer sets; 12 of the 20 primer sets were had an average CoV below 10%.

Table 12: Average Intra-Run Coefficient of Variation (CoV) for 20 Primer Sets

Gene Name	Average CoV ± SD (%)	Gene Name	Average CoV ± SD (%)
hspX	13.7 ± 10.4	*Rv1103c*	4.41 ± 3.49
Rv3551	13.4 ± 8.32	*atpD*	5.50 ± 4.40
tgs1	14.9 ± 9.96	Rv2141c	8.02 ± 6.24
icl1	6.10 ± 3.89	*qcrC*	4.13 ± 2.33
narK2	14.6 ± 9.31	*mmaA2*	7.34 ± 2.23
virS	13.4 ± 14.9	*ctaD*	5.64 ± 4.34
mce3C	14.2 ± 9.21	*nuoL*	7.69 ± 8.85
ltp2	12.8 ± 10.8	*mce1A*	4.94 ± 4.11
Rv3180c	16.7 ± 11.2	*nuoB*	5.08 ± 4.13
ppsA	4.93 ± 3.42	*rpsL*	4.31 ± 2.32

4.2.4 Gene expression of *M. tuberculosis* H37Rv under selected growth conditions

The transcriptome of *Mtb* bacilli in sputum is unique, compared with established, *in vitro* growth conditions, particularly aerobic, exponential growth (Garton et al., 2008). Therefore, the use of 7H9 or Sautons media as a base medium would be inappropriate for the study of the sputum transcriptome in an *in vitro* setting.

As an initial experiment (4.2.1 – Experimental Strategy), gene expression of *Mtb* H37Rv for the 20 target genes was examined under selected alternative growth conditions and compared to the sputum transcriptome. Cultures were first grown to exponential growth phase as described in 2.3.2, and transferred to the selected growth media in triplicate. Conical flasks were used for the preparation of cultures. The cultures were:

- Grown to exponential growth phase (control).
- Grown to stationary phase (constant optical density for three consecutive days).
- Exposed to PBS for 4 hours at 37°C, shaking at 100rpm.
- Exposed to complete RPMI medium for 4 hours at 37°C, shaking at 100rpm.
- Exposed to final concentration of 0.5% (w/v) SDS for 1 hour at 37°C, shaking at 100rpm.
- Exposed to artificial saliva at 37°C, shaking at 100rpm.

RNA was extracted and purified from these cultures as described in Chapter 3. The RNA was reverse transcribed into cDNA and diluted 1:4. Quantitative PCRs were performed in

triplicate for each of the extracted samples for the 20 selected genes and 4 housekeeping genes.

The results were analysed using the Pfaffl (relative quantification) method. The quantitative PCR and biological triplicates were averaged. Gene expression fold changes for stationary phase, PBS, RPMI-1640, SDS and artificial saliva exposed cultures were calculated in relation to exponential growth phase cultures based on Equation 11.

Equation 11: Calculation of Fold Changes

$$Fold\ Change\ for\ Selected\ Gene = \frac{(Test\ Culture\ Gene)\ /\ (Test\ Culture\ NF4)}{(Control\ Culture\ Gene)\ /\ (Control\ Culture\ NF4)}$$

To examine the relationship between the gene expression of the 20 target genes for each growth condition against the sputum microarray gene expression, Spearman's Rank Correlation was used for all samples, as the microarray gene expression data had non-Gaussian distribution. P-values, R values and 95% confidence intervals are displayed in Table 13.

Table 13: Correlation Values for Single Growth Condition Gene Expression versus Sputum Microarray

Sample	P-Value	Spearman R-Value (95% Confidence Interval)
Stationary Phase / Exponential Phase	0.0109	0.5563 (0.1370 – 0.8065)
PBS / Exponential Phase	0.1258	0.3539 (-0.1191 – 0.6960)
RPMI / Exponential Phase	0.7477	0.07678 (-0.3907 – 0.5128)
SDS / Exponential Phase	0.5886	0.1287 (-0.3453 – 0.5504)
Artificial Saliva / Exponential Phase	0.1660	0.3222 (-0.1542 – 0.6770)

Correlation results showed a statistically significant relationship between the gene expression of the stationary phase culture and the sputum microarray gene expression, with an approximate p-value of <0.05. However, while the relationship between the two is significant, as the Spearman's Rank Test is non-parametric, a positive correlation only signals a monotonic function (any function that preserves a given order) between the two variables; the test does not necessarily signify a linear relationship.

The p-values were non-significant and the R values for the other 4 growth conditions are close to zero. This suggests no relationship between the gene expression of *Mtb* exposed to any of these 4 growth conditions and the sputum transcriptome.

As the Spearman's Rank Correlation values did not necessarily show a linear relationship, the gene expression for each of the growth conditions was examined further. The expression of the 20 target genes is displayed visually in Figure 25. The expression pattern for the stationary phase cultures showed mainly upregulated genes; 4 out of 10 genes upregulated in sputum showed greater than 2 fold (1 $\log_2$) upregulation in stationary phase, based on their 95% confidence intervals. None of the genes downregulated in sputum were similarly repressed in stationary phase culture. Thus, while there was a statistical relationship between the expression in stationary phase culture and sputum, only 6 out of 20 genes appeared to have matching expression as compared to the sputum transcriptome.

The gene expression pattern for the other 4 growth conditions (PBS, RPMI, SDS and Artificial Saliva) confirmed the Spearman's Rank correlation result showing no relationship with the sputum transcriptome.

As AFBs in sputum showed a wide range of lipid body production, samples were also taken from each of the six cultures for microscopic analysis and stained with LipidTOX. Fluorescent images were taken and blinded with Image J, using the methods described in Chapter 2 (Bell, 2010). Lipid bodies (LB) were hand counted from blinded images; a minimum of 100 bacilli were counted for LB analysis. As images from three samples, stationary phase, SDS exposure and artificial saliva, featured significant clumping, fluorescence analysis could not be performed.

The LB results for all the growth conditions are shown in Figure 26. These included the number of cells analysed, number of positive cells and the frequency of LB-positive cells observed (as a percentage). LBs were observed in all samples including Exponential Phase culture with the highest frequency of LB-positive cells found in Stationary Phase culture.

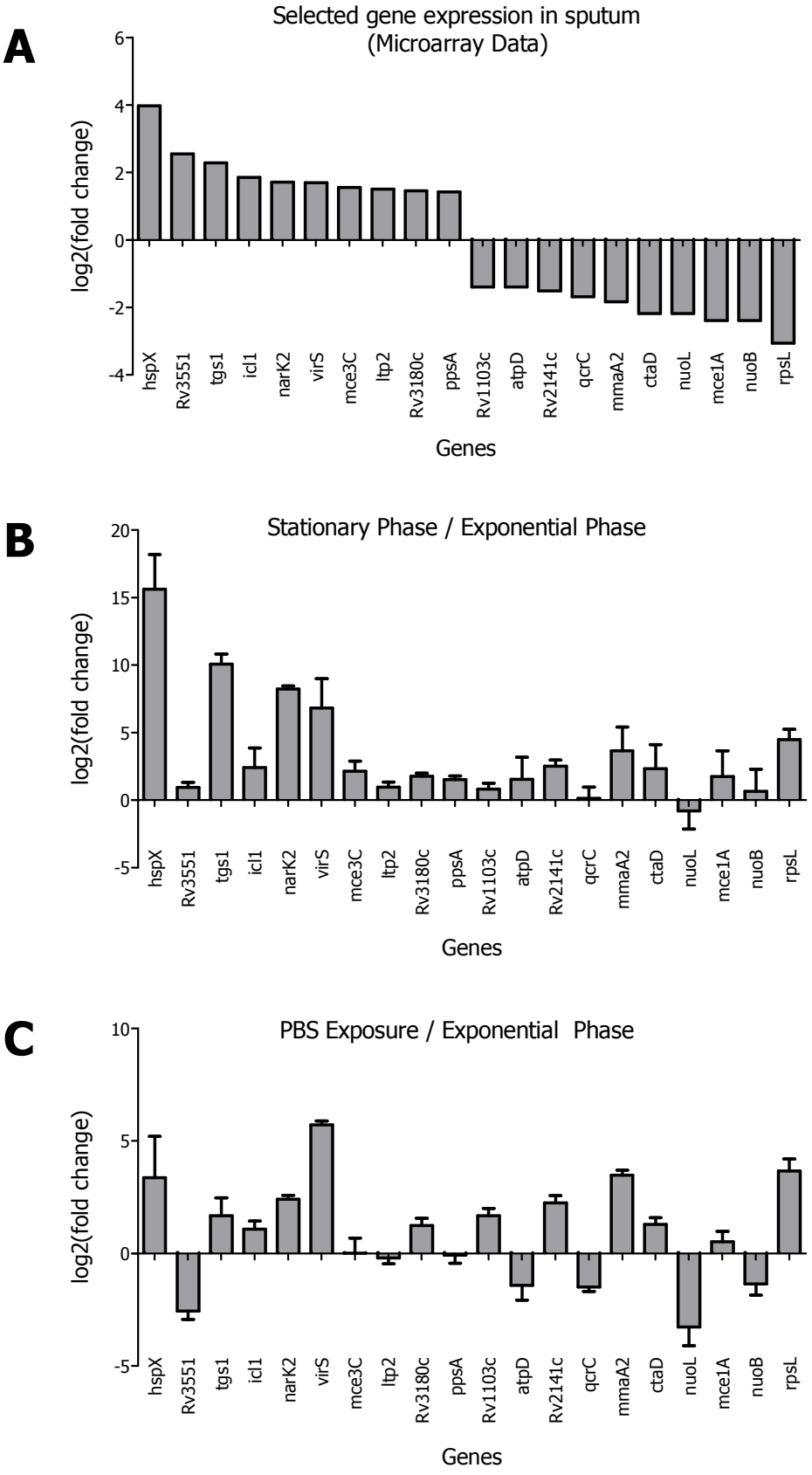
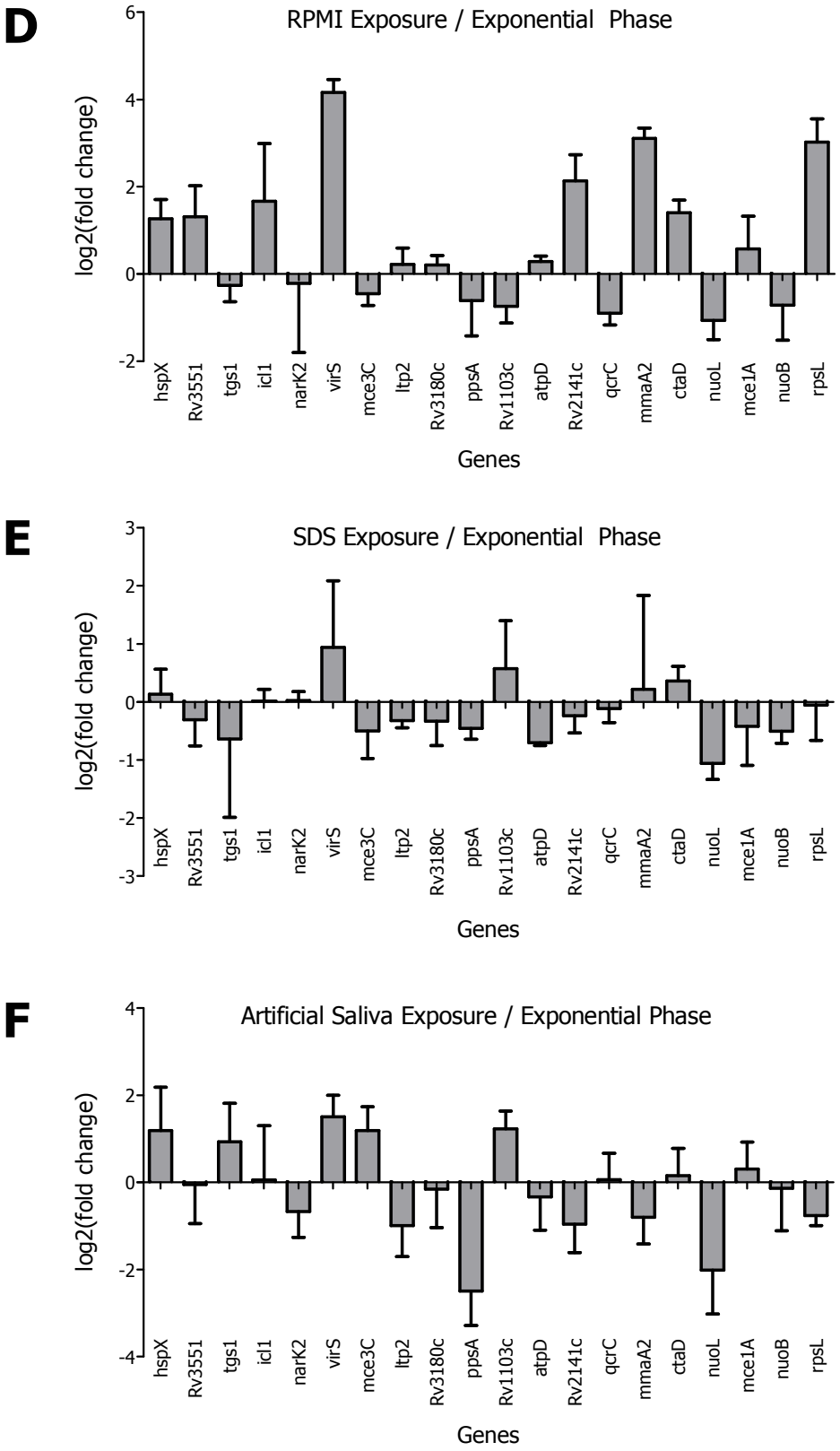
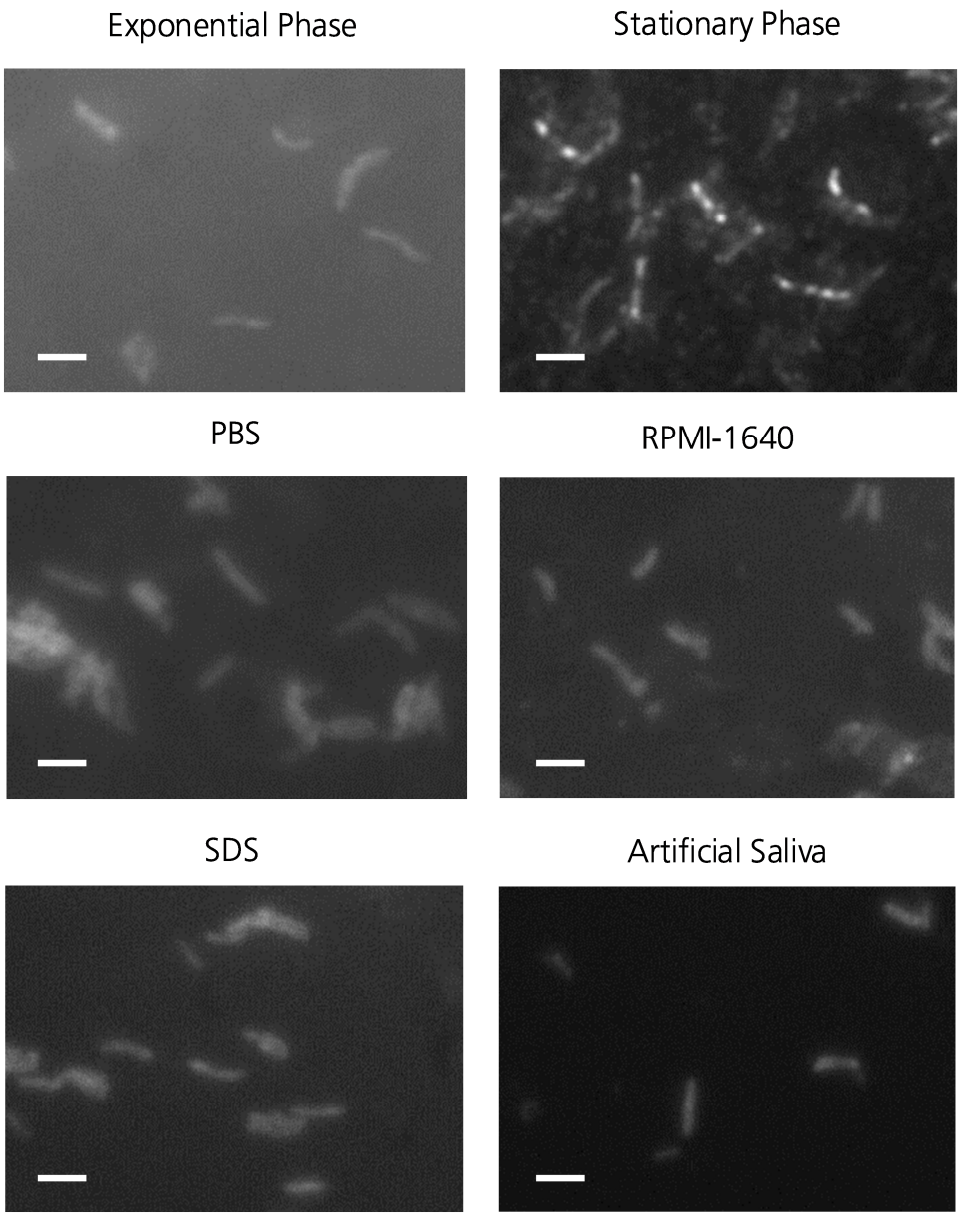


Figure 25: *M. tuberculosis* Gene Expression for 20 Target Genes under Selected Growth Conditions (Caption Continued Overleaf)



Bar graphs display gene expression of 20 selected target genes in *Mtb* for A) Sputum Microarray, B) Aerobic Stationary Phase, C) PBS exposure, D) RPMI exposure, E) 0.5% (w/v) SDS Exposure and F) Artificial Saliva Exposure; all fold changes calculated against aerobic exponential growth.



Sample	#LB +ve (# Cells total)	LB %
Exponential Phase (Control)	60 (283)	21.2
Stationary Phase	645 (1183)	54.5
PBS	68 (480)	14.2
RPMI	52 (431)	12.1
SDS	24 (165)	14.6
Artificial Saliva	47 (156)	30.1

Figure 26: Lipid Body Analysis of *M. tuberculosis* under selected growth conditions

Fluorescence images for control (exponential phase), stationary phase, PBS, RPMI, SDS and artificial saliva are displayed above. The table displays the number of cells analysed and lipid body positive% is shown. Bar is 2µm.

4.2.5 Gene expression of *M. tuberculosis* H37Rv under combined growth conditions

The gene expression of Stationary Phase *Mtb* culture correlated significantly to sputum expression; as the sputum expression showed slow-growth expression signals, this result was not unexpected. However, there were a number of issues associated with stationary phase growth (discussed further in 4.3.3.2). First, due to the non-Gaussian distribution of the Stationary Phase gene expression, Spearman's Rank Correlation was used. A significant correlation demonstrated a relationship between the two variables; however, this relationship may not have be linear. Examination of the stationary phase gene expression results directly (Figure 25) showed upregulation of the majority of the target genes, but none of the 10 downregulated genes in sputum were similarly repressed in stationary phase. Furthermore, while initial results suggested that stationary phase growth was the most closely related growth condition, the time frame required for cultures to reach stationary phase was >7 days; however, as introduced previously, it was postulated that this time frame is unrealistic in an *in vivo* setting.

Thus, while stationary phase culture does show some similar expression to the sputum transcriptome, further experiments were carried out on the gene expression of *Mtb* H37Rv under a combination of growth conditions. Instead of stationary phase growth in 7H9 media, PBS and RPMI were selected as the base media for further experiments (discussed in 4.3.3.2).

Cultures were grown to exponential growth phase in conical flasks. The cultures were then exposed to a combination of growth conditions. Duplicate exponential *Mtb* cultures were prepared in 7H9 media, PBS or complete RPMI medium. Table 14 shows the stimuli added to the test cultures and the relevant control cultures, using the methods described in Chapter 2. Further to the addition of single stimuli to 7H9, PBS and/or RPMI, a combination of multiple stimuli were added to PBS or RPMI; these samples are subsequently referred to as 'PBS/RPMI multi-stimulus (30ml tube/flask) cultures'.

Nitric oxide was chosen as a stimulus as exposure has been shown to induce an *Mtb* dormancy program (Voskuil et al., 2003). Nitric oxide inhibited mitochondrial and bacterial respiration and initiated a response similar to that induced by the reduction of oxygen in the *in vitro* dormancy model latency, as reviewed earlier (Voskuil et al., 2003). Furthermore, clinical tuberculosis patients may expel increased levels of nitric oxide, suggesting that greater levels are produced during infection (Van Beek et al., 2011).

Table 14: Stimuli added for Stimuli Test Cultures and Experimental Control Cultures

Test Culture	Control Culture
Nitric oxide: cultures resuspended in RPMI-1640 or PBS with added spermine NOnoate solution (1h and 4h time-points)	7H9 broth with added spermine hydrocholoride (1h and 4h time-points)
Oleic acid: 7H9 broth with added oleic acid (in BSA solution) (4h time-point)	7H9 broth with added BSA solution (4h time-point)
Cholesterol: cultures resuspended in 7H9 pre-prepared with cholesterol (in 100% (w/v) tyloxypol/ethanol solution) (4h time-point)	Control cultures resuspended in 7H9 pre-prepared with 100% (w/v) tyloxypol/ethanol solution (4h time-point)
Static Incubation in conical flasks: 7H9 broth in conical flasks, incubated at 37°C, statically (4h time-point)	7H9 broth in conical flasks, incubated at 37°C, shaking at 100rpm (4h time-point)
Static Incubation in 30ml centrifuge tubes: 7H9 broth in 30ml centrifuge tubes, incubated at 37°C, statically (4h time-point)	7H9 broth in conical flasks, incubated at 37°C, shaking at 100rpm (4h time-point)
Multi-stimulus: cultures resuspended in PBS or RPMI-1640 in conical flasks or 30ml centrifuge tubes with added: ▪ Cholesterol (in 100% (w/v) tyloxypol/ethanol solution (pre-prepared) ▪ Oleic acid in BSA ▪ Spermine NOnoate ▪ Static incubation at 37°C (4h time-point)	Cultures resuspended in fresh 7H9 broth in conical flasks with added: ▪ 100% (w/v) tyloxypol/ethanol solution (pre-prepared) ▪ BSA ▪ Spermine hydrochloride ▪ Shaking incubation at 100rpm at 37°C (4h time-point)

The sputum transcriptome showed upregulation of both lipid and cholesterol utilisation genes. Thus the addition of both oleic acid (a monosaturated omega-9 fatty acid found in various animal and vegetable lipids) and cholesterol to the media was made. The availability of oleic acid or cholesterol may allow *in vitro Mtb* to induce genes required to accumulate and/or utilise the lipid/cholesterol. Finally, as genes involved in aerobic respiration were repressed in sputum, static incubation was selected as the final stress/stimulus. Static incubation formed an oxygen gradient and hypoxia in culture (Bacon et al., 2004) and may lead to the downregulation of these aerobic respiration genes.

Following exposure to stimuli, RNA was extracted and purified through the method described previously. The RNA for each of the samples was reverse transcribed into cDNA and diluted 1:4. Quantitative PCRs were performed in triplicate for each of the extracted samples for the 20 selected genes and the 4 housekeeping genes. The results were normalised to a 4 gene

normalising factor and analysed using the Pfaffl (relative quantification) method. Quantitative PCR triplicates and biological duplicates were averaged. Gene expression fold changes for the test samples were calculated against the control cultures, using Equation 11.

The gene expression results for biological replicates for each growth condition and combination of stimuli are displayed in the heat map in Figure 27. The dendrogram shows the replicates clustered hierarchically according to expression profile using the complete linkage method. The sample replicates clustered closely; this clustering demonstrated that the biological replicates for a particular growth condition had the most similar gene expression as compared to other growth conditions. The exception to the clustering of replicates was for the RPMI multi-stimulus (30ml tubes) cultures. The replicates did not cluster together, but individually clustered to the PBS multi-stimulus (flasks) cultures and PBS multi-stimulus (30ml tubes) cultures.

In addition to the clustering of replicates, PBS and RPMI multi-stimulus cultures all clustered closely. It was notable that the RPMI multi-stimulus (flasks) cultures samples were the exception and clustered separately from the other multi-stimuli samples. Other groups of samples that also clustered closely were the cholesterol and oleic acid samples and the nitric oxide and static incubation samples.

Figure 28 is a dendrogram that clustered the sputum microarray gene expression with the gene expression of each growth condition (using the average of the biological replicates). Using multi-scale bootstrap resampling (with 10,000 replicates) and the complete linkage method, the sputum gene expression clustered most closely to the PBS and RPMI-1640 multi-stimulus (30ml tubes) cultures 81% (approximately unbiased probability value) of the time. The probability value (termed 'p-value') of a cluster is a value between 0 and 1 (expressed as a percentage), which indicates how strongly the cluster is supported by data; this so-called 'p-value' of 0.95 or 95% was deemed significant (Suzuki and Shimodaira, 2006). After those two samples, the sputum gene expression clustered with the nitric oxide exposed samples and statically incubated samples.

For comparison, the clustering of gene expression samples normalised to *sigA* was examined (results not displayed). Using an identical clustering method, the sputum gene expression clustered most closely to PBS and RPMI-1640 multi-stimulus cultures in both flasks and 30ml tubes, 79% (approximately unbiased p-value) of the time.

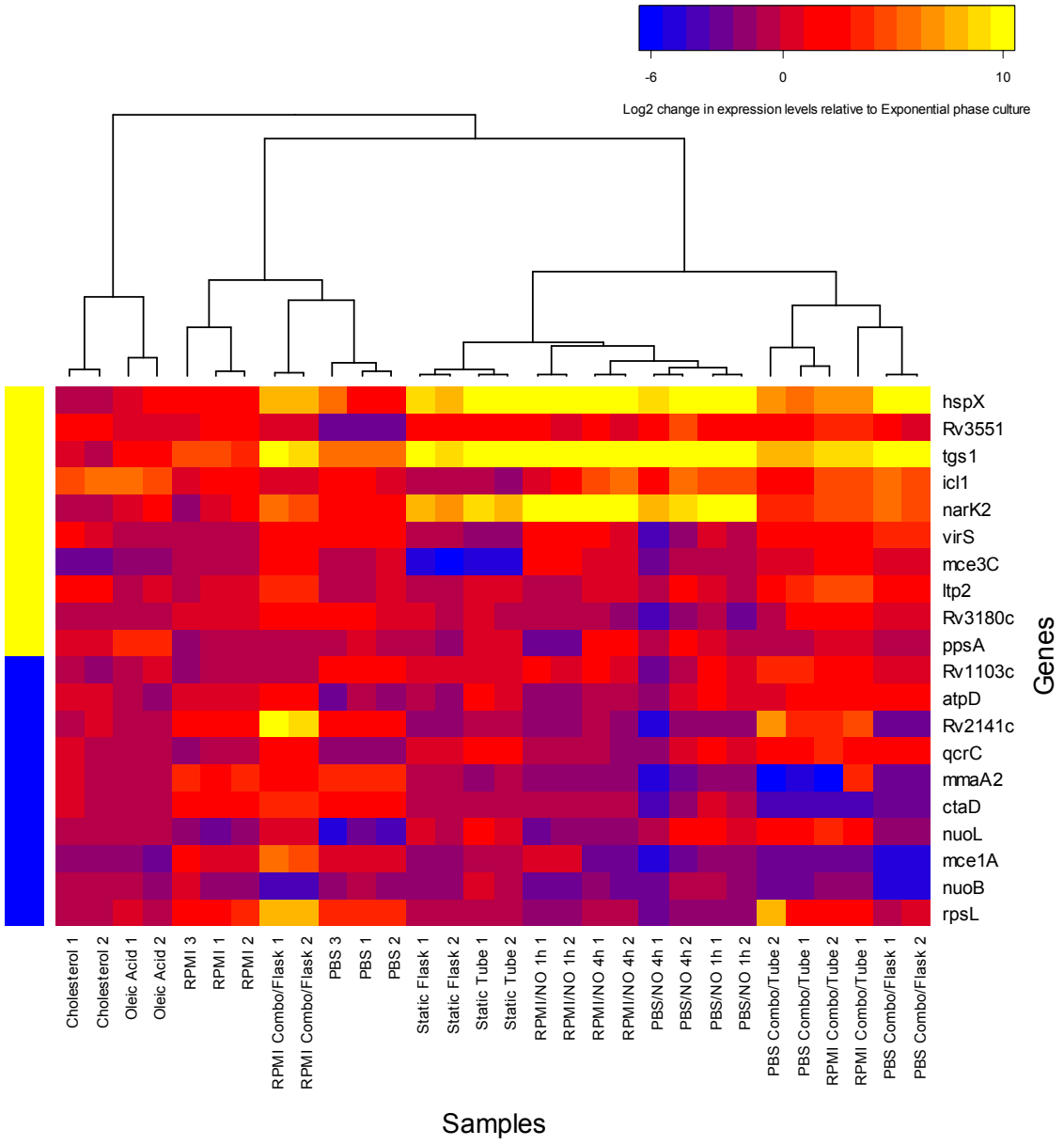


Figure 27: Heat Map showing Fold Changes for Sample Replicates

The heat map displays the expression of 20 target genes (selected based on sputum expression: 10 genes upregulated and 10 genes downregulated) in *Mtb* H37Rv under multiple growth conditions and stimuli. Gene expression/fold change is calculated relative to exponential phase *Mtb* H37Rv. Multi-stimulus cultures are labelled as 'Combo', i.e. combination. The legend displays the range of fold changes. The vertical bar on the left demonstrates the expected upregulated and downregulated genes (yellow and blue, respectively). Samples are hierarchically clustered according expression profile using the complete linkage method.

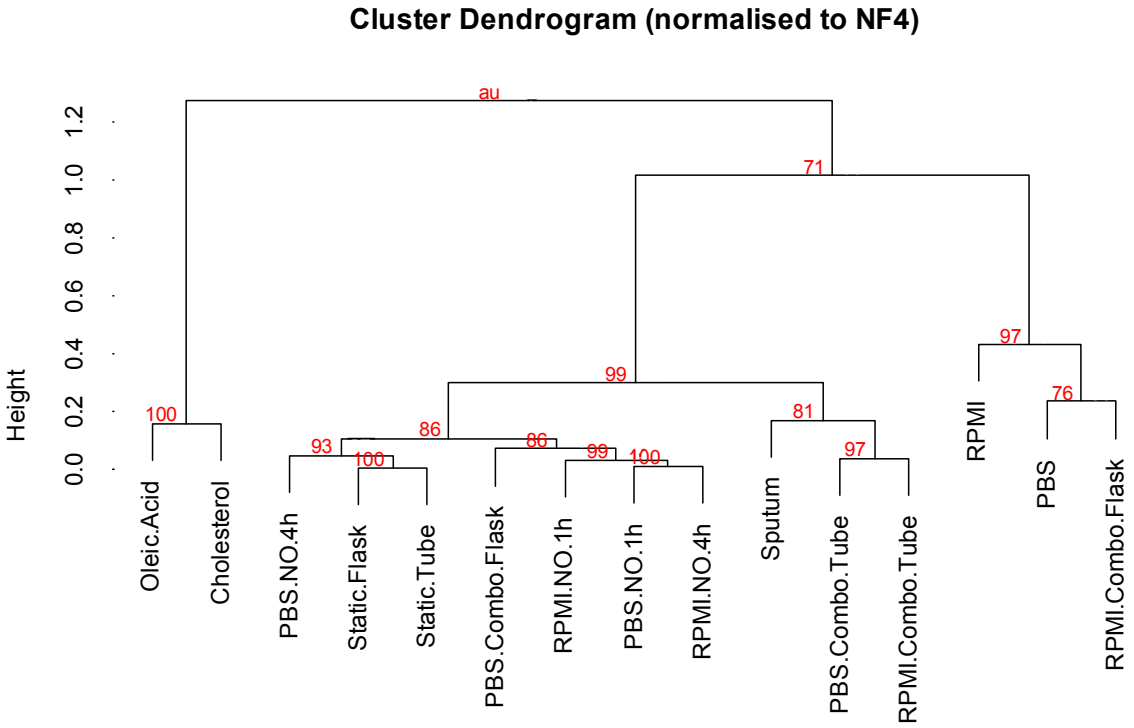


Figure 28: NF4 Cluster Dendrogram

The dendrogram shows hierarchical clustering of the gene expression from 14 samples against sputum gene expression. The samples represent the average gene expression of duplicate (PBS and RPMI are in triplicates) biological samples, normalised to NF4 normalisation factor. Multi-stimulus cultures are labelled 'Combo', i.e. combination.

Clustering was based upon the correlation of expression profile using the complete linkage method. AU (approximate unbiased) probability values are displayed in red and were computed using multiscale bootstrap resampling. For comparison, a dendrogram for gene expression normalised to *sigA* is shown in Appendix 4 (Supplemental Data).

For both normalisation factor and *sigA* normalised samples, sputum microarray gene expression clustered to certain variations of PBS and RPMI multi-stimulus cultures. However, as cluster analysis can vary greatly, the approximately unbiased values were not significant and the cluster analysis did not take error/standard deviations into account, the association between these cultures/growth conditions with the sputum gene expression was examined further using correlation, scatter plots and bar graphs of the gene expression.

A Spearman's rank test of correlation was used to examine the relationship between the sputum expression and four samples: PBS multi-stimulus or RPMI multi-stimulus cultures in conical flasks or 30ml tubes. Figure 29 shows scatter plots comparing these 4 samples to sputum gene expression. The test of correlation and scatter plots did not take sample error into account.

Three of the four test samples showed significant correlation with sputum expression; RPMI multi-stimulus (flask) culture was the only sample that showed no correlation ($p=0.7301$), confirming the differential expression of this sample. The two samples with the best correlation had high, similar R values and the gene expression of these samples is shown in Figure 30; the dotted lines show 1 $\log_2$ and -1 $\log_2$, signifying a 2-fold upregulation and downregulation, respectively.

With regards to the 10 upregulated genes, the results suggested that the RPMI multi-stimulus sample has more similarly upregulated genes (above 2-fold) than the corresponding sample in PBS. For the 10 downregulated genes, the results suggested that the PBS multi-stimulus sample had more similarly repressed genes to sputum as the equivalent RPMI sample.

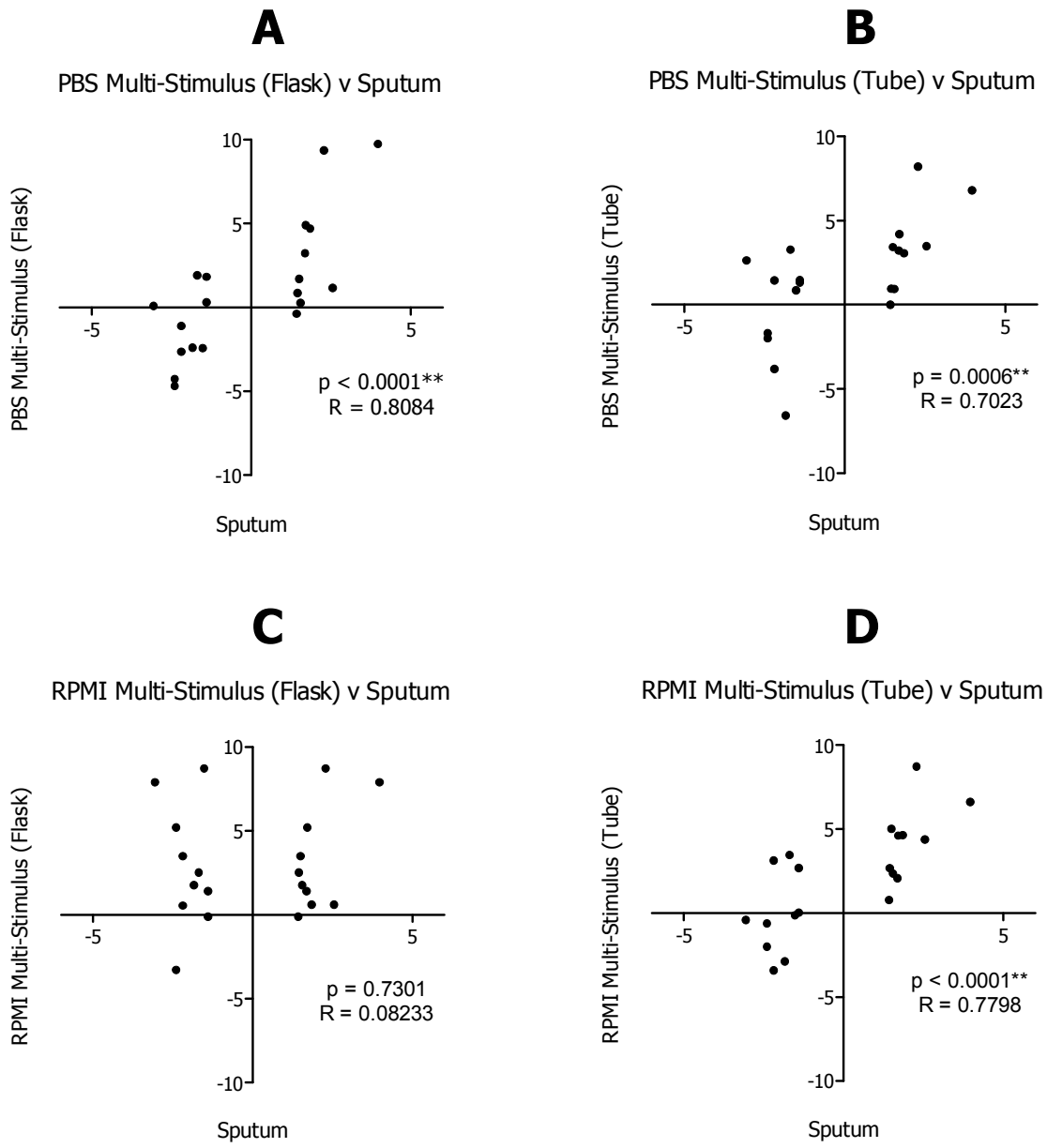


Figure 29: Scatter Plots comparing PBS/RPMI + Stimuli vs. Sputum Microarray Gene Expression

Scatter plots display gene expression for *Mtb* H37Rv multi-stimulus cultures versus Sputum Microarray gene expression. (A) and (B) show *Mtb* PBS multi-stimulus cultures in conical flasks and 30ml tubes, respectively. (C) and (D) show *Mtb* RPMI multi-stimulus cultures in conical flasks and 30ml tubes, respectively. All samples were also exposed to oleic acid, cholesterol, nitric oxide and static incubation. Spearman's Rank Correlation R and p-values are displayed for each graph.

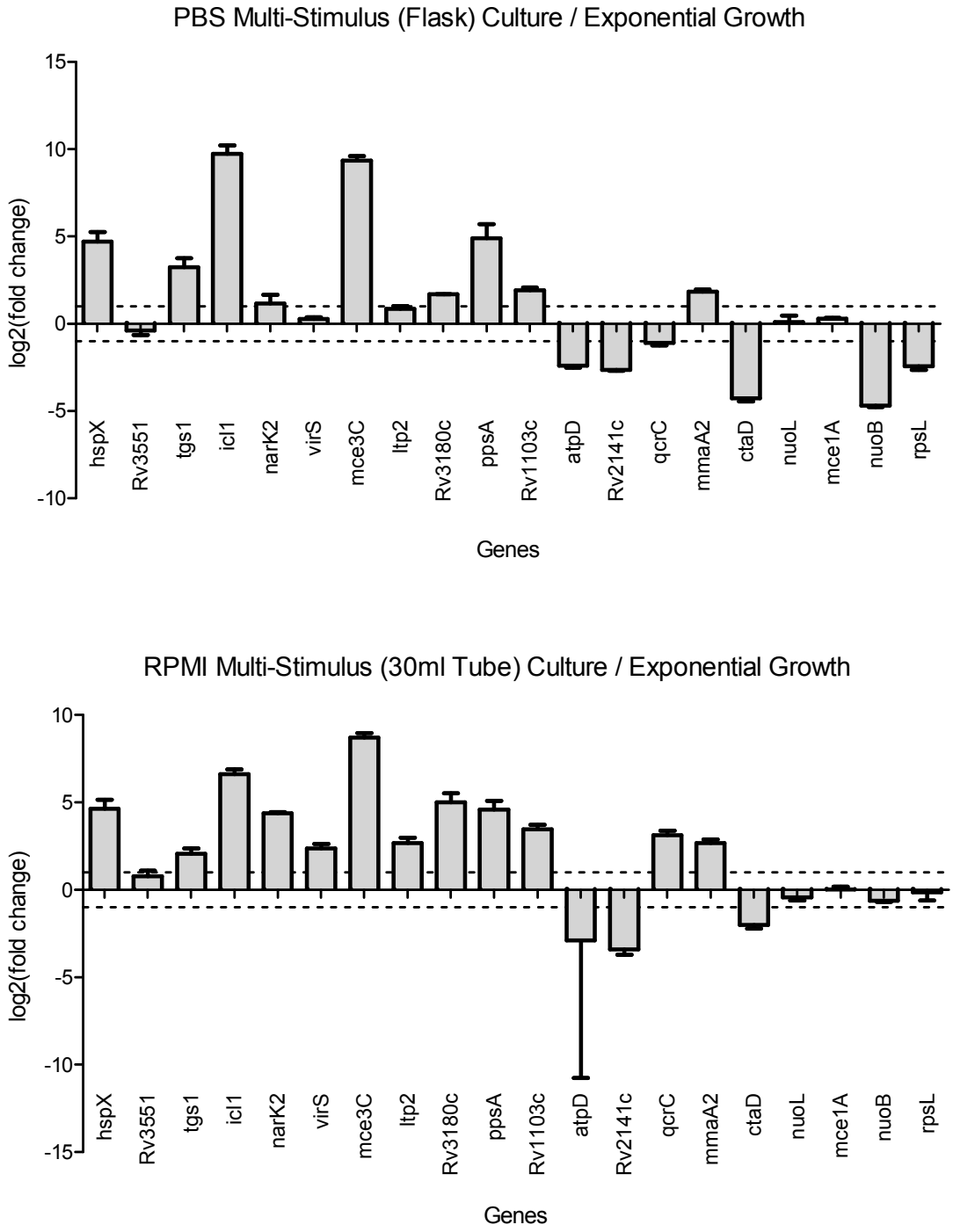


Figure 30: Gene Expression for Samples with Highest Correlation

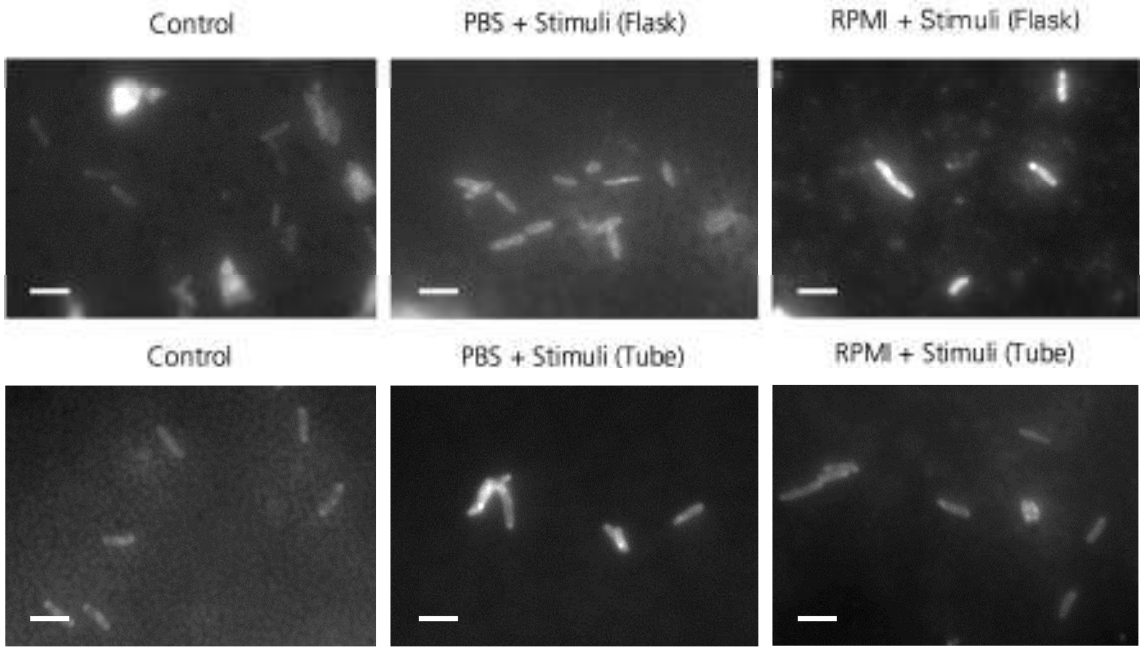
Bar graphs display gene expression of 20 selected target genes in *Mtb* for the two samples with the highest correlation to the sputum microarray data. The top graph displays *Mtb* H37Rv in PBS with additional exposure to oleic acid, cholesterol, nitric oxide and static incubation (in conical flasks). The bottom graph displays *Mtb* H37Rv in RPMI with oleic acid, cholesterol, nitric oxide and static incubation (in 30ml universal tubes).

As for the original 5 growth conditions (plus exponential phase culture), samples were also taken from each of the four multi-stimulus cultures and controls for microscopic analysis. The samples were stained with LipidTOX. Fluorescence images were taken of the test and control cultures. Images were blinded and hand counted for lipid bodies; a minimum of 100 bacilli were counted per sample.

The median fluorescence intensity per cell was also analysed. Fluorescent images were blinded and analysed with ImageJ and R Project Statistical Package using the methods described in Materials and Methods; as for lipid body counting, a minimum of 100 bacilli were analysed per sample. The lipid body percentages and fluorescence intensity values for test cultures and controls are listed in Figure 31.

The frequency of LB-positive cells and median fluorescence intensities varied greatly over the 2 experiments. The frequency of LB-positive cells was highest in the RPMI multi-stimulus (flask) samples. Overall, in both experiments the frequency of LB-positive cells in the RPMI samples was higher than the controls. However, the frequency varied greatly in the PBS samples, as the frequency was lower than the control in conical flasks and higher in 30ml tubes.

The median fluorescence intensities also varied greatly between experiments. In Experiment 1 (conical flasks), the median fluorescence intensities were significantly higher in the PBS and RPMI-1640 (plus multiple stimuli) samples ($p<0.0001$*, Mann-Whitney test), while the opposite was true for Experiment 2 (30ml tubes), with the control sample having a significantly higher median fluorescence intensity ($p<0.0001$*, Mann-Whitney test). The frequency of LB-positive cells did not appear to match the median fluorescence intensity values.



Sample	# of Cells	Median Fluorescence	Interquartile Range	Ratio (vs. Control)	# LB +ve (# Cells)	LB %
Experiment 1 (Flasks)						
Control	424	50522	30851 – 73984	-	105 (387)	27.1
PBS Multi-Stimulus	609	134383	95058 – 165520	2.66	70 (474)	14.8
RPMI Multi-Stimulus	435	167495	108872 - 269661	3.32	241 (504)	47.8
Experiment 2 (Tube)						
Control	481	80506	59920 – 146445	-	100 (562)	17.8
PBS Multi-Stimulus	325	57018	42073 – 75247	0.71	143 (442)	32.4
RPMI Multi-Stimulus	285	53548	45281 - 62695	0.67	102 (331)	30.8

Figure 31: Fluorescence and Lipid Body Analysis of *M. tuberculosis* in PBS/RPMI and Stimuli

Fluorescence images for control and PBS/RPMI multi-stimulus cultures are displayed above. The table displays the number of cells analysed, median fluorescence intensities per call and interquartile ranges for the fluorescence analysis. The ratio compares the median fluorescent intensity of the growth conditions to exponential phase growth (control). For lipid body analysis, the number of cells analysed and lipid body positive% is shown. Bar is 2µm.

4.3 Discussion

4.3.1 Selection of Upregulated and Downregulated Genes

The gene expression of aerobic, *in vitro* growth conditions is far different from the gene expression seen in bacilli extracted from human sputum. Ten upregulated and ten downregulated genes were selected from the sputum microarray data. Genes were selected if they either reflected the microarray gene expression trends or if they were of particular interest. The function and selection of these 20 genes was introduced in Paragraph 4.1.2.

The selection of 20 genes had both benefits and limitations. Ideally, microarray analysis would be used to monitor the expression changes of a large number of genes. However the technique is complex and impractical for a preliminary experiment (Schena, 2003), and as more growth conditions are analysed, the cost becomes prohibitive. Furthermore, more sensitive RT-PCR results are generally used to validate key microarray results (Groen, 2001). The analysis of only 20 genes allowed for the analysis and screening of a number of growth conditions, a requirement of this study design.

A disadvantage of the current study was the nature of gene selection; the genes selected were based upon our interpretations of the sputum transcriptome results; this selection may have led to bias and was thus likely to alter the selection of growth conditions and stimuli. Furthermore, the selection of 20 genes out of the whole sputum transcriptome is an inherent limitation. The *Mtb* genome contains approximately 4000 annotated genes (Cole et al., 1998). Moreover, the sputum transcript had 182 significantly induced genes and 334 significantly repressed genes, making the selection of genes representative of the expression trends more difficult.

Hence, as the genes analysed was limited to 20, a number of genes that have been studied extensively could not be included. Members of *echA* and *fadE* families of genes were upregulated in sputum; these genes are thought to be involved in the β-oxidation of fatty acids and could have been selected as alternative lipid metabolism genes. *Mce3C and mce1A,* members of the *mce* family of genes were selected for this study; in addition to those two genes, *mce2C* was upregulated and is also associated with *Mtb* virulence (Gioffre et al., 2005). A range of ribosomal function genes were downregulated in sputum; it may have been appropriate to select an extra target ribosomal function gene, as opposed to 1 of the 5 aerobic respiration genes studied.

The PE and PPE families of genes are named for the proline-glutamic acid and proline-proline-glutamic acid motifs that they contain, respectively (Voskuil et al., 2004). A number of

these genes are upregulated in sputum (Garton et al., 2008). While the exact function of the gene families is unknown, the families may have a role in pathogenesis by promoting antigenic diversity, allowing the bacilli to evade the immune system (Voskuil et al., 2004). The genes are not globally regulated as one group, but are thought to change in response to changes in *in vivo* microenvironments (Voskuil et al., 2004). Thus, the expression of these genes may be specific to the sputum environment.

Intriguingly, the gene (Rv3875, synonyms: *esxA, esat-6*) encoding for the 6kDa early secretory antigenic target (ESAT-6) protein was found to be downregulated in sputum. ESAT-6 is a member of a family of small proteins; 4 other genes encoding for proteins in this family were similarly repressed. ESAT-6 has been found to be a strong T-cell antigen; members of the ESAT-6 family of proteins have been shown to be deleted in some strains of *Mtb* (Brodin et al., 2004). It has been proposed that these deletions may have resulted from immune pressure and allow bacilli to escape host immune defences. There is however, evidence that ESAT-6 may play a role in pathogenesis of tuberculosis as mutations that abolish the production or secretion of the protein attenuate the bacteria in various animal models (Brodin et al., 2004).

While the upregulated and downregulated genes were not selected for this study, they could be included in future as an extension of the current study.

4.3.2 Validation of qPCR Assay Primers

The validation of qPCR assay primers and reactions is vital prior to the use of them in experimental assays. The amplification of a single band of the correct size must be confirmed. Further, although efficiency correction is used (as described in 3.3.2), a higher PCR amplification efficiency is more appropriate to ensure efficient amplification of the target product/gene.

Gel electrophoresis was used to analyse the resultant PCR product for each primer set. Each primer set was confirmed to amplify only a single band. Further, the amplified bands were of expected size; thus, the correct product is more likely to be amplified. Finally, the fluorescence melt curves of the PCR products were examined. The melt curves confirmed only one product was amplified for each primer set; the product melt temperatures were all above the fluorescence acquisition temperature during the PCR reaction, confirming the correct acquisition temperature.

LinRegPCR was also used to analyse the efficiencies of each primer set for the given cycling conditions. An amplification efficiency cut-off of 80% (1.80) was selected. All primer sets had average efficiencies above this cut-off.

Intra-assay variation was assessed by calculating the CoV of the 3 PCR assay replicates from multiple PCR runs using each of the 20 primer sets. The average CoV, expressed as a percentage, was less than 15% for all 20 primer sets, with 12 of the 20 primer sets having a CoV of less than 10%. Thus, within each PCR assay, the variation in the technical replicates is limited. Furthermore, the actual variation in certain primer sets may be lower than the values calculated here. The relative quantities used to calculate the CoV were expressed as a fraction of 1. For specific genes with low expression values (e.g. *hspX* in exponential phase culture), the mean relative quantity is close to zero, giving a CoV that approaches infinity and resulting in a high average CoV, when in actual fact the variation may be low. Inter-assay variation was not assessed, as each PCR used different cDNA samples.

With the confirmation of the amplification of a single product of the correct/expected size, high amplification efficiency and small variation within each PCR assay, the 20 primer sets for the target genes were used for all further gene expression analysis.

4.3.3 Gene expression of *M. tuberculosis* H37Rv under selected growth conditions

4.3.3.1 Selection of 5 preliminary growth conditions

Five growth conditions/stimuli (stationary phase, PBS, RPMI-1640, SDS and artificial saliva), as compared to aerobic, exponential growth, were selected for the preliminary gene expression experiment. Aerobic, exponential growth was the control growth condition, as previously used for the sputum gene expression (Garton et al., 2008). All other growth conditions were compared to the gene expression of these cultures.

Stationary phase culture, RPMI-1640, PBS and SDS were previously used to examine the candidate housekeeping genes and were again analysed, along with artificial saliva for the expression of the selected target genes. The growth conditions and reasons for their selection were introduced in Paragraph 4.1.3.

While stationary phase culture, PBS and RPMI-1640 were easily replicable, disadvantages of SDS and artificial saliva were their variability and representation of *in vivo* conditions. Firstly, SDS was added to a final volume of 0.5% (w/v); this concentration can be infinitely altered

and small changes may alter the overall gene expression. A similar disadvantage occurs with artificial saliva, as numerous compositions exist in the literature (Shellis, 1978). In addition to this, while a proportion of genes were documented to be similarly regulated following both SDS exposure and growth in macrophages, SDS may not fully replicate *in vivo* damage to the mycobacterial cell wall (Schnappinger et al., 2003). This issue is paralleled for artificial saliva, as it may not be fully representative of human saliva. Moreover, the influence of saliva on sputum composition and the resultant gene expression is unknown and is likely to vary between patients.

As it had been postulated that cells may become persistent over shorter time periods, cells were exposed to the growth media for time periods of 4 hours with the exception of stationary phase culture. Furthermore, in addition to the hypothesis, a time course of 4h was desirable as it featured a good balance of gene expression changes, minimisation of cell clumping and allowed for experiments to be conducted within the time constraints imposed by the C3 laboratory code of conduct.

Nitric oxide, oleic acid, cholesterol and static incubation were not included in the preliminary growth conditions, but were analysed in subsequent experiments and are discussed below. Other notable growth conditions/stimuli not included in this study included phosphate deprivation, reactive oxygen species (ROS) and low pH. Furthermore, certain cellular environments not analysed in this study may play a part in the gene expression seen in sputum, but may not be considered 'growth conditions' per se. These growth conditions and cellular environments may contribute to the sputum transcriptome and are discussed in Chapter 6.

As the composition of tuberculous sputum is still unknown, the non-tuberculous sputum was also considered for use as a preliminary growth condition. However, there are difficulties associated with the use of sputum. The addition of *Mtb* bacilli to the sputum is a problem; the cell density at which to infect the sputum may be critical, plus the addition of any media used to transfer the bacilli may alter the sputum environment. Furthermore, there is difficulty obtaining large quantities of sputum. The volume of sputum available from a single patient is generally limited and variable; for this reasons, samples must be pooled together from multiple patients. Pooling of sputum may however, eliminate some of the variability in the composition of different sputum samples.

Sputum is obtained from hospital patients with other morbidities, often COPD or cystic fibrosis; as disease states alter the macromolecular composition and biophysical properties of sputum, it is unknown how this would influence the gene expression (Voynow and Rubin, 2009). While, sputum can be obtained from healthy volunteers through induction, this has

further ethical implications (Belda et al., 2000). Furthermore, in a study involving COPD patients, the induction of sputum has been shown to alter the viability of the eukaryotic cells found in sputum (Bhowmik et al., 1998).

4.3.3.2 Stationary Phase Gene Expression correlates to the Sputum Transcriptome

When compared to microarray data, of the five growth conditions, the stationary phase cultures appeared to have the most similar expression. The p-value of the Spearman's rank correlation was significant and the value suggested that there was a relationship between the gene expression of stationary phase culture and sputum and the relationship is not due to random sampling. None of the other samples showed significant correlation to the sputum transcriptome.

Stationary phase gene expression had a large upregulation of the dosR regulated genes. This may be due to the cells entering a state of slow growth. However, as the stationary phase cultures reached a high OD, this may have led to clumping and reduced oxygen availability to the bacilli, causing a hypoxic upregulation of the regulon.

However, as described in 4.2.5, a significant Spearman's rank correlation only signifies a monotonic relationship between the two variables; the relationship may not be linear. Closer examination revealed that four and zero of the upregulated and downregulated genes were similarly regulated, respectively. Also, stationary phase was judged to have occurred when the optical density of the *Mtb* H37Rv culture had stabilised for three consecutive days. This stabilisation of the optical density required a time course of greater than 10 days for this to occur. Of the five growth conditions, stationary phase was the only growth condition that exceeded a 4h exposure. This was an exception to the proposed hypothesis that in the human body, the advent of host-derived stresses on *Mtb* may cause the bacilli to enter a persistence phase over a shorter time period. The time required for *Mtb* to enter non-replicating persistence *in vivo* is still unknown; there is no consensus in past *in vitro* studies; some studies required short timeframes, i.e. hours, to induce persistence (e.g. Taneja et al., 2010, Voskuil et al., 2003, Wayne and Sohaskey, 2001), while others required longer time periods, i.e. days to months (e.g. Hampshire et al., 2004). Regardless of this debate, the length of time required for *Mtb* to reach stationary phase contradicted the original theory presented in 4.1.2.

For these reasons, analysis of stationary phase culture was not continued and PBS and RPMI-1640 media were selected as the base media for future experiments. These two media were selected on the following basis: firstly, with the exception of stationary phase culture, the

gene expression of the other four specific growth media/conditions showed no relationship to the sputum gene expression. Secondly, both PBS and RPMI-1640 are established media with specific formulations and have been used in past studies; as RPMI-1640 is commonly used for human tissue cell culture, it may provide growth conditions more similar to *in vivo* growth. On the contrary, SDS may be used at any concentration and there are many formulations of artificial saliva in the literature (Shellis, 1978, van Ruth et al., 2001). Finally, PBS and RPMI-1640 are good base media for the addition of extra stimuli, particularly as they are isotonic, non-toxic to cells and possess buffering systems to maintain a stable pH. As media, artificial saliva was untested, while SDS damages bacterial cell walls and is toxic in higher concentrations (Pang et al., 2007).

4.3.3.3 Lipid body production was highest in Stationary Phase Culture

LBs in acid-fast bacilli have been shown to be a universal feature of smear-positive tuberculous sputum; LB-positive cells have been visualised in 82 unique sputum samples (Garton et al., 2002, Garton et al., 2008). The frequency ranged from 3% to 86%, with a mean of 45% (SD 20%). As LB-positive cells were a predominant feature in sputum, it was postulated that *in vitro Mtb* cultures with similar gene expression as bacilli in sputum should also produce lipid bodies phenotypically. Thus, the frequency of LB-positive cells was examined for the 5 growth conditions and exponential phase growth.

Stationary phase culture was shown to have the highest frequency of LB-positive cells in the results; this was likely to be associated with the upregulation of *tgs1*, a DGAT involved in the synthesis of triacylglycerol, the main component of LBs. As discussed, events leading to this *tgs1* induction may result from increased cell clumping leading to reduced oxygen tension and upregulation of the dosR regulon.

The LB content was reduced (compared to control) for the samples in PBS, RPMI-1640 and SDS. In PBS and SDS, this may be due to the nature of the medium; PBS is composed of salt and a phosphate-buffer system, while the SDS 'medium' consisted of distilled water and SDS. Thus, *Mtb* in both PBS and SDS may undergo a nutrient starvation stress; triacylglycerol may be utilised instead of sugars, although, *icl1* was only found to be upregulated in PBS. The reduced LB content in RPMI-1640 was more surprising, as the medium is richer and contains fatty acids (from FCS supplementation).

4.3.4 Gene expression of *M. tuberculosis* H37Rv under combined growth conditions

4.3.4.1 Multiple growth conditions and the Sputum Transcriptome

As the gene expression of *Mtb* H37Rv for the initially selected growth conditions did not adequately match the sputum gene expression, further analysis was performed on *Mtb* cultures in 7H9, PBS or RPMI-1640 (4.2.5) with the addition of a number of alternative stimuli. These stimuli included nitric oxide, oleic acid, cholesterol and static incubation and a combination of all 4 additional stimuli. The clustering of biological replicates was examined first. All the sample replicates clustered closely with the exception of RPMI multi-stimulus (30ml tube) cultures; these replicates each clustered to one of the PBS multi-stimulus samples. The nature of this clustering demonstrated that the gene expression between biological replicates was similar; furthermore, the clustering of the PBS/RPMI multi-stimulus samples may indicate a similar gene expression across the 20 genes.

When cluster analysis was performed comparing each set of samples (averaged biological replicates), the PBS and RPMI multi-stimulus (30ml tube) samples clustered most closely to sputum, when normalised to NF4. However, as the samples clustered 81% and 78% of the time for NF4 and *sigA*, respectively, this was not significant. When the gene expression was normalised to *sigA*, there was similar clustering; however, sputum expression clustered to all the multi-stimulus samples (PBS and RPMI) in both 30ml tubes and conical flasks.

The complete-link clustering method was used to generate the dendrogram in Figure 28. In complete-link clustering, the similarity of two clustered samples is based upon the similarity of their most dissimilar members, i.e. the distance between clusters is based upon the maximum distance between a pair of objects in each cluster sample (Everett et al., 2001). This is in contrast to single-link clustering, where the similarity of two clusters is based on the similarity of their two most similar members. Both cluster assessments use measurements based on a single pair of results within each sample and can thus result in undesirable clusters being formed; in complete-link analysis, for example, clustering is overly influenced by outliers within a sample (Everett et al., 2001).

Complete-link analysis avoids the chaining phenomenon associated with single-link clustering. In single-link analysis, clustering may occur due to single elements being close together, despite the majority of the elements in each cluster being disparate (Everett et al., 2001). Therefore, in practice, complete-link analysis tends to be of greater use and was selected for this cluster analysis.

However, as the clustering was not significant and did not take error into account, the gene expression correlation between sputum and selected cultures was also analysed. As the PBS and RPMI multi-stimulus (30ml tubes) samples clustered most closely to sputum, these were included in the analysis. The equivalent samples in conical flasks were also analysed as the gene expression of these samples clustered closely to sputum when normalised to *sigA*. Individually, the gene expression of three of the four growth conditions correlated significantly to the sputum expression.

Two samples, PBS multi-stimulus (flasks) and RPMI multi-stimulus (30ml tubes) had similar correlation p-values and Spearman's R values. For the 10 upregulated genes, the results suggested that the RPMI sample had more similarly upregulated genes (above 2-fold) than the PBS sample. The opposite was true for the 10 downregulated genes, as the results suggested that the PBS sample had more similarly repressed genes.

It was noted that the RPMI multi-stimulus (flask) cultures clustered separately from the other multi-stimulus cultures. Furthermore, this was the only multi-stimulus culture with gene expression that did not correlate with the sputum transcriptome. The reason for these differences was not established but considering the similar gene expression of the other 3 samples, was most likely a problem with regards to the RPMI multi-stimulus (flask) cultures. As the RNA from the biological replicates was extracted on the same day, there may have been a problem with regards to the stimulus reagents or increased transcript degradation for one of the housekeeping genes.

While the gene expression of these PBS/RPMI multi-stimulus samples correlated significantly, there were still differences in gene expression. Again, the scatter plots did not take error into account, and unfortunately, as the gene expression was performed in duplicate, the experiment did not have enough statistical power to compare each gene individually. The results do suggest that a number of growth conditions are required to reproduce the sputum transcriptome. The base medium used (PBS or RPMI) does not appear to have as great an effect on gene expression as the additional stimuli. Further, specific genes may be under the control of multiple stimuli, as gene expression under various growth conditions may overlap considerably, e.g. nitric oxide and hypoxia.

It is interesting to note that the upregulation of the genes in the *in vitro* growth conditions tended to be higher than in sputum. For example, *icl1* is 3.61-fold upregulated in sputum, but >512-fold and >64-fold upregulated in the PBS and RPMI multi-stimulus samples, respectively. The same was observed in certain dosR regulated genes; in the PBS sample, expression of *hspX* and *tgs1* (but not *narK2*) were greater than in sputum. In the RPMI sample, *hspX* and *narK2* expression (but not *tgs1* expression) was higher than in sputum. This suggests

that the level of exposure to the various stimuli may be lower *in vivo*, or the combination being used is incorrect. For example, nitric oxide and static incubation (through the generation of an oxygen gradient) both induce dosR regulated genes; although *Mtb* is thought to encounter both nitric oxide and hypoxia *in vivo,* this may not contribute to the gene expression of the sputum samples (Kendall et al., 2004, Voskuil et al., 2003).

The regulation of the three dosR regulon genes is interesting. While the three genes were upregulated in a number of samples, as compared to aerobic, exponential growth, the degree of this induction varied with between the genes. There are a number of possible reasons behind this. It is possible that the longevity of the transcripts vary *in vivo* and are degraded by different methods and rates. Furthermore, despite the dosR regulation, the transcription of certain genes may be influenced by other external factors. For example, toxins as part of TA systems can change translation of genes through the cleavage of mRNA transcripts (Ramage et al., 2009). Finally, there is overlap between the expression changes of certain stimuli (e.g. nitric oxide and hypoxia); this overlap may result in the differential upregulation of the dosR regulated genes (Voskuil et al., 2003).

In addition to the gene expression of *Mtb* multi-stimulus cultures, the transcription of genes following exposure to each of the individual stimuli (oleic acid, cholesterol, nitric oxide and static incubation) was also examined. The notable gene expression changes are discussed here. Again, it is important to note that the experiment was only performed in biological duplicates; thus, each individual gene could not be examined statistically.

Oleic acid was selected to be added as a stimulus because of the upregulation of lipid utilisation genes in *Mtb* in sputum; the addition of oleic acid may induce similar genes. Further, the addition of oleic acid has been associated with increased frequency of LB-positive cells in *Mtb*, a prominent feature in cells found in sputum (Sherratt, 2008). The gene expression findings following oleic acid addition did confirm these expectations. The most highly induced gene was *icl1*, which encodes an enzyme involved in lipid metabolism (reviewed in Chapter 1). Furthermore, *ppsA* was also found to be upregulated, which suggests that the oleic acid may have been channeled into the synthesis of PDIM. It was interesting to note, that this *ppsA* upregulation was not a feature of any of the multi-stimulus samples.

As for lipid utilisation genes, certain genes involved in the metabolism of cholesterol were also upregulated in sputum (Garton et al., 2008). The most notable gene expression following cholesterol supplementation was the increase of 4 genes: *icl1, ltp2, ppsA* and Rv3551. The upregulation of these genes was expected. As for the addition of oleic acid, the increase in *icl1* transcription indicated an increase in lipid metabolism, possibly due to the catabolism of the cholesterol side chains. Furthermore, *ltp2* and Rv3551 were both found to be induced;

ltp2 has been identified as essential for cholesterol metabolism (Chang et al., 2009), whilst Rv3551 is a member of a cholesterol catabolic gene cluster (Van der Geize et al., 2007).

The gene expression of *Mtb* in PBS and RPMI following exposure to nitric oxide confirmed previous laboratory results and studies in the literature (Sherratt, 2008, Voskuil et al., 2003). As previously shown, nitric oxide induces genes belonging to the dosR regulon, including *hspX, tgs1* and *narK2* (Voskuil et al., 2003), and the upregulation of these genes was seen in this study. *Icl1* was also upregulated in these nitric oxide exposed samples. While *icl1* was not noted to be upregulated in previous studies, a similar upregulation was noted in *Mtb* cultures using either PBS or RPMI only (Ohno et al., 2003, Voskuil et al., 2004). Complete RPMI medium contains FCS, and in turn, fatty acids; therefore, in RPMI, this *icl1* induction may be expected. This upregulation is more surprising in PBS and under these conditions, *Mtb* may use internal lipid stores for the generation of energy. However, it is worth noting that *icl1* was not noted to be induced in previous nutrient starvation studies using PBS (Betts et al., 2002).

Of the genes downregulated in sputum, Rv2141c was expected to be repressed as this had been previously shown by Ohno et al. (2003). In this study, the result was confirmed as Rv2141c was downregulated >2 fold in each of the nitric oxide exposed samples. With regards to the ribosomal genes, despite the documented growth arrest of *Mtb* by nitric oxide (Voskuil et al., 2003), this did not result in downregulation of *rpsL* in this study.

The statically incubated samples (in both conical flasks and 30ml tubes) again led to the upregulation of the dosR regulated genes: *narK2*, *tgs1* and *hspX*. This confirmed the result demonstrated by Kendall et al. (2004), who showed induction of the dosR regulation following static incubation for 30min. In addition to the dosR regulation, it was notable that all the samples that were statically incubated in conical flasks did not produce the expected downregulation of aerobic respiration genes. This was likely due to the large headspace volume in conical flasks; combined with large surface area available for diffusion, the bacilli were likely able to undergo aerobic respiration. The static incubation was thus repeated in 30ml centrifuge flasks. Again, this did not produce a downregulation in the aerobic respiration genes. Despite the smaller headspace, it may be that the 4h incubation period is not long enough to adequately lower the oxygen content in the culture media and force cells to anaerobically respire.

The addition of an oxygen scavenger maybe necessary if the sputum transcriptome is to be replicated in the imposed 4h experiment. Taneja et al. (2010) showed that physiological levels of ascorbic acid/vitamin C scavenge oxygen, reduce saturation to <30% and rapidly induce the dosR regulon. Whether this is representative of the conditions *in vivo* and vitamin C is present in the granuloma remains to be confirmed.

Finally, the stimulus controlling certain genes from the 20 targets remained undetermined. These genes included *mce3C, virS,* Rv3180c, *mmaA2* and *mce1A*. For these genes, expression was variable throughout the tested growth conditions. For example, Rv3180c had >2-fold induction under PBS and both RPMI multi-stimulus cultures, but not in the PBS multi-stimulus cultures; the growth condition responsible for induction of Rv3180c was therefore difficult to establish. Further studies must be performed to fully understand the stimulus or stimuli that underlie the expression of these genes.

Overall, the gene expression results presented in this section showed similar gene expression between *Mtb* in sputum, and specific *in vitro* cultures: stationary phase cultures and cultures exposed to PBS/RPMI with multiple stimuli. These results also suggest that, over a short period of time, multiple stimuli are required to produce the sputum phenotype. However, a number of possible *in vivo* conditions remained untested and the overall impact of the up- or down-regulated genes on the phenotype of *Mtb* in sputum remains to be confirmed.

4.3.4.2 Fluorescence and Lipid Body Content varies when multiple stimuli are used

As discussed in 4.2.5, LB-positive cell frequency was also examined in the samples exposed to multiple stimuli. Furthermore, samples were analysed for fluorescence intensity. Studies by Bell et al. (2010) have shown a correlation between the fluorescence intensity of *Mtb* and *M. smegmatis* stained with LipidTOX neutral lipid stain with the total triacylglycerol content (the main component of LBs).

The results were highly variable between the 2 experiments. The frequency of LB-positive cells in the 1st experiment (in conical flasks) was found to be higher for the RPMI samples vs. control, but lower for the PBS samples vs. control. In the 2nd experiment, frequency of LB-positive cells was higher for the PBS and RPMI samples vs. control. The high frequency of LB-positive cells was expected, as previous results have shown LB frequency in *Mtb* to be increased in response to the addition of nitric oxide and oleic acid (Sherratt, 2008). Overall, clumping did not appear to be a factor; the addition of tyloxypol, a detergent, to the cultures (through the addition of cholesterol) may have contributed to this.

When the median fluorescence intensity of the cells was analysed in the 1st experiment, the intensity was significantly higher for both sets of samples as compared to the control (using a Mann-Whitney test). However, the opposite was true in the 2nd experiment, with the control fluorescence intensity being significantly higher.

The wide range in median fluorescence intensity may be explained by the heavy background staining observed; the addition of oleic acid may be the cause of this. Oleic acid, being a neutral lipid, stains intensely with the LipidTOX fluorescence stain; the presence of oleic acid in these samples would explain the background staining observed in some of these samples. Furthermore, the intense staining of oleic acid may limit the availability of LipidTOX stain for the *Mtb* bacilli. Finally, the use of fluorescence intensity to analyse the lipid content in cells is largely automated, and non-cell objects may be included in the analysis accidently; this is especially true in samples with high background staining.

The addition of cholesterol may also affect the staining of cells. Mukherjee et al. (2007) showed that fluorescence emission, intensity, polarisation and lifetime of Nile Red, a lipophilic stain, varies with the cholesterol content of cell membrane. It is possible that the use of cholesterol as a stimulus may also affect LipidTOX staining and alter both the background and cell staining.

4.3.4.3 H37Rv may not be an appropriate strain to study

It was originally postulated that the gene expression of *Mtb* in sputum signalled cells adapted for transmission; however, an alternative explanation for this adaptation, if the hypothesis is correct, is strain variation.

As a well characterised strain of *Mtb,* H37Rv was an ideal baseline strain with which to conduct the studies. However, as a laboratory strain, there is some question as to how representative it is of the clinical *Mtb* isolates isolated from sputum. There is also a question of the variation within the H37Rv strains used in laboratories worldwide; a recent study by Ioerger et al. (2010) carried out whole-genome sequencing on six strains of H37Rv maintained in laboratories worldwide. The researchers found a number of polymorphisms unique to individual strains as well as several shared polymorphisms (Ioerger et al., 2010). Even more significantly, a frameshift mutation was found in a merocerosic acid synthase gene (*mas*), which resulted in two of the strains to be deficient in the synthesis of PDIM (Ioerger et al., 2010). As the expression of a gene involved in PDIM biosynthesis was assessed in this study, mutations such as this may seriously alter the results.

In addition to variations within H37Rv, the overall phenotype of certain *Mtb* strains may differ. It has been suggested that the characteristics of the W-Beijing strains of *Mtb* may play a part in the success of the strain globally, by conferring properties of increased virulence or transmission properties; this success is reflected by global studies where Beijing strains account for approximately 10% of all tuberculosis cases (Bifani et al., 2002). Reed et al.

(2007) previously showed constitutive upregulation of the dosR regulon in *Mtb* clinical isolates from the W-Beijing family of strains; this upregulation is not seen in laboratory strains of *Mtb*. The dosR regulon has been shown to assist in the in metabolic homeostasis and to allow *Mtb* to resume growth after being in being in a state of anaerobic or nitric oxide-induced non-respiring state (Leistikow et al., 2010). Furthermore, this family of strains was shown to accumulate large quantities of triacylglycerol in *in vitro* aerobic culture (Reed et al., 2007). The utilisation of triacylglycerol has been shown in *M. bovis* BCG to be required for the regrowth of hypoxic, non-replicating cultures (Low et al., 2009). Both these characteristics may play a part in *Mtb* transmission.

While *Mtb* bacilli in sputum and the Beijing strains share both dosR regulon induction and triacylglycerol accumulation, alternative mechanisms may underpin the increased virulence seen in the Beijing strains. The Beijing strains have differences in protein and lipid structures that may affect their interaction with the human immune system (Parwati et al., 2010). For example, the Beijing strains produce the biologically active lipid, polyketide synthase-derived phenolic glycolipid, which can inhibit the release of pro-inflammatory mediators (Parwati et al., 2010).

As the Beijing strains were unavailable to this lab at the time of the study, the baseline gene expression of another virulent strain of *Mtb,* CH, was compared to the sputum transcriptome (results included in Appendix 3 – Supplemental Data). *Mtb* CH was a strain responsible for a tuberculosis outbreak in Leicester, UK involving 254 cases; as 77 active cases were notified within 1 year, the progression to symptomatic disease from this strain was greater than the 5-10% lifetime risk normally quoted (Newton et al., 2006).

The baseline gene expression of *Mtb* CH showed no correlation with the sputum gene expression ($p=0.2746$, Spearman's Rank Correlation). In fact, the gene expression of the CH strain was similar to that of laboratory strain H37Rv, although both strains were grown to exponential phase in aerobic conditions. If CH does overexpress virulence and transmission related genes as compared to H37Rv, it may not do so constitutively, but instead, only in reaction to the stresses during infection. It may be more appropriate to expose the CH strain to multiple stimuli as previously performed with H37Rv.

Intriguingly, when the median fluorescence and frequency of LB-positive cells was analysed, they were found to be higher in *Mtb* CH as compared to H37Rv. Using a Mann-Whitney test, the median fluorescence intensity for CH was significantly higher than H37Rv for both biological replicates ($p<0.0001$**). This indicates that there may be higher baseline neutral lipid content in the clinical CH strain.

Overall, It remains to be determined if the production of lipid bodies is related to the virulence and/or transmission adaptation of *Mtb* strains.

4.4 Conclusions

The underlying aims of the work presented in this chapter were to identify potential stimuli responsible for the sputum transcriptome and to replicate the gene expression of *Mtb* in sputum in an *in vitro* setting. This was performed using an approach involving the analysis and comparison of individual and combinations of *in vitro* growth conditions. The specific findings of the studies in this chapter are as follows:

- No single growth condition tested exactly replicates the sputum gene expression of the 20 target genes.
- Gene expression of *Mtb* in stationary phase correlated positively to the gene expression in sputum and produced increased LBs *in vitro*; however, none of the 10 genes downregulated in sputum were similarly repressed in stationary phase.
- *Mtb* cultures in PBS or RPMI and a combination of all 4 additional stimuli provided the best correlation of gene expression to the sputum transcriptome, but again, did not show exact replication of the transcriptome.
- As observed in sputum, the multi-stimulus cultures produced increased LBs (compared to control) for 3 of the 4 multi-stimulus cultures.

Overall, while the gene expression of the multi-stimulus cultures did correlate significantly, further studies must be conducted before concrete conclusions may be drawn from this study.

Chapter 5

The Role of Lipid Bodies during Transmission Stress and Macrophage Infection

5.1 Introduction

Tuberculosis is transmitted almost exclusively through the aerosol route. While airborne, *M. tuberculosis* (*Mtb*) may encounter a number of environmental stresses, including desiccation, radiation, oxygen and pollution (Cox, 1989). If the mycobacteria survive and are deposited in the lung, host phagocytic cells are encountered in the initial step of the infection process (Li et al., 2002). As with any bacterium, *Mtb* transcriptionally adapts to differing environments, leading to the translation of proteins that serve a variety of functions. *Mtb* may express genes that encode proteins required for survival in an airborne environment, and to the stresses associated with aerosol transmission. Furthermore, at the time bacilli encounter phagocytes, they express virulence determinants that may be regulated by the environmental conditions within the host (Li et al., 2002).

As introduced in Chapter 1, *Mtb* in sputum differentially express a number of genes and produce lipid bodies (LBs). As triacylglycerol is the major constituent of LBs, its increased production or accumulation may play a role in the adaptation of *Mtb* to transmission including resistance to stresses. Such a finding was observed in *Rhodococcus opacus*, where the formation of LBs was linked to improved desiccation survival (Alvarez et al., 2004). Based on these results, it was proposed that acid-fast bacilli in sputum may be in a transmission-adapted state.

The overall role of LBs is still unknown; while *Mtb* is thought to utilise fatty acids as a source of energy during human infection, relatively little is known about their role in the transmission of tuberculosis (Cole et al., 1998). Further, it remains to be seen if the production of LBs provokes the phenotype seen in sputum, or if they are merely a secondary phenomenon or marker of such a phenotype.

In this chapter, the role that LBs and triacylglycerol accumulation may play in the survival of *Mtb* to two specific transmission stresses, ultraviolet light and desiccation, and its effect on binding of bacilli to macrophages is explored. The work in this chapter also continues the work to replicate the sputum transcriptome in Chapter 4, by examining the effect of these gene expression changes on macrophage binding.

5.1.1 Aerosol Transmission

Aerosols are defined as 'suspensions in air (or in a gas) of solid or liquid particles, small enough that they remain airborne for prolonged periods because of their low settling velocity' (Tellier, 2006). Bioaerosols contain living organisms such as bacteria, viruses or fungi and are

variable, complex and either natural or man-made in origin (Srikanth et al., 2008). A number of factors can influence the travel and settling of bioaerosols, including both physical properties of the aerosol and environmental characteristics. Physical properties include the size, density and shape of the droplets or particles, while the environmental characteristics include the air current vector, relative humidity and temperature (Srikanth et al., 2008).

Bioaerosols form in the respiratory tract, upon inspiration and expiration, as a result of changes in the airflow in and out of the lungs (Fiegel et al., 2006). These changes in airflow create wave-like disturbances that lead to the creation of droplets. Normal, tidal breathing, likely accounts for the majority of exhaled bioaerosols over the period of the day; however, it has been shown in studies that coughing, talking and sneezing produce more droplets than normal breathing (Fiegel et al., 2006). However, there is some evidence of an association between cough frequency and tuberculous aerosol production (Fennelly et al., 2004).

The particles in a bioaerosol generally range from 0.3 to 100μm in diameter (Srikanth et al., 2008). The diameters at which particles exhibit aerosol behaviour also correspond to the sizes that are deposited in the lower respiratory tract; many authors use a ≤5μm cut-off in for aerosols (Tellier, 2006). Particles ≥6μm are more likely to be trapped in the upper respiratory tract and particles >20μm are rarely transmitted to the lower respiratory tract (Tellier, 2006). However, the use of a cut-off is arbitrary, as deposition in the lower respiratory tract may still occur at diameters >5μm; furthermore, expelled particles rapidly shrink in size through evaporation (referred to as droplet nuclei) and increase the number of particles that behave as aerosols (Tellier, 2006). Airborne pathogens in larger particles can also accumulate on external surfaces; this deposition can lead to disease transmission by physical contact (Fiegel et al., 2006).

Only one study has reported the direct isolation and size quantification of *Mtb* aerosols generated *in vivo* by patients with pulmonary tuberculosis (Fennelly et al., 2004). Fennely et al. (2004) found that the majority of the viable particles in the cough-generated aerosols were in the immediately respirable size range, ranging from 0.65 to 4.7μm in diameter. In studies of *in vitro* generated aerosols, Lever et al. demonstrated that particle sizes vary with the suspension medium; the particles ranged from 0.65 to 3.3μm when nebulised in artificial saliva and 0.65 to >7.0μm when in saline.

Bacteria in aerosols face a number of environmental stresses through the course of infection; the ability of bacteria to initiate and spread disease depends on how well they are able to maintain their survival and infectivity during transmission (Cox, 1989). All airborne microbes in bioaerosols generated from liquid suspensions undergo desiccation, which is the predominant stress (Srikanth et al., 2008). Other environmental stresses vary on an individual

basis; these include temperature, relative humidity, oxygen, ozone, radiation and pollutants (Cox, 1989).

Currently, little information exists about the stability and viability of *Mtb* in aerosols, with which one could quantify the risk of transmission. Two studies attempted to examine mycobacterial survival in aerosols but their results were disparate. Loudon et al. (1969) reported 50% bacterial survival after 6h while Lever et al. (2000) reported much lower survival, at 50% after 5 minutes and only 10% after 30 minutes. This variability may be down to differences in experimental conditions, such as relatively humidity. Lever et al. (2000) also examined aerosol survival for *M. avium* and *Mycobacterium intracellulare* and found similar survival rates to *Mtb*. This is interesting, given the importance of airborne transmission to the virulence of *Mtb*, and the differences in the abilities of the three strains studied to cause human disease. Aerosol studies of *M. bovis* led to results that fell between the two *Mtb* studies, maintaining ~94% viability after 10 minutes, ~10% viability after 6 hours, and <1% viability after 12 hours. However, these differences in survival may be expected and limited conclusions can be inferred from them due to differences in experimental design, including bacterial strain, relative humidity and possibly temperature (Lever et al., 2000).

The composition of the medium in which the bacilli are suspended was not considered in the three studies on mycobacterial aerosols listed above. These made use of artificially generated aerosols in distilled water, saline and artificial saliva (Gannon et al., 2007, Lever et al., 2000, Loudon et al., 1969). It is thought that the droplets form from the airway lining fluid, that also contains lung mucus, surfactant material and a number of non-volatile soluble elements (Fiegel et al., 2006). The contribution of saliva and lower lung secretions to mycobacterial aerosols has not been studied.

5.1.1.1 Ultraviolet Radiation

Ultraviolet (UV) light is an environmental radiation stress to which *Mtb* in aerosols may be exposed (Cox, 1989). UV light is an environmental parameter and varies depending on location; however, UV light is also routinely used for the decontamination of surfaces in labs and hospitals (Collins, 1971). There are three classifications of UV light, based on wavelength: Ultraviolet A (UVA) from 400nm to 315nm, Ultraviolet B (UVB) from 315nm to 280nm and Ultraviolet C (UVC) from 280nm to 100nm (Vazquez and Hanslmeier, 2006). UVC is the most energetic UV light, having the shortest wavelength, and is germicidal; it is thus used in UV sterilisation methods (Vazquez and Hanslmeier, 2006).

UV light in addition to other energetic forms of radiation (e.g. gamma and X-rays) induce free-radical mediated reactions, leading to nucleic acid, protein, lipid and membrane damage (Cox, 1989). This damage includes the formation of thymine-thymine dimers, DNA-protein and protein-protein cross-links and the fragmentation or polymerisation of amino acids and sugars (Cox, 1989). These effects may be observed in bacteria in aqueous environments, but the damage is exacerbated by desiccation and oxygen stress (Cox, 1989).

Bacteria use various mechanisms to repair damage induced by UV light. These mechanisms include direct reversal of damage by a photolyase (photoreactivation), removal of the damaged base by a DNA glycosylase (base excision repair), incision of DNA adjacent to a damaged area by an endonuclease (UV-damage endonuclease) or removal of a complete oligonucleotide containing the damage (nucleotide excision repair) (Goosen and Moolenaar, 2008).

A number of these repair mechanisms have been documented in Mycobacteria. Genes encoding proteins for base excision repair, nucleotide excision repair and SOS (a global response to DNA damage) repair, as well as genes involved in the repair of oxidation and alkylation damage, have been identified in *Mtb* (Dos Vultos et al., 2009). Experimentally, David et al. (1971) showed that *Mtb* is capable of photoreactivation after inactivation from UV light. Furthermore, a nucleotide excision repair gene, *uvrB*, was required by *Mtb* resistance to both UV light and ROS *in vitro*; deletion of *uvrB* led to attenuation of *Mtb* in mice (Darwin and Nathan, 2005). Intriguingly, the upregulation of these DNA repair mechanisms was not a feature of the sputum transcriptome (Garton et al., 2008).

5.1.1.2 Desiccation Stress

To desiccate is defined as 'to remove the moisture from'; desiccation tolerance is the ability of cells to undergo this process of near absolute dehydration without dying (Billi and Potts, 2002, Oxford Dictionaries, 2010). Truly desiccated cells have insufficient residual water to maintain metabolic enzyme-catalysed reactions and therefore, desiccated cells do not grow (Billi and Potts, 2002). Hence, the main features of desiccation tolerance are arrest of cellular metabolism, time spent in this arrested state and then subsequent recovery of metabolic processes (Potts, 2001).

Studies aimed specifically at the survival of tubercle bacilli under desiccation have been limited, with the majority of work being done before the first World War; the summarised work concluded that 'dried sputum in rooms protected from abundant light has occasionally been found to contain virulent tubercle bacilli for as long as ten months' (Smith, 1942). More

recent studies have shown that *Mtb* bacilli can remain viable for up to a week when dried as an aerosol on glass in physiological saline (Potts, 1994).

A number of cellular targets are affected during bacterial desiccation. The presence of water in cells influences the physiological processes such as gene expression, membrane structure and stability and folding and assembly of proteins (Billi and Potts, 2002). Reduced hydration of proteins causes dysfunction of enzymes and electron transport chains, which can lead to the accumulation of free radicals; free radicals cause lipid peroxidation, protein denaturation and DNA mutations (Billi and Potts, 2002). Nucleic acids in particular are a major target of desiccation-induced damage as DNA damage may arise from chemical modification (alkylation or oxidation), base removal (e.g. depurination), cross-linking or radiation (ionising or UV) (Potts, 1994). The effects of cell damage from desiccation are greatest when the cell is in its fully hydrated state (Billi and Potts, 2002).

Desiccation tolerant cells may employ a number of mechanisms to curtail the damage induced by desiccation stress. It is thought that stress proteins serve as molecular chaperones to protect other proteins from denaturation and to aid their rehydration (Billi and Potts, 2002). Furthermore, in light of the oxidative stress caused by desiccation, proteins involved in oxygen scavenging mechanisms may have a role in increasing tolerance to dehydration (Billi and Potts, 2002). In studies in *E. coli,* the upregulation of control mechanisms for oxygen scavenging enzymes such as superoxide dismutases, catalase, glutathione reductase and alkyl hydroperoxide reductase has been shown in response to desiccation (Potts, 1994).

Eukaryotic anhydrobiotes feature the intracellular accumulation of trehalose in response to desiccation which is of particular interest in *Mtb* (Billi and Potts, 2002). Harland et al. (2008) demonstrated that trehalose 6,6' dimycolate (TDM), a major component of mycobacterial cell membranes, conveys desiccation resistance to model mycobacterial membranes. TDM itself is a glycolipid that presents as a highly-insoluble wax-like substance that is a strong immunoadjuvant (Hunter et al., 2006). TDM monolayers had the ability to be dehydrated and rehydrated without loss of integrity; the addition of as little as 25% TDM to phospholipid membranes confers this desiccation resistance (Harland et al., 2008). An extension of these studies demonstrated that other, synthetic trehalose glycolipids also conferred this desiccation resistance in lipid monolayers. This observation suggests that molecular interactions with the trehalose headgroup is the protective factor in this resistance (Harland et al., 2009). In the 'water replacement hypothesis', it is postulated that trehalose (and also sucrose) hydrogen bond membrane phospholipids and proteins, preventing membrane damage and inhibiting protein denaturation (Billi and Potts, 2002).

TDM is a molecule composed a trehalose headgroup attached to mycolic acid chains. The synthesis of mycolic acids is a multistep process (Takayama et al., 2005). The first step involves the synthesis of C24-C26 and C16 saturated straight chain fatty acids by FAS-1 (Bhatt et al., 2007). The C24-C26 chain creates the α-alkyl branch of mycolic acids. The C16 fatty acids produced by FAS-1 are transferred to Co-enzyme A; this C16 acyl-CoA provides the substrate for FAS-2 to produce the C56 meromycolate chain of the mycolic acids (Bhatt et al., 2007).

In addition to mycolic acid synthesis, C16, as well as C18, C20 and longer fatty acids are often incorporated into triacylglycerol. These fatty acids are either produced by FAS-1 or are obtained from the environment or culture medium; during *Mtb* infection, these fatty acids can enter the FAS-1 pathway and be used for the synthesis of mycolic acids (Takayama et al., 2005).

Overall, *Mtb* in aerosols and droplet nuclei undergo both desiccation and UV stress during transmission; these bacilli implement physiological, structural and molecular changes to survive this dehydration and exposure to radiation (Potts, 2001). However, the mechanisms and pathways which aid survival are still poorly understood in prokaryotes and it appears that a global stress response, which may include the formation of LBs, could aid survival to both stresses. In this context, bacilli displaying global stress resistance will promote survival that is essential to transmission. A study in *Rhodococcus* lends support to this theory, as the production of LBs led to increased desiccation tolerance (Alvarez et al., 2004). To wit, as the original hypothesis set out in this thesis proposed that LB-containing bacilli in sputum are transmission adapted, *Mtb* cultures with increased LB production were examined for tolerance to transmission stresses.

5.1.2 *M. tuberculosis* Binding to Macrophages

The main pathogenesis of tuberculosis was reviewed in Chapter 1, including the mechanisms by which *Mtb* is recognised and bound by host macrophages. A number of *Mtb* cell surface molecules interact with host cells. In particular, cells of the innate immune system identify invading microorganisms through the recognition of conserved molecular patterns in the pathogen. In mycobacteria, these molecular patterns may include DNA, lipoproteins, glycolipids and mycolic acids (Sundaramurthy and Pieters, 2007). It is interesting to note that some of these molecules are also involved in the virulence of *Mtb* and interact with the host immune response. For example, TDM is capable of inducing granuloma formation in the host, and inhibiting the trafficking of phagocytosed bacilli to acidic vesicles in the macrophage

(Rajni et al., 2011). Another cell wall component, lipoarabinomannan (LAM) allows *Mtb* to survive in the host by helping to inhibit T-cell proliferation and bactericidal activities of macrophages (Rajni et al., 2011). Other cell wall components, including mycolic acids and pthhthiocerol dimycocerosate (PDIM) also contribute to *Mtb* virulence (McKinney, 2000).

The sputum transcriptome featured the differential expression of a number of genes that may be involved in modification of mycobacterial cell wall components. Three polyketide synthase genes were upregulated in sputum; *pks5* and *pks9* encode proteins involved in the synthesis of methyl-branched fatty acids and *ppsA* encodes for a protein involved in PDIM biosynthesis (Garton et al., 2008). Furthermore, three genes, *mmaA2/mmaA3* and *cmaA2,* that encode for methoxy-mycolic acid synthases and a mycolic acid cyclopropane synthase, respectively, were downregulated in sputum; the three genes are involved in the modification of mycolic acids (Garton et al., 2008). The impact of the up or downregulation of these genes in sputum is yet to be elucidated.

The presence of triacylglycerol in the mycobacterial cell wall is debated. Ortalo-Magne et al. (1996) stated that triacylglycerol is a component of the *Mtb* cell envelope, but is not surface exposed. However, other studies have concluded that triacylglycerol is not a cell wall component as it has never been conclusively detected in cell wall (Brennan, 2003). Despite the controversy in the location of triacylglycerol in the cell wall, *Mtb* is known to accumulate triacylglycerol in the form of LBs and it has been suggested that cells may accumulate this triacylglycerol to ensure there is high intracellular fatty acid concentration required to synthesise components of the mycobacterial cell wall (Caceres et al., 2009).

The components of the mycobacterial cell wall and their interactions with human host cells play an important role in the transmission of tuberculosis. It remains to be determined if the cell wall components of *Mtb* bacilli in sputum differ from *in vitro* and other *in vivo* growth and if so, how these differences influence the transmission of these cells.

As for the transmission stresses, and in keeping with the hypothesis, *Mtb* cultures with increased LB production were examined for changes to macrophage binding.

5.1.3 Aims

In line with the hypothesis set out at the beginning of this thesis, the aims of the studies in this chapter were to examine the impact of LB accumulation on *Mtb* and its survival to transmission-related stresses and binding to macrophages:

- Examine the impact of two LB producing cultures, *tgs1* overexpression and nitric oxide exposure on the survival of *Mtb* to ultraviolet light and desiccation stress.
- Assess changes in *Mtb* binding to macrophages secondary to LB accumulation using the *tgs1* overexpressing strain.
- Examine the binding of the putative sputum phenotype *Mtb* cultures (Chapter 4) to macrophages.

5.2 Materials and Methods

5.2.1 Transmission Stresses

5.2.1.1 Ultraviolet Light Stress

Mtb was exposed to ultraviolet light in 12-well cell culture plates. Cultures of *Mtb* were prepared as described in Materials and Methods.

Prior to exposure, the optical density of the culture was taken, and the culture was diluted to a final optical density of 0.2 in 7H9 medium. 200µl of each culture to be tested was added to 2 wells of a 12-well cell culture plate (Becton Dickinson). The centre wells of the 12-well plate were used to maintain consistent ultraviolet exposure.

A UV Stratalinker (Stratagene) was set up in a Class I microbiological safety cabinet. The plates containing the *Mtb* culture was exposed to 2 x 10^5 and 4 x 10^5 ergs of ultraviolet light in the Stratalinker. A control plate of culture was included for each experiment. Post-UV exposure, the culture from each well was transferred to a screw-top micro-centrifuge tube. CFUs were enumerated from the culture from each well as described in 2.3.2.2.

5.2.1.2 Desiccation Stress

The desiccation stress experiment was based upon previous desiccation studies by Archuleta et al. on (2002) *M. avium*.

Mtb was desiccated on sterile coverslips. Nineteen millimetre glass coverslips were sterilised by soaking in industrial methylated spirits and were flamed over a Bunsen burner. Sterilised coverslips were stored in sterile Petri dishes prior to use.

A roller bottle of *Mtb* culture was prepared as described above. The culture was incubated at 37°C for 4-5 days and growth was monitored by measurement of optical density. Lipid body production was induced through exposure of nitric oxide as described in 2.3.2.4 or *tgs1* overexpression. Following lipid body induction, a volume of culture was centrifuged for 25 minutes. The cell pellet was washed once in 0.1% (w/v) Tween-80, centrifuged again for 25 minutes and resuspended in sterile distilled water at an optical density of 1.0. A control sample was prepared in the same manner. 100µl aliquots of sample were dropped onto the sterile coverslips. The samples were allowed to dry in a Class II microbiological safety cabinet.

Following drying, coverslips were resuspended in triplicate at each selected time-point. A single coverslips was transferred to a 50ml screw-top centrifuge tube containing 10ml 7H9 broth. The tube was vortex mixed for 2-3 minutes to adequately resuspend the desiccated *Mtb* bacilli.

CFU counts were performed on the lipid body positive and control cell suspensions, at a time-point immediately after drying/desiccation, and at all subsequent time-points.

5.2.2 Macrophage Infection Assays

5.2.2.1 Cell Line

The THP-1 adult leukaemia monocyte cell line was used for macrophage infection assays with *Mtb* strains. The cell line was acquired from the American Type Culture Collection (ATCC) and Dr. Bernard Burke (University of Leicester).

5.2.2.2 Storage of M. tuberculosis in RPMI-1640 for Infection

Mtb cultures used in preliminary macrophage infections were frozen in complete RPMI-1640 growth medium. *Mtb* was grown to exponential phase (OD 0.6-0.8) at 37°C, shaking at 100rpm. The cultures were centrifuged in 1ml volumes at 8000 × g and resuspended in complete RPMI-1640; this was repeated and the cell pellets were resuspended in a final volume of 1ml complete RPMI-1640. The 1ml aliquots were frozen at -80°C. One aliquot was retained and enumerated by CFU counting.

5.2.2.3 Culture of THP-1 cells

THP-1 cells were maintained in ATCC-formulated RPMI-1640 media supplemented with FCS, Penicillin/Streptomycin and ITS supplement as described in 2.2.2.

Monocytes were grown from frozen cell stocks in 1ml aliquots. The frozen aliquot was defrosted in water at 37°C. The volume of the aliquot was added to 9ml of pre-incubated complete RPMI-1640 medium in a 10ml screw-top centrifuge tube. The cells were centrifuged at 1000 × g for 10 minutes at 37°C. The supernatant was removed, and the cells were resuspended in 10ml of fresh, preincubated, complete RPMI-1640 medium. The cells

were transferred to a 75cm^2 tissue culture flask (Nalgene) and incubated statically at 37°C, 5% CO_2 and humidity to saturation.

THP-1 cell culture flasks were split and fed (through the addition of fresh medium) twice weekly (Monday and Friday). Cells were counted on a haemocytometer and diluted to a density of 2-4 × 10^5 cells/ml with fresh RPMI-1640 medium. The cells were subcultured when they reached a density of approximately 8 × 10^5 cells/ml. The cell concentration was not allowed to exceed 1 × 10^6 cells/ml. Cells were maintained in 75cm^2 cell culture flasks and incubated statically at 37°C, 5% CO_2 and humidity to saturation.

Frozen stocks were prepared through the centrifugation of growing cell culture and resuspension in complete RPMI-1640 medium supplemented with 5% (v/v) dimethyl sulphoxide (DMSO). Cells were first frozen in 1ml aliquots at -80°C using a "Mr. Frosty" Cryo 1°C Freezing Container (Nalgene), before being transferred to liquid nitrogen vapour phase for long term storage.

5.2.2.4 Preparation of THP-1 cells for infection

THP-1 cells were prepared for infection from actively dividing THP-1 cells culture, as described in 5.2.2.3. Cell cultures were transferred from culture flasks to a 50ml centrifuge tube and centrifuged at 1000 × g for 10 minutes. The supernatant was removed and the pellet was resuspended in 10-20ml of fresh, pre-incubated, complete RPMI-1640 media. The cells were counted in a haemocytometer.

The cell suspension was diluted to a cell density of 4 × 10^5 cells/ml and was supplemented with phorbol 12-myristate 13-acetate (PMA) to a final concentration of 100µM. The cells were plated out in 24-well cell culture plates (Nalgene). Each well was filled with 1ml cell suspension/PMA, giving a cell density of 4 × 10^5 cells/well.

The plates were incubated statically at 37°C, 5% CO_2 and humidity to saturation, for 48 hours to differentiate and adhere the THP-1 cells.

Following incubation and adherence of the THP-1 cells, the RPMI-1640 media was removed from each of the wells. The monolayers were washed once with Dulbecco's PBS (Sigma-Aldrich) and overlaid with 1ml of fresh, complete RPMI-1640 medium. The cells were re-incubated at 37°C, 5% CO_2 and humidity to saturation. The monolayers were then ready for infection with *Mtb*.

5.2.2.5 Infection of THP-1 cells with M. tuberculosis for Analysis of Binding

Erlenmeyer flasks of *Mtb* culture were prepared as described above. The culture was incubated at 37°C for 4-5 days and growth was monitored by measurement of optical density.

Optical density of the culture was taken on the day of the experiment. The culture was transferred to a 50ml screw-top centrifuge tube and centrifuged for 10 minutes at 2500 × g. The supernatant (7H9 media) was removed and the *Mtb* pellet was resuspended in complete RPMI-1640 media at a calculated OD of 0.2.

Following resuspension, a volume of cell suspension was added to each well of the 24-well cell culture plate containing THP-1 cell monolayers, as prepared above. Wells were infected in triplicate for each of the cultures tested. The volume of cell suspension added was to an estimated multiplicity of infection (MOI) of 1:1 and/or 10:1. The wells were mixed and incubated for 2 hours statically at 37°C, 5% CO_2 and humidity to saturation. CFU counts were performed for each of the cell suspensions as a control (as described previously).

Following incubation, plates were removed from the incubator, and the media overlaying each THP-1 cell monolayer was removed. The monolayers were washed three times with Dulbecco's PBS to remove unbound *Mtb* bacilli, and were overlaid with 900µl fresh, complete RPMI-1640 media. The plates were re-incubated for 2 hours, at which point, monolayers were lysed and *Mtb* cells were enumerated.

5.2.2.6 Lysing of monolayers and enumeration of M. tuberculosis cells

The THP-1 cell monolayers were lysed following 2 hours incubation for analysis of binding. To each of the triplicate wells, 100µl of 0.5% (w/v) SDS was added to lyse the monolayers. The final concentration of SDS in the wells was 0.05% (w/v). The supernatant was mixed by pipetting and left for 10 minutes at room temperature. The supernatant was mixed again by pipetting to ensure that the monolayers were thoroughly lysed. The lysed monolayer/supernatant was transferred to a new 2ml micro-centrifuge tube.

The lysed monolayers were serially diluted immediately and CFU counts were performed on the same day.

5.3 Results

5.3.1 M. tuberculosis Lipid Body Production *In Vitro*

As introduced in Chapter 1, the importance of LBs *in vivo* has yet to be determined. The unique phenotype of *Mtb* in sputum and its association with the consistent production of LBs led to the hypothesis that LB-positive *Mtb* bacilli possess a transmission adapted phenotype (Garton et al., 2008). Thus, the impact of the induction of LBs on the survival of *Mtb* to UV radiation and desiccation, and on binding to macrophages was assessed.

As outlined in Chapter 1, LB may occur as a result of the metabolic state that the bacillus is in. Thus, to determine if LBs were the main causative factor for the sputum phenotype, or merely an epiphenomenon or marker of a metabolic state, two independent methods for the induction of LBs *in vitro* were used. These two methods included the overexpression of *tgs1* and nitric oxide exposure.

In *Mtb*, *tgs1* was overexpressed through the introduction of the gene in the pSMT3 mycobacterium-*E. coli* plasmid, and was performed by Natalie Garton at the University of Leicester (Garton, 2011). The pSMT3-*tgs1* overexpressing strain was shown to only upregulate *tgs1* and confirmed by Natalie Garton to contain 3.5-fold greater triacylglycerol as compared to control (Garton, 2011). The results of this overexpressing strain, including the assessment of growth, lipid body production and fluorescence analysis of the *tgs1* overexpressing strain are shown and discussed in Appendix 3.

This strain was particularly useful for the assessment of LB production on *Mtb;* as only *tgs1* was overexpressed, this led to an increase in triacylglycerol content and, in turn, LB production. As the gene expression of other genes was not altered compared to the control strain (see Appendix 3), the production of LBs was isolated and studied without the influence of confounding factors.

Mtb exposed to nitric oxide was also examined. Nitric oxide exposure has been shown to induce LB production in *Mtb* and also leads to a host of other relevant changes (Sherratt, 2008). These changes include the induction of the dosR regulon and bacterial growth arrest. Upregulation of dosR regulated genes and the presence of slow growth transcription signals were prominent features of the sputum transcriptome (Garton et al., 2008). Furthermore, nitric oxide may play a role in the control of infection and as such, exhaled nitric oxide is detectable in a number of respiratory conditions, including asthma, other diseases causing airway inflammation and tuberculosis (Van Beek et al., 2011). Both nitric oxide exposure and the *tgs1* overexpressing strain were used for the main studies.

5.3.2 Transmission Stress: Ultraviolet Light Exposure

Ultraviolet light is a stress encountered by bacterial aerosols in the environment; furthermore, it is commonly used in hospital settings for bacterial disinfection (Collins, 1971). The impact of *tgs1* overexpression or nitric oxide exposure in *Mtb* on survival to UV light was assessed.

Preliminary experiments were performed prior to the start of the study. The exact output of the UV Stratalinker was not measured as the study aim was to directly compare the survival of two distinct samples of *Mtb*. UV Exposure colour change strips (UV FastCheck™, UV Process Supply, Inc.) were used to confirm the reproducibility of the UV Stratalinker settings. The strips change colour following a radiation dose and allow for quantification of an overall radiation dose; similar doses of radiation were noted for 5 replicate exposures (results not displayed).

The survival of *Mtb* H37Rv exposed to UV light at a range of energies was tested to choose the appropriate level for the main study. The total energy of each exposure was measured in ergs, where 1 erg is equivalent to 1×10^{-7} joules. Duplicate *Mtb* H37Rv samples were stressed with UV light at energy levels ranging from 100,000 ergs to 400,000 ergs. CFU counts were performed and the percentage survival was calculated. The results are displayed in Table 15. For the main study, *Mtb* cultures were exposed to two UV energies, 200,000 (approximately <50% survival) and 400,000 ergs (<1% survival).

Table 15: Survival of *M. tuberculosis* H37Rv to a range of UV light exposure

Energy of UV Exposure	Percentage Survival
100,000 ergs	92.11 ± 9.3
150,000 ergs	48.2 ± 22.6
200,000 ergs	40.3 ± 11.4
300,000 ergs	20.6 ± 2.9
350,000 ergs	6.8 ± 3.0
400,000 ergs	0.69 ± 0.003

For the main study, to assess the impact of *tgs1* overexpression and lipid accumulation on survival to ultraviolet light exposure, two experiments were performed. In the first experiment,

duplicate *Mtb* H37Rv cultures were grown to exponential phase. One culture was exposed to nitric oxide for 4h as described in Materials and Methods. The second culture was exposed to spermine hydrochloride for 4h as a control. Both cultures were exposed in duplicate to ultraviolet light in the Stratalinker UV Crosslinker. CFU counts were performed in duplicate and the survival of each strain calculated.

The first experiment was performed in duplicate. Results are displayed in Figure 32. Analysis of survival between nitric oxide exposed cultures and control cultures showed no significant difference in either experiment at both 200,000 ergs (p=0.3340, 0.4305, unpaired t-test) and 400,000 ergs (p=0.3935, 0.7820, unpaired t-test).

For the second experiment, the *Mtb* pSMT3-*tgs1* overexpressing strain and the pSMT3 plasmid control were grown to exponential phase. Both cultures were then exposed in duplicate to ultraviolet light in the Stratalinker UV Crosslinker. CFU counts were performed in duplicate and the survival of each strain calculated.

This second experiment was performed in triplicate. Results are displayed in Figure 33. Analysis of survival between pSMT3-*tgs1* overexpressing strain cultures and pSMT3 plasmid control cultures showed no significant difference in all three experiments at 200,000 ergs (p=0.9979, 0.8396 and 0.2320, unpaired t-test). At 400,000 ergs, the comparison of cultures is not significant for 2 of the 3 experiments (p=0.7345 and 0.1496, unpaired t-test). For the third experiment, the survival for pSMT3-*tgs1* was significantly higher than the pSMT3 plasmid control (p=0.0357*, unpaired t-test).

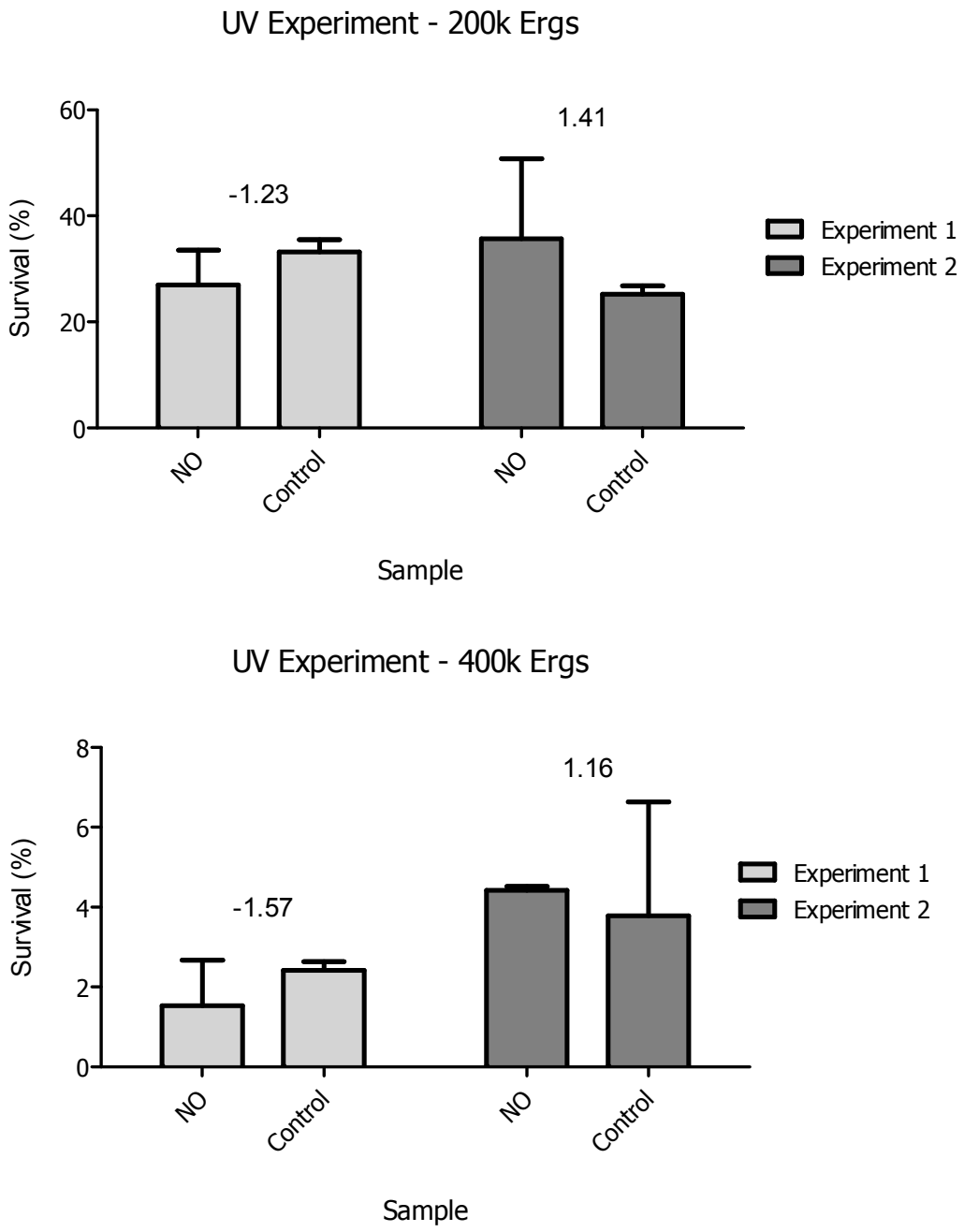


Figure 32: UV Exposure Survival of Nitric Oxide Exposed *M. tuberculosis* H37Rv vs Control

The bar graphs display the effect of ultraviolet light exposure on the survival of *Mtb* H37Rv exposed to nitric oxide versus control. The top and bottom graphs represent survival after 200,000 and 400,000 ergs of ultraviolet light exposure, respectively. Each experiment represents two replicates, each consisting of duplicate CFU counts; error bars show standard deviation. The number above each experiment displays the fold difference of survival of the nitric oxide exposed samples versus control samples.

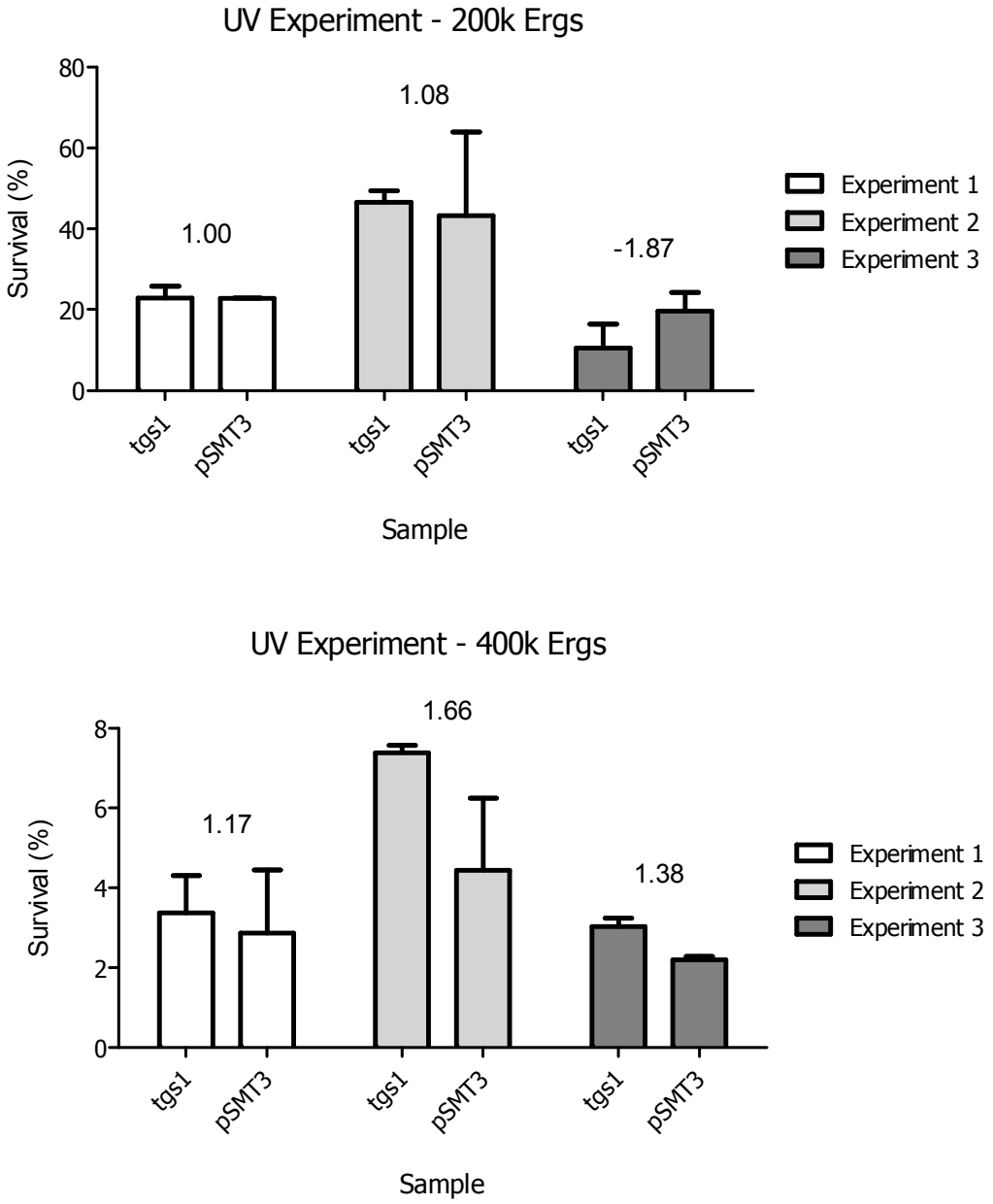


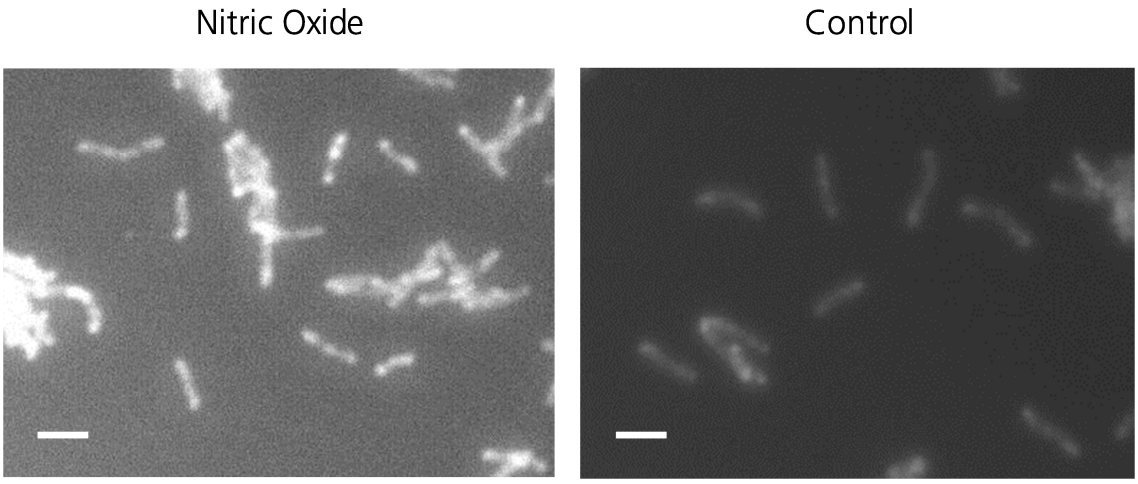
Figure 33: UV Exposure Survival of *M. tuberculosis* pSMT3-*tgs1* vs Control

The bar graphs display the effect of ultraviolet light exposure on the survival of *Mtb* pSMT3-*tgs1* overexpressing strain vs. pSMT3 plasmid control. The top and bottom graphs represent survival after 200,000 and 400,000 ergs of ultraviolet light exposure, respectively. Each experiment represents two replicates, each consisting of duplicate CFU counts; error bars show standard deviation. The number above each experiment displays the fold difference of survival of the *tgs1* overexpressing strain versus pSMT3 control strain.

Samples were taken from each culture for microscopic analysis. Nitric oxide exposed and control cultures were stained with Nile Red lipophilic stain. Work performed by Andrew Bell at the University of Leicester demonstrated that LipidTOX is a more appropriate stain as compared to Nile Red; LipidTOX is specific for neutral lipids, while Nile Red is not neutral lipid specific. Further, LipidTOX is more stable, as Nile Red fluorescence tends to bleach quickly following exposure to light (Bell, 2010). Finally, LipidTOX fluorescence has been demonstrated to correlate with the triacylglycerol content in mycobacteria; Nile Red fluorescence does not show this correlation (Bell, 2010). Thus, the pSMT3-*tgs1* and control cultures were subsequently stained with LipidTOX neutral lipid stain.

Lipid bodies were enumerated for the cultures in the UV exposure experiments. Fluorescent images were taken of the test and control cultures. Images were blinded and hand counted for lipid bodies; a minimum of 100 cells were counted per sample. The frequency of LB-positive cell percentages between the NO-exposed culture and control are listed in Figure 34. The percentages for the pSMT3-*tgs1* overexpressing strain and control are shown in Figure 35.

The median fluorescence intensity per cell was analysed for the *tgs1* overexpressing strains and control strains used in the UV exposure experiments. Fluorescent images were blinded and analysed with ImageJ and R Project Statistical Package, using the methods described in paragraph 2.4.5 (Bell, 2010). A minimum of 100 cells were analysed per sample. The results are shown in Figure 34. Fluorescence intensity was not analysed for the nitric oxide exposure versus control experiments as culture samples were stained with Nile Red.



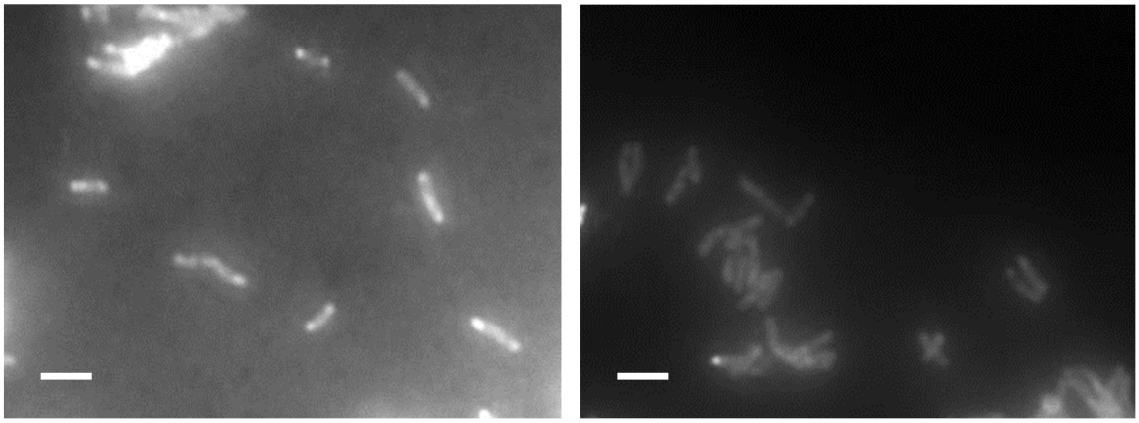
Experiment	LB-positive Cells	Total Cells	Lipid Body %	Ratio
Nitric Oxide Exposed Culture – 1	82	198	41.4	2.04
Control – 1	41	203	20.2	-
Nitric Oxide Exposed Culture – 2	65	172	37.8	3.12
Control – 2	20	165	12.1	-

Figure 34: Lipid Body Percentage Values for UV Experiment (Nitric Oxide exposed vs. Control) Cultures

The figure shows example fluorescence images of nitric oxide exposed *Mtb* cultures and unexposed control cultures. The table displays the number of LB-positive cells, total number of cells analysed, percentage of LB-positive cells in each culture. The ratio compares the percentage of LB-positive cells in the nitric oxide exposed culture versus control. Bar is 2μm.

pSMT3-*tgs1* Overexpression

pSMT3 Control



Sample	# of Cells	Median Fluorescence	Interquartile Range	Ratio (vs. Control)	# LB +ve (# Cells)	LB %	LB % Ratio
Experiment 1							
pSMT3-*tgs1*	336	368501	282431 - 446397	1.58	140 (329)	42.6	1.81
pSMT3 Control	221	233494	170692 – 304750	-	42 (179)	23.5	-
Experiment 2							
pSMT3-*tgs1*	137	147626	87142 – 207052	2.18	77 (214)	36.0	1.91
pSMT3 Control	172	67484	52568 – 86357	-	30 (160)	18.8	-
Experiment 3							
pSMT3-*tgs1*	229	193780	141770 – 272398	6.7	76 (224)	33.9	2.11
pSMT3 Control	163	28904	13996 – 83245	-	31 (192)	16.1	-

Figure 35: Fluorescence Analysis and Lipid Body Percentage Values for UV Experiment (pSMT3-tgs1 overexpressing strain and pSMT3 control) Cultures

The figure shows example fluorescence images of *Mtb tgs1* overexpressing strain cultures and control cultures. The table displays the number of cells analysed for fluorescence, median fluorescence intensity, interquartile range and the ratio comparing the fluorescence between the *tgs1* overexpressing strain and control. The table also displays LB data, including the number of LB-positive cells, total number of cells analysed, percentage of LB-positive cells in each culture. The LB % ratio compares the percentage of LB-positive cells in the *tgs1* overexpressing strain versus Control. Bar is 2μm.

The frequency of LB-positive cells was greater in both the nitric oxide exposed samples and the *tgs1*-overexpressing strain as compared to the respective controls. A Mann-Whitney test was used to compare the median fluorescence intensity per cell for the *Mtb* pSMT3-*tgs1* cultures versus pSMT3 control cultures. The median fluorescence intensity was significantly higher for the pSMT3-*tgs1* overexpressing strain in all three experiments ($p<0.0001^{**}$).

The effect of the presence of LB-positive cells on the survival to UV light was also assessed. The percentage of LB-positive cells was compared to percentage of cells surviving following UV light exposure for the *tgs1* overexpressing strain, nitric oxide exposed cultures and both sets of control cultures. The relationship between UV light survival and percentage of LB-positive cells was compared in Figure 36. As the presence of LBs and their effect on *Mtb* survival was examined (regardless of the method of induction), both sets of experiments (nitric oxide exposure and *tgs1* overexpressing strain) were included in the figure.

There appears to be no correlation between the ratios of survival and % of LB-positive cells for test samples versus control at both of the UV light exposures (200,000 and 400,000 ergs). A Spearman's rank test of correlation showed no correlation at either of the two exposures ($p=0.7850$ at 200,000 ergs and $p=0.9184$ at 400,000 ergs). The *tgs1* overexpressing strain experiments were also examined individually. A Spearman's rank test of correlation continued to show no correlation between survival percentage and LB-positive cell percentage. The nitric oxide exposure UV experiments were not examined individually as only 4 data points were available.

A comparison of the ratio of survival versus the ratio of LB production between the LB producing cultures (*tgs1* overexpressing strain and nitric oxide exposure) also revealed no relationship between production of LBs and survival to UV light (Results not displayed).

In addition to the UV experiments above, an experiment looking at the survival of *Mtb* pSMT3-*tgs1* versus control to a combination of both UV light and desiccation stress (see below) was attempted. *Mtb* was desiccated onto glass coverslips for 4 hours, and exposed to a 200,000 erg dose of UV radiation. After 4 hours, desiccated bacteria were resuspended and CFU counts were performed to enumerate the viable bacilli. However, no viable bacilli were recovered from these CFU counts, possibly due to the poor sensitivity of this experimental method. This is discussed further below.

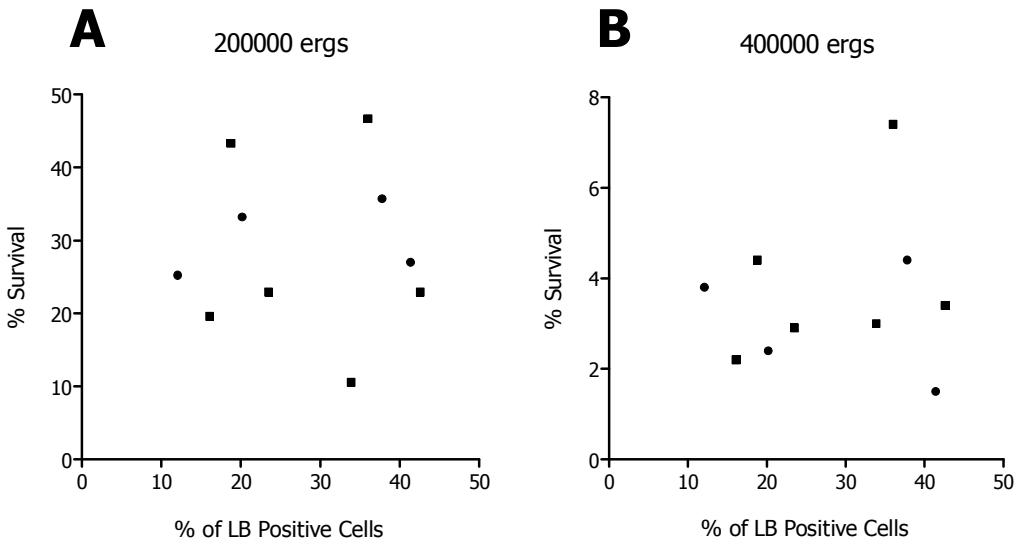


Figure 36: Percentage of Surviving *M. tuberculosis* cells following UV light exposure vs. Percentage of Lipid Body Positive Cells

The figures are scatter plots comparing the survival percentage of *Mtb* to UV light and the percentage of LB-positive cells, for the test samples (nitric oxide exposure and *tgs1*-overexpressing strain) and controls. Graphs A and B represent the two UV light exposures (200,000 ergs and 400,000 ergs), respectively. The points from the nitric oxide exposure experiments and *tgs1* overexpressing strain experiments are represented by circles and squares, respectively.

5.3.3 Transmission Stress: Desiccation

As earlier reviewed, desiccation stress is experienced by all airborne microbes (Cox, 1989). Desiccation is the most severe water deficit stress, as the removal of cell-bound water imposes structural, physiological and biochemical stress upon the cell (Billi and Potts, 2002).

Preliminary experiments were conducted to refine and standardise the desiccation procedure. The studies were based upon previous desiccation studies performed by Archuleta et al. (2002) Growth media were tested for its use in the desiccation studies. *Mtb* H37Rv in Sautons broth, 7H9 broth and PBS was dried onto sterilised glass coverslips. The samples in PBS dried fully, leaving a layer of bacilli and salt on the coverslip. However, the samples in Sautons and 7H9 broths were shown to have ineffective drying; this was likely due to the glycerol and Tween-80 used in both Sautons and 7H9 broth. Thus, the media used for the desiccation of *Mtb* was PBS.

The coverslips were tested to analyse the effectiveness of sterilisation. Glass coverslips were sterilised as described in 5.2.1.2 and placed in 10ml of Youmans broth in 50ml centrifuge tubes. The coverslip/broth was incubated for 1 week and growth noted. Five replicates were performed, and no bacterial growth was noted. Finally, microscopic examination on coverslips before and after resuspension of desiccated bacteria confirmed that no bacilli remained post-resuspension.

Following the preliminary studies, experiments were conducted compare the impact of lipid body lipid accumulation on the survival of *Mtb* to desiccation stress. As for ultraviolet light exposure, two sets of experiments were performed and made use of either the *Mtb tgs1* overexpressing strain or nitric oxide exposure.

In the first experiment, duplicate *Mtb* H37Rv cultures were grown to exponential phase. One culture was exposed to nitric oxide for 4h as described in Materials and Methods. The second culture was exposed to spermine hydrochloride for 4h as a control. Both cultures were dried onto coverslips and three coverslips were resuspended at time 0h, 4h and 24h as described in 5.2.1.2. CFU counts were performed in duplicate for each coverslip and the survival of each strain calculated.

For the second experiment, the *Mtb* pSMT3-*tgs1* overexpressing strain and the pSMT3 plasmid control were grown to exponential phase. Both cultures were dried onto coverslips and three coverslips were resuspended at time 0h, 4h and 24h as described. CFU counts were performed in duplicate for each coverslip and the survival of each strain calculated. Significance of strain survival differences were analysed by unpaired t-test.

Both sets of experiments were run in triplicate. Figure 37 shows the results for the survival of nitric oxide exposed cultures versus control. Figure 38 shows the survival results for the pSMT3-*tgs1* overexpressing strain cultures versus control.

At time 0h for the nitric oxide exposure versus control experiments, the survival for the control is significantly higher in one experimental replicate (p=0.0004*). The difference in survival was not significant for the other time 0h experimental replicate (p=0.6110; the third replicate was contaminated. For the other time-points, the control survival was significantly higher in two of the three experimental replicates at time 4h (p=0.0208*, 0.0004* and 0.6234) and time 24h (p=0.0303*, 0.0107* and 0.2267).

For the three experiments comparing pSMT3-*tgs1* and control, the control survival is significantly higher in one of the three experimental replicates at time 0h (p=0.0058*, 0.9235 and 0.1182) and time 4h (p=0.0226*, 0.1227 and 0.7183). At time 24h, there was no

significant difference between strain survival in any of the experimental replicates (p=0.5807, 0.9124 and 0.2298).

Attempts were made to measure desiccation experiment using the most probable number (MPN) method of enumeration to determine if the desiccated cells required liquid media to be revived. The desiccation experiment was repeated using the pSMT3-*tgs1* overexpressing strain and control; the MPN method was used, as opposed to CFU counts. However, the desiccation method involves the resuspension of coverslips (100µl of culture) in 10ml fresh media, diluting the initial culture 1:100 (v/v). This dilution rendered the MPN method to be not sensitive enough for the detection of bacilli and hence, results could not be obtained.

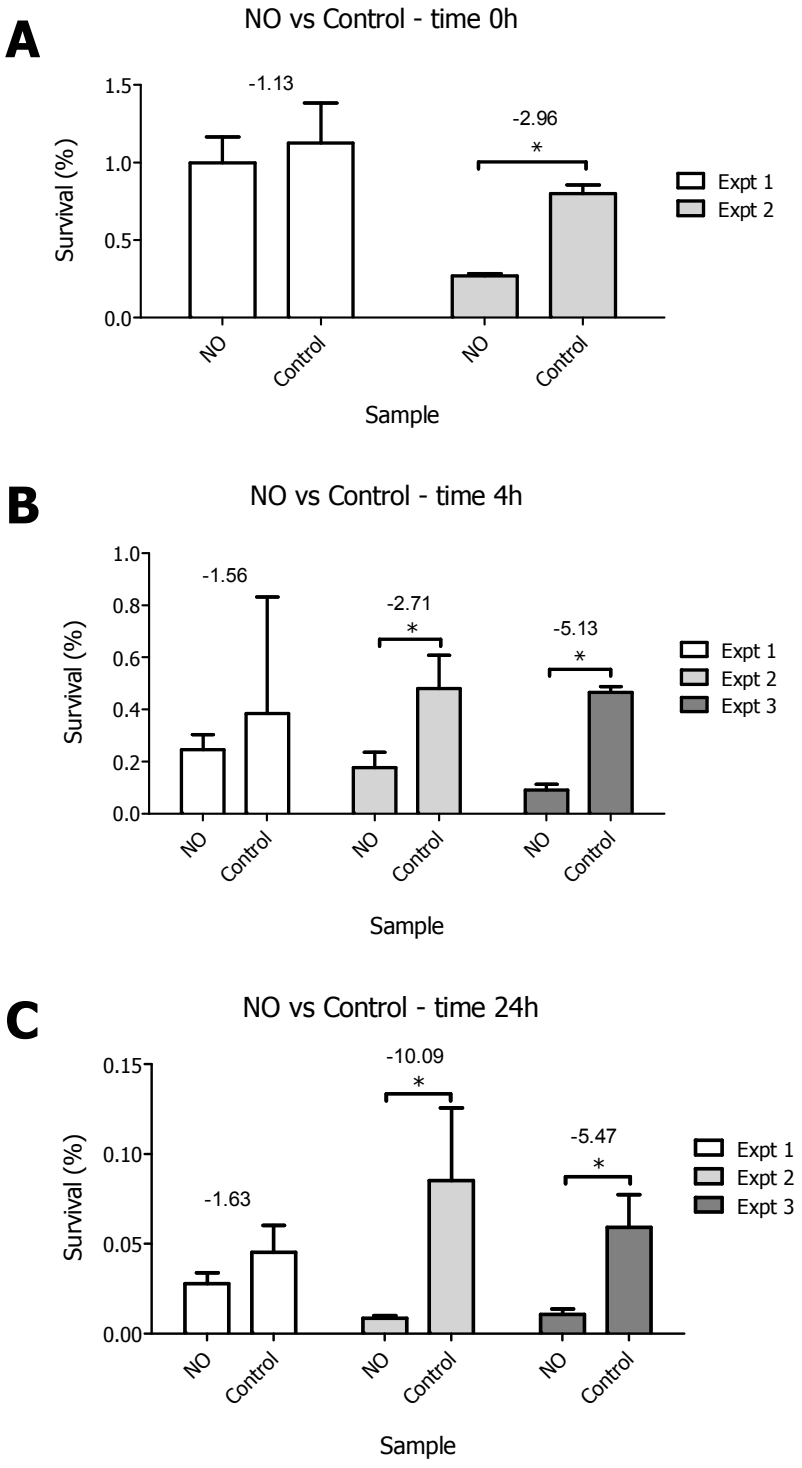


Figure 37: *M. tuberculosis* tolerance to Desiccation Stress – Nitric Oxide Exposure vs. Control

The graphs display tolerance of *Mtb* to desiccation stress over a period of 24 hours. Graphs A, B and C compare the percentage survival of *Mtb* exposed to nitric oxide versus control, at 0h, 4h and 24h after desiccation, respectively. Each experiment represents three replicates; error bars display standard deviation. Unpaired t-tests were used to compare the experiments at each time-point; stars designate significant results. The number above each experiment displays the fold difference of survival of the nitric oxide exposed samples versus control samples.

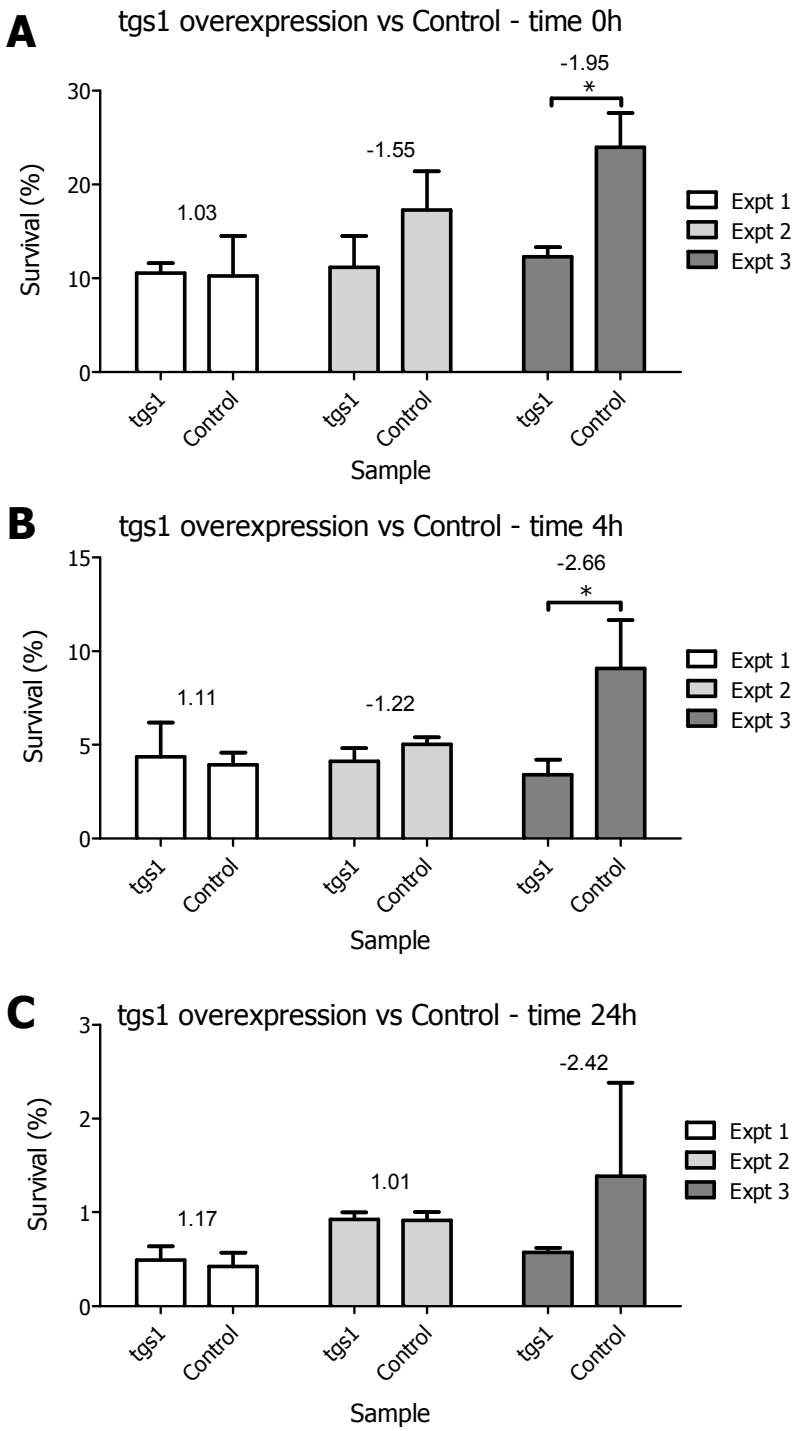


Figure 38: *M. tuberculosis* tolerance to Desiccation Stress – *tgs1* Overexpressing strain vs. Control

The graphs display tolerance of *Mtb* to desiccation stress over a period of 24 hours. Graphs A, B and C compare the percentage survival of *Mtb* pSMT3-*tgs1* overexpressing strain versus pSMT3 control strain, at 0h, 4h and 24h after desiccation, respectively. Each experiment represents three replicates; error bars display standard deviation. Unpaired t-tests were used to compare the experiments at each time-point; stars designate significant results. The number above each experiment displays the fold difference of survival of the *tgs1* overexpressing strain versus control.

Samples were also taken from each culture for microscopic analysis. As for the Ultraviolet light exposure, nitric oxide exposed and control cultures were stained with Nile Red lipophilic stain, while the pSMT3-*tgs1* and control cultures were stained with LipidTOX neutral lipid stain.

Fluorescent images were taken of the test and control cultures. Images were blinded and hand counted for lipid bodies; a minimum of 100 bacilli were counted per sample. The frequency of LB-positive cell percentages between the nitric oxide exposed culture and controls are listed in Table 16. The LB percentages for the pSMT3-*tgs1* overexpressing strain and control are shown in Table 17.

Table 16: Desiccation: Difference in Lipid Body Percentages between Test and Control

Experiment	LB-positive Cells	Total Cells	Lipid Body %	Ratio
Nitric Oxide Exposed Culture – 1	42	140	30.0	2.94
Control – 1	14	137	10.2	-
Nitric Oxide Exposed Culture – 2	65	160	40.6	2.86
Control – 2	16	113	14.2	-
Nitric Oxide Exposed Culture – 3	55	125	44.0	3.70
Control – 3	16	135	11.9	-

The median fluorescence intensity per cell was analysed for the *tgs1* overexpressing strains and control strains used in the UV exposure experiments. Fluorescent images were blinded and analysed with ImageJ and R Project Statistical Package; a minimum of 100 bacilli were analysed per sample. The results are shown in Table 17. Fluorescence intensity was not analysed for the nitric oxide exposure versus control experiments as culture samples were stained with Nile Red.

Table 17: Fluorescence Analysis and Lipid Body Percentage Values for Desiccation Experiment (pSMT3-tgs1 overexpressing strain and pSMT3 control)

Sample	# of Cells	Median Fluorescence	Interquartile Range	Ratio (vs. Control)	# LB +ve (# Cells)	LB %	LB % Ratio
Experiment 1							
pSMT3-*tgs1*	216	327219	242440 – 425067	1.61	104 (203)	51.2	2.32
pSMT3 control	218	203436	110,679 – 302732	-	46 (208)	22.1	-
Experiment 2							
pSMT3-*tgs1*	313	367104	252756 – 562686	1.75	146 (253)	57.7	2.33
pSMT3 control	274	209488	149551 – 271300	-	54 (218)	24.8	-
Experiment 3							
pSMT3-*tgs1*	362	222310	157629 – 288260	1.20	179 (290)	61.7	3.59
pSMT3 control	321	185212	137915 - 258420	-	47 (273)	17.2	-

The frequency of LB-positive cells was greater in both the nitric oxide exposed samples and the *tgs1*-overexpressing strain as compared to the respective controls.

A Mann-Whitney test was used to compare the median fluorescence intensity per cell for the *Mtb* pSMT3-*tgs1* cultures versus pSMT3 control cultures. The median fluorescence intensity was significantly higher for the pSMT3-*tgs1* overexpressing strain in all three experiments ($p<0.0001$** for Experiments 1 and 2, $p=0.0004$* for Experiment 3).

The effect of the presence of LB-positive cells on desiccation survival was also assessed. The percentage of LB-positive cells was compared to percentage of cells surviving following desiccation for the *tgs1* overexpressing strain, nitric oxide exposed cultures and both sets of control cultures. The relationship between desiccation survival and percentage of LB-positive cells was compared in Figure 39. As the presence of LBs and their effect on *Mtb* survival was examined (regardless of the method of induction), both sets of experiments (nitric oxide exposure and *tgs1* overexpressing strain) were included in the figure.

There appeared to be no correlation between the survival of *Mtb* and the percentage of LB-positive cells at any of the three desiccation time-points (times 0h, 4h and 24h). A Spearman's rank test of correlation showed no correlation at any of the 3 time-points ($p=0.4918$ at t0,

p=0.5731 at t4 and p=0.3545 at t24). It was notable that the nitric oxide exposed samples and controls had very low percentages of survival as opposed to the *tgs1* overexpressing strain and control samples. Due to this difference, the experiments were also examined individually; however, for each type of experiment (nitric oxide exposure and *tgs1* overexpression), a Spearman's rank test of correlation continued to show no relationship between survival percentage and LB-positive cell percentage. The nitric oxide exposure desiccation experiments at t0 were not examined individually as only 4 data points were available.

A comparison of the ratio of survival versus the ratio of LB production between the LB producing cultures (*tgs1* overexpressing strain and nitric oxide exposure) also revealed no relationship between production of LBs and survival to desiccation stress (Results not displayed).

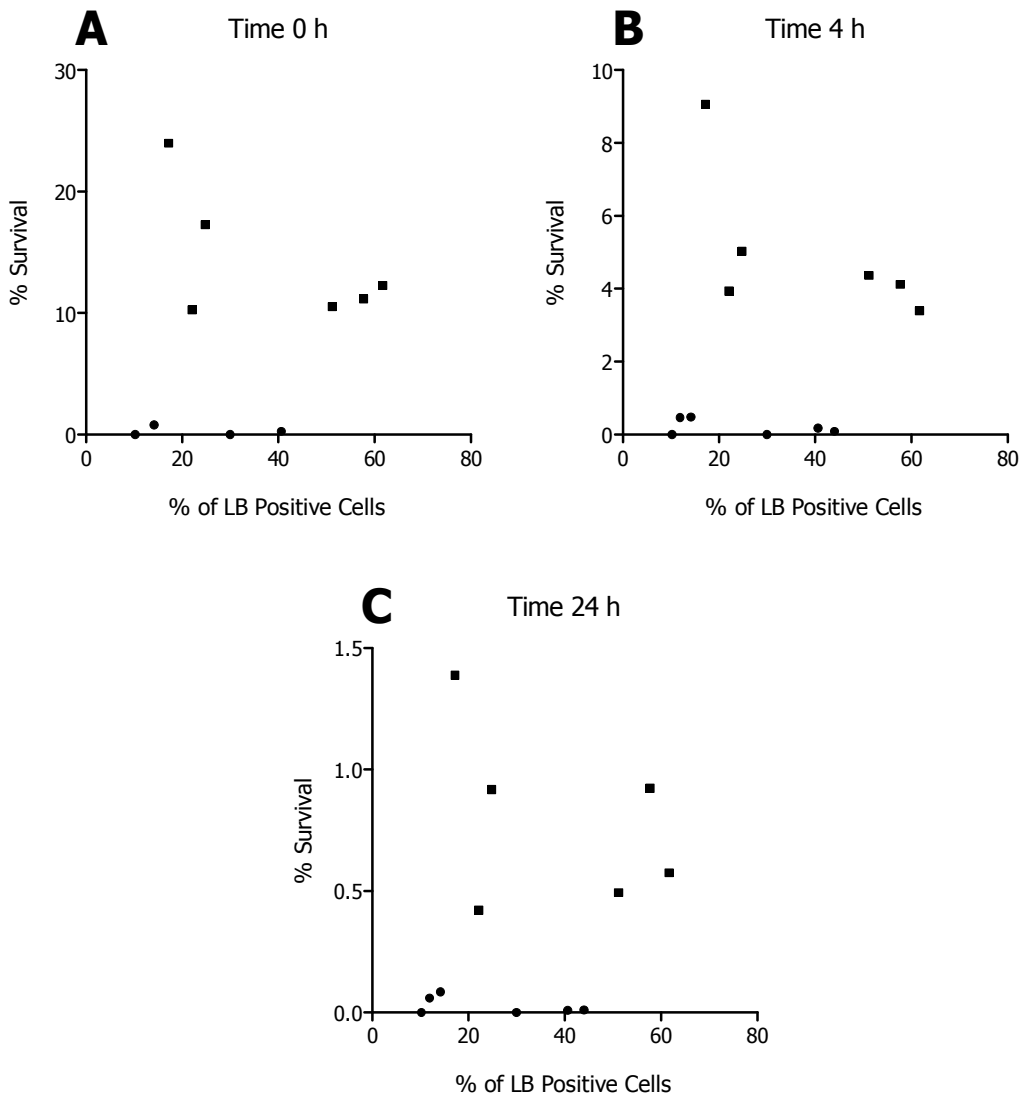


Figure 39: Percentage of Surviving *M. tuberculosis* Cells Following Desiccation vs. Percentage of Lipid Body Positive Cells

The figures are scatter plots comparing the survival percentage of *Mtb* to desiccation stress and the percentage of LB-positive cells, for the test samples (nitric oxide exposure and *tgs1*-overexpressing strain) and controls. Graphs A, B and C represent the three desiccation time-points (time 0h, 4h and 24h), respectively. The points from the nitric oxide exposure experiments and *tgs1* overexpressing strain experiments are represented by circles and squares, respectively.

5.3.4 Macrophage Binding

Macrophage infection is central in the pathogenesis of *Mtb* pathogenesis in humans. Binding and internalisation of *Mtb* bacilli by macrophages during primary infection is an important early step in the pathogenesis of tuberculosis (Fenton et al., 2005). As the early steps in the infection of hosts are vital in the transmission of tuberculosis, early intracellular growth of *Mtb* in macrophages was examined in preliminary experiments.

Preliminary experiments were performed using frozen *Mtb* strains (pSMT3-*tgs1* and control). The intracellular growth of the pSMT3-*tgs1* overexpressing and control strains was assessed over a period of 48 hours using frozen *Mtb* strains (results not displayed); this was performed as a preliminary experiment prior to the macrophage binding experiments. Macrophages were infected at an MOI of 0.1 for 2 hours, before being washed to remove unbound bacilli. Following phagocytosis of the previously frozen *Mtb* by THP-1 cells, a proportion of internalised bacteria died over the first 24 hours; the death of *Mtb* bacilli after internalisation was not noted in studies making use of actively growing cultures (Parish and Stoker, 2010). An assessment of survival also revealed a proportion of *Mtb* bacilli stored in RPMI-1640 died whilst stored at -80°C; strains had approximately 78% survival immediately after freezing and dropped to approximately 50% viability after one week. This change in viability in the frozen strains led to inaccuracies in MOI. Furthermore, the addition of dead bacilli to the macrophage monolayers may have an unknown biological effect.

Thus, as a proportion of frozen cultures died in storage and displayed poor binding to THP-1 macrophages, the decision was made to use actively growing, un-frozen *Mtb* cultures for further experiments. The intracellular growth of the *Mtb* strains was not studied further and experiments were designed to focus on the assessment of differences in the strain binding to macrophages.

In the preliminary experiments, an assessment of the initial growth of *Mtb* cultures in macrophages was made. Alternatively, for the main study, the assessment of *Mtb* macrophage binding was made. This decision was made as at the time of intial aerosolisation of *Mtb,* to the point in which bacilli encounter alveolar cells, *Mtb* express virulence determinants, which may be regulated by environmental conditions (Li et al., 2002). The environmental conditions within macrophages are different, and constitute a stage in the pathogenesis of tuberculosis after which initial infection has occurred.

To examine the influence of LBs on macrophage binding, the pSMT3-*tgs1* and control strains were assessed. Also, the macrophage binding of the *Mtb* H37Rv cultures in RPMI-1640 were exposed to multiple stimuli (see Chapter 4), which have gene expression more similar to the

sputum transcriptome, was assessed. For the purposes of this chapter, the *Mtb* H37Rv culture in RPMI-160 and exposed to nitric oxide, oleic acid, cholesterol and static incubation is referred to as the 'multi-stimulus culture'. *Mtb* cultures exposed to only nitric oxide were not assessed for their binding to macrophages.

In this study, experiments made use of human macrophages (THP-1 cell line), as *Mtb* is a pathogen exclusive to humans. Murine macrophages were not utilised as, compared to human macrophages, they exhibit differences in the *in vitro* interactions with *Mtb* and other intracellular pathogens (Reviewed by Kusner, 2005). These differences translate to the *in vivo* pathophysiology, where murine and rodent species are resistant to diseases caused by intracellular pathogens, such as tuberculosis, legionellosis, etc., while humans are susceptible to these diseases (Kusner, 2005).

For the comparison of the pSMT3-*tgs1* and control strains, *Mtb* pSMT3-*tgs1* overexpressing strain and pSMT3 control cultures were grown to exponential phase, washed and resuspended in RPMI-1640. THP-1 cell monolayers were prepared and infected in triplicate as described in 5.2.2.4. CFU counts were performed in duplicate for the cultures to determine MOI. Following infection and incubation, monolayers were lysed and CFU counts were performed in duplicate for each well. MOI and macrophage binding was calculated and compared. The experiment at two MOIs: twice at a low MOI and once at a higher MOI.

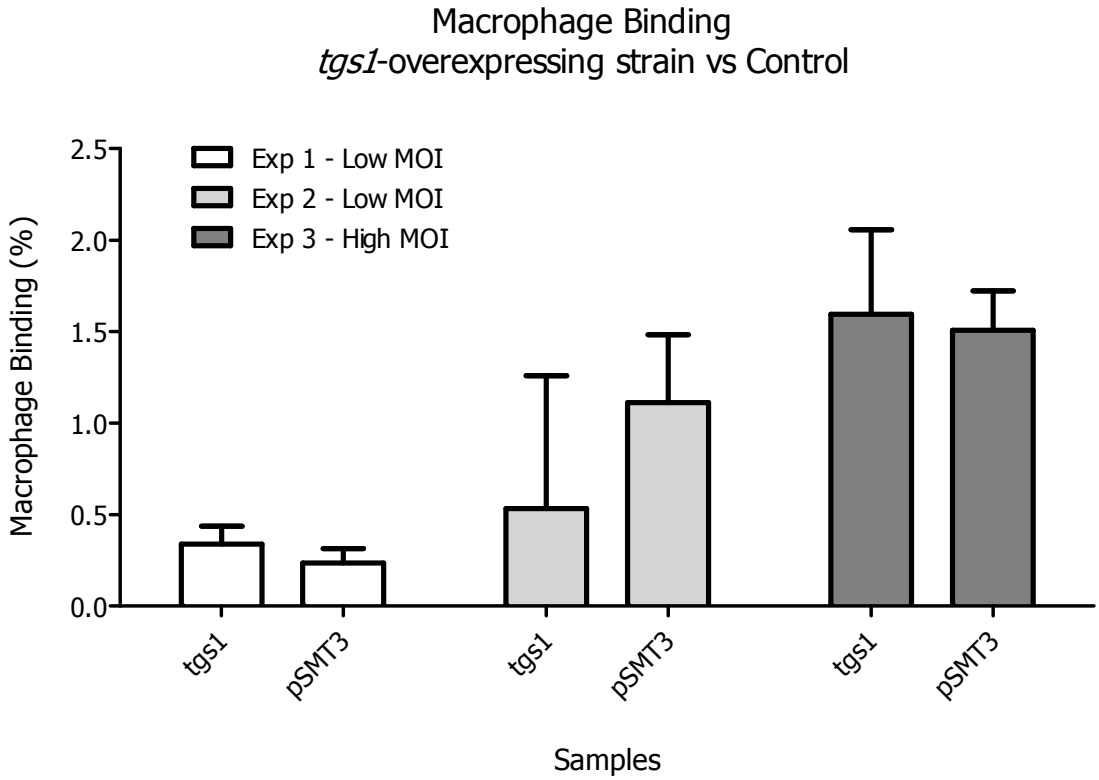
The binding percentages for *Mtb* pSMT3-*tsg1* strain versus pSMT3 control are shown in Figure 40. The MOIs were similar for both strains. The binding percentage results show no significant difference in binding between the two strains at low MOI ($p=0.2276$ and 0.2853) and high MOI ($p=0.7834$) using an unpaired t-test.

As a continuation of the work in Chapter 4, the gene expression sample that correlated best with sputum microarray gene expression was also examined. While there was variation in the microscopy data from these samples (discussed in 4.3.4.2), the samples currently represent the closest *in vitro* model of *Mtb* in sputum and were thus studied in this macrophage binding model. The two *Mtb* cultures that correlated best were PBS (with multiple stimuli and conical flasks) and RPMI-1640 (with multiple stimuli and universal tubes). Compared with sputum gene expression, correlation was significant and R^2 values were 0.573 and 0.566, respectively. As the correlation of both growth conditions to sputum gene expression was simliar, the *Mtb* culture in RPMI-1640 was chosen (referred to as 'multi-stimulus culture'); this decision was made on the basis that the THP-1 macrophages are cultivated in the same growth media.

The binding experiment made use of the same cultures analysed in Chapter 4. Two *Mtb* H37Rv cultures were washed and resuspended in complete RPMI-1640 media containing cholesterol in tyloxapol/ethanol and further exposed to nitric oxide (spermine NOnoate) and oleic acid. The cultures were incubated at 37°C statically for 4 hours. The controls were duplicate exponential phase culture in 7H9 media exposed to spermine hydrochloride, tyloxapol/ethanol and 37°C shaking incubation. The four cultures were used to infect THP-1 cell monolayers in triplicate, as for the other strains. Two MOIs were used per experiment.

The MOIs and macrophage binding percentages are displayed in Figure 41. The MOI appears to be twice as high for the control culture compared to the multi-stimulus culture. At the lower MOI, the macrophage binding percentage for the multi-stimulus culture is significantly higher than the control culture ($p<0.0001$**, unpaired t-test). At the higher MOI, the result similar; the macrophage binding percentage for the RPMI-1640 test culture is significantly higher than control ($p<0.0001$**, unpaired t-test).

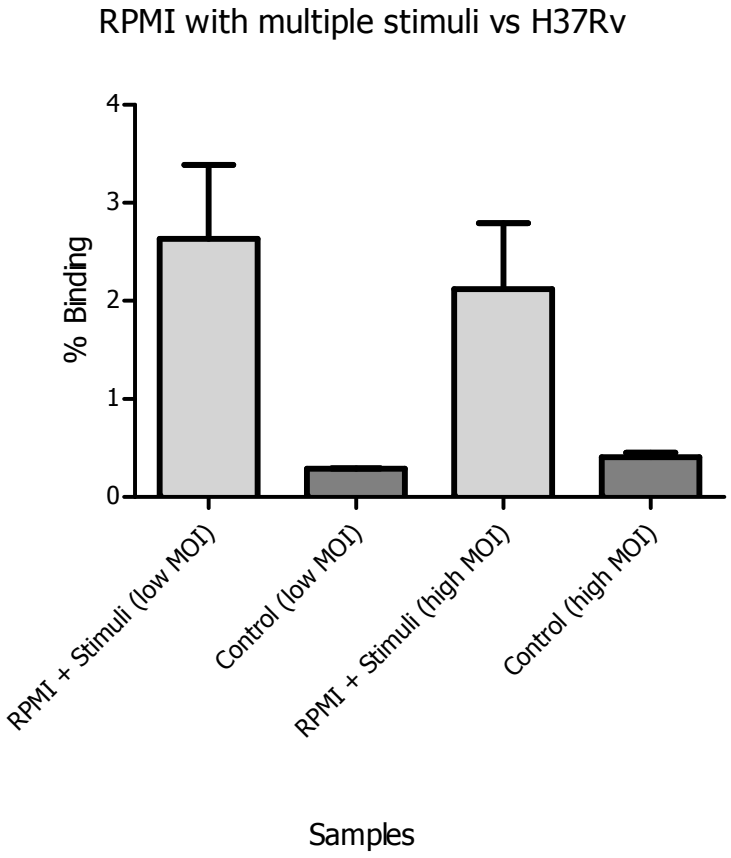
Finally, the binding of *Mtb* in sputum to macrophages was attempted. Sputum was obtained from tuberculous patients; aerobically grown *Mtb* H37Rv was used as the control. PANTA antibiotic mixture (Becton-Dickinson) was added to the sputum samples, as per the instructions (Lee, 2007). The sputum sample and control were decontaminated as described by Lee (2007) and used for infection in a macrophage binding experiment as described. However, following incubation of the CFU counts, no results were obtained due to widespread contamination from other bacteria present in sputum.



Sample	% Binding ± SD	MOI
pSMT3-*tgs1* overexpressing (Low MOI/Exp 1)	0.23 ± 0.08	3.84
pSMT3 Control (Low MOI/Exp 1)	0.34 ± 0.15	5.14
pSMT3-*tgs1* overexpressing (Low MOI/Exp 2)	1.11 ± 0.37	3.20
pSMT3 Control (Low MOI/Exp 2)	0.53 ± 0.76	3.78
pSMT3-*tgs1* overexpressing (High MOI/Exp 3)	1.51 ± 0.22	64.96
pSMT3 Control (High MOI/Exp 3)	1.59 ± 0.46	64.9

Figure 40: Macrophage binding of *M. tuberculosis* pSMT3-*tgs1* vs. Control

The comparison in macrophage binding of the *Mtb* pSMT3-*tgs1* overexpressing strain versus pSMT3 plasmid control strain is displayed. Each experiment represents three replicates. Error bars display standard deviation. The table shows numerical values and MOI for each experiment.



Sample	% Binding ± SD	MOI
Multi-Stimulus Culture (Low MOI)	2.64 ± 0.84	4.10
H37Rv Control (Low MOI)	0.29 ± 0.07	8.20
Multi-Stimulus Culture (High MOI)	2.12 ± 0.58	41.03
H37Rv Control (High MOI)	0.34 ± 0.15	82.01

Figure 41: Macrophage Binding of *M. tuberculosis* H37Rv in Multi-Stimulus Culture vs. Control

The comparison in macrophage binding of the *Mtb* H37Rv multi-stimulus culture versus control strain is displayed. Each experiment represents three replicates; error bars display standard deviation. The table shows numerical values and MOI for each experiment.

Samples were also taken from each culture for microscopic analysis to confirm an increased production of LB and increased fluorescence intensity in the *tgs1* overexpressing strain compared to control. The cultures were stained with LipidTOX neutral lipid stain. Fluorescent images were taken of the test and control cultures. Fluorescent images were blinded and the median fluorescence intensity per cell was analysed with ImageJ and R Project Statistical Package. Images were hand counted for LBs; a minimum of 100 bacilli were analysed/counted per sample for both fluorescence intensity and lipid body counting. The differences in lipid body percentages between the test cultures pSMT3-*tgs1* overexpression and controls are listed in Table 18.

Table 18: Fluorescence Analysis and Lipid Body Percentage Values for Macrophage Binding Experiments between Test and Control Cultures

Sample	# of Cells	Median Fluorescence	Interquartile Range	Ratio (vs. Control)	# LB +ve (# Cells)	LB %	LB % Ratio
Experiment 1							
pSMT3-*tgs1*	228	194236	141862 - 273515	5.75	100 (194)	51.5	2.28
pSMT3 control	276	33771	25401 - 64720	-	52 (230)	22.6	-
Experiment 2							
pSMT3-*tgs1*	332	366451	281760 - 444666	1.62	115 (233)	49.4	2.04
pSMT3 control	198	226617	169967 – 283719	-	54 (223)	24.2	-
Experiment 3							
pSMT3-*tgs1*	280	212201	151691 – 315964	3.14	150 (240)	62.5	3.22
pSMT3 control	170	67484	52603 - 85968	-	51 (263)	19.4	-

The frequency of LB-positive cells was demonstrated to be greater in the *tgs1*-overexpressing strain as compared to control, in all three cultures used in the binding experiments. Further, the median fluorescence intensity per cell for the *Mtb* pSMT3-*tgs1* cultures versus pSMT3 control cultures was compared using a Mann-Whitney test. The median fluorescence intensity was confirmed to be significantly higher for the pSMT3-*tgs1* overexpressing strain in all three experiments ($p<0.0001$**).

For the multi-stimulus cultures and controls, lipid body counts and median fluorescence intensity values were analysed in Chapter 4. As previously discussed, the percentage of LB-positive cells was higher in the multi-stimulus cultures as compared to control. However, paradoxically, the median fluorescence intensity of the multi-stimulus cultures was found to be significantly lower than the control.

5.4 Discussion

5.4.1 M. tuberculosis Lipid Body Production *In Vitro*

Tgs1 overexpression and nitric oxide exposure were used in this study for the induction of LB production. Both sets of samples were particularly useful for the purposes of this study. The overexpression of *tgs1* (further discussed in Appendix 3) produced LBs in *Mtb* in isolation; this allowed for the impact of increased LB production to be examined without the contribution of other phenotypic or transcriptional changes.

Alternatively, *Mtb* cultures exposed to nitric oxide also produce LBs. However, this production occurred secondarily to a host of transcriptional and phenotypic changes. As mentioned in 5.3.1, these changes include the induction of the dosR regulon and bacterial growth arrest, and coincide with the sputum transcriptome findings (Garton et al., 2008). Hence, the use of nitric oxide exposed *Mtb* cultures allowed for the study of increased LB production as a secondary feature.

As described in each of the 3 experiments (UV light exposure, desiccation and macrophage binding), the induction of LB production over the controls was described in both *tgs1* overexpression and nitric oxide exposure. It is important to note however, that LB production was not isolated to these samples. There was underlying LB production in the control samples, and any induction in LB production was on top of the normal physiological LB level. This production of LBs in control samples was not documented in previous studies (Sherratt, 2008). As LBs are counted by individual users, this interpretation of LBs may account for this observed difference; however, control sample and experimental sample images were blinded prior to counting in this study. With a certain percentage of LBs being naturally produced, the task of differentiating the natural *Mtb* level of survival to stresses and any survival advantage conferred by additional LBs is made more difficult.

The reason for LB production in vitro is intriguing, as the 7H9 medium used to grow control *Mtb* strains contains carbohydrates. One possibility is the addition of Polysorbate 80 (Tween-80) to the 7H9 medium. Polysorbate 80 is derived from polyethoxylated sorbitan and oleic acid and in culture, acts as a non-ionic surfactant, the detergent effects of which prevent cellular clumping. However, the presence of a bacterial lipase may release fatty acids into culture. Exogenous fatty acid sources are potentially bactericidal, due to their detergent like actions (Kondo and Kanai, 1972). The proposed model of LB formation in mycobacteria in Chapter 1 demonstrates that these exogenous fatty acids may be incorporated into LBs as a protective mechanism. Observationally, *M. smegmatis* treated with Polysorbate 80 quickly resulted in increased LB production (Garton, 2011).

5.4.2 Transmission Stress: Ultraviolet Light Exposure

For *Mtb* and other aerosolised microbes, UV light is the most commonly encountered radiation form (Collins, 1971). While the overexpression of *tgs1* or the exposure to nitric oxide in *Mtb* may not be directly associated with the induction of genes involved in bacterial survival to UV light, both sets of samples were associated with increased LB production. Thus, the impact of this LB production and the possible transmission adaptation of LB-positive cells was tested through the exposure of *Mtb* to UV light.

5.4.2.1 Tgs1 overexpression and Nitric Oxide exposure do not increase survival of M. tuberculosis to UV light

In 5.3.2, the results show no significant difference in survival to UV light exposure for the *Mtb* pSMT3-*tgs1* overexpressing strain versus control or nitric oxide exposed *Mtb* versus control, respectively. This result was repeated in duplicate experiments for the nitric oxide exposed cultures, and in triplicate experiments for the pSMT3-*tgs1* overexpressing cultures.

Analysis of fluorescence images confirmed a higher frequency of LB-positive cells in both the pSMT3-*tgs1* overexpressing strain and nitric oxide exposed cultures, as compared to control, suggesting the accumulation of triacylglycerol in these samples. Median fluorescence was only analysed for the *tgs1* overexpressing strain, and confirmed higher median fluorescence intensity than the control strain, suggesting a higher level of neutral lipid.

There appeared to be no relationship between the survival *Mtb* to UV light and percentage of LB-positive cells as analyses showed no correlation at both 200,000 and 400,000 ergs. There was also no correlation when the *tgs1* overexpressing strain experiments were examined individually. Unfortunately, there were only 4 data points available for the nitric oxide exposure UV experiments and therefore, correlation with LB positivity was not examined. In addition to the percentage survival, the ratio of survival between the test samples and controls was analysed; again, there was no correlation between this ratio, and the ratio of LB-positive cell percentage.

Overall, these results suggest that the accumulation of LBs in *Mtb* does not confer any UV survival advantage. Thus, while a feature of *Mtb* in sputum, LBs may reflect a greater phenotypic change than the simple upregulation of triacylglycerol production. As outlined in

5.1.1.1, upregulation of a number of repair genes is likely involved in the increased survival to UV radiation.

One intriguing feature from these experiments was that the dose required to inactivate the *Mtb* cultures was higher than in previous published studies. The amount of UV light required to inactivate *Mtb* did vary between these studies. David et al. (1971) showed that a UV dose of 56,700 ergs was required to reduce the number of viable bacilli by 1 log. However, a similar study showed a lower dose (approximately 48,000 ergs) led to a 2-log reduction in viable bacilli (Collins, 1971). In this study, the dose required for a 2-log reduction in viability was nearly 10-fold greater. The reasons for this may be due to variations in experimental design, including the UV lamp set-up and the manner by which *Mtb* cultures are exposed to the radiation. This difference is interesting, but does not affect the overall result, as the study directly compared the survival of two *Mtb* cultures and did not involve the measurement of UV doses required for inactivation.

The likelihood of *Mtb* encountering the bactericidal levels of UV light used in this study is naturally difficult to elicit. The level of UV light reaching the Earth's surface varies based upon geography and atmospheric characteristics (Vazquez and Hanslmeier, 2006). UVC is the most bactericidal of the three classes of UV light (Vazquez and Hanslmeier, 2006). However, the environmental contribution of UVC is limited; although measurements of the UV radiation that reaches the Earth's surface vary, approximately 95% is UVA. The majority of UVB and UVC is blocked by the atmosphere and the bactericidal effect of these classes of radiation, particularly UVC may be limited (Vazquez and Hanslmeier, 2006). The UV light produced by the Stratalinker lamp has a wavelength of 254nm, and hence is classified as UVC. However, whilst not as effective as UVC, UVA has been shown to inhibit the growth of *E. coli* in a dose-dependent manner (Daronch et al., 2007). Thus, the overall environmental effect of the three classes of UV light on *Mtb* is largely undefined.

While the levels of environmental UVB and UVC are low and *Mtb* can undergo photorepair, rather than acting as an individual stress (which may only be bactericidal at levels higher than those naturally encountered), UV light may act as part of a greater spectrum of stresses (Cox, 1989). In addition to UV radiation, other factors that affect bacterial survival in aerosols include desiccation, oxygen stress, temperature, etc. (Cox, 1989). These already 'stressed' cells may have a lower tolerance to UV light. This may be evidenced by the lower UV exposure required to kill *Mtb* H37Rv that had been pre-stressed through the exposure to nitric oxide. Furthermore, no viable *Mtb* bacilli were found following exposure to a combination of two stresses, UV light and desiccation; when each stress was tested individually, viable cells were detected.

As described in 5.3.2, a combined UV light and desiccation experiment was performed, but no viable bacilli were found through CFU counting. When conducted individually, the UV light experiment (200,000 ergs) and desiccation experiments (see below) both returned viable bacilli. This may suggest that the combination of two stresses led to the decrease in viability of the *Mtb* bacilli. While the Miles and Misra method for CFU enumeration is very sensitive, the neat and 10^{-1} dilutions could not be analysed, as the resuspension method of the desiccated bacteria involved a 1:100 (v/v) dilution and should be studied further. Thus, due to the unique logistics of this experiment, the CFU count method may be rendered too insensitive for the purposes of this experiment.

Overall, the results presented from the UV experiments suggest that an increase in LBs alone does not impart resistance to UV light on *Mtb* strains. As discussed, LBs were unlikely to be directly involved in the resistance to UV light, but may indicate the transmission adaptation to the cell. It appears that LBs alone are not the causative agent of a transmission-adapted phenotype, and it is likely that a global response involving a number of other genes/enzymes may be required.

5.4.3 Transmission Stress: Desiccation

For the transmission of airborne disease indoors, desiccation is the predominant stress faced by aerosolised microorganisms (Cox, 1989). The study compared the impact of *tgs1* overexpression or nitric oxide exposure and subsequent LB production on the survival of *Mtb* to desiccation stress.

5.4.3.1 Tgs1 overexpression and Nitric Oxide exposure do not increase desiccation resistance in M. tuberculosis

In 5.3.3, the results show no significant desiccation survival advantage for the *Mtb* pSMT3-*tgs1* overexpressing strain versus control or nitric oxide exposed *Mtb* versus control, respectively. In fact, the pSMT3-*tgs1* overexpressing strain had significantly lower survival than control in one out of three replicates in all three time-points (0h, 4h and 24h). Again, the control sample had significantly higher survival than a nitric oxide exposed culture for one out of two experimental replicates at the 0h time-point, and two out of three replicates at the 4h and 24h time-points.

The LB analysis did confirm a higher frequency of LB-positive cells in both the pSMT3-*tgs1* overexpressing strain and nitric oxide exposed cultures, as compared to control, suggesting the accumulation of triacylglycerol in these samples. The fluorescence analysis was only analysed for the *tgs1* overexpressing strain, and confirmed higher median fluorescence intensity than the control strain, suggesting a higher level of neutral lipid.

There appeared to be no relationship between the survival *Mtb* to desiccation and percentage of LB-positive cells as analyses showed no correlation at any of the three time-points (t0, t4 and t24). A Spearman's rank correlation test was also performed for each of the experiment types. Individually, there was also no correlation between survival and LB-positive cell percentage for either the *tgs1* overexpressing strain or nitric oxide exposure experiments. A notable feature was that the samples from the nitric oxide exposure experiments (plus controls) had lower survival compared to the *tgs1* overexpression samples. Taking this into account, the survival and LB cell positivity of test samples and controls were compared as ratios. However, there continued to be no relationship between the survival and presence of LB-positive cells.

Overall, the results suggest that the increase in intracellular triacylglycerol (stored as LBs) alone does not coincide with increased desiccation tolerance. In fact, the exposure to nitric oxide appears to increase the susceptibility of the bacilli to the stress. Furthermore, the control strain had significantly higher survival than the *tgs1* overexpression strain in one of the biological triplicates.

Thus, the expression of *tgs1* and the subsequent increase in intracellular triacylglycerol may not play a role in desiccation resistance, or other enzymes are required in a global response. The reasons for this significantly lower survival of the test samples compared to the respective controls are unknown. If TDM imparts desiccation resistance on *Mtb*, the upregulation of *tgs1* may direct fatty acids away from TDM synthesis and towards the production of triacylglycerol instead. In the nitric oxide exposed samples, the combination of the RNI stress imposed by nitric oxide, in addition to the desiccation stress may to the reduction in *Mtb* viability.

Despite the observation in *Rhodococcus* that formation of LBs was associated with increased survival to desiccation, this was not a singular change (Alvarez et al., 2004). Other cell changes included reduction in metabolic activity, biosynthesis of various osmolytes, adjustments to the cell wall, formation of short-fragmenting cells and the production of an extracellular 'slime' covering the surface of colonies (Alvarez et al., 2004). While the increase of LBs in *Mtb* was addressed through the upregulation of *tgs1* in both test cultures, and the reduction in metabolic activity has been demonstrated following nitric oxide exposure, other

genes and enzymes are likely to be required for *Mtb* desiccation tolerance (Voskuil et al., 2003).

While fatty acids, stored as triacylglycerol, can be funnelled into the FAS-I pathway or be used in the production of mycolic acids, other genes are required for the fatty acids to be liberated from the glycerol backbone. The changes in LB content displayed by *Mtb* suggest the existence of either intracellular or extracellular lipases (Cotes et al., 2008). As reviewed in Chapter 1, the sequencing of the *Mtb* genome revealed that, for bacteria, there are a large number of enzymes involved in the degradation of lipids (Cotes et al., 2008). The annotated genome contains 24 genes that may encode for lipolytic enzymes; one such lipase, LIPY (product of Rv3097c), belonging to the hormone-sensitive lipase family, has been shown to be induced under conditions of starvation to utilise stored triacylglycerol in *Mtb* (Deb et al., 2006).

Lipases may be involved in bacterial pathogenicity, a theory suggested based upon results in other bacteria, such as *Vibrio cholerae* and *Pseudomonas aeruginosa* (Cotes et al., 2008). Enzymes purified from these bacteria have been found to inhibit phagocytic functions of alveolar macrophages and modulate the release of inflammatory mediators by cells of the immune system (Reviewed by Cotes et al., 2008). The influence of these lipases on the *Mtb* transmission and desiccation remains to be determined. Furthermore, with the exception of the studies by Deb et al. (2006) using nutrient stress, the conditions inducing *Mtb* utilisation of triacylglycerol stores remains undefined.

In addition to lipases, isocitrate lyase (ICL) may play a role in the desiccation response. A study on the gene expression of *Bradyrhizobium japonicum* showed upregulation of *icl* 11.7-fold and 31.9-fold after 4h and 24h of desiccation stress, respectively (Cytryn et al., 2007). As reviewed in Chapter 1, ICL is a glyoxylate shunt enzyme; the glyoxylate shunt is required for carbohydrate synthesis and supply of Krebs cycle intermediates when fatty acids are the principle carbon source (Russell, 2001). Furthermore, ICL has been shown to be required for infection in mice and the establishment of late stage infection (Munoz-Elias and McKinney, 2005). Whether ICL is involved in the global desiccation response in *Mtb* or during transmission remains to be seen; however, upregulation of *icl1* is seen in the sputum transcriptome (Garton et al., 2008).

Unfortunately, previous studies examining desiccation survival of *Mtb* aren't comparable to this current study. This study directly compared two strains and looked at the survival of bacilli (as a percentage) over a fixed period of time. Previous studies examined the total length of time that *Mtb* bacilli were detectable after desiccation (Potts, 1994).

Survival was also lower than studies looking at mycobacterial survival in aerosols. A major difference between this study and the aerosol survival studies was the volume of culture used; in principle, the aerosol cultures should have reached a desiccated state faster. However, it appeared that in these mycobacterial aerosol studies, the samples were not fully desiccated; aerosols were produced and maintained in a rotating drum, at a high relative humidity (74%); Hence, the rotating drum prevented settling and humidity reduced the likelihood of dehydration (Lever et al., 2000). Louden et al. (1969) demonstrated 50% bacterial viability after 6h in aerosol, while Lever et al. (2000) reported 10% survival after 30 minutes. This study and the aerosol studies are thus not directly comparable.

While the results are not directly comparable to previous studies, it is notable that the experimental design may still be improved upon. The choice to desiccate *Mtb* on glass coverslips may differ from the natural desiccation of aerosols *ex vivo*, including the cell density, timescale of the experiment and the drying medium.

The cell density of *Mtb* culture on each coverslip was much higher than what would be experienced by hosts in the environment. However, clinical studies using guinea pigs, have demonstrated that infection may result from a single, aerosolised bacillus. In these studies, the presence of surrounding cells may influence the mycobacterial response to desiccation. For example, cells may undergo altered cell signalling, or the increased cell density may offer a protective effect against drying. The experimental protocol must be refined to account for these differences.

The time required for the cultures to dry was approximately 2 hours; this timescale occurred due to the 100µl volume of culture that was desiccated on each coverslip. The 100µl volume of culture was required to maintain sensitivity of the experiment following the resuspension. As the majority of aerosols generated by clinical tuberculosis patients ranged from 0.65 to 4.7µm in diameter, the desiccation occurs over a timespan of seconds or minutes, as opposed to hours (Fennelly et al., 2004). The response by bacteria to osmotic stress is considered a consequence of the initial stages of air drying, when water is sufficient to allow some bacterial growth and adaptation (Billi and Potts, 2002). However, once the culture or aerosol has dried fully, the desiccation process is ongoing up until the point of rehydration. Hence, the overall impact of this increased drying time on the survival of *Mtb* remains unknown.

Further, the drying took place in PBS solution, as Sautons and 7H9 media were inappropriate for use in the drying medium. This use of PBS introduced a further starvation stress upon the bacilli. The imposition of this stress alters *Mtb* gene expression and may alter the survival of the desiccated bacilli. However, as introduced in 5.1.1, research on the actual composition of mycobacterial aerosols is limited. The liquid that the bacilli are suspended in may originate

from a variety of sources, most likely the airway surface liquid, which includes mucus, surfactant and dissolved solutes (Fiegel et al., 2006). There may be some contribution of saliva to the aerosol medium, and artificial saliva was considered as a possible desiccation medium. However, the use of PBS may be appropriate until more research is performed on the aerosol liquid medium.

Finally, the humidity and temperature were not controlled in the experiment. These factors may account for the variations in survival to desiccation, as experiments were performed on different days. However, these variations are unlikely to affect the overall study, as the survival of two samples was directly compared; both samples are simultaneously exposed to the prevailing atmospheric conditions.

Overall, the results presented from the desiccation experiments suggest that an increase in LBs alone does not impart desiccation resistance on *Mtb* strains. Furthermore, this production may in fact have a detrimental effect on the overall survival of bacilli. If LBs do play a role in survival to desiccation, a global response involving a number of other genes/enzymes may be required.

5.4.4 Macrophage Binding

5.4.4.1 Overexpression of tgs1 does not alter binding to macrophages.

The comparison of *Mtb* pSMT3-*tgs1* overexpressing strain to control demonstrated that there was no significant difference between the two strains in binding to THP-1 macrophages. The results were similar in 3 different experiments at two MOIs. LB analysis was performed and confirmed a higher frequency of LB-positive cells in the pSMT3-*tgs1* overexpressing strain, as compared to control, suggesting the accumulation of triacylglycerol in these samples. Fluorescence analysis also confirmed higher median fluorescence intensity than the control strain, suggesting a higher neutral lipid content.

Thus, while Cáceres et al. (2009) suggested that *Mtb* may accumulate triacylglycerol to ensure there is the high intracellular fatty acid concentration required for the synthesis of cell wall components, the upregulation of *tgs1* alone in *Mtb* does not appear to alter the binding of bacilli to macrophages. In a similar scenario as described with the desiccation experiments, the fatty acids in triacylglycerol may be unavailable without a number of other enzymes; it is likely that there is a global gene expression change during the transmission of tuberculosis.

A number of factors could be improved upon in these experiments. Firstly, non-frozen inocula were used for the main experiments, as in preliminary experiments, previously frozen cultures were inappropriate for use. However, the use of non-frozen inocula was associated with other difficulties, in particular, with enumeration of cell density. As the enumeration of *Mtb* via CFU counts take 2-3 weeks, the inocula were estimated based upon the culture OD. This led to the macrophages being infected at an MOI higher than expected; MOIs were estimated to be 1 and 10, with the resultant MOI being approximate 3-5 and 60, respectively. However, the MOIs for each strain were comparable.

In addition, the percentage of bacilli binding in these experiments appears to be lower than in previous studies (Li et al., 2002). Li et al. (2002) demonstrated a wide range in the percentage of bacilli binding/uptake, with variation following exposure of *Mtb* to pH changes, anaerobiosis and temperature. In media at pH 7.2 and 37°C, they showed binding of approximately 18 ± 6% of bacilli. This discrepancy between the binding percentages from this study and Li et al. (2002) may be due to experimental variation.

This study made use of a THP-1 monocyte cell line (differentiated to macrophages) as opposed to monocyte-derived macrophages (MDMs) (Li et al., 2002). MDMs are human-derived, and may vary from patient to patient. Furthermore, as discussed, MOIs in this study were higher than expected; Li et al. (2002) used MOIs of 1 and 10 for the binding/uptake of *Mtb* by MDMs.

Another method that may be appropriate for examination of the binding of *Mtb to* macrophages is employed by other studies; these studies use microscopy as opposed to CFU counting (Stokes and Doxsee, 1999). Binding was assessed by counting the percentage of macrophages binding at least one bacillus and also the total number of bacilli associated with 100 macrophages (Stokes and Doxsee, 1999).

5.4.4.2 Macrophage binding is significantly higher for M. tuberculosis cultures exposed to multiple stresses.

The binding of the *Mtb* multi-stimulus cultures (see Chapter 4) was compared to control cultures. The RPMI-1640 multi-stimulus cultures were chosen over the equivalent cultures in PBS. While the correlation of the PBS gene expression to the sputum transcriptome was slightly higher than RPMI-1640, both were comparable, with similar R values. Notably, the binding of the *Mtb* multi-stimulus culture was significantly higher than the control culture. The experiments were performed at two MOIs, and binding was approximately 10-fold greater for both MOIs.

While the results are promising, one cannot draw many conclusions from the increased binding shown by the multi-stimulus *Mtb* culture. First of all, the analysis of the gene expression in Chapter 4, while correlating significantly, did not match that of the original sputum transcriptome. In addition to this, the exact changes responsible for this increased binding are unknown; it is unknown if the increased macrophage binding is due to a single factor, or a number of changes. However, analysis of these changes leading to this increased binding is beyond the scope of this study.

It is also difficult to comment on the LB and fluorescence analysis. There was large variation in frequency of LB-positive cells and the median fluorescence intensity in the multi-stimulus cultures using either PBS or RPMI and was discussed previously in Chapter 4. The percentage of LB-positive cells was higher in the multi-stimulus cultures as compared to the controls. However, the opposite was true for the median fluorescence intensity, and this was found to be significantly higher in the control samples. This was unexpected, as both nitric oxide and static incubation induce the dosR regulon (including *tgs1*) and may lead to triacylglycerol production (Bacon et al., 2004, Voskuil et al., 2003). The reasons for this staining pattern have been discussed previously in Chapter 4 (4.3.4.2).

It is worth noting that while the multi-stimulus cultures had both increased macrophage binding and an increased percentage of LB-positive cells versus control, it is unlikely that the LBs alone are the main factor leading to this increased binding. As discussed for the *tgs1* overexpressing strain results, the induction of LB production did not alter the binding of *Mtb* to macrophages. Furthermore, the multi-stimulus cultures had a host of other transcriptional changes in addition to the induction of *tgs1* and production of LBs.

The problem of MOI estimation was also a feature of this experiment. The MOI for the control was approximately twice as high as for the cultures in RPMI-1640 (plus stimuli). As a larger number of bacilli was used to infect the macrophages for the control, this may falsely lower the percentage of bacilli that were bound to macrophages; however, this does not account for the binding percentage of the cultures in RPMI-1640 (plus stimuli) being approximately 10-fold greater than the control. Furthermore, examination of the absolute number of bacilli bound to macrophages was significantly higher for the RPMI-1640 (plus stimuli) culture as compared to the control sample at the lower MOI ($p=0.0023$, unpaired t-test) and the higher MOI ($p<0.0001$, unpaired t-test). While the overall impact of the difference in MOI between the samples appears to be limited, the experiment should be repeated to confirm the results. In addition to the problems estimating MOI, there was lower *Mtb* binding than in previous studies (Li et al., 2002) and were discussed above.

While there were differences in the experimental procedures and the percentage of bacilli bound by macrophages, the results from Li et al. (2002) lend support the theory that bacilli in sputum are transmission adapted. The authors performed experiments to determine if environmental conditions thought to be encountered in the lung cavity (including hyperosmolarity, acidic pH and low oxygen tension) would influence the binding/uptake of *Mtb* by macrophages (Li et al., 2002). While hyperosmolarity did not affect *Mtb* binding/uptake, both acidic pH and anaerobiosis increased binding/uptake by four- to six-fold as compared to control (Li et al., 2002). The results by Li et al. (2002) demonstrate that host conditions encountered *in vivo* by *Mtb* can lead to changes that affect macrophage binding. These same conditions may apply to *Mtb* in expectorated sputum.

Finally, one of the aims of more closely replicating the sputum transcriptome *in vitro* was to use these cultures for further experiments, such as the assessment of macrophage binding above. Ideally, bacilli extracted directly from AFB-positive sputum would be used; this was attempted, but results were not obtained due to contamination.

Contamination is a problem affecting any study involving sputum, due to the normal flora present in the respiratory tract (Farzan, 1990). The addition of PANTA antibiotic mixture helps, but may not completely solve this contamination problem. In addition to contamination, with actual smear positive sputum samples, it is difficult to accurately quantify the number of viable bacilli/ml of sputum without CFU counts, a process that takes 2-3 weeks. Thus, the volume of decontaminated sputum required to infect macrophages at an adequate MOI is a major concern with regards to future experiments. These problems encountered in the experimental design must be rectified before it is repeated.

Looking at both sets of macrophage experiments, it appears that as for resistance to UV light and desiccation, the production of LBs in *Mtb* alone does not impact on the binding of bacilli to macrophages.

5.5 Conclusions

The relationship between lipid bodies and the survival of *Mtb* to transmission-related stresses and binding to macrophages was examined in this chapter. The specific findings of the studies in this chapter are as follows:

- The production of lipid bodies by *Mtb* occurs naturally during *in vitro* growth and was observed in all *Mtb* H37Rv samples grown in 7H9 medium and the percentage of lipid body positive cells in these samples ranged from approximately 10% to 25%.
- The production of lipid bodies in *Mtb,* using two independent methods showed no significant difference in survival to the control cultures against both desiccation and UV light stress. The exposure of cultures to nitric oxide may actually result in significantly decreased survival of bacilli to desiccation, as demonstrated in 2 of 3 biological replicates. Furthermore, a comparison of the *Mtb* pSMT3-*tgs1* overexpressing strain and control showed no significant difference in the binding of bacilli to THP-1 human macrophages.
- There was no correlation between the production of lipid bodies and survival of bacilli in either the UV or desiccation stress experiments.
- The *Mtb* RPMI multi-stimulus (30ml tubes) cultures (see Chapter 4) had significantly greater binding of bacilli to THP-1 human macrophages, as compared to control cultures.

Taken broadly, no evidence has been found to suggest that LBs possess enhanced transmission adaptation in the test systems studied here.

Chapter 6

General Discussion, Future Work and Conclusions

6.1 General Discussion

Insights on *M. tuberculosis* (*Mtb*) in expectorated sputum are a product of recent research. Studies by Garton et al. demonstrated a unique *Mtb* phenotype in sputum, with notable characteristics including slowly or non-replicating bacilli and the ubiquitous presence of LB-positive cells (Garton et al., 2002, Garton et al., 2008). Based upon these results, the authors postulated that bacilli in sputum display traits that underpin the transmission of tuberculosis. Concurrent studies have also elicited the molecular basis of triacylglycerol synthesis and growth conditions inducing the production of LBs in mycobacteria. However, studies have not directly examined the relationship of LBs to transmission of disease. In light of the findings presented in the literature prior to the commencement of this study, the research presented in this thesis was conducted with the aim of replicating the *Mtb* sputum transcriptome in an *in vitro* setting and examining the significance of LBs with regard to aspects of transmission.

6.1.1 What is the best housekeeping gene for *M. tuberculosis*?

In Chapter 3, the development of a qRT-PCR assay for *Mtb* was described. Through the course of this study, it was noted that *sigA* and 16S, two conventionally used housekeeping genes in the literature, were deemed to have among the lowest stability of the 11 candidates. The results demonstrated the use of four housekeeping genes, *thyA*, *dfrA*, *aroA*, and *polA,* in the form of a normalising factor (NF4) for the normalisation of target gene expression. While it was concluded that NF4 was the most appropriate normalisation method and was used for subsequent experiments, the reasons behind *sigA* and 16S being lower stability housekeeping genes were not discussed and may arise from their function.

SigA is an essential principal sigma factor which is a bacterial transcription initiation factor that enables specific binding of RNA polymerase to gene promoters (Wu et al., 2004). While principal sigma factors primarily regulate housekeeping genes, some regulate virulence factors and may be induced in response to stress (Wu et al., 2004). Interestingly, *sigA* overexpression enhances growth of *Mtb* in macrophages and murine lungs, while point mutations within the *sigA* gene lead to marked attenuation in virulence in guinea pigs (Wu et al., 2004). Hence, it appears that the sigA protein may have some role to play in the transcription of virulence factors (Wu et al., 2004). On the other hand, 16S rRNA is a component of the 30S ribosomal subunit of prokaryotic ribosomes, and thus is involved in the translation of proteins (Schaechter et al., 2006). The 16S subunit has a structural role, and also binds upstream to the AUG start codon on mRNA (Schaechter et al., 2006). As both *sigA* and

16S are constitutively expressed, they have long been considered good housekeeping genes for the normalisation of gene expression.

However, as briefly alluded to, the low stability displayed by these two genes in the study present here may arise from these very functions. Under optimal growth conditions, as displayed in normal, *in vitro* aerobic growth conditions, *sigA* is likely to be expressed highly. Furthermore, the *sigA* transcript has been demonstrated to be very stable, with little degradation 40 minutes after cell treatment with rifampicin (Hu and Coates, 1999). Hence, both these factors under optimal conditions may lead to higher *sigA* expression. As the cell reaches stationary phase, Hu and Coates (1999) reported that *sigA* transcription did not change. However, other experimental studies demonstrated a contradictory result and as cells reach stationary phase or bacterial growth is inhibited, the ratio of *sigA* to 16S declined (Timm et al., 2003). This suggests that either *sigA* is expressed in lower quantities during this period, or 16S expression increases. In the case of *sigA* and its regulation of basic cellular function, this reduction in optimal growth may result in its decreased regulation. What is evident is that the expression levels of one or both of these genes is changing, when in theory, as both have been used as housekeeping genes, the ratio should remain constant throughout all growth conditions or stresses.

16S had previously been thought to be constitutively regulated as it is expressed in all growth conditions for the biogenesis of new ribosomes (Thorrez et al., 2008). However, as with *sigA*, ribosomal RNA, including 16S, has been shown to be differentially regulated under different growth conditions (Thorrez et al., 2008). Furthermore, rRNA may have different regulatory mechanisms as compared to mRNA, further increasing variability and decreasing stability in expression.

Taken together, *sigA* and 16S have differing functions, both of which may lead to variable expression under different growth conditions and underlie the low stability demonstrated in Chapter 3. Through the use of a single housekeeping gene for the normalisation of gene expression data, differential expression of the housekeeping gene is translated directly into variation in the target gene quantities. As the purpose of normalisation is to control for variations in starting genetic material and to reduce the effect of experimental error by standardising variations in qPCR cycling performance and enzyme efficiencies, this differential expression of *sigA* and 16S is inappropriate. In contrast, NF4 uses a combination of the most stable housekeeping genes, helping correct for this variability (Vandesompele et al., 2002). Thus, the use of a single housekeeping gene for the normalisation of *Mtb* studies; in fact, the best genes for the normalisation of *Mtb* gene expression may vary from study to study.

6.1.2 What stresses underlie the nature of *M. tuberculosis* bacilli in sputum?

The nature of the bacillus in sputum is undoubtedly complex and poorly understood. In this study, nitric oxide, static incubation, cholesterol and oleic acid were identified as potential stresses contributing to the sputum transcriptome. Through the gene expression studies, a number of stimuli were tested and analysed. The studies were designed to include a broad overview of the stimuli/stresses that may be encountered by *Mtb*. Obviously a limitless number of stimulus combinations could be included; however, the sputum gene expression was used to select the relevant possibilities. In this process, a number of major stresses were excluded: phosphate limitation, ROS and acid stress. These stresses still represent potential stimuli inducing gene expression and are discussed here in addition to other potential cell environments underlying the *Mtb* sputum transcriptome.

First, as previously introduced in Chapter 1, *Mtb* may encounter a phosphate-limited environment with the macrophage phagosome; Rifat et al. (2009) listed a set of 92 upregulated and 58 downregulated genes in response to 24 and 72h of phosphate starvation. In sputum, only 6 of upregulated genes and 8 of the downregulated genes were similarly regulated, respectively (Garton et al., 2008). More importantly, these similarly regulated genes did not include the phosphate-specific transport operon *pstS3-pstC2-pstA1* or the *SenX3-RegX3* regulatory system, which has been implicated in the control of this operon (Parish et al., 2003b, Rifat et al., 2009). This suggested that phosphate limitation was not a major influence on the sputum gene expression.

Next, the use of ROS was also excluded from the study. At the time this study was performed, Jang et al. (2009) had identified the induction of iron acquisition genes in response to ROS (in *M. bovis* BCG); however, a similar induction was not found in sputum. Furthermore, the protective SOD genes (*sodA* and *sodC*), catalase gene *katG,* and the *mel2* locus (encoding luciferase-like genes) were all not upregulated in sputum. Thus, the impact of ROS on the sputum transcriptome was deemed to be limited.

Third, acidic pH is a stress likely encountered by *Mtb in vivo*. However, of the 81 genes identified by Fisher et al. (2002) to be upregulated in response to acidic conditions, only 5 were similarly induced in sputum. On the contrary, *virS* expression was induced in sputum and encodes for a protein essential for the induction of the *mymA* operon in *Mtb*. The *mymA* operon is upregulated under acidic conditions and within macrophages, and may encode proteins involved in the modification of cell envelope fatty acids (Singh et al., 2003). As it was unknown if acidic conditions were the cause of the *virS* upregulation in sputum, and other genes associated with growth in acidic conditions were not upregulated, this stress was excluded from this study.

These three stresses discussed above were excluded on the basis of gene expression from previous studies. However, it is impossible to fully exclude their contribution towards the gene expression in sputum, as experimental conditions may vary considerably from *in vivo* conditions.

Overall, it may be too simplistic to examine individual growth conditions. There were some similarities between the sputum transcriptome and the gene expression of *Mtb* within macrophages (Schnappinger et al., 2003). The combination of dosR regulon activation, utilisation of lipids and the slow- or non-growth signature in sputum were similar to the previously published experimental results from these macrophage infections (Garton et al., 2008). As such, there may be a significant intracellular component to *Mtb* bacilli in sputum that may not be seen on microscopy due to the destruction of cells during the sputum decontamination process and as staining procedures were aimed specifically at identifying AFBs. A number of potential *in vivo* intracellular environments have been proposed during the course of this study.

Two studies by Kolattukudy et al. demonstrated an *in vitro Mtb* dormancy model and consecutively, a hypoxic macrophage dormancy model, which may aid the understanding of *Mtb* in sputum (Daniel et al., 2011, Deb et al., 2009). The *in vitro* dormancy model made use of stresses thought to be encountered *in vivo*, including low oxygen (5%), high CO_2 (10%), reduced nutrient medium (10% Dubos medium) and acidic pH (5.0) (Deb et al., 2009). These conditions led to accumulation of intracellular lipid, loss of acid-fastness and expression of lipid biosynthetic and utilisation and dormancy associated genes in *Mtb*, including *dosR, hspX, tgs1* and *icl1* (Deb et al., 2009). This dormancy model displays certain characteristics similar to bacilli in sputum, but further analysis of the microarray data is required to compare this dormancy model to sputum expression.

In the second study, the authors showed that *Mtb* adopted a similar phenotype to their *in vitro* dormancy model while within hypoxic, lipid-loaded macrophages (Daniel et al., 2011). As with the dormancy model, bacilli accumulated LBs and upregulated lipid and dormancy genes, including *dosR, hspX, icl1* and *tgs1*. (Daniel et al., 2011). This adaptation towards a persistence phenotype in lipid loaded cells had previously been identified in foamy macrophages. In the proposed model of latent tuberculosis (Chapter 1), foamy macrophages have been shown to provide a key role in sustaining persistent bacilli and contributing to tissue pathology that may cause cavitation and the release and transmission of bacilli (Peyron et al., 2008). Peyron et al. (2008) demonstrated that *Mtb* engulfed by foamy macrophages acquired a dormant phenotype which included the accumulation of host lipid into LBs. A similar persistent state and LB formation is also seen in *Mtb* in adipose cells, although the

adipose tissue environment is more relevant to extra-pulmonary tuberculosis (Neyrolles et al., 2006).

Mtb has also been demonstrated to interact with other cells, including neutrophils, human respiratory cells and fibroblasts, although the contribution of the bacilli contained in these cells to sputum is unknown (Hernandez-Pando et al., 2000). Eum et al. (2010) demonstrated that neutrophils, and not macrophages, may be the predominant host cell type infected with *Mtb* in clinical tuberculosis patients. Neutrophils were more numerous than macrophages in sputum, bronchoaveolar lavage and cavity contents (Eum et al., 2010). Also, more intracellular bacilli were found in neutrophils than macrophages. These results alter the conventional models of pathogenesis in which macrophages are the predominant phagocytic cell, and it is unknown how these interactions influence the gene expression and phenotype of *Mtb*, as well as the progression of disease.

These intracellular environments present alternatives to the macrophage and will undoubtedly lead to variations in *Mtb* gene expression, due to the advent of multiple host stresses upon the phagocytosed bacilli. However, at this time, there is limited knowledge as to the gene expression within these alternative environments. In addition, it is possible that the sputum transcriptome may actually consist of the gene expression from a combination of these cellular environments.

In addition to growth conditions, as introduced in Chapter 4, toxin-antitoxin (TA) systems are prominent in *Mtb* and may play a prominent role in the sputum transcriptome. The unstable antitoxin component acts as a transcriptional repressor to the stable toxin; environmental stresses inactivating the antitoxin lead to increased cellular toxin concentrations (Ramage et al., 2009). Systems are grouped based upon their function; in *Mtb*, the majority of TA systems fall into 3 groups: *vapBC, mazEF* and *relBE* (Ramage et al., 2009). Each of the toxins in these systems, VapC, MazF and RelE have been shown to act as ribonucleases and inhibit translation (Gupta, 2009). For example, RelE cleaves translating mRNA in a sequence specific manner, with a preference for stop codons (UAG>UAA>UGA), codons adjacent to the start codon and codons with a G or C in the third position (Korch et al., 2009). However, these codons are found extensively throughout both the upregulated and downregulated target genes analysed in this study. This suggests that in addition to the ribonuclease activity, a number of other variables must contribute to the inhibition of specific transcripts to attain differential gene expression in *Mtb*.

With regards to the sputum transcriptome, the significant upregulation or downregulation of both toxin and antitoxin pairs was not noted, although, as reviewed in 4.1.2, a putative toxin was found to be upregulated and one antitoxin and four toxins were found to be significantly

downregulated (Garton et al., 2008). The growth conditions eliciting the differential expression of these toxins or antitoxins remain un-elicited following this study. Excitingly, the recent analysis of the Rv1102c-Rv1103c TA system, the antitoxin (Rv1103c) of which was downregulated in sputum, showed that increased regulation of Rv1102c (toxin) in *M. smegmatis* induced a growth arrest and increased the number of persister cells (Han et al., 2010). As to the conditions generating differential TA systems expression, only systems belonging to the *RelBE* family have thus far been demonstrated to be expressed in broth grown *Mtb* and during infection of human macrophages (Korch et al., 2009).

6.1.3 What role do Lipid Bodies play in *M. tuberculosis* transmission?

An aim of this study was to examine the impact of the presence of LBs in *Mtb* on transmission of disease. As both gene expression changes and LB production were characteristic of the proposed transmission adapted *Mtb* sputum phenotype, it was useful to identify if LBs played any role in adaptation to transmission, or were merely co-existing phenomenon, secondary to the phenotype. From the results outlined in Chapter 5, however, there appeared to be no link between the production of LBs and the survival to desiccation and UV radiation or to macrophage binding.

Based upon such results, it would be reasonable to conclude that LBs are merely an epiphenomenon, resulting from the gene expression or phenotype of the *Mtb* bacillus. This fits with the proposed model of LB formation introduced in Chapter 1 which suggests that LBs are produced as a result of a reduction in mycobacterial growth rate. As discussed in 1.4.1.1, foamy macrophage and adipose cells have been proposed as possible environments of dormant bacilli *in vivo*. In both cellular environments, *Mtb* acquires a non-replicating persister-like state and the accumulation of intracellular lipid (Neyrolles et al., 2006, Peyron et al., 2008). Again, the studies by Kolattukudy et al. demonstrated that a multiple-stress dormancy model accumulated lipid and *Mtb* in lipid-loaded macrophages became dormant and utilised host triacylglycerol to form LBs (Daniel et al., 2011, Deb et al., 2009). These four studies add further support to this LB formation model.

As discussed, the poor survival to transmission stresses displayed by the LB producing culture may stem from the fact that a number of genes may be required in addition to *tgs1*. Genes that may be required for survival to desiccation and UV radiation were highlighted in Chapter 5.

At the same time as this study, Ag85A was described by Elamin et al. (2011). TDM has been shown to confer desiccation resistance in model mycobacterial membranes (Harland et al.,

2008). Enzymes of the antigen 85 complex, Ag85A, B, and C, exclusively synthesise TDM; the enzymes are acyltransferases that convert TMM (trehalose 6'-monomycolate) to TDM (Takayama et al., 2005). However, Elamin et al. revealed that Ag85A (*fbpA*/Rv3804c) was a novel *Mtb* diacylglycerol acyltransferase not previously presented amongst the 15 *tgs1* genes (Daniel et al., 2004, Elamin et al., 2011). Induction of Ag85A in *Mtb* led to the accumulation of triacylglycerol and LB production. Further, the expression of Rv3804c in *M. smegmatis* led to an increase in the concentration of glycolipids TDM and TMM in the cell (Elamin et al., 2011). Hence, the results suggest that Ag85A has the ability to modify the synthesis of triacylglycerol whilst maintaining acyltransferase function (Elamin et al., 2011). These results are intriguing and suggest that there may be a link in the production of LBs, triacylglycerol and TDM in *Mtb* (Elamin et al., 2011).

Although not analysed, it was unlikely that induction of Ag85A to produce trehalose dimycolate was a feature of the *tgs1* overexpressing strain, based on the gene expression of the 20 analysed target genes (Appendix 3). Furthermore, two separate *Mtb* gene expression studies following nitric oxide exposure did not reveal upregulation of Ag85A (Ohno et al., 2003, Voskuil et al., 2003). It is also notable that Ag85A is not upregulated in the expression of sputum gene expression (Garton et al., 2008). Thus, the contribution of Ag85A to the production of triacylglycerol and lipid bodies, and the conditions that lead to its regulation, remains to be defined.

Furthermore, the expression of lipases is required for the liberation of fatty acids from triacylglycerol and these fatty acids may subsequently be used in other metabolic functions *in vivo*. LipY (Paragraph 1.7.2) has already been identified as the main lipase upregulated in *Mtb* during nutrient starvation (Deb et al., 2006). LipY also possesses a PE domain, suggesting that the lipase may interact with the immune system, as other PE proteins may have a role in immune evasion (Mishra et al., 2008). Thus far, this has been the first and only study to demonstrate the hydrolysis of triacylglycerol through the action of a *Mtb* lipase. LipR, however, has been proposed to have a role in tuberculosis transmission and pathogenesis as, in addition to hydrolysis of triacylglycerol, lipases may also modify the cell envelope (Sheline et al., 2009). Epidemiological studies have demonstrated that large sequence polymorphisms (LSPs) interrupting *LipR* were more likely to be linked to epidemiological clusters (Sheline et al., 2009). Epidemiological clusters are more commonly seen between friends/relatives than more distant contacts. Strains with impaired tuberculosis transmission require more extensive contact between the source and host individuals to achieve transmission; therefore, these strains are more likely to form epidemiologically linked clusters than wild-type strains. These findings suggest that *LipR* may be associated with transmission of tuberculosis but must still be confirmed experimentally (Sheline et al., 2009).

Both Ag85A and lipases suggest that alternative genes could link cell lipid content and transmission, but require still require further study. Also, despite the recent evidence in the literature describing LB production by non-replicating *Mtb* bacilli in a variety of environments, a direct link between LBs and transmission does not appear to exist. If the *Mtb* bacilli in sputum are in fact transmission adapted, any advantage conferred upon these bacilli is not due solely to the formation of LBs. Instead, LBs are more likely to appear as a phenomenon secondary to the phenotype actualised by *Mtb* in sputum.

6.1.4 What role do Resuscitation Promoting Factors (Rpfs) play?

An aspect of *Mtb* that may have an impact on the results presented in this study is the role of resuscitation promoting factors (Rpfs). Rpfs are a family of proteins, secreted by *Mtb*, that stimulate mycobacterial growth (Mukamolova et al., 2010). Cells that are dependent on Rpf are unculturable via normal techniques, and cannot resume normal growth until they are resuscitated.

Mukamolova et al. (2010) demonstrated that in 20 of 25 AFB-smear positive sputum samples tested, 80-99.99% of the *Mtb* bacilli belonged to this Rpf-dependent group. Of the remaining 5 samples, 2 had no response to Rpf, and 3 samples had 65.63-66.70% Rpf-dependent cells (Mukamolova et al., 2010). Thus, in 23 out of 25 samples, Rpf-dependent cells comprised the majority of the population.

The body of evidence supporting the importance of Rpf *in vivo* is growing. Rpfs were first demonstrated in *Micrococcus luteus*. *M. luteus* can enter a dormant state after prolonged storage and starvation and these cultures lose the ability to grow on agar plates. It was demonstrated that these dormant cells could be resuscitated from dormancy to re-establish growth through exposure to non-dormant *M. luteus* (Votyakova et al., 1994). The effect was shown to be due to the secretion of a Rpf protein (Mukamolova et al., 1998, Mukamolova et al., 1999). A study in *M. smegmatis* showed that suboptimal growth conditions, attained through alterations in growth temperature, medium and oxygen tension, were required for cells to enter a conditionally non-culturable state requiring Rpf resuscitation (Shleeva et al., 2004). The state may represent an adaptive response of bacteria in environmental conditions that may compromise normal metabolism.

Mtb has been shown to have five *rpf* gene homologues, encoding proteins RpfA-RpfE with very similar characteristics and properties of *M. luteus* Rpf. These *Mtb* Rpfs have been shown to promote the growth of *M. luteus, M. smegmatis, and M. bovis* (Mukamolova et al., 2002). The deletion of any of the single *rpf* genes from *Mtb* has no effect on growth *in vitro* and *in*

vivo in acutely infected mice and global gene expression profiling suggests that the *rpf* genes in *Mtb* serve wholly or partially overlapping functions (Downing et al., 2004, Tufariello et al., 2004). Despite the redundancy of the five *rpf* genes during normal *in vitro* and *in vivo* growth, Tufariello et al. (2006) demonstrated that RpfC (*rpfC*/Rv1009) displayed delayed reactivation from chronic infection in a murine model of persistence, while having no growth defects *in vitro* or during *in vivo* acute infection. Other studies examined the effect of double and triple deletion mutants. Russell-Goldman et al. (2008) showed that a Δ*rpfAB* knockout exhibited impaired growth in primary mouse macrophages, induced higher levels of proinflammatory cytokines and displayed defects in chronic tuberculosis rectivation. Biketov et al. (2007) noted that *rpf* double and triple mutants were attenuated in their ability to disseminate to mouse lungs after intraperitoneal infection and were defective in their ability to re-grow after immunosuppression. Furthermore, inactivation of three genes generated mutants that could not resuscitate in culture *in vitro* and were significantly attenuated in mouse infections (Downing et al., 2005). Most recently, Kana et al. (2008) found that deletion of all 5 *rpf* genes from *Mtb* also results in failure of resuscitation *in vitro*, delayed colony formation, hypersensitivity to detergents and attenuation in mice.

Several studies have also examined the expression of *rpf* genes. There is evidence that mycobacterial *rpf* genes are expressed *in vivo* in the lungs of mice acutely infected with virulent *Mtb* (Tufariello et al., 2004). In addition, Rpf protein expression can also be detected in human tissues infected with *Mtb*. Davies et al. (2008) detected Rpf using anti-Rpf antibodies in the immediate vicinity of the AFBs in the necrotic centres of granulomas, as well as in the giant cells and macrophages surrounding the necrotic areas.

The function of Rpf proteins may be related to their muralytic activity, with its biological function resulting from direct or indirect cleaving of bacterial peptidoglycan bonds. A recent study has demonstrated that mycobacterial RpfB interacts with a peptidoglycan hydrolase (Hett et al., 2007). Peptidoglycan hydrolases are involved in the final stages of cell division, in which the peptdoglycan layer which connects the daughter cells is cleaved. Deletion of the peptidoglycan hydrolase gene *ripA* in *M. smegmatis* led to a decrease in bacterial growth. Hence, Rpf may be involved in the regulation of the late stages of cell division, possibly in the withdrawal from a dormant state, although how peptidoglycan hydrolysis causes changes in growth of the bacteria is unknown (Hett et al., 2008, Keep et al., 2006).

As discussed in Chapter 1, a relationship has been developed between the production of LBs and the metabolic activity of *Mtb* bacilli. The production of LBs seems to occur as the rate of replication slows, as in non-replicating persistence and disease latency. The exact physiological and metabolic state of *Mtb* bacilli during this latent infection remains

controversial, but the observations described above strengthen the idea that bacterial dormancy is at some point involved during tuberculosis infection and that Rpf proteins appear to act together to promote resuscitation from this dormancy.

The fact that Rpf dependent cells comprise the majority of expectorated cells in sputum adds another aspect to the assertion that the characteristic phenotype of sputum reflects a transmission adaptation and represents a trait that was not examined in the experiments presented in this thesis. If cells in sputum require Rpf to become culturable, similar conditions may be required following exposure to the transmission related stresses, desiccation and UV radiation. Although a possible explanation for the low survival of the LB inducing cultures, this remains to be confirmed experimentally.

As to the impact of Rpfs on transmission, a few questions still remain. For instance, which cells secrete the Rpfs? As up to 99.99% in sputum required Rpf for resuscitation, the actively growing population might be limited and these replicating cells may be more prone injury or death from transmission stresses. Thus, of the cells reaching the host, the population of Rpf-dependent cells may dwarf the actively replicating proportion. It is unknown if these active cells are able to produce Rpf capable of resuscitating the dependent cells. Furthermore, the upregulation of *rpf* genes in sputum was not noted (Garton et al., 2008). As *Mtb* Rpf was able to activate other dormant bacteria *in vitro*, an intriguing, if less likely, prospect is if the presence of certain bacteria in the natural host microbiota produce factors capable of resuscitating transmitted *Mtb*.

Also, are other proteins required for the resuscitation of cells in sputum? The use of culture supernatant led to enhanced resuscitation of cells as compared to recombinant Rpf proteins (Shleeva et al., 2004). This opens the possibility that another component (e.g. protein or peptide) in addition to Rpf may aid revival of non-culturable cells. *Mtb* contains a variety of peptidoglycan hydrolases, and the influence of these alternative enzymes cannot be excluded. Finally, there's a question of what leads to the variation in Rpf-dependent cells in sputum. A large proportion of Rpf-dependent cells were demonstrated in 20 of the 25 samples examined by Mukamolova et al. (2010). However, no Rpf-dependent cells were demonstrated in 2 samples, and approximately 65% in the remaining 3 samples. It remains to be determined if this variation is due to *in vivo* variables, or due to a cell associated inhibitory factor noted by Mukamolova et al. (2010).

Altogether, the association between Rpfs, non-replicating persistence and bacilli in expectorated sputum further complicates the overall understanding of the sputum phenotype, and must be studied further.

6.2 Future Work

Through the work in this thesis, it has been possible to demonstrate *in vitro* cultures that display gene expression that correlates with the sputum transcriptome, and express a phenotype that leads to an increase in macrophage binding. The multi-stimulus culture work presented in Chapter 4 was established through a unique study, in an attempt to recreate the sputum transcriptome *in vitro* through the assessment of the gene expression of multiple growth conditions. Future work concerning this study may be used to further improve the study design and refine the results presented in this thesis. While the final version of the *in vitro* multi-stimulus cultures used in this study had gene expression that correlated significantly with the sputum transcriptome, closer examination revealed that there were still a number of differences between the two samples. As the experiment was only performed in biological duplicate, repeating the experiment would prove wise and allow for the expression of individual genes to be analysed statistically.

As briefly alluded to in 6.1.2, it appears that a number of alternative growth conditions may be required to extend the current study and fully replicate the transcriptome in sputum. This may be performed through a combination of approaches. Firstly, a greater understanding of the pathogenesis of tuberculosis is necessary. This includes the process of expectoration, i.e. the *in vivo* source of bacteria in sputum, the means by which these bacilli enter sputum and the final composition of tuberculous sputum. A further comprehension of how this process of expectoration relates to the production of aerosols, which transmit tuberculosis is also needed. In addition, research into the pathogenesis of tuberculosis will inevitably lead to further insights into the host stresses encountered by bacilli. Secondly, in addition to understanding the pathogenesis, determining the function of specific genes will be important. The sputum transcriptome featured 51 upregulated and 94 downregulated conserved hypothetical proteins. Comprehending the nature of these and other genes is fundamental to the transcriptome. Further to the gene expression analysis of a variety of growth conditions and stimuli as described in this study, a number of alternative stimuli were described in 6.1.2, and may be required to more closely replicate the transcriptome. Finally, a limitation of current study was the number of genes selected for analysis. After further refinement of the multi-stimulus cultures, the use of microarray techniques, or the analysis of a greater number of genes through RT-PCR may be performed to examine genome wide expression changes. Taken together, as both genomic changes and environmental factors contribute to phenotype, both aspects must be researched in more detail.

The direct examination of the characteristics of the bacilli is required to confirm that the gene expression analysed *in vitro* translates to a similar phenotype. These characteristics include

acid-fast staining, Rpf-dependency and antibiotic tolerance. As previously mentioned, the staining characteristics of *Mtb* bacilli may be altered during dormancy and latent infection. Furthermore, variations in the acid-fast staining (using the Ziehl-Neelson method) have been noted of cells in sputum and this is confirmed by observational evidence from this laboratory (Shapiro and Hanscheid, 2008). As Rpf-dependent cells dominate the population of bacilli in sputum, the multi-stimulus cultures could be assessed for their state of culturability (Mukamolova et al., 2010). Direct assessment of the replication of bacilli may also be used; this may be performed through the incorporation of radiolabelled thymidine to determine the rate of DNA synthesis and cell turnover or radiolabelled uracil to assess the rate of transcription. Rpf-dependent cells in sputum also displayed antibiotic tolerance to rifampicin and it would be appropriate to assess this characteristic (Mukamolova et al., 2010).

An experiment using *Mtb* added to sputum was suggested in Chapter 4. It remains possible that the *Mtb* phenotype and transcriptome results from the components found in sputum, rather than *in vivo* growth conditions. By spiking non-tuberculous sputum with laboratory strain *Mtb* H37Rv, this theory could be tested through the analysis of gene expression and, subsequently, exposure to transmission stresses.

It has been demonstrated that the production of LBs in *Mtb* confer neither resistance to transmission related stresses, nor an increase in binding to macrophages. However, as discussed in 6.1.4, the role of Rpfs in sputum may alter this finding. A rather simple experiment to assess the contribution of Rpfs to cell resuscitation following transmission stresses (desiccation or UV light), would be to compare CFU counts to both Most Probable Number (MPN) counts in both fresh 7H9 medium (control) and fresh culture supernatant in order to observe for any differences in the obtained counts. Differences between the CFU count and culture-supernatant MPN (but not the control) may indicate cell resuscitation.

As indicated, a number of other proteins may contribute to the survival of *Mtb* to transmission-related stresses, including Ag85A, lipY and ICL. For example, *icl1* (encoding ICL) has been shown to be upregulated in *Bradyrhizobium japonicum* in response to desiccation stress (Cytryn et al., 2007). Assessment of this and other lipid genes may elicit their role to these transmission stresses.

Lastly, the macrophage experiments revealed that the multi-stimulus culture had significantly higher macrophage binding as compared to control. However, as this was only performed with biological duplicates, a repeat experiment is necessary to confirm the results. A comparison of the binding once the multi-stimulus culture has been refined, in addition to the testing the macrophage binding of *Mtb* extracted from sputum is an alternative.

A logical progression from the macrophage binding/infectivity studies is the study of *Mtb* infectivity in an animal infection model. While the binding of *Mtb* to macrophages represents an initial step in the pathogenesis of tuberculosis, it does not represent the final proportion of bacilli that are able to establish an infection. The use of animal models allows for infection and disease progression to be analysed *in vivo*. Appropriate models for the study of bacterial virulence and infectivity have been established in the literature.

Williams et al. (2005) established an *in vivo* assay in guinea pigs to compare the infectivity of *Mtb* isolates via the aerosol infection route, in terms of their ability to cause a pulmonary infection and disseminate to the spleen. Guinea pigs are an ideal candidate for *in vivo* infection, as they reproduce key aspects of human disease, particularly pulmonary lesions that histologically resemble human tubercles and extra-pulmonary dissemination (Williams et al., 2005). Bishai et al. (1999) also demonstrated a *Mtb* rabbit inhalation model to assess the virulence of *Mtb* strains by Lurie's tubercle count method. As for guinea pigs, tuberculosis in rabbits closely resembles tuberculosis seen in immunocompetent humans. The assessment via Lurie's tubercle count method allows for the direct assessment of visible tubercles. A more virulent bacillus will be more resistant to destruction by alveolar macrophages and the host immune response, thus producing more grossly visible lesions (Bishai et al., 1999). Further assessment of virulence may be obtained through the sizes of pulmonary tubercles and the number of culturable bacteria from the lung.

Both of these animal models would be appropriate for future studies using the multi-stimulus cultures and would examine the virulence and infection process beyond the initial binding to alveolar macrophages.

6.3 Conclusions

The work presented in this thesis has fulfilled the main objectives of the study which were to examine the impact of lipid bodies in *Mtb* in relation to aspects of transmission and to investigate the *Mtb* sputum transcriptome in an *in vitro* setting.

The specific findings of each chapter have been summarised previously. The findings of this report are as follows:

- The best housekeeping gene for a given experiment may vary and the conventional housekeeping genes *sigA* and 16S were inappropriate for the normalisation of these studies. The variability in *sigA* and 16S gene expression may occur secondarily to their function. Alternatively, a normalising factor consisting of four genes was found to be more stable and used for the studies presented.
- Multiple stimuli underlie the transcriptome of *Mtb* in sputum, as no obvious single growth condition or stimulus fully replicates the sputum transcriptome in an *in vitro* setting.
- The gene expression of *Mtb* cultured in PBS/RPMI medium and exposed to four stimuli (nitric oxide, oleic acid, cholesterol and static incubation) correlated most accurately to the sputum transcriptome. However, the pattern of expression did not match exactly, and the exact combination of stimuli is yet to be fully elucidated.
- The induction of lipid body production in *Mtb* using two independent methods did not coincide with increased survival of bacteria to transmission related stresses and as such, lipid bodies represent an epiphenomenon of the *Mtb* sputum phenotype. The role of resuscitation promoting factors was not examined in these experiments, but recent evidence suggests that they may play a role in the transmission adaptation of *Mtb* in sputum.
- A proportion of *Mtb* bacilli in aerobically, exponentially growing culture contained lipid bodies, in contrast with previously documented studies. The presence of lipid body containing cells in these cultures demonstrated that lipid bodies play a role in optimal growth and again, represent a secondary characteristic to the *Mtb* sputum phenotype.

Mtb cultures exposed to RPMI and a combination of four stimuli (nitric oxide, oleic acid, cholesterol and static incubation) led to increased macrophage binding. This lends support to the hypothesis that *Mtb* in sputum are transmission adapted, although the mechanisms underlying this result remain to be determined.

Appendices

Appendix 1: Primers Used in this Study

Genome- Directed Primers for RT-PCR

Primer Name	Sequence	Primer Name	Sequence
mtGDP1	5′ - CGG CCA GC - 3′	mtGDP20	5′ - GAT CGG C - 3′
mtGDP2	5′ - CGG CGG CG - 3′	mtGDP21	5′ - CGC CGC G - 3′
mtGDP3	5′ - CGC CGC CG - 3′	mtGDP22	5′ - CGG TGG C - 3′
mtGDP4	5′ - CGT CGG CG - 3′	mtGDP23	5′ - CAC CGT C - 3′
mtGDP5	5′ - CGC CGG CG - 3′	mtGDP24	5′ - GCG GCC A - 3′
mtGDP6	5′ - CGG CCG CG - 3′	mtGDP25	5′ - CAC CGG C - 3′
mtGDP7	5′ - CGC CGT CG - 3′	mtGDP26	5′ - GCG GCC G - 3′
mtGDP8	5′ - CGG CGT CG - 3′	mtGDP27	5′ - CCG CGC C - 3′
mtGDP9	5′ - CGG CGA CC - 3′	mtGDP28	5′ - GCC CAG C - 3′
mtGDP10	5′ - CGG CGA TG - 3′	mtGDP29	5′ - TCG GCC A - 3′
mtGDP11	5′ - CGT CGT CG - 3′	mtGDP30	5′ - CCC GGC G - 3′
mtGDP12	5′ - GGC CGC CG - 3′	mtGDP31	5′ - CAC CAG C - 3′
mtGDP13	5′ - CGC CAC CG - 3′	mtGDP32	5′ - CAC CGC C - 3′
mtGDP14	5′ - GCA GCA GC - 3′	mtGDP33	5′ - CGG GCC G - 3′
mtGDP15	5′ - CGG TGC CG - 3′	mtGDP34	5′ - CGC TGA C - 3′
mtGDP16	5′ - CAC CGC G - 3′	mtGDP35	5′ - GGT GTT G - 3′
mtGDP17	5′ - GTC GCC G - 3′	mtGDP36	5′ - ACG CAG C - 3′
mtGDP18	5′ - GTC GAC G - 3′	mtGDP37	5′ - ACC GGA C - 3′
mtGDP19	5′ - CGC CAG C - 3′		

Primer sequences taken from Rachman et al. (2006).

SYBR Green PCR primers for selected Housekeeping genes

Gene Name	Primer	Primer Sequence	Amp (bp)	T_m (°C)
gyrB	Forward Reverse	5′ – CCG GGT GGC CTG GTG GAC TTC – 3′ 5′ – CCT CGT GGG TGT TGA TGG TGT TGG – 3′	178	64.2 63.9
lysA	Forward Reverse	5′ – CAC CGC GCA CGA GGA CCA GAA – 3′ 5′ – CGC GTA GCA GGC CGA TGA CAC – 3′	185	64.9 62.3
polA	Forward Reverse	5′ – GCA GCG CGG CCG ACA TCA TC – 3′ 5′ – CCG TAG CCC ACC GAC ACC TCC AG – 3′	214	65.3 65.3
fbpB	Forward Reverse	5′ – GGG CCG CGA CAT CAA GGT TCA GTT – 3′ 5′ – GCG CAG GCC GTC GAG CAG ATA AA – 3′	76	65.7 65.9
ftsZ	Forward Reverse	5′ – CCG AGG ACG CCA AGG ACG AGA T – 3′ 5′ – GGG CGA TGC TGG CGA CGA C – 3′	120	63.2 63.2
sigA	Forward Reverse	5′ – GAG ATC GGC CAG GTC TAC GGC GTG – 3′ 5′ – CTG ACA TGG GGG CCC GCT ACG TTG – 3′	160	65.1 66.0
dfrA	Forward Reverse	5′ – GGA GAT CAC CAT GGG GCA CAC GA – 3′ 5′ – GCT TGG CGG CTC AGT ACG ACA TTT – 3′	111	64.6 62.3
thyA	Forward Reverse	5′ – ATT CCC GGC GCA TCA TCG TGT C – 3′ 5′ – AGG TCG GCG CTG CGT TGG TAG A – 3′	139	64.5 64.5
aroA	Forward Reverse	5′ – GCT GTC CGC GGC ATC GTT CAC – 3′ 5′ – GCA GCA TCG CCG CCG TCA T – 3′	98	64.3 63.7
mtrA	Forward Reverse	5′ – CGT GGA TGT GGT GCT GGG TCT GG – 3′ 5′ – CTC GGC GGG TTC GTC GTC GTT – 3′	124	65.1 64.9
16S	Forward Reverse	5′ – GAA ACT GGG TCT AAT ACC – 3′ 5′ – ATC TCA GTC CCA GTG TGG – 3′ (Garton et al., 2008)	173	46.6 53.7

Acqusition Temperatures:

polA: 87°C
aroA: 86°C
All other primer sets: 82°C

SYBR Green PCR primers for selected *M. tuberculosis* genes

Gene Name	Primer	Primer Sequence	Amp (bp)	T_m (°C)
icl1	Forward Reverse	5′ – GCG GTG CGG AGG TGC TGT GG – 3′ 5′ – AGG CTC TGG TCG GGG TAG GTG – 3′	181	66.4 66.6
narK2	Forward Reverse	5′ – CGC CGC CGG AGG TGT GGT – 3′ 5′ – TGC CGC GGC GAC GAT TCC – 3′	143	64.2 66.5
mce3C	Forward Reverse	5′ – CGG GCG TGC TGG TGG AGT TCA – 3′ 5′ – AGG GCG TCG GGC AGT TGG TAG G – 3′	187	64.7 65.9
hspX	Forward Reverse	5′ – ACT CCG GCC CAC CTT CGA CA – 3′ 5′ – AGC ACC TAC CGG CAG CGA CA – 3′	235	63.6 64.0
ppsA	Forward Reverse	5′ – ATC ATC GCC AAC CGC CTC TCG T – 3′ 5′ – AAT ACC GCC GGG GAC AAC AAC AAA – 3′	176	63.2 61.4
virS	Forward Reverse	5′ – CCT CAT CCG CGC CAC CAA CC – 3′ 5′ – ACG AAC CCG GCC AGC GAC AT – 3′	144	63.8 64.5
tgs1	Forward Reverse	5′ – AAC GAA GAC CAG TTA TTC GAG – 3′ 5′ – CTC ATA CTT TCA TCG GAG AGC – 3′	197	55.0 54.9
Rv3180c	Forward Reverse	5′ – CGG CCC GCA ATC ACG ACC TAA C – 3′ 5′ – CAA TGC GCT GGC CAG ATG AAC G – 3′	179	62.8 61.5
ltp2	Forward Reverse	5′ – AGG CGG CCA TCG TCG GTA TCG – 3′ 5′ – CAT CGG CCA ACG CAT CCA ACA – 3′	102	64.0 60.9
Rv3551	Forward Reverse	5′ – GCT CGG CGG GCA AGG TCA AG – 3′ 5′ – CCG GCG CGA ATA GGC AGG AA – 3′	190	64.0 62.7
mce1A	Forward Reverse	5′ – GGG TGG AGA TCG GGC GGG TAG A – 3′ 5′ – TCG GGT TTT TCG GCG TGG TCA A – 3′	186	64.9 61.9
mmaA2	Forward Reverse	5′ – CGA GTG TTG CTG GCG GGA TGG – 3′ 5′ – CGG GCG GCA GGA TTT TGT GG – 3′	133	63.3 62.0
rpsL	Forward Reverse	CAGCTGGTCCGCAAGGGTCGTC GAGTTCGGCTTCTTCGGAGTGGTG	125	64.6 62.0
Rv1103c	Forward Reverse	5′ – CGC TTG CCC GAT GAG ATT GT – 3′ 5′ – CCA GCG CCC GCA GCA CGA C – 3′	91	58.1 67.4
atpD	Forward Reverse	5′ – GCC GGT CGG TGA GGG TGT GAA – 3′ 5′ – CGG CGG CTT GCG GTG AAT C – 3′	112	64.3 62.4
Rv2141c	Forward Reverse	5′ – TGG TGG GCA TGA TGA TTG TGG TTG – 3′ 5′ – TTG CCG CCG TGC TCC TCG TC – 3′	106	60.1 65.3
qcrC	Forward Reverse	5′ – CGG TGG CGA CCT TCA CGT AG – 3′ 5′ – GCG CCC GCC AAT GAA C – 3′	130	60.4 58.3
ctaD	Forward Reverse	5′ – GCG TCG CGT GCA TAA GCT TTT – 3′ 5′ – GCC GTG CAT GGT GAA CAA CTG – 3′	128	59.2 59.6

SYBR Green PCR primers for selected *M. tuberculosis* genes Continued

Gene Name	Primer	Primer Sequence	Amp (bp)	T_m (°C)
nuoB	Forward Reverse	5′ – CGA GAA GGT GGC GGG CTA TGT – 3′ 5′ – GAA CCG CGC AAT GTC AAA CCT – 3′	118	62.0 59.6
nuoL	Forward Reverse	5′ – GCT GCG CCA AGG ACG ACA TCA A – 3′ 5′ – CGC GCC GGA CCC AAG GAA TAG – 3′	167	62.6 62.9

Acqusition Temperatures:

All primer sets: 82°C

Appendix 2: Derivation of PCR Quantification Strategies

The derivation of the 'delta delta Ct' method starts again with the basic mathematical equation for PCR amplification as displayed in Equation 12.

Equation 12: PCR Amplification

$$N_C = N_0 \cdot (E+1)^C$$

A threshold cycle (*Ct*) is again calculated for the reactions; N_C becomes a constant N_t, where N_t is the threshold number of target molecules, N_0 is the initial number of target molecules, E_N is the amplification efficiency, Ct_N is the threshold constant for the target molecule, and K_N is a constant. This is displayed in Equation 13 (1). A similar equation for a housekeeping/reference gene is displayed in Equation 13 (2), where R has replaced N to signify the reference gene (Livak and Schmittgen, 2001).

Dividing N_t by R_t gives the expression displayed in Equation 13 (3).

Equation 13: Derivation of Delta Delta Ct – Part 1

$$N_t = N_0 \cdot (E_N + 1)^{C_{t,N}} = K_N \quad (1)$$

$$R_t = R_0 \cdot (E_R + 1)^{C_{t,R}} = K_R \quad (2)$$

$$\frac{N_t}{R_t} = \frac{N_0 \cdot (E_N+1)^{C_{t,N}}}{R_0 \cdot (E_R+1)^{C_{t,R}}} = \frac{K_N}{K_R} = K \quad (3)$$

Assuming that the amplification efficiencies of the target and reference gene are the same, $E_N = E_R = E$, and Equation 13 (3) can be rearranged as Equation 14 (1). This equation can be rearranged further into Equation 14 (2), where N_N is equal to the normalised amount of target (N_0/R_0) and ΔCt is the difference in threshold cycles for the target and reference ($Ct_N - Ct_R$). Equation 14(2) can be rearranged again to give Equation 14(3) (Livak and Schmittgen, 2001).

Equation 14: Derivation of Delta Delta Ct – Part 2

$$\frac{N_o}{R_0} \cdot (E+1)^{C_{t,N} - C_{t,R}} = K \quad \text{(A)}$$

$$N_N \cdot (E+1)^{\Delta C_t} = K \quad \text{(B)}$$

$$N_N = K \cdot (E+1)^{-\Delta C_t} \quad \text{(C)}$$

The final step in the derivation is to divide the N_N for any sample q by the N_N for the control sample (*cs*), as displayed in Equation 15 (Livak and Schmittgen, 2001).

Equation 15: Derivation of Delta Delta Ct – Part 3

$$\frac{N_{N,q}}{N_{N,cs}} = \frac{K \cdot (E+1)^{-\Delta C_{t,q}}}{K \cdot (E+1)^{-\Delta C_{t,cs}}} = (E+1)^{-\Delta\Delta C_t}$$

Where $-\Delta\Delta C_t = -(\Delta C_{t,q} - \Delta C_{t,cb})$

This method of relative expression therefore compares two samples as a ratio. It makes the assumption of perfect efficiency. In that scenario, $E = 1$ and the final equation is derived (Equation 16) (Livak and Schmittgen, 2001):

Equation 16: Delta Delta Ct Method of Relative Expression

$$Ratio = 2^{-[\Delta C_{t\,sample} - \Delta C_{t\,control}]}$$

$$Ratio = 2^{-\Delta\Delta C_t}$$

Appendix 3: *tgs1* Overexpressing Strain

Introduction

Triacylglycerol accumulation in *M. tuberculosis* has been observed in sputum samples from clinical patients and is described as an essential step in dormancy (Deb et al., 2009, Garton et al., 2008). While the diacylglycerol acyltransferase *tgs1* (Rv3130c) has been identified as the most active triacylglycerol synthase in *M. tuberculosis* (Daniel et al., 2004), *in vitro* upregulation has largely been studied under conditions of stress (Sirakova et al., 2006). To identify the phenotype of *M. tuberculosis* with increased *tgs1* expression, a *tgs1* overexpressing strain of *M. tuberculosis* was developed by Natalie Garton (Garton, 2011).

The assessment and discussion of the *M. tuberculosis* pSMT3-*tgs1* overexpressing strain, used in the studies discussed in Chapter 5, is displayed below.

Results

Overexpression of *tgs1* (Rv3130c)

The *Mtb* pSMT3-*tgs1* overexpressing strain was developed by Natalie Garton at the University of Leicester (Garton, 2011). The *tgs1* gene was transformed into *M. tuberculosis* H37Rv on the mycobacterium-*Escherichia Coli* shuttle vector pSMT3; the pSMT3 plasmid was stably maintained in *M. tuberculosis* H37Rv as a control. The vector pSMT3 is a high copy number plasmid that expresses mycobacterial proteins at many multiples of the regular, endogenous level in the wild-type strain (Harth et al., 2004). The plasmid was maintained within *M. tuberculosis* through the addition of hygromycin to the growth medium.

Gene Expression of pSMT3-tgs1 vs pSMT3 control strain

Gene expression for the *M. tuberculosis* pSMT3-*tgs1* overexpressing strain was compared to the pSMT3 control strain to confirm overexpression of the *tgs1* gene. The two strains were grown to exponential growth phase in conical flasks. RNA was extracted and purified from these cultures, reverse transcribed into cDNA and diluted 1:4. Quantitative PCRs were performed in triplicate for each of the extracted samples for the 20 selected genes and 4 housekeeping genes.

The results were analysed using the Pfaffl (relative quantification) method. The quantitative PCR and biological triplicates were averaged. Gene expression fold changes for the pSMT3-tgs1 overexpressing strain were calculated in relation to pSMT3 plasmid control cultures; the gene expression is displayed visually in Figure 42.

The expression of 20 target genes was examined. Expression of *tgs1* in the overexpression strain was >2000 fold greater as compared to the pSMT3 control strain. For the other 19 target genes, expression was similar for both the overexpression strain and the pSMT3 control strain. The gene expression confirmed the overexpression of the *tgs1* gene.

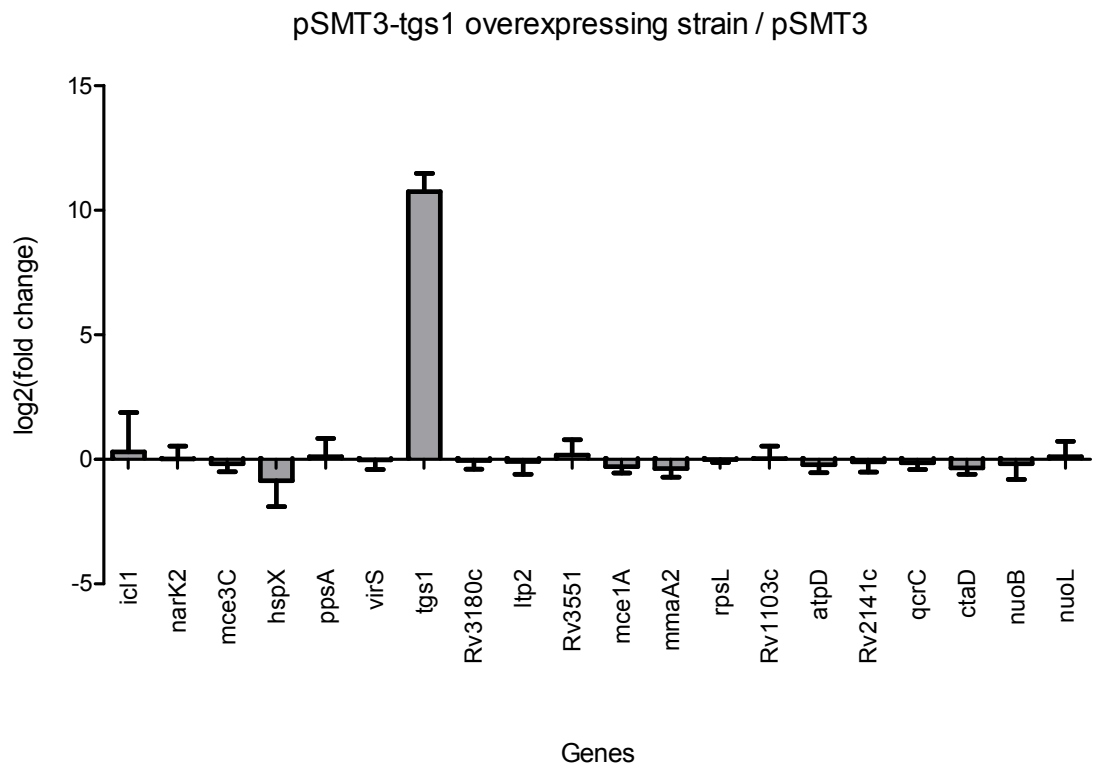


Figure 42: Gene Expression of *M. tuberculosis* pSMT3-*tgs1* overexpressing strain

The bar graph displays the gene expression of 20 target genes in exponential phase *M. tuberculosis* pSMT3-*tgs1* overexpression strain calculated against exponential phase *M. tuberculosis* pSMT3 plasmid control strain. The genes were normalised to NF4 (normalising factor - *aroA/dfrA/polA/thyA*). Error bars show standard deviation.

Growth Curves and Microscopy

The growth of the *M. tuberculosis* pSMT3-*tgs1* strain and control strain was assessed against the *M. tuberculosis* H37Rv laboratory strain. The three strains were grown in triplicate in 7H9 growth medium at a starting OD of 0.05. The cultures were maintained at 37°C, shaking at 100rpm. The ODs were recorded once daily for a total of 10 days. The average growth curves with standard deviation are shown in Figure 43.

Statistical significance between the curves in Figure 43 was assessed using a permutation test developed by Thompson, R. and Smyth, G. (2010) and used previously in Elso et al. (2004). The test compared the OD between two groups at each time-point, over the course of the experiment. Results were obtained by simulating the pairwise comparisons between two groups of growth curves; the culture replicates were randomly allocated to each of two groups and the mean was recalculated for 10000 data sets permuted in this manner (Thompson and Smyth, 2010). The p-value is the proportion of permutations where the mean is greater in absolute value than the mean of the original data set (Elso et al., 2004, Thompson and Smyth, 2010). This statistical test is automated on the Walter+Eliza Hall Institute of Medical Research Bioinformatics website (Thompson and Smyth, 2010). Pairwise comparisons were performed between the three *M. tuberculosis* strains, resulting in 3 pairwise tests and p-values. Statistical comparison of the growth curves following 10,000 permutations demonstrated no significant difference between the growth curves.

Fluorescence and LB analysis of the pSMT3-*tgs1* overexpressing strain and control is shown in Table 19. The fluorescence was analysed using automated cell counting as developed by Andrew Bell (Bell, A., Unpublished Data). At least 100 cells were analysed for both fluorescence and LB analysis.

The frequency of LB-positive cells was noted to be higher in the pSMT3-*tgs1* overexpressing strain as compared to the control strain. Further, a Mann-Whitney test confirmed that the median fluorescence intensity pSMT3-*tgs1* overexpressing strain is significantly higher than that of the pSMT3 control ($p<0.0001$*).

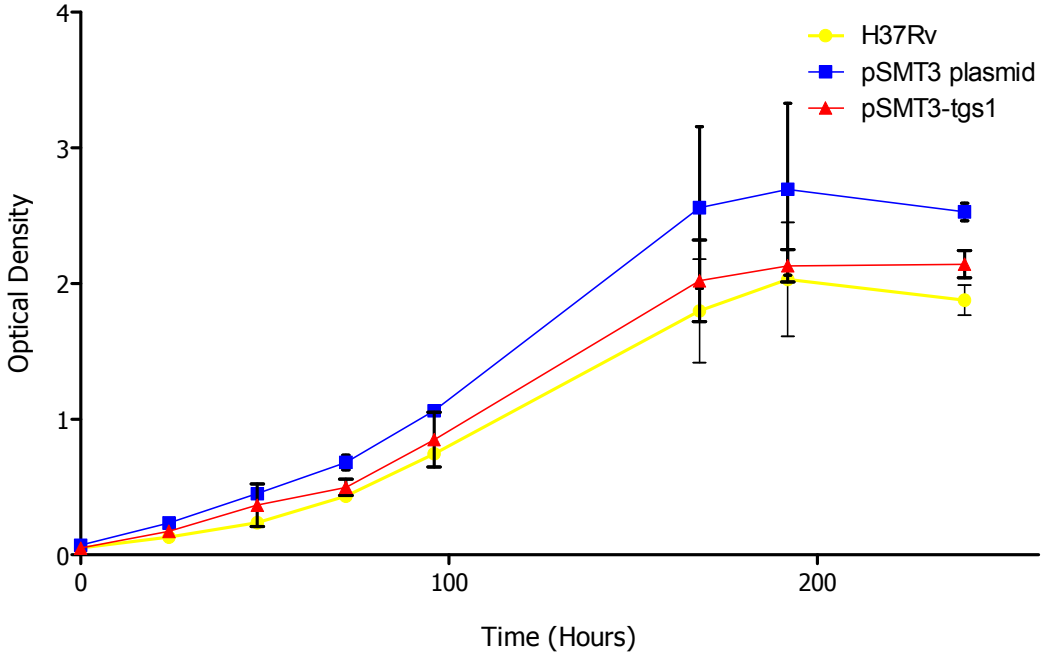


Figure 43: Growth Curves for *M. tuberculosis* pSMT3-*tgs1* overexpressing strain, pSMT3 control and H37Rv

The bar graph displays average growth curves for *M. tuberculosis* H37Rv (yellow), pSMT3-*tgs1* (red) and pSMT3 plasmid control (blue). The x-axis displays time in hours and the y-axis displays optical density (OD). OD time-points are averaged over 3 biological replicates; error bars display standard deviation. The growth curves were compared using a permutation test.

Table 19: Fluorescence Analysis of *M. tuberculosis* pSMT3-*tgs1* overexpressing strain vs. pSMT3 control

Sample	# of Cells	Median Fluorescence	Interquartile Range	Ratio (vs. Control)	# LB +ve (# Cells)	LB %
pSMT3-*tgs1*	286	212663	152567 – 318027	6.25	164 (303)	54.13
pSMT3 control	279	34033	25456 – 63661	-	40 (296)	13.5

Discussion

Overexpression of *tgs1* (Rv3130c)

The *M. tuberculosis* pSMT3-*tgs1* overexpressing strain was developed by Natalie Garton at the University of Leicester (2011). The strain made use of the pSMT3 *E. coli*-Mycobacterium shuttle plasmid that harbours a hygromycin resistance gene and the *M. bovis* hsp60 promoter (Harth et al., 2004).

Overexpression of tgs1 and M. tuberculosis pattern of growth

The gene expression of the 20 target genes (from Chapter 4) was analysed for the *M. tuberculosis* pSMT3-*tgs1* overexpressing strain, compared to the control plasmid strain. The gene expression changes appeared to be limited to *tgs1;* there did not appear to be any gene expression difference between the two strains for the other 19 genes. *Tgs1* was found to be expressed greater than 2000-fold higher in the *tgs1* overexpressing strain as compared to control.

The gene expression of the pSMT3-*tgs1* was expected, based upon the nature of the pSMT3 plasmid. The pSMT3 plasmid is a high-copy number plasmid, that when used, expressed mycobacterial proteins at many multiples of their endogenous level in the wild-type strain (Harth et al., 2004). The *M. bovis* hsp60 promoter is a heat shock protein promoter, controlling the expression of mycobacterial antigen, and stable under a number of stress conditions (Batoni et al., 1998).

As the induction of *tgs1* was confirmed in the *tgs1* overexpressing strain and as differences in the growth of strains may affect future experiments, growth curves of the OD of the pSMT3-*tgs1* overexpressing strain, pSMT3 control strain, and H37Rv laboratory strain were analysed for differences. A permutation test for the comparison of growth curves was selected for the analysis of the growth curves. This was more appropriate for the analysis of this dataset than individual t-tests at each time-point; separate t-tests at each time-point would lose statistical power, and would not address the whole time period of growth (Thompson and Smyth, 2010).

Despite the increased expression of *tgs1* and translation of the tgs1 enzyme, as well as the increased production of triacylglycerol (see below), the strain of *M. tuberculosis* was able to cope with these demands and maintain a growth rate comparable to both the control strain

and H37Rv. It is worth noting that the measurement of OD was used to construct these growth curves; the viability of cells cannot be inferred from OD.

Frequency of lipid body positive cells and fluorescence intensity is increased in tgs1 overexpressing strain

The frequency of LB-positive cells and the median fluorescence intensity of the bacilli were analysed for both *M. tuberculosis* pSMT3-*tgs1* and pSMT3 control strains. For the analysed cultures, the frequency of LB-positive cells was approximately 4-fold higher in the pSMT3-*tgs1* strain over the control. Furthermore, the median fluorescence intensity of the pSMT3-*tgs1* bacilli was significantly higher than the control strain, as analysed by a Mann-Whitney test.

As LBs are composed of triacylglycerol, the increase in the frequency of LB-positive cells suggested that the upregulation of *tgs1* and LB formation are associated. Furthermore, LipidTOX stains neutral lipids; thus, the higher intensity of fluorescence displayed by the *tgs1* overexpressing stain suggests a higher neutral lipid content in those bacilli, compared to the control strain. Work by Garton (2011) confirmed this increased triacylglycerol accumulation: both the *M. tuberculosis* pSMT3-*tgs1* strain and control were labelled with 14C acetate for 24h and non-polar lipids were extracted. The equivalent activity of extracts was loaded on thin-layer chromatography (TLC) and the *tgs1* overexpressing strain was found to contain approximately 3.5 fold greater triacylglycerol than control (Garton, 2011).

There were no background problems for either the overexpressing strain or the control strain, as compared to *M. tuberculosis* exposed to multiple growth conditions (Chapter 4).

Conclusions

The overall results suggested that the *M. tuberculosis* pSMT3-*tgs1* did upregulate the expression of *tgs1*, a DGAT, which subsequently led to an increased accumulation of triacylglycerol. Both strains were used in further experiments (Chapter 5).

Appendix 4: Supplemental Data

Figure 19: Selection of the most suitable reference genes for normalisation of *M. tuberculosis* expression samples (Data)

Number	Gene	M Value	Number	Gene	M Value
1	*thyA*	0.52	**7**	*gyrB*	1.06
2	*dfrA*	0.64	**8**	*mtrA*	1.08
3	*aroA*	0.73	**9**	*sigA*	1.35
4	*polA*	0.85	**10**	16S	1.50
5	*lysA*	0.88	**11**	*fbpB*	1.66
6	*ftsZ*	1.00			

Figure 23: Selected Gene Expression of *M. tuberculosis* in sputum (Microarray Data)

Upregulated		Downregulated	
Gene	**Fold Change**	**Gene**	**Fold Change**
hspX	15.69	*Rv1103c*	-2.63
Rv3551	5.85	*atpD*	-2.63
tgs1	4.86	Rv2141c	-2.86
icl1	3.61	*qcrC*	-3.23
narK2	3.27	*mmaA2*	-3.57
virS	3.23	*ctaD*	-4.54
mce3C	2.93	*nuoL*	-4.55
ltp2	2.83	*mce1A*	-5.26
Rv3180c	2.73	*nuoB*	-5.26
ppsA	2.68	*rpsL*	-8.33

Figure 25: *M. tuberculosis* Gene Expression for 20 Target Genes under Selected Growth Conditions (Caption Continued Overleaf)**(Data)**

Upregulated Genes					
Sample	***hspX***	**Rv3551**	***tgs1***	***icl1***	***narK2***
Stationary	15.6 ± 2.56	0.93 ± 0.39	10.1 ± 0.74	2.41 ± 1.44	8.22 ± 0.22
PBS	3.36 ± 1.84	-2.55 ± 0.38	1.67 ± 0.80	1.09 ± 0.35	2.42 ± 0.16
RPMI-1640	1.27 ± 0.44	1.31 ± 0.71	-0.26 ± 0.37	1.67 ± 1.32	-0.22 ± 1.58
SDS	0.13 ± 0.43	-0.30 ± 0.45	-0.64 ± 1.35	0.018 ± 0.20	0.023 ± 0.15
Artificial Saliva	1.19 ± 0.99	-0.053 ± 0.89	0.94 ± 0.88	0.057 ± 1.24	-0.67 ± 0.59
Sample	***virS***	***mce3C***	***ltp2***	**Rv3180c**	***ppsA***
Stationary	6.83 ± 2.17	2.14 ± 0.74	0.97 ± 0.37	1.76 ± 0.24	1.51 ± 0.26
PBS	5.72 ± 0.18	0.017 ± 0.67	-0.19 ± 0.27	1.25 ± 0.32	-0.074 ± 0.35
RPMI-1640	4.17 ± 0.29	-0.45 ± 0.27	0.22 ± 0.37	0.21 ± 0.22	-0.61 ± 0.81
SDS	0.94 ± 1.15	-0.50 ± 0.48	-0.32 ± 0.13	-0.33 ± 0.42	-0.46 ± 0.19
Artificial Saliva	1.51 ± 0.49	1.19 ± 0.55	-0.99 ± 0.71	-0.16 ± 0.89	-2.50 ± 0.79
Downregulated Genes					
Sample	**Rv1103c**	***atpD***	**Rv2141c**	***qcrC***	***mmaA2***
Stationary	0.82 ± 0.42	1.53 ± 1.64	2.51 ± 0.47	0.13 ± 0.83	3.64 ± 1.77
PBS	1.67 ± 0.33	-1.41 ± 0.66	2.25 ± 0.32	-1.49 ± 0.21	3.48 ± 0.22
RPMI-1640	-0.74 ± 0.38	0.29 ± 0.12	2.13 ± 0.60	-0.90 ± 0.27	3.11 ± 0.24
SDS	0.57 ± 0.83	-0.70 ± 0.047	-0.24 ± 0.29	-0.11 ± 0.24	0.22 ± 1.62
Artificial Saliva	1.23 ± 0.41	-0.33 ± 0.77	-0.96 ± 0.64	0.062 ± 0.60	-0.80 ± 0.61
Sample	***ctaD***	***nuoL***	***mce1A***	***nuoB***	***rpsL***
Stationary	2.33 ± 1.77	-0.80 ± 1.35	1.75 ± 1.90	0.64 ± 1.65	4.48 ± 0.76
PBS	1.30 ± 0.29	-3.28 ± 0.82	0.53 ± 0.46	-1.35 ± 0.49	3.66 ± 0.54
RPMI-1640	1.41 ± 0.29	-1.06 ± 0.45	0.58 ± 0.75	-0.72 ± 0.80	3.02 ± 0.53
SDS	0.36 ± 0.25	-1.06 ± 0.28	-0.42 ± 0.67	-0.50 ± 0.21	-0.050 ± 0.61
Artificial Saliva	0.15 ± 0.63	-2.01 ± 1.01	0.31 ± 0.62	-0.13 ± 0.98	-0.77 ± 0.22

Labels: 'Stationary' = Stationary Phase Culture

Figure 27: Heat Map showing Fold Changes for Sample Replicates (Data)

Values displayed are Log_2 Values ± Standard Deviation.

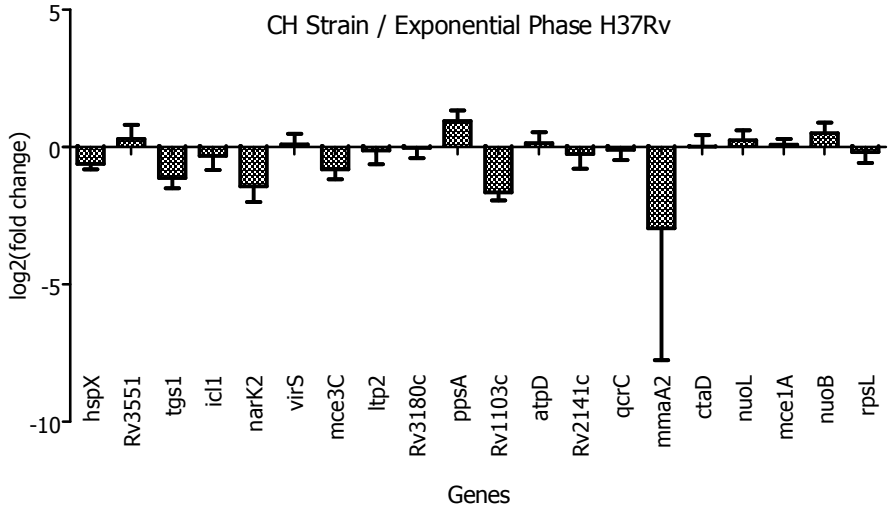
Upregulated Genes					
Sample	***hspX***	**Rv3551**	***tgs1***	***icl1***	***narK2***
Cholesterol	-0.67 ± 0.30	2.36 ± 0.71	0.14 ± 0.33	4.66 ± 1.10	-0.38 ± 0.62
Oleic Acid	0.66 ± 1.07	0.55 ± 0.17	2.09 ± 1.35	4.63 ± 0.86	0.90 ± 0.48
Static (Flask)	8.29 ± 0.66	2.00 ± 0.24	9.48 ± 0.83	-0.38 ± 0.69	6.78 ± 0.34
Static (Tube)	9.85 ± 0.86	2.18 ± 0.74	11.83 ± 0.73	-0.58 ± 0.81	7.74 ± 0.97
PBS + NO 1h	11.82 ± 0.56	2.32 ± 0.37	12.06 ± 0.49	4.68 ± 0.37	9.18 ± 0.25
PBS + NO 4h	12.84 ± 0.36	0.99 ± 0.56	11.90 ± 0.25	0.98 ± 0.29	10.30 ± 0.22
RPMI + NO 1h	9.72 ± 1.43	3.83 ± 1.67	10.03 ± 1.36	4.07 ± 1.62	7.92 ± 1.29
RPMI + NO 4h	11.73 ± 0.28	1.02 ± 0.40	12.12±0.085	5.29 ± 0.61	9.71 ± 0.17
PBS Multi (Flask)	9.74 ± 0.49	1.15 ± 0.50	9.36 ± 0.25	4.70 ± 0.56	4.89 ± 0.81
PBS Multi (Tube)	6.29 ± 0.47	3.00 ± 0.15	7.71 ± 0.094	2.57 ± 0.34	3.70 ± 0.41
RPMI Multi (Flask)	7.90 ± 0.18	0.61 ± 0.27	8.72 ± 0.44	0.61 ± 0.27	5.20 ± 0.45
RPMI Multi (Tube)	6.09 ± 0.068	3.86 ± 0.20	8.21 ± 0.12	4.12 ± 0.29	4.09 ± 0.29
Sample	***virS***	***mce3C***	***ltp2***	**Rv3180c**	***ppsA***
Cholesterol	0.87 ± 0.16	-2.25 ± 0.63	1.69 ± 0.52	-0.29 ± 0.59	0.73 ± 0.47
Oleic Acid	-0.78 ± 0.23	-1.82 ± 0.17	-0.13 ± 0.43	-0.35 ± 0.24	3.04 ± 0.20
Static (Flask)	-0.83 ± 0.41	-0.86 ± 5.22	-0.39 ± 0.48	0.039 ± 0.62	-0.99 ± 0.28
Static (Tube)	-1.10 ± 0.42	-4.21 ± 5.20	0.71 ± 0.50	0.094 ± 0.62	0.74±0.0039
PBS + NO 1h	0.39 ± 0.73	-0.45 ± 0.26	0.28 ± 0.46	-1.42 ± 0.85	0.12 ± 0.23
PBS + NO 4h	1.90 ± 0.25	2.12 ± 0.22	-0.64 ± 0.26	-0.42 ± 0.25	-2.18 ± 0.18
RPMI + NO 1h	-2.22 ± 1.24	-1.83 ± 1.58	0.52 ± 1.68	-2.45 ± 1.10	0.24 ± 1.52
RPMI + NO 4h	1.00 ± 0.40	0.74 ± 0.21	0.59 ± 0.40	-1.10 ± 0.47	1.73 ± 0.13
PBS Multi (Flask)	3.23 ± 0.53	0.26 ± 0.13	1.68±0.0095	0.86 ± 0.15	-0.38 ± 0.27
PBS Multi (Tube)	2.73 ± 0.35	0.44 ± 0.46	2.94 ± 0.19	0.47 ± 1.09	-0.48±0.099
RPMI Multi (Flask)	1.41 ± 0.32	1.77 ± 0.60	3.49 ± 0.14	2.52 ± 0.50	-0.11±0.015
RPMI Multi (Tube)	1.55 ± 0.25	1.84 ± 0.17	4.49 ± 0.30	2.15 ± 0.11	0.26 ± 0.29

Labels: 'Static' = Static Incubation; 'Multi' = Multi-Stimulus Culture

Downregulated Genes					
Sample	**Rv1103c**	***atpD***	**Rv2141c**	***qcrC***	***mmaA2***
Cholesterol	-0.70 ±0.20	0.22 ±0.54	-0.093 ±0.71	0.010 ±0.22	-0.17 ± 0.24
Oleic Acid	-0.088±0.23	-1.10 ± 0.42	-0.21±0.087	-0.52 ± 0.23	-0.24 ± 0.20
Static (Flask)	0.29 ±0.21	-0.85 ±0.24	-1.35 ±0.053	0.36 ±0.28	-0.88 ± 0.13
Static (Tube)	0.80 ± 0.081	0.96 ± 0.58	-0.63 ± 0.32	1.23±0.080	-0.88 ± 0.24
PBS + NO 1h	0.92 ± 0.31	0.71 ± 0.74	-1.42 ± 0.10	0.95 ± 0.33	-1.63 ± 0.41
PBS + NO 4h	0.91 ± 0.48	-1.20 ± 0.15	-1.68 ± 0.24	-0.73 ± 0.17	-1.23 ± 0.13
RPMI + NO 1h	-1.23 ± 1.13	-0.39 ± 1.41	-3.21 ± 1.74	-0.40 ± 1.20	-3.29 ± 1.30
RPMI + NO 4h	1.01 ± 0.54	-0.64 ± 0.48	-1.01 ±0.048	-0.71 ± 0.59	-1.55 ± 0.23
PBS Multi (Flask)	0.29 ± 0.052	1.82 ± 0.15	-2.44 ± 0.21	1.91 ± 0.17	-2.40 ± 0.13
PBS Multi (Tube)	0.84 ± 0.17	0.97 ± 0.24	0.36 ± 1.65	2.78 ± 0.055	-7.08 ± 4.63
RPMI Multi (Flask)	-0.11±0.015	1.41 ± 0.32	8.72 ± 0.44	2.52 ± 0.50	1.77 ± 0.60
RPMI Multi (Tube)	-0.50 ± 0.19	2.17 ± 0.18	-0.66 ± 0.25	2.94 ± 0.28	-3.42 ± 7.66
Sample	***ctaD***	***nuoL***	***mce1A***	***nuoB***	***rpsL***
Cholesterol	0.040 ± 0.19	-0.59 ± 0.68	-1.72±0.051	-0.70 ± 0.25	-0.38 ± 0.29
Oleic Acid	-0.19 ± 0.30	-0.53 ± 0.38	-1.86 ± 0.77	-1.24 ± 0.56	-0.065±0.23
Static (Flask)	-0.58±0.055	-0.15 ± 0.25	-1.79 ± 0.23	-1.24 ± 0.33	-0.24 ± 0.22
Static (Tube)	-0.87 ± 0.10	1.59 ± 0.92	-0.78 ± 0.16	-0.19 ± 0.47	-0.31 ± 0.43
PBS + NO 1h	0.15 ± 0.39	0.87 ± 0.31	-1.63 ± 0.40	-1.00 ± 0.66	-1.56 ± 0.19
PBS + NO 4h	-0.36 ± 0.25	-2.11 ± 0.26	0.15 ± 0.080	-2.23 ± 0.50	-1.32 ± 0.15
RPMI + NO 1h	-2.11 ± 1.46	0.54 ± 1.24	-3.05 ± 1.40	-1.56 ± 1.30	-1.97 ± 1.17
RPMI + NO 4h	-0.69 ± 0.39	-1.69 ± 0.11	-2.25 ± 0.28	-2.14 ± 0.61	-0.76 ± 0.26
PBS Multi (Flask)	-2.66±0.040	-1.11 ± 0.14	-4.28 ± 0.16	-4.70±0.075	0.091 ± 0.37
PBS Multi (Tube)	-4.30 ± 0.26	0.96 ± 0.30	-2.47±0.085	-2.17 ± 0.23	2.41 ± 3.44
RPMI Multi (Flask)	3.49 ± 0.14	0.55 ± 0.41	5.20 ± 0.45	-3.29 ± 0.39	7.90 ± 0.18
RPMI Multi (Tube)	-3.93±0.079	2.61 ± 0.37	-2.52±0.073	-1.15 ± 0.21	-0.94±0.053

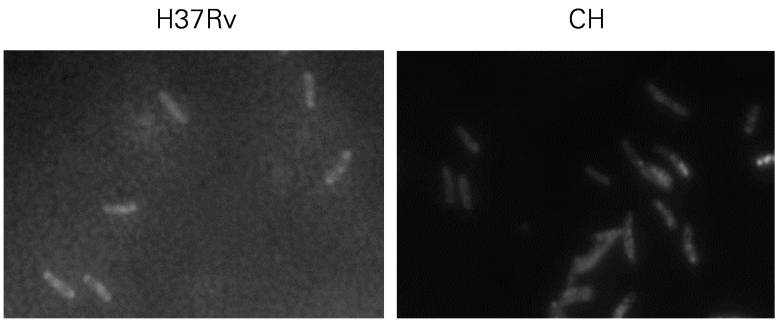
Labels: 'Static' = Static Incubation; 'Multi' = Multi-Stimulus Culture

Gene Expression of *M. tuberculosis* CH



The bar graph displays the gene expression of exponential phase *Mtb* CH calculated against exponential phase *Mtb* H37Rv laboratory strain. The correlation of *Mtb* CH gene expression does not correlate significantly with the sputum microarray data (p=0.2746, R=-0.2567).

Fluorescence and Lipid Body Analysis for *M. tuberculosis* CH versus H37Rv



Sample	# of Cells	Median	Interquartile Range	Ratio (vs. Control)	# LB +ve (# Cells)	LB %	Ratio (vs. Control
H37Rv	522	135080	112843 - 168022	-	123 (618)	19.9	-
CH	367	171282	132990 - 216527	1.27	174 (516)	33.7	1.69

Fluorescence images for *M. tuberculosis* CH (clinical strain) and H37Rv (laboratory strain) are displayed above. The table displays the number of cells analysed, median fluorescence intensities per call and interquartile ranges. The ratio compares the median fluorescent intensity of the growth conditions to exponential phase growth (control). For lipid body analysis, the number of cells analysed and lipid body positive% is shown.

References

ADEREM, A. & UNDERHILL, D. M. 1999. Mechanisms of phagocytosis in macrophages. *Annu Rev Immunol,* 17**,** 593-623.

AHMAD, S. 2011. Pathogenesis, immunology, and diagnosis of latent Mycobacterium tuberculosis infection. *Clin Dev Immunol,* 2011**,** 814943.

ALVAREZ, H. M., SILVA, R. A., CESARI, A. C., ZAMIT, A. L., PERESSUTTI, S. R., REICHELT, R., KELLER, U., MALKUS, U., RASCH, C., MASKOW, T., MAYER, F. & STEINBUCHEL, A. 2004. Physiological and morphological responses of the soil bacterium Rhodococcus opacus strain PD630 to water stress. *FEMS Microbiol Ecol,* 50**,** 75-86.

ALVAREZ, H. M. & STEINBUCHEL, A. 2002. Triacylglycerols in prokaryotic microorganisms. *Appl Microbiol Biotechnol,* 60**,** 367-76.

AMBION INCORPORATED. 2007a. *DNase I Demystified* [Online]. Ambion Incorporated. Available: http://www.ambion.com/techlib/tips/dnase1demystified.html [Accessed 11/01 2010].

AMBION INCORPORATED. 2007b. *Technical Bulletin #176: Avoiding DNA Contamination in RT-PCR* [Online]. Ambion Incorporated. Available: http://www.ambion.com/techlib/tb/tb_176.html [Accessed 11/01 2010].

AMBION INCORPORATED. 2007c. *Technical Bulletin #181: Methods to Remove DNA Contamination from RNA Samples* [Online]. Ambion Incorporated. Available: http://www.ambion.com/techlib/tb/tb_181.html [Accessed 11/01 2010].

ANDERSEN, C. L., JENSEN, J. L. & ORNTOFT, T. F. 2004. Normalization of real-time quantitative reverse transcription-PCR data: a model-based variance estimation approach to identify genes suited for normalization, applied to bladder and colon cancer data sets. *Cancer Res,* 64**,** 5245-50.

ARANAZ, A., COUSINS, D., MATEOS, A. & DOMINGUEZ, L. 2003. Elevation of Mycobacterium tuberculosis subsp. caprae Aranaz et al. 1999 to species rank as Mycobacterium caprae comb. nov., sp. nov. *Int J Syst Evol Microbiol,* 53**,** 1785-9.

ARCHULETA, R. J., MULLENS, P. & PRIMM, T. P. 2002. The relationship of temperature to desiccation and starvation tolerance of the Mycobacterium avium complex. *Arch Microbiol,* 178**,** 311-4.

ARGYROU, A., VETTING, M. W., ALADEGBAMI, B. & BLANCHARD, J. S. 2006. Mycobacterium tuberculosis dihydrofolate reductase is a target for isoniazid. *Nat Struct Mol Biol,* 13**,** 408-13.

ARRUDA, S., BOMFIM, G., KNIGHTS, R., HUIMA-BYRON, T. & RILEY, L. W. 1993. Cloning of an M. tuberculosis DNA fragment associated with entry and survival inside cells. *Science,* 261**,** 1454-7.

AZIZ, M. A., WRIGHT, A., LASZLO, A., DE MUYNCK, A., PORTAELS, F., VAN DEUN, A., WELLS, C., NUNN, P., BLANC, L. & RAVIGLIONE, M. 2006. Epidemiology of antituberculosis drug resistance (the Global Project on Anti-tuberculosis Drug Resistance Surveillance): an updated analysis. *Lancet,* 368**,** 2142-54.

BACON, J., JAMES, B. W., WERNISCH, L., WILLIAMS, A., MORLEY, K. A., HATCH, G. J., MANGAN, J. A., HINDS, J., STOKER, N. G., BUTCHER, P. D. & MARSH, P. D. 2004. The influence of reduced oxygen availability on pathogenicity and gene expression in Mycobacterium tuberculosis. *Tuberculosis (Edinb),* 84**,** 205-17.

BARER, M. R. & GARTON, N. J. 2009. Mycobacterial lipid bodies and the chemosensitivity and transmission of tuberculosis. *In:* TIMMIS, K. N. (ed.) *Handbook of Hydrocarbon and Lipid Microbiology: Microbial Interactions with Hydrocarbons, Oils Fats and Related Hydrophobic Substrates and Products.* Amsterdam: Springer Verlag.

BATONI, G., MAISETTA, G., FLORIO, W., FREER, G., CAMPA, M. & SENESI, S. 1998. Analysis of the Mycobacterium bovis hsp60 promoter activity in recombinant Mycobacterium avium. *FEMS Microbiol Lett,* 169**,** 117-24.

BEGGS, C. B., NOAKES, C. J., SLEIGH, P. A., FLETCHER, L. A. & SIDDIQI, K. 2003. The transmission of tuberculosis in confined spaces: an analytical review of alternative epidemiological models. *Int J Tuberc Lung Dis,* 7**,** 1015-26.

BELDA, J., LEIGH, R., PARAMESWARAN, K., O'BYRNE, P. M., SEARS, M. R. & HARGREAVE, F. E. 2000. Induced sputum cell counts in healthy adults. *Am J Respir Crit Care Med,* 161**,** 475-8.

BELL, A. J. 2010. *RE: Fluorescent Analysis of Mycobacteria.* Type to LEE, S. M.

BETTS, J. C., LUKEY, P. T., ROBB, L. C., MCADAM, R. A. & DUNCAN, K. 2002. Evaluation of a nutrient starvation model of Mycobacterium tuberculosis persistence by gene and protein expression profiling. *Mol Microbiol,* 43**,** 717-31.

BHATT, A., MOLLE, V., BESRA, G. S., JACOBS, W. R., JR. & KREMER, L. 2007. The Mycobacterium tuberculosis FAS-II condensing enzymes: their role in mycolic acid biosynthesis, acid-fastness, pathogenesis and in future drug development. *Mol Microbiol,* 64**,** 1442-54.

BHOWMIK, A., SEEMUNGAL, T. A., SAPSFORD, R. J., DEVALIA, J. L. & WEDZICHA, J. A. 1998. Comparison of spontaneous and induced sputum for investigation of airway inflammation in chronic obstructive pulmonary disease. *Thorax,* 53**,** 953-6.

BIFANI, P. J., MATHEMA, B., KUREPINA, N. E. & KREISWIRTH, B. N. 2002. Global dissemination of the Mycobacterium tuberculosis W-Beijing family strains. *Trends Microbiol,* 10**,** 45-52.

BIKETOV, S., POTAPOV, V., GANINA, E., DOWNING, K., KANA, B. D. & KAPRELYANTS, A. 2007. The role of resuscitation promoting factors in pathogenesis and reactivation of Mycobacterium tuberculosis during intra-peritoneal infection in mice. *BMC Infect Dis,* 7**,** 146.

BILLI, D. & POTTS, M. 2002. Life and death of dried prokaryotes. *Res Microbiol,* 153**,** 7-12.

BISHAI, W. R., DANNENBERG, A. M., JR., PARRISH, N., RUIZ, R., CHEN, P., ZOOK, B. C., JOHNSON, W., BOLES, J. W. & PITT, M. L. 1999. Virulence of Mycobacterium tuberculosis CDC1551 and H37Rv in rabbits evaluated by Lurie's pulmonary tubercle count method. *Infect Immun,* 67**,** 4931-4.

BLOCH, H. & SEGAL, W. 1956. Biochemical differentiation of Mycobacterium tuberculosis grown in vivo and in vitro. *J Bacteriol,* 72**,** 132-41.

BOSHOFF, H. I. & BARRY, C. E., 3RD 2005. Tuberculosis - metabolism and respiration in the absence of growth. *Nat Rev Microbiol,* 3**,** 70-80.

BRADDOCK, M., SCHWACHTGEN, J. L., HOUSTON, P., DICKSON, M. C., LEE, M. J. & CAMPBELL, C. J. 1998. Fluid Shear Stress Modulation of Gene Expression in Endothelial Cells. *News Physiol Sci,* 13**,** 241-246.

BRANDLI, O. 1998. The clinical presentation of tuberculosis. *Respiration,* 65**,** 97-105.

BRENNAN, P. J. 2003. Structure, function, and biogenesis of the cell wall of Mycobacterium tuberculosis. *Tuberculosis (Edinb),* 83**,** 91-7.

BRENNAN, P. J. & NIKAIDO, H. 1995. The envelope of mycobacteria. *Annu Rev Biochem,* 64**,** 29-63.

BRODIN, P., ROSENKRANDS, I., ANDERSEN, P., COLE, S. T. & BROSCH, R. 2004. ESAT-6 proteins: protective antigens and virulence factors? *Trends Microbiol,* 12**,** 500-8.

BROSCH, R., GORDON, S. V., MARMIESSE, M., BRODIN, P., BUCHRIESER, C., EIGLMEIER, K., GARNIER, T., GUTIERREZ, C., HEWINSON, G., KREMER, K., PARSONS, L. M., PYM, A. S., SAMPER, S., VAN SOOLINGEN, D. & COLE, S. T. 2002. A new evolutionary scenario for the Mycobacterium tuberculosis complex. *Proc Natl Acad Sci U S A,* 99**,** 3684-9.

BRZOSTEK, A., PAWELCZYK, J., RUMIJOWSKA-GALEWICZ, A., DZIADEK, B. & DZIADEK, J. 2009. Mycobacterium tuberculosis is able to accumulate and utilize cholesterol. *J Bacteriol,* 191**,** 6584-91.

BUSTIN, S. A., BENES, V., NOLAN, T. & PFAFFL, M. W. 2005. Quantitative real-time RT-PCR--a perspective. *J Mol Endocrinol,* 34**,** 597-601.

CACERES, N., TAPIA, G., OJANGUREN, I., ALTARE, F., GIL, O., PINTO, S., VILAPLANA, C. & CARDONA, P. J. 2009. Evolution of foamy macrophages in the pulmonary granulomas of experimental tuberculosis models. *Tuberculosis (Edinb),* 89**,** 175-82.

CAMINERO, J. A. 2010. Multidrug-resistant tuberculosis: epidemiology, risk factors and case finding. *Int J Tuberc Lung Dis,* 14**,** 382-90.

CAMPBELL, V. W. & JACKSON, D. A. 1980. The effect of divalent cations on the mode of action of DNase I. The initial reaction products produced from covalently closed circular DNA. *J Biol Chem,* 255**,** 3726-35.

CANETTI, G. 1955a. Growth of the tubercle bacillus in the tuberculosis lesion. *The Tubercle Bacillus in the Pulmonary Lesion of Man.* New York: Springer Publishing.

CANETTI, G. 1955b. *The Tubercle Bacillus in the Pulmonary Lesion of Man,* New York, Springer Publishing Company, Inc.

CARDONA, P. J. 2009. A dynamic reinfection hypothesis of latent tuberculosis infection. *Infection,* 37**,** 80-6.

CHANG, J. C., MINER, M. D., PANDEY, A. K., GILL, W. P., HARIK, N. S., SASSETTI, C. M. & SHERMAN, D. R. 2009. igr Genes and Mycobacterium tuberculosis cholesterol metabolism. *J Bacteriol,* 191**,** 5232-9.

CHATTERJEE, D. 1997. The mycobacterial cell wall: structure, biosynthesis and sites of drug action. *Curr Opin Chem Biol,* 1**,** 579-88.

CHINI, V., FOKA, A., DIMITRACOPOULOS, G. & SPILIOPOULOU, I. 2007. Absolute and relative real-time PCR in the quantification of tst gene expression among methicillin-resistant Staphylococcus aureus: evaluation by two mathematical models. *Lett Appl Microbiol,* 45**,** 479-84.

CIRILLO, S. L., SUBBIAN, S., CHEN, B., WEISBROD, T. R., JACOBS, W. R., JR. & CIRILLO, J. D. 2009. Protection of Mycobacterium tuberculosis from reactive oxygen species conferred by the mel2 locus impacts persistence and dissemination. *Infect Immun,* 77**,** 2557-67.

COLE, S. T. 2002. Comparative and functional genomics of the Mycobacterium tuberculosis complex. *Microbiology,* 148**,** 2919-28.

COLE, S. T., BROSCH, R., PARKHILL, J., GARNIER, T., CHURCHER, C., HARRIS, D., GORDON, S. V., EIGLMEIER, K., GAS, S., BARRY, C. E., 3RD, TEKAIA, F., BADCOCK, K., BASHAM, D., BROWN, D., CHILLINGWORTH, T., CONNOR, R., DAVIES, R., DEVLIN, K., FELTWELL, T., GENTLES, S., HAMLIN, N., HOLROYD, S., HORNSBY, T., JAGELS, K., KROGH, A., MCLEAN, J., MOULE, S., MURPHY, L., OLIVER, K., OSBORNE, J., QUAIL, M. A., RAJANDREAM, M. A., ROGERS, J., RUTTER, S., SEEGER, K., SKELTON, J., SQUARES, R., SQUARES, S., SULSTON, J. E., TAYLOR, K., WHITEHEAD, S. & BARRELL, B. G. 1998. Deciphering the biology of Mycobacterium tuberculosis from the complete genome sequence. *Nature,* 393**,** 537-44.

COLLINS, F. M. 1971. Relative susceptibility of acid-fast and non-acid-fast bacteria to ultraviolet light. *Appl Microbiol,* 21**,** 411-3.

COOK, G. C. & ZUMLA, A. I. (eds.) 2002. *Manson's Tropical Diseases,* London: Saunders.

CORBETT, E. L., MARSTON, B., CHURCHYARD, G. J. & DE COCK, K. M. 2006. Tuberculosis in sub-Saharan Africa: opportunities, challenges, and change in the era of antiretroviral treatment. *Lancet,* 367**,** 926-37.

COSMA, C. L., SHERMAN, D. R. & RAMAKRISHNAN, L. 2003. The secret lives of the pathogenic mycobacteria. *Annu Rev Microbiol,* 57**,** 641-76.

COTES, K., BAKALA N'GOMA J, C., DHOUIB, R., DOUCHET, I., MAURIN, D., CARRIERE, F. & CANAAN, S. 2008. Lipolytic enzymes in Mycobacterium tuberculosis. *Appl Microbiol Biotechnol,* 78**,** 741-9.

COUSINS, D. V., BASTIDA, R., CATALDI, A., QUSE, V., REDROBE, S., DOW, S., DUIGNAN, P., MURRAY, A., DUPONT, C., AHMED, N., COLLINS, D. M., BUTLER, W. R., DAWSON, D., RODRIGUEZ, D., LOUREIRO, J., ROMANO, M. I., ALITO, A., ZUMARRAGA, M. & BERNARDELLI, A. 2003. Tuberculosis in seals caused by a novel member of the Mycobacterium tuberculosis complex: Mycobacterium pinnipedii sp. nov. *Int J Syst Evol Microbiol,* 53**,** 1305-14.

COX, C. S. 1989. Airborne bacteria and viruses. *Sci Prog,* 73**,** 469-99.

CUNNINGHAM, A. F. & SPREADBURY, C. L. 1998. Mycobacterial stationary phase induced by low oxygen tension: cell wall thickening and localization of the 16-kilodalton alpha-crystallin homolog. *J Bacteriol,* 180**,** 801-8.

CYTRYN, E. J., SANGURDEKAR, D. P., STREETER, J. G., FRANCK, W. L., CHANG, W. S., STACEY, G., EMERICH, D. W., JOSHI, T., XU, D. & SADOWSKY, M. J. 2007. Transcriptional and physiological responses of Bradyrhizobium japonicum to desiccation-induced stress. *J Bacteriol,* 189**,** 6751-62.

DAFFE, M. & DRAPER, P. 1998. The envelope layers of mycobacteria with reference to their pathogenicity. *Adv Microb Physiol,* 39**,** 131-203.

DANIEL, J., DEB, C., DUBEY, V. S., SIRAKOVA, T. D., ABOMOELAK, B., MORBIDONI, H. R. & KOLATTUKUDY, P. E. 2004. Induction of a novel class of diacylglycerol acyltransferases and triacylglycerol accumulation in Mycobacterium tuberculosis as it goes into a dormancy-like state in culture. *J Bacteriol,* 186**,** 5017-30.

DANIEL, J., MAAMAR, H., DEB, C., SIRAKOVA, T. D. & KOLATTUKUDY, P. E. 2011. Mycobacterium tuberculosis uses host triacylglycerol to accumulate lipid droplets and acquires a dormancy-like phenotype in lipid-loaded macrophages. *PLoS Pathog,* 7**,** e1002093.

DARONCH, M., RUEGGEBERG, F. A., DICKINSON, D. & WATAHA, J. C. Year. Biological Effect of UVA/Blue Irradiation on bacteria. *In:* International Association for Dental Research, 2007 New Orleans.

DARWIN, K. H. & NATHAN, C. F. 2005. Role for nucleotide excision repair in virulence of Mycobacterium tuberculosis. *Infect Immun,* 73**,** 4581-7.

DAVID, H. L., JONES, W. D., JR. & NEWMAN, C. M. 1971. Ultraviolet light inactivation and photoreactivation in the mycobacteria. *Infect Immun,* 4**,** 318-9.

DAVIES, A. P., DHILLON, A. P., YOUNG, M., HENDERSON, B., MCHUGH, T. D. & GILLESPIE, S. H. 2008. Resuscitation-promoting factors are expressed in Mycobacterium tuberculosis-infected human tissue. *Tuberculosis (Edinb),* 88**,** 462-8.

DE JONG, B. C., HILL, P. C., AIKEN, A., JEFFRIES, D. J., ONIPEDE, A., SMALL, P. M., ADEGBOLA, R. A. & CORRAH, T. P. 2007. Clinical presentation and outcome of tuberculosis patients infected by M. africanum versus M. tuberculosis. *Int J Tuberc Lung Dis,* 11**,** 450-6.

DEB, C., DANIEL, J., SIRAKOVA, T. D., ABOMOELAK, B., DUBEY, V. S. & KOLATTUKUDY, P. E. 2006. A novel lipase belonging to the hormone-sensitive lipase family induced under starvation to utilize stored triacylglycerol in Mycobacterium tuberculosis. *J Biol Chem,* 281**,** 3866-75.

DEB, C., LEE, C. M., DUBEY, V. S., DANIEL, J., ABOMOELAK, B., SIRAKOVA, T. D., PAWAR, S., ROGERS, L. & KOLATTUKUDY, P. E. 2009. A novel in vitro multiple-stress dormancy model for Mycobacterium tuberculosis generates a lipid-loaded, drug-tolerant, dormant pathogen. *PLoS ONE,* 4**,** e6077.

DELOGU, G., SANGUINETTI, M., POSTERARO, B., ROCCA, S., ZANETTI, S. & FADDA, G. 2006. The hbhA gene of Mycobacterium tuberculosis is specifically upregulated in the lungs but not in the spleens of aerogenically infected mice. *Infect Immun,* 74**,** 3006-11.

DHARMADHIKARI, A. S. & NARDELL, E. A. 2009. Transmission of *Mycobacterium tuberculosis*. *In:* SCHAAF, H. S. & ZUMLA, A. (eds.) *Tuberculosis: A Comprehensive Clinical Reference.* Philadelphia: W.B. Saunders.

DOLGANOV, G. M., WOODRUFF, P. G., NOVIKOV, A. A., ZHANG, Y., FERRANDO, R. E., SZUBIN, R. & FAHY, J. V. 2001. A novel method of gene transcript profiling in airway biopsy homogenates reveals increased expression of a Na+-K+-Cl- cotransporter (NKCC1) in asthmatic subjects. *Genome Res,* 11**,** 1473-83.

DONAHUE, M. 2002. "Spare the cough, spoil the airway:" back to the basics in airway clearance. *Pediatr Nurs,* 28**,** 107-11.

DOS VULTOS, T., MESTRE, O., TONJUM, T. & GICQUEL, B. 2009. DNA repair in Mycobacterium tuberculosis revisited. *FEMS Microbiol Rev,* 33**,** 471-87.

DOWNING, J. F., PASULA, R., WRIGHT, J. R., TWIGG, H. L., 3RD & MARTIN, W. J., 2ND 1995. Surfactant protein a promotes attachment of Mycobacterium tuberculosis to alveolar macrophages during infection with human immunodeficiency virus. *Proc Natl Acad Sci U S A,* 92**,** 4848-52.

DOWNING, K. J., BETTS, J. C., YOUNG, D. I., MCADAM, R. A., KELLY, F., YOUNG, M. & MIZRAHI, V. 2004. Global expression profiling of strains harbouring null mutations reveals that the five rpf-like genes of Mycobacterium tuberculosis show functional redundancy. *Tuberculosis (Edinb),* 84**,** 167-79.

DOWNING, K. J., MISCHENKO, V. V., SHLEEVA, M. O., YOUNG, D. I., YOUNG, M., KAPRELYANTS, A. S., APT, A. S. & MIZRAHI, V. 2005. Mutants of Mycobacterium tuberculosis lacking three of the five rpf-like genes are defective for growth in vivo and for resuscitation in vitro. *Infect Immun,* 73**,** 3038-43.

DU TOIT, L. C., PILLAY, V. & DANCKWERTS, M. P. 2006. Tuberculosis chemotherapy: current drug delivery approaches. *Respir Res,* 7**,** 118.

DYE, C. 2006. Global epidemiology of tuberculosis. *Lancet,* 367**,** 938-40.

DYE, C., FLOYD, K. & UPLEKER, M. 2008. Global tuberculosis control: surveillance, planning, financing: WHO report 2008. Geneva: World Health Organization.

DYE, C., LONNROTH, K., JARAMILLO, E., WILLIAMS, B. G. & RAVIGLIONE, M. 2009. Trends in tuberculosis incidence and their determinants in 134 countries. *Bull World Health Organ,* 87**,** 683-91.

ELAMIN, A. A., STEHR, M., SPALLEK, R., ROHDE, M. & SINGH, M. 2011. The Mycobacterium tuberculosis Ag85A is a novel diacylglycerol acyltransferase involved in lipid body formation. *Mol Microbiol,* 81**,** 1577-92.

ELBEIN, A. D., PAN, Y. T., PASTUSZAK, I. & CARROLL, D. 2003. New insights on trehalose: a multifunctional molecule. *Glycobiology,* 13**,** 17R-27R.

ELSO, C. M., ROBERTS, L. J., SMYTH, G. K., THOMSON, R. J., BALDWIN, T. M., FOOTE, S. J. & HANDMAN, E. 2004. Leishmaniasis host response loci (lmr1-3) modify disease severity through a Th1/Th2-independent pathway. *Genes Immun,* 5**,** 93-100.

ERNST, J. D. 1998. Macrophage receptors for Mycobacterium tuberculosis. *Infect Immun,* 66**,** 1277-81.

ESPINAL, M. A. 2003. The global situation of MDR-TB. *Tuberculosis (Edinb),* 83**,** 44-51.

EUM, S. Y., KONG, J. H., HONG, M. S., LEE, Y. J., KIM, J. H., HWANG, S. H., CHO, S. N., VIA, L. E. & BARRY, C. E., 3RD 2010. Neutrophils are the predominant infected phagocytic cells in the airways of patients with active pulmonary TB. *Chest,* 137**,** 122-8.

EVERETT, B. S., LANDAU, S. & LEESE, M. 2001. *Cluster Analysis,* West Sussex, United Kingdom, Hodder Arnold.

FARZAN, S. 1990. Cough and Sputum Production. *In:* WALKER, H. W., HALL, W. D. & HURST, J. W. (eds.) *Clinical Methods: The History, Physical and Laboratory Examinations*. 3rd ed. Boston, USA: Butterworth-Heinemann.

FENNELLY, K. P., MARTYNY, J. W., FULTON, K. E., ORME, I. M., CAVE, D. M. & HEIFETS, L. B. 2004. Cough-generated aerosols of Mycobacterium tuberculosis: a new method to study infectiousness. *Am J Respir Crit Care Med,* 169**,** 604-9.

FENTON, M. J., RILEY, L. W. & SCHLESINGER, L. S. 2005. Receptor-Mediated Recognition of *Mycobacterium tuberculosis* by Host Cells. *In:* COLE, S. T. (ed.) *Tuberculosis and the Tubercle Bacillus*. Washington, D.C.: ASM Press.

FERGUSON, J. S. & SCHLESINGER, L. S. 2000. Pulmonary surfactant in innate immunity and the pathogenesis of tuberculosis. *Tuber Lung Dis,* 80**,** 173-84.

FIEGEL, J., CLARKE, R. & EDWARDS, D. A. 2006. Airborne infectious disease and the suppression of pulmonary bioaerosols. *Drug Discov Today,* 11**,** 51-7.

FISHER, M. A., PLIKAYTIS, B. B. & SHINNICK, T. M. 2002. Microarray analysis of the Mycobacterium tuberculosis transcriptional response to the acidic conditions found in phagosomes. *J Bacteriol,* 184**,** 4025-32.

FLYNN, J. L. & CHAN, J. 2001. Tuberculosis: latency and reactivation. *Infect Immun,* 69**,** 4195-201.

FOL, M., CHAUHAN, A., NAIR, N. K., MALONEY, E., MOOMEY, M., JAGANNATH, C., MADIRAJU, M. V. & RAJAGOPALAN, M. 2006. Modulation of Mycobacterium tuberculosis proliferation by MtrA, an essential two-component response regulator. *Mol Microbiol,* 60**,** 643-57.

FRIEDEN, T. R., STERLING, T. R., MUNSIFF, S. S., WATT, C. J. & DYE, C. 2003. Tuberculosis. *Lancet,* 362**,** 887-99.

GABRIELSSON, B. G., OLOFSSON, L. E., SJOGREN, A., JERNAS, M., ELANDER, A., LONN, M., RUDEMO, M. & CARLSSON, L. M. 2005. Evaluation of reference genes for studies of gene expression in human adipose tissue. *Obes Res,* 13**,** 649-52.

GANDHI, N. R., MOLL, A., STURM, A. W., PAWINSKI, R., GOVENDER, T., LALLOO, U., ZELLER, K., ANDREWS, J. & FRIEDLAND, G. 2006. Extensively drug-resistant tuberculosis as a cause of death in patients co-infected with tuberculosis and HIV in a rural area of South Africa. *Lancet,* 368**,** 1575-80.

GANNON, B. W., HAYES, C. M. & ROE, J. M. 2007. Survival rate of airborne Mycobacterium bovis. *Res Vet Sci,* 82**,** 169-72.

GARTON, N. J. 2011. Lipid bodies in Mycobacteria and *M. tuberculosis* pSMT3-*tgs1* overexpressing strain. Leicester: University of Leicester.

GARTON, N. J., CHRISTENSEN, H., MINNIKIN, D. E., ADEGBOLA, R. A. & BARER, M. R. 2002. Intracellular lipophilic inclusions of mycobacteria in vitro and in sputum. *Microbiology,* 148**,** 2951-8.

GARTON, N. J., WADDELL, S. J., SHERRATT, A. L., LEE, S. M., SMITH, R. J., SENNER, C., HINDS, J., RAJAKUMAR, K., ADEGBOLA, R. A., BESRA, G. S., BUTCHER, P. D. & BARER, M. R. 2008. Cytological and transcript analyses reveal fat and lazy persister-like bacilli in tuberculous sputum. *PLoS Med,* 5**,** e75.

GERDES, K., CHRISTENSEN, S. K. & LOBNER-OLESEN, A. 2005. Prokaryotic toxin-antitoxin stress response loci. *Nat Rev Microbiol,* 3**,** 371-82.

GIOFFRE, A., INFANTE, E., AGUILAR, D., SANTANGELO, M. P., KLEPP, L., AMADIO, A., MEIKLE, V., ETCHECHOURY, I., ROMANO, M. I., CATALDI, A., HERNANDEZ, R. P. & BIGI, F. 2005. Mutation in mce operons attenuates Mycobacterium tuberculosis virulence. *Microbes Infect,* 7**,** 325-34.

GOKULAN, K., RUPP, B., PAVELKA, M. S., JR., JACOBS, W. R., JR. & SACCHETTINI, J. C. 2003. Crystal structure of Mycobacterium tuberculosis diaminopimelate decarboxylase, an essential enzyme in bacterial lysine biosynthesis. *J Biol Chem,* 278**,** 18588-96.

GOMES, M. S., PAUL, S., MOREIRA, A. L., APPELBERG, R., RABINOVITCH, M. & KAPLAN, G. 1999. Survival of Mycobacterium avium and Mycobacterium tuberculosis in acidified vacuoles of murine macrophages. *Infect Immun,* 67**,** 3199-206.

GOMEZ, J. E., CHEN, J. M. & BISHAI, W. R. 1997. Sigma factors of Mycobacterium tuberculosis. *Tuber Lung Dis,* 78**,** 175-83.

GOMEZ, J. E. & MCKINNEY, J. D. 2004. M. tuberculosis persistence, latency, and drug tolerance. *Tuberculosis (Edinb),* 84**,** 29-44.

GOODFELLOW, M. & MINNIKIN, D. E. 1980. Definition of the genus *Mycobacterium vis a vi* other allied taxa. *In:* KUBICA, G. P., WAYNE, L. G. & GOOD, R. C. (eds.) *1954-1979: Twenty-Five Years of Mycobacterial Taxonomy*. Atlanta: United States Department of Health, Education and Welfare, Center for Disease Control and Prevention.

GOOSEN, N. & MOOLENAAR, G. F. 2008. Repair of UV damage in bacteria. *DNA Repair (Amst),* 7**,** 353-79.

GROEN, A. K. 2001. The pros and cons of gene expression analysis by microarrays. *J Hepatol,* 35**,** 295-6.

GUPTA, A. 2009. Killing activity and rescue function of genome-wide toxin-antitoxin loci of Mycobacterium tuberculosis. *FEMS Microbiol Lett,* 290**,** 45-53.

GUPTA, S., JAIN, S. & TYAGI, A. K. 1999. Analysis, expression and prevalence of the Mycobacterium tuberculosis homolog of bacterial virulence regulating proteins. *FEMS Microbiol Lett,* 172**,** 137-43.

GURR, M. I. & JAMES, A. T. 1975. *Lipid Biochemistry: An Introduction,* London, Chapman and Hall.

HAILE, Y., BJUNE, G. & WIKER, H. G. 2002. Expression of the mceA, esat-6 and hspX genes in Mycobacterium tuberculosis and their responses to aerobic conditions and to restricted oxygen supply. *Microbiology,* 148**,** 3881-6.

HALL-STOODLEY, L., WATTS, G., CROWTHER, J. E., BALAGOPAL, A., TORRELLES, J. B., ROBISON-COX, J., BARGATZE, R. F., HARMSEN, A. G., CROUCH, E. C. & SCHLESINGER, L. S. 2006. Mycobacterium tuberculosis binding to human surfactant proteins A and D, fibronectin, and small airway epithelial cells under shear conditions. *Infect Immun,* 74**,** 3587-96.

HAMBURG, M. A. & FRIEDEN, T. R. 1994. Tuberculosis transmission in the 1990s. *N Engl J Med,* 330**,** 1750-1.

HAMPSHIRE, T., SONEJI, S., BACON, J., JAMES, B. W., HINDS, J., LAING, K., STABLER, R. A., MARSH, P. D. & BUTCHER, P. D. 2004. Stationary phase gene expression of Mycobacterium tuberculosis following a progressive nutrient depletion: a model for persistent organisms? *Tuberculosis (Edinb),* 84**,** 228-38.

HAN, J. S., LEE, J. J., ANANDAN, T., ZENG, M., SRIPATHI, S., JAHNG, W. J., LEE, S. H., SUH, J. W. & KANG, C. M. 2010. Characterization of a chromosomal toxin-antitoxin, Rv1102c-Rv1103c system in Mycobacterium tuberculosis. *Biochem Biophys Res Commun,* 400**,** 293-8.

HARLAND, C. W., BOTYANSZKI, Z., RABUKA, D., BERTOZZI, C. R. & PARTHASARATHY, R. 2009. Synthetic trehalose glycolipids confer desiccation resistance to supported lipid monolayers. *Langmuir,* 25**,** 5193-8.

HARLAND, C. W., RABUKA, D., BERTOZZI, C. R. & PARTHASARATHY, R. 2008. The Mycobacterium tuberculosis virulence factor trehalose dimycolate imparts desiccation resistance to model mycobacterial membranes. *Biophys J,* 94**,** 4718-24.

HARTH, G., MASLESA-GALIC, S. & HORWITZ, M. A. 2004. A two-plasmid system for stable, selective-pressure-independent expression of multiple extracellular proteins in mycobacteria. *Microbiology,* 150**,** 2143-51.

HEID, C. A., STEVENS, J., LIVAK, K. J. & WILLIAMS, P. M. 1996. Real time quantitative PCR. *Genome Res,* 6**,** 986-94.

HELLYER, T. J., DESJARDIN, L. E., HEHMAN, G. L., CAVE, M. D. & EISENACH, K. D. 1999. Quantitative analysis of mRNA as a marker for viability of Mycobacterium tuberculosis. *J Clin Microbiol,* 37**,** 290-5.

HERNANDEZ-PANDO, R., JEYANATHAN, M., MENGISTU, G., AGUILAR, D., OROZCO, H., HARBOE, M., ROOK, G. A. & BJUNE, G. 2000. Persistence of DNA from Mycobacterium tuberculosis in superficially normal lung tissue during latent infection. *Lancet,* 356**,** 2133-8.

HETT, E. C., CHAO, M. C., DENG, L. L. & RUBIN, E. J. 2008. A mycobacterial enzyme essential for cell division synergizes with resuscitation-promoting factor. *PLoS Pathog,* 4**,** e1000001.

HETT, E. C., CHAO, M. C., STEYN, A. J., FORTUNE, S. M., DENG, L. L. & RUBIN, E. J. 2007. A partner for the resuscitation-promoting factors of Mycobacterium tuberculosis. *Mol Microbiol,* 66**,** 658-68.

HICKMAN, C., MACDONALD, K. L., OSTERHOLM, M. T., SCHECTER, G. F., ROYCE, S., VUGIA, D. J., PROCTOR, M. E., BUR, S. & DWYER, D. 1995. Exposure of passengers and flight crew to Mycobacterium tuberculosis on commercial aircraft, 1992-1995. *MMWR Morb Mortal Wkly Rep,* 44**,** 137-40.

HONER ZU BENTRUP, K. & RUSSELL, D. G. 2001. Mycobacterial persistence: adaptation to a changing environment. *Trends Microbiol,* 9**,** 597-605.

HOPEWELL, P. C. & JASMER, R. M. 2005. Overview of Clinical Tuberculosis. *In:* COLE, S. T., EISENACH, K. D., MCMURRAY, D. N. & JACOBS JR, W. R. (eds.) *Tuberculosis and the Tubercle Bacillus*. Washington,DC: ASM Press.

HOWARD, S. T. & BYRD, T. F. 2000. The rapidly growing mycobacteria: saprophytes and parasites. *Microbes Infect,* 2**,** 1845-53.

HU, Y. & COATES, A. R. 1999. Transcription of two sigma 70 homologue genes, sigA and sigB, in stationary-phase Mycobacterium tuberculosis. *J Bacteriol,* 181**,** 469-76.

HU, Y., MOVAHEDZADEH, F., STOKER, N. G. & COATES, A. R. 2006. Deletion of the Mycobacterium tuberculosis alpha-crystallin-like hspX gene causes increased bacterial growth in vivo. *Infect Immun,* 74**,** 861-8.

HUBERTS, P. & MIZRAHI, V. 1995. Cloning and sequence analysis of the gene encoding the DNA polymerase I from Mycobacterium tuberculosis. *Gene,* 164**,** 133-6.

HUNTER, R. L., OLSEN, M., JAGANNATH, C. & ACTOR, J. K. 2006. Trehalose 6,6'-dimycolate and lipid in the pathogenesis of caseating granulomas of tuberculosis in mice. *Am J Pathol,* 168**,** 1249-61.

INSTITUT PASTEUR. 2004. *TubercuList Web Server* [Online]. Paris, France: Institut Pasteur. Available: http://genolist.pasteur.fr/TubercuList/ [Accessed 14/01 2010].

IOERGER, T. R., FENG, Y., GANESULA, K., CHEN, X., DOBOS, K. M., FORTUNE, S., JACOBS, W. R., JR., MIZRAHI, V., PARISH, T., RUBIN, E., SASSETTI, C. & SACCHETTINI, J. C. 2010. Variation among genome sequences of H37Rv strains of Mycobacterium tuberculosis from multiple laboratories. *J Bacteriol,* 192**,** 3645-53.

JANG, H. J., NDE, C., TOGHROL, F. & BENTLEY, W. E. 2009. Microarray analysis of Mycobacterium bovis BCG revealed induction of iron acquisition related genes in response to hydrogen peroxide. *Environ Sci Technol,* 43**,** 9465-72.

JARLIER, V. & NIKAIDO, H. 1994. Mycobacterial cell wall: structure and role in natural resistance to antibiotics. *FEMS Microbiol Lett,* 123**,** 11-8.

JEFFERY, P. K. & LI, D. 1997. Airway mucosa: secretory cells, mucus and mucin genes. *Eur Respir J,* 10**,** 1655-62.

KALSCHEUER, R. & STEINBUCHEL, A. 2003. A novel bifunctional wax ester synthase/acyl-CoA:diacylglycerol acyltransferase mediates wax ester and triacylglycerol biosynthesis in Acinetobacter calcoaceticus ADP1. *J Biol Chem,* 278**,** 8075-82.

KANA, B. D., GORDHAN, B. G., DOWNING, K. J., SUNG, N., VOSTROKTUNOVA, G., MACHOWSKI, E. E., TSENOVA, L., YOUNG, M., KAPRELYANTS, A., KAPLAN, G. & MIZRAHI, V. 2008. The resuscitation-promoting factors of Mycobacterium tuberculosis are required for virulence and resuscitation from dormancy but are collectively dispensable for growth in vitro. *Mol Microbiol,* 67**,** 672-84.

KAPRELYANTS, A. S. & KELL, D. B. 1993. Dormancy in Stationary-Phase Cultures of Micrococcus luteus: Flow Cytometric Analysis of Starvation and Resuscitation. *Appl Environ Microbiol,* 59**,** 3187-96.

KEEP, N. H., WARD, J. M., ROBERTSON, G., COHEN-GONSAUD, M. & HENDERSON, B. 2006. Bacterial resuscitation factors: revival of viable but non-culturable bacteria. *Cell Mol Life Sci,* 63**,** 2555-9.

KENDALL, S. L., BURGESS, P., BALHANA, R., WITHERS, M., TEN BOKUM, A., LOTT, J. S., GAO, C., UHIA-CASTRO, I. & STOKER, N. G. 2010. Cholesterol utilization in mycobacteria is controlled by two TetR-type transcriptional regulators: kstR and kstR2. *Microbiology,* 156**,** 1362-71.

KENDALL, S. L., MOVAHEDZADEH, F., RISON, S. C., WERNISCH, L., PARISH, T., DUNCAN, K., BETTS, J. C. & STOKER, N. G. 2004. The Mycobacterium tuberculosis dosRS two-component system is induced by multiple stresses. *Tuberculosis (Edinb),* 84**,** 247-55.

KENDALL, S. L., WITHERS, M., SOFFAIR, C. N., MORELAND, N. J., GURCHA, S., SIDDERS, B., FRITA, R., TEN BOKUM, A., BESRA, G. S., LOTT, J. S. & STOKER, N. G. 2007. A highly conserved transcriptional repressor controls a large regulon involved in lipid degradation in Mycobacterium smegmatis and Mycobacterium tuberculosis. *Mol Microbiol,* 65**,** 684-99.

KENNEY, T. J. & CHURCHWARD, G. 1996. Genetic analysis of the Mycobacterium smegmatis rpsL promoter. *J Bacteriol,* 178**,** 3564-71.

KHOMENKO, A. G. 1987. The variability of Mycobacterium tuberculosis in patients with cavitary pulmonary tuberculosis in the course of chemotherapy. *Tubercle,* 68**,** 243-53.

KIM, M. J., WAINWRIGHT, H. C., LOCKETZ, M., BEKKER, L. G., WALTHER, G. B., DITTRICH, C., VISSER, A., WANG, W., HSU, F. F., WIEHART, U., TSENOVA, L., KAPLAN, G. & RUSSELL, D. G. 2010. Caseation of human tuberculosis granulomas correlates with elevated host lipid metabolism. *EMBO Mol Med,* 2**,** 258-74.

KIRKSEY, M. A., TISCHLER, A. D., SIMEONE, R., HISERT, K. B., UPLEKAR, S., GUILHOT, C. & MCKINNEY, J. D. 2011. Spontaneous phthiocerol dimycocerosate-deficient variants of Mycobacterium tuberculosis are susceptible to gamma interferon-mediated immunity. *Infect Immun,* 79**,** 2829-38.

KONDO, E. & KANAI, K. 1972. The lethal effect of long-chain fatty acids on mycobacteria. *Jpn J Med Sci Biol,* 25**,** 1-13.

KORCH, S. B., CONTRERAS, H. & CLARK-CURTISS, J. E. 2009. Three Mycobacterium tuberculosis Rel toxin-antitoxin modules inhibit mycobacterial growth and are expressed in infected human macrophages. *J Bacteriol,* 191**,** 1618-30.

KUMAR, A., BOSE, M. & BRAHMACHARI, V. 2003. Analysis of expression profile of mammalian cell entry (mce) operons of Mycobacterium tuberculosis. *Infect Immun,* 71**,** 6083-7.

KUSNER, D. J. 2005. Mechanisms of mycobacterial persistence in tuberculosis. *Clin Immunol,* 114**,** 239-47.

LEE, S. M. 2007. *Studies on Mycobacterial Lipid Bodies on Patients with Tuberculosis in The Gambia*. Intercalated BSc, University of Leicester.

LEISTIKOW, R. L., MORTON, R. A., BARTEK, I. L., FRIMPONG, I., WAGNER, K. & VOSKUIL, M. I. 2010. The Mycobacterium tuberculosis DosR regulon assists in metabolic homeostasis and enables rapid recovery from nonrespiring dormancy. *J Bacteriol,* 192**,** 1662-70.

LEVER, M. S., WILLIAMS, A. & BENNETT, A. M. 2000. Survival of mycobacterial species in aerosols generated from artificial saliva. *Lett Appl Microbiol,* 31**,** 238-41.

LI, Y. J., PETROFSKY, M. & BERMUDEZ, L. E. 2002. Mycobacterium tuberculosis uptake by recipient host macrophages is influenced by environmental conditions in the granuloma of the infectious individual and is associated with impaired production of interleukin-12 and tumor necrosis factor alpha. *Infect Immun,* 70**,** 6223-30.

LIVAK, K. J. & SCHMITTGEN, T. D. 2001. Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. *Methods,* 25**,** 402-8.

LOUDON, R. G., BUMGARNER, L. R., LACY, J. & COFFMAN, G. K. 1969. Aerial transmission of mycobacteria. *Am Rev Respir Dis,* 100**,** 165-71.

LOUDON, R. G. & SPOHN, S. K. 1969. Cough frequency and infectivity in patients with pulmonary tuberculosis. *Am Rev Respir Dis,* 99**,** 109-11.

LOW, K. L., RAO, P. S., SHUI, G., BENDT, A. K., PETHE, K., DICK, T. & WENK, M. R. 2009. Triacylglycerol utilization is required for regrowth of in vitro hypoxic nonreplicating Mycobacterium bovis bacillus Calmette-Guerin. *J Bacteriol,* 191**,** 5037-43.

LUTFALLA, G. & UZE, G. 2006. Performing quantitative reverse-transcribed polymerase chain reaction experiments. *Methods Enzymol,* 410**,** 386-400.

MACMICKING, J. D., NATHAN, C., HOM, G., CHARTRAIN, N., FLETCHER, D. S., TRUMBAUER, M., STEVENS, K., XIE, Q. W., SOKOL, K., HUTCHINSON, N. & ET AL. 1995. Altered responses to bacterial infection and endotoxic shock in mice lacking inducible nitric oxide synthase. *Cell,* 81**,** 641-50.

MADIGAN, M. T., MARTINKO, J. M. & PARKER, J. (eds.) 1997. *Brock Biology of Microorganisms,* Upper Saddle River, New Jersey: Prentice Hall International, Inc. .

MAKINOSHIMA, H. & GLICKMAN, M. S. 2005. Regulation of Mycobacterium tuberculosis cell envelope composition and virulence by intramembrane proteolysis. *Nature,* 436**,** 406-9.

MALHOTRA, V., SHARMA, D., RAMANATHAN, V. D., SHAKILA, H., SAINI, D. K., CHAKRAVORTY, S., DAS, T. K., LI, Q., SILVER, R. F., NARAYANAN, P. R. & TYAGI, J. S. 2004. Disruption of response regulator gene, devR, leads to attenuation in virulence of Mycobacterium tuberculosis. *FEMS Microbiol Lett,* 231**,** 237-45.

MANGANELLI, R., DUBNAU, E., TYAGI, S., KRAMER, F. R. & SMITH, I. 1999. Differential expression of 10 sigma factor genes in Mycobacterium tuberculosis. *Mol Microbiol,* 31**,** 715-24.

MANGANELLI, R., VOSKUIL, M. I., SCHOOLNIK, G. K. & SMITH, I. 2001. The Mycobacterium tuberculosis ECF sigma factor sigmaE: role in global gene expression and survival in macrophages. *Mol Microbiol,* 41**,** 423-37.

MARTIN, E., DAVIS, K., BIAN, K., LEE, Y. C. & MURAD, F. 2000. Cellular signaling with nitric oxide and cyclic guanosine monophosphate. *Semin Perinatol,* 24**,** 2-6.

MATTHEWS, J. L., CHUNG, M. & MATYAS, R. J. 2002. Persistent DNA contamination in competitive RT-PCR using cRNA internal standards: identity, quantity, and control. *Biotechniques,* 32**,** 1412-4, 1416-7.

MCKINNEY, J. D. 2000. In vivo veritas: the search for TB drug targets goes live. *Nat Med,* 6**,** 1330-3.

MCKINNEY, J. D. & GOMEZ, J. E. 2003. Life on the inside for Mycobacterium tuberculosis. *Nat Med,* 9**,** 1356-7.

MCKINNEY, J. D., HONER ZU BENTRUP, K., MUNOZ-ELIAS, E. J., MICZAK, A., CHEN, B., CHAN, W. T., SWENSON, D., SACCHETTINI, J. C., JACOBS, W. R., JR. & RUSSELL, D. G. 2000. Persistence of Mycobacterium tuberculosis in macrophages and mice requires the glyoxylate shunt enzyme isocitrate lyase. *Nature,* 406**,** 735-8.

MILLER, L. P., CRAWFORD, J. T. & SHINNICK, T. M. 1994. The rpoB gene of Mycobacterium tuberculosis. *Antimicrob Agents Chemother,* 38**,** 805-11.

MINER, M. D., CHANG, J. C., PANDEY, A. K., SASSETTI, C. M. & SHERMAN, D. R. 2009. Role of cholesterol in Mycobacterium tuberculosis infection. *Indian J Exp Biol,* 47**,** 407-11.

MINNIKIN, D. E., KREMER, L., DOVER, L. G. & BESRA, G. S. 2002. The methyl-branched fortifications of Mycobacterium tuberculosis. *Chem Biol,* 9**,** 545-53.

MISHRA, K. C., DE CHASTELLIER, C., NARAYANA, Y., BIFANI, P., BROWN, A. K., BESRA, G. S., KATOCH, V. M., JOSHI, B., BALAJI, K. N. & KREMER, L. 2008. Functional role of the PE domain and immunogenicity of the Mycobacterium tuberculosis triacylglycerol hydrolase LipY. *Infect Immun,* 76**,** 127-40.

MITCHISON, D. A. 1979. Basic mechanisms of chemotherapy. *Chest,* 76**,** 771-81.

MITCHISON, D. A. 2004. The search for new sterilizing anti-tuberculosis drugs. *Front Biosci,* 9**,** 1059-72.

MITCHISON, D. A. 2005. The diagnosis and therapy of tuberculosis during the past 100 years. *Am J Respir Crit Care Med,* 171**,** 699-706.

MOVAHEDZADEH, F., COLSTON, M. J. & DAVIS, E. O. 1997. Determination of DNA sequences required for regulated Mycobacterium tuberculosis RecA expression in response to DNA-damaging agents suggests that two modes of regulation exist. *J Bacteriol,* 179**,** 3509-18.

MUKAMOLOVA, G. V., KAPRELYANTS, A. S., YOUNG, D. I., YOUNG, M. & KELL, D. B. 1998. A bacterial cytokine. *Proc Natl Acad Sci U S A,* 95**,** 8916-21.

MUKAMOLOVA, G. V., KORMER, S. S., KELL, D. B. & KAPRELYANTS, A. S. 1999. Stimulation of the multiplication of Micrococcus luteus by an autocrine growth factor. *Arch Microbiol,* 172**,** 9-14.

MUKAMOLOVA, G. V., TURAPOV, O., MALKIN, J., WOLTMANN, G. & BARER, M. R. 2010. Resuscitation-promoting factors reveal an occult population of tubercle Bacilli in Sputum. *Am J Respir Crit Care Med,* 181**,** 174-80.

MUKAMOLOVA, G. V., TURAPOV, O. A., YOUNG, D. I., KAPRELYANTS, A. S., KELL, D. B. & YOUNG, M. 2002. A family of autocrine growth factors in Mycobacterium tuberculosis. *Mol Microbiol,* 46**,** 623-35.

MUKHERJEE, S., RAGHURAMAN, H. & CHATTOPADHYAY, A. 2007. Membrane localization and dynamics of Nile Red: effect of cholesterol. *Biochim Biophys Acta,* 1768**,** 59-66.

MUNOZ-ELIAS, E. J. & MCKINNEY, J. D. 2005. Mycobacterium tuberculosis isocitrate lyases 1 and 2 are jointly required for in vivo growth and virulence. *Nat Med,* 11**,** 638-44.

MUNOZ-ELIAS, E. J. & MCKINNEY, J. D. 2006. Carbon metabolism of intracellular bacteria. *Cell Microbiol,* 8**,** 10-22.

MURRY, J. P., PANDEY, A. K., SASSETTI, C. M. & RUBIN, E. J. 2009. Phthiocerol dimycocerosate transport is required for resisting interferon-gamma-independent immunity. *J Infect Dis,* 200**,** 774-82.

MUTTUCUMARU, D. G., ROBERTS, G., HINDS, J., STABLER, R. A. & PARISH, T. 2004. Gene expression profile of Mycobacterium tuberculosis in a non-replicating state. *Tuberculosis (Edinb),* 84**,** 239-46.

NEWTON, S. M., SMITH, R. J., WILKINSON, K. A., NICOL, M. P., GARTON, N. J., STAPLES, K. J., STEWART, G. R., WAIN, J. R., MARTINEAU, A. R., FANDRICH, S., SMALLIE, T., FOXWELL, B., AL-OBAIDI, A., SHAFI, J., RAJAKUMAR, K., KAMPMANN, B., ANDREW, P. W., ZIEGLER-HEITBROCK, L., BARER, M. R. & WILKINSON, R. J. 2006. A deletion defining a common Asian lineage of Mycobacterium tuberculosis associates with immune subversion. *Proc Natl Acad Sci U S A,* 103**,** 15594-8.

NEYROLLES, O., HERNANDEZ-PANDO, R., PIETRI-ROUXEL, F., FORNES, P., TAILLEUX, L., BARRIOS PAYAN, J. A., PIVERT, E., BORDAT, Y., AGUILAR, D., PREVOST, M. C., PETIT, C. & GICQUEL, B. 2006. Is adipose tissue a place for Mycobacterium tuberculosis persistence? *PLoS ONE,* 1**,** e43.

OHL, F., JUNG, M., XU, C., STEPHAN, C., RABIEN, A., BURKHARDT, M., NITSCHE, A., KRISTIANSEN, G., LOENING, S. A., RADONIC, A. & JUNG, K. 2005. Gene expression studies in prostate cancer tissue: which reference gene should be selected for normalization? *J Mol Med,* 83**,** 1014-24.

OHNO, H., ZHU, G., MOHAN, V. P., CHU, D., KOHNO, S., JACOBS, W. R., JR. & CHAN, J. 2003. The effects of reactive nitrogen intermediates on gene expression in Mycobacterium tuberculosis. *Cell Microbiol,* 5**,** 637-48.

ORMEROD, L. P. 2005. Multidrug-resistant tuberculosis (MDR-TB): epidemiology, prevention and treatment. *Br Med Bull,* 73-74**,** 17-24.

ORTALO-MAGNE, A., LEMASSU, A., LANEELLE, M. A., BARDOU, F., SILVE, G., GOUNON, P., MARCHAL, G. & DAFFE, M. 1996. Identification of the surface-exposed lipids on the cell envelopes of Mycobacterium tuberculosis and other mycobacterial species. *J Bacteriol,* 178**,** 456-61.

OSWALDO CRUZ INSTITUTE. 2010. *Department of Biochemistry and Molecular Biology: M. tuberculosis housekeeping genes* [Online]. Rio de Janeiro. Available: http://www.dbbm.fiocruz.br/genome/mycobac/tubhousekeep.html [Accessed 14/01 2010].

OXFORD DICTIONARIES 2010. Oxford Dictionary of English. 3rd ed. Oxford: Oxford University Press.

PANDEY, A. K. & SASSETTI, C. M. 2008. Mycobacterial persistence requires the utilization of host cholesterol. *Proc Natl Acad Sci U S A,* 105**,** 4376-80.

PANG, X. & HOWARD, S. T. 2007. Regulation of the alpha-crystallin gene acr2 by the MprAB two-component system of Mycobacterium tuberculosis. *J Bacteriol,* 189**,** 6213-21.

PANG, X., VU, P., BYRD, T. F., GHANNY, S., SOTEROPOULOS, P., MUKAMOLOVA, G. V., WU, S., SAMTEN, B. & HOWARD, S. T. 2007. Evidence for complex interactions of stress-associated regulons in an mprAB deletion mutant of Mycobacterium tuberculosis. *Microbiology,* 153**,** 1229-42.

PARISH, T., SMITH, D. A., KENDALL, S., CASALI, N., BANCROFT, G. J. & STOKER, N. G. 2003a. Deletion of two-component regulatory systems increases the virulence of Mycobacterium tuberculosis. *Infect Immun,* 71**,** 1134-40.

PARISH, T., SMITH, D. A., ROBERTS, G., BETTS, J. & STOKER, N. G. 2003b. The senX3-regX3 two-component regulatory system of Mycobacterium tuberculosis is required for virulence. *Microbiology,* 149**,** 1423-35.

PARISH, T. & STOKER, N. (eds.) 2010. *Mycobacterium tuberculosis Protocols,* New York: Humana Press.

PARISH, T. & STOKER, N. G. 2002. The common aromatic amino acid biosynthesis pathway is essential in Mycobacterium tuberculosis. *Microbiology,* 148**,** 3069-77.

PARK, H. D., GUINN, K. M., HARRELL, M. I., LIAO, R., VOSKUIL, M. I., TOMPA, M., SCHOOLNIK, G. K. & SHERMAN, D. R. 2003a. Rv3133c/dosR is a transcription factor that mediates the hypoxic response of Mycobacterium tuberculosis. *Mol Microbiol,* 48**,** 833-43.

PARK, T., YI, S. G., KANG, S. H., LEE, S., LEE, Y. S. & SIMON, R. 2003b. Evaluation of normalization methods for microarray data. *BMC Bioinformatics,* 4**,** 33.

PARRISH, N. M., DICK, J. D. & BISHAI, W. R. 1998. Mechanisms of latency in Mycobacterium tuberculosis. *Trends Microbiol,* 6**,** 107-12.

PARWATI, I., VAN CREVEL, R. & VAN SOOLINGEN, D. 2010. Possible underlying mechanisms for successful emergence of the Mycobacterium tuberculosis Beijing genotype strains. *Lancet Infect Dis,* 10**,** 103-11.

PENNING, L. C., VRIELING, H. E., BRINKHOF, B., RIEMERS, F. M., ROTHUIZEN, J., RUTTEMAN, G. R. & HAZEWINKEL, H. A. 2007. A validation of 10 feline reference genes for gene

expression measurements in snap-frozen tissues. *Vet Immunol Immunopathol,* 120**,** 212-22.

PEYRON, P., VAUBOURGEIX, J., POQUET, Y., LEVILLAIN, F., BOTANCH, C., BARDOU, F., DAFFE, M., EMILE, J. F., MARCHOU, B., CARDONA, P. J., DE CHASTELLIER, C. & ALTARE, F. 2008. Foamy macrophages from tuberculous patients' granulomas constitute a nutrient-rich reservoir for M. tuberculosis persistence. *PLoS Pathog,* 4**,** e1000204.

PFAFFL, M. W. 2006. Relative quantification. *In:* DORAK, T. (ed.) *Real-time PCR.* 1st ed. London: Taylor & Francis.

PFAFFL, M. W. & HAGELEIT, M. 2001. Validities of mRNA quantification using recombinant RNA and recombinant DNA external calibration curves in real-time RT-PCR. *Biotechnol Lett,* 23**,** 275-282.

PFAFFL, M. W., HORGAN, G. W. & DEMPFLE, L. 2002. Relative expression software tool (REST) for group-wise comparison and statistical analysis of relative expression results in real-time PCR. *Nucleic Acids Res,* 30**,** e36.

PFAFFL, M. W., TICHOPAD, A., PRGOMET, C. & NEUVIANS, T. P. 2004. Determination of stable housekeeping genes, differentially regulated target genes and sample integrity: BestKeeper--Excel-based tool using pair-wise correlations. *Biotechnol Lett,* 26**,** 509-15.

POTTS, M. 1994. Desiccation tolerance of prokaryotes. *Microbiol Rev,* 58**,** 755-805.

POTTS, M. 2001. Desiccation tolerance: a simple process? *Trends Microbiol,* 9**,** 553-9.

PRIMM, T. P., ANDERSEN, S. J., MIZRAHI, V., AVARBOCK, D., RUBIN, H. & BARRY, C. E., 3RD 2000. The stringent response of Mycobacterium tuberculosis is required for long-term survival. *J Bacteriol,* 182**,** 4889-98.

PROVVEDI, R., BOLDRIN, F., FALCIANI, F., PALU, G. & MANGANELLI, R. 2009. Global transcriptional response to vancomycin in Mycobacterium tuberculosis. *Microbiology,* 155**,** 1093-102.

QUAST, T. M. & BROWNING, R. F. 2006. Pathogenesis and clinical manifestations of pulmonary tuberculosis. *Dis Mon,* 52**,** 413-9.

RACHMAN, H., LEE, J. S., ANGERMANN, J., KOWALL, J. & KAUFMANN, S. H. 2006. Reliable amplification method for bacterial RNA. *J Biotechnol,* 126**,** 61-8.

RADONIC, A., THULKE, S., MACKAY, I. M., LANDT, O., SIEGERT, W. & NITSCHE, A. 2004. Guideline to reference gene selection for quantitative real-time PCR. *Biochem Biophys Res Commun,* 313**,** 856-62.

RAJNI, RAO, N. & MEENA, L. S. 2011. Biosynthesis and Virulent Behavior of Lipids Produced by Mycobacterium tuberculosis: LAM and Cord Factor: An Overview. *Biotechnol Res Int,* 2011**,** 274693.

RAMAGE, H. R., CONNOLLY, L. E. & COX, J. S. 2009. Comprehensive functional analysis of Mycobacterium tuberculosis toxin-antitoxin systems: implications for pathogenesis, stress responses, and evolution. *PLoS Genet,* 5**,** e1000767.

RAMAKERS, C., RUIJTER, J. M., DEPREZ, R. H. & MOORMAN, A. F. 2003. Assumption-free analysis of quantitative real-time polymerase chain reaction (PCR) data. *Neurosci Lett,* 339**,** 62-6.

REED, M. B., GAGNEUX, S., DERIEMER, K., SMALL, P. M. & BARRY, C. E., 3RD 2007. The W-Beijing lineage of Mycobacterium tuberculosis overproduces triglycerides and has the DosR dormancy regulon constitutively upregulated. *J Bacteriol,* 189**,** 2583-9.

RENGARAJAN, J., BLOOM, B. R. & RUBIN, E. J. 2005. Genome-wide requirements for Mycobacterium tuberculosis adaptation and survival in macrophages. *Proc Natl Acad Sci U S A,* 102**,** 8327-32.

RENGARAJAN, J., SASSETTI, C. M., NARODITSKAYA, V., SLOUTSKY, A., BLOOM, B. R. & RUBIN, E. J. 2004. The folate pathway is a target for resistance to the drug para-aminosalicylic acid (PAS) in mycobacteria. *Mol Microbiol,* 53**,** 275-82.

RIFAT, D., BISHAI, W. R. & KARAKOUSIS, P. C. 2009. Phosphate depletion: a novel trigger for Mycobacterium tuberculosis persistence. *J Infect Dis,* 200**,** 1126-35.

RILEY, R. L. & O'GRADY, F. 1961. *Airborne Infection,* New York, Macmillan.

ROBERTS, D. M., LIAO, R. P., WISEDCHAISRI, G., HOL, W. G. & SHERMAN, D. R. 2004. Two sensor kinases contribute to the hypoxic response of Mycobacterium tuberculosis. *J Biol Chem,* 279**,** 23082-7.

ROHDE, K. H., ABRAMOVITCH, R. B. & RUSSELL, D. G. 2007. Mycobacterium tuberculosis invasion of macrophages: linking bacterial gene expression to environmental cues. *Cell Host Microbe,* 2**,** 352-64.

RUIJTER, J. M., RAMAKERS, C., HOOGAARS, W. M., KARLEN, Y., BAKKER, O., VAN DEN HOFF, M. J. & MOORMAN, A. F. 2009. Amplification efficiency: linking baseline and bias in the analysis of quantitative PCR data. *Nucleic Acids Res,* 37**,** e45.

RUSSELL-GOLDMAN, E., XU, J., WANG, X., CHAN, J. & TUFARIELLO, J. M. 2008. A Mycobacterium tuberculosis Rpf double-knockout strain exhibits profound defects in reactivation from chronic tuberculosis and innate immunity phenotypes. *Infect Immun,* 76**,** 4269-81.

RUSSELL, D. G. 2001. Mycobacterium tuberculosis: here today, and here tomorrow. *Nat Rev Mol Cell Biol,* 2**,** 569-77.

RUSSELL, D. G. 2003. Phagosomes, fatty acids and tuberculosis. *Nat Cell Biol,* 5**,** 776-8.

RUSSELL, D. G. 2007. Who puts the tubercle in tuberculosis? *Nat Rev Microbiol,* 5**,** 39-47.

RUSTAD, T. R., HARRELL, M. I., LIAO, R. & SHERMAN, D. R. 2008. The enduring hypoxic response of Mycobacterium tuberculosis. *PLoS ONE,* 3**,** e1502.

RUTLEDGE, R. G. & COTE, C. 2003. Mathematics of quantitative kinetic PCR and the application of standard curves. *Nucleic Acids Res,* 31**,** e93.

SALKIN, D. & WAYNE, L. G. 1956. The bacteriology of resected tuberculous pulmonary lesions. I. The effect of interval between reversal of infectiousness and subsequent surgery. *Am Rev Tuberc,* 74**,** 376-87.

SAVIOLA, B. 2010. All Stressed Out: Mycobacterial Responses to Stress. *In:* MENDEZ-VILAS, A. (ed.) *Current Research, Technology and Education Topics in Applied Microbiology and Microbial Biotechnology*. Badajoz, Spain: Formatex Research Center.

SCHAECHTER, M., INGRAHAM, J. L. & NEIDHARDT, F. C. 2006. *Microbe,* Washington, D.C., ASM Press.

SCHENA, M. 2003. *Microarray Analysis,* Hoboken, New Jersey, Wiley-Blackwell.

SCHLESINGER, L. S., HULL, S. R. & KAUFMAN, T. M. 1994. Binding of the terminal mannosyl units of lipoarabinomannan from a virulent strain of Mycobacterium tuberculosis to human macrophages. *J Immunol,* 152**,** 4070-9.

SCHNAPPINGER, D., EHRT, S., VOSKUIL, M. I., LIU, Y., MANGAN, J. A., MONAHAN, I. M., DOLGANOV, G., EFRON, B., BUTCHER, P. D., NATHAN, C. & SCHOOLNIK, G. K. 2003. Transcriptional Adaptation of Mycobacterium tuberculosis within Macrophages: Insights into the Phagosomal Environment. *J Exp Med,* 198**,** 693-704.

SCHOREY, J. S., CARROLL, M. C. & BROWN, E. J. 1997. A macrophage invasion mechanism of pathogenic mycobacteria. *Science,* 277**,** 1091-3.

SCHWEIZER, E. & HOFMANN, J. 2004. Microbial type I fatty acid synthases (FAS): major players in a network of cellular FAS systems. *Microbiol Mol Biol Rev,* 68**,** 501-17.

SENARATNE, R. H., SIDDERS, B., SEQUEIRA, P., SAUNDERS, G., DUNPHY, K., MARJANOVIC, O., READER, J. R., LIMA, P., CHAN, S., KENDALL, S., MCFADDEN, J. & RILEY, L. W. 2008. Mycobacterium tuberculosis strains disrupted in mce3 and mce4 operons are attenuated in mice. *J Med Microbiol,* 57**,** 164-70.

SHAPIRO, H. M. & HANSCHEID, T. 2008. Fuchsin fluorescence in Mycobacterium tuberculosis: the Ziehl-Neelsen stain in a new light. *J Microbiol Methods,* 74**,** 119-20.

SHELINE, K. D., FRANCE, A. M., TALARICO, S., FOXMAN, B., ZHANG, L., MARRS, C. F., BATES, J. H., CAVE, M. D. & YANG, Z. 2009. Does the lipR gene of tubercle bacilli have a role in tuberculosis transmission and pathogenesis? *Tuberculosis (Edinb),* 89**,** 114-9.

SHELLIS, R. P. 1978. A synthetic saliva for cultural studies of dental plaque. *Arch Oral Biol,* 23**,** 485-9.

SHERMAN, D. R., VOSKUIL, M., SCHNAPPINGER, D., LIAO, R., HARRELL, M. I. & SCHOOLNIK, G. K. 2001. Regulation of the Mycobacterium tuberculosis hypoxic response gene encoding alpha -crystallin. *Proc Natl Acad Sci U S A,* 98**,** 7534-9.

SHERRATT, A. L. 2008. *Lipid Bodies in Mycobacteria*. PhD, University of Leicester.

SHI, L., SOHASKEY, C. D., KANA, B. D., DAWES, S., NORTH, R. J., MIZRAHI, V. & GENNARO, M. L. 2005. Changes in energy metabolism of Mycobacterium tuberculosis in mouse lung and under in vitro conditions affecting aerobic respiration. *Proc Natl Acad Sci U S A,* 102**,** 15629-34.

SHIMONO, N., MORICI, L., CASALI, N., CANTRELL, S., SIDDERS, B., EHRT, S. & RILEY, L. W. 2003. Hypervirulent mutant of Mycobacterium tuberculosis resulting from disruption of the mce1 operon. *Proc Natl Acad Sci U S A,* 100**,** 15918-23.

SHINNICK, T. M., VODKIN, M. H. & WILLIAMS, J. C. 1988. The Mycobacterium tuberculosis 65-kilodalton antigen is a heat shock protein which corresponds to common antigen and to the Escherichia coli GroEL protein. *Infect Immun,* 56**,** 446-51.

SHLEEVA, M., MUKAMOLOVA, G. V., YOUNG, M., WILLIAMS, H. D. & KAPRELYANTS, A. S. 2004. Formation of 'non-culturable' cells of Mycobacterium smegmatis in stationary phase in response to growth under suboptimal conditions and their Rpf-mediated resuscitation. *Microbiology,* 150**,** 1687-97.

SIMPSON, J. L., TIMMINS, N. L., FAKES, K., TALBOT, P. I. & GIBSON, P. G. 2004. Effect of saliva contamination on induced sputum cell counts, IL-8 and eosinophil cationic protein levels. *Eur Respir J,* 23**,** 759-62.

SINGH, A., JAIN, S., GUPTA, S., DAS, T. & TYAGI, A. K. 2003. mymA operon of Mycobacterium tuberculosis: its regulation and importance in the cell envelope. *FEMS Microbiol Lett,* 227**,** 53-63.

SIRAKOVA, T. D., DUBEY, V. S., DEB, C., DANIEL, J., KOROTKOVA, T. A., ABOMOELAK, B. & KOLATTUKUDY, P. E. 2006. Identification of a diacylglycerol acyltransferase gene involved in accumulation of triacylglycerol in Mycobacterium tuberculosis under stress. *Microbiology,* 152**,** 2717-25.

SMITH, C. R. 1942. Survival of Tubercle Bacilli. *American Review of Tuberculosis,* 45**,** 334-345.

SMITH, I. 2003. Mycobacterium tuberculosis pathogenesis and molecular determinants of virulence. *Clin Microbiol Rev,* 16**,** 463-96.

SOHASKEY, C. D. & WAYNE, L. G. 2003. Role of narK2X and narGHJI in hypoxic upregulation of nitrate reduction by Mycobacterium tuberculosis. *J Bacteriol,* 185**,** 7247-56.

SOUAZE, F., NTODOU-THOME, A., TRAN, C. Y., ROSTENE, W. & FORGEZ, P. 1996. Quantitative RT-PCR: limits and accuracy. *Biotechniques,* 21**,** 280-5.

SREEVATSAN, S., PAN, X., STOCKBAUER, K. E., CONNELL, N. D., KREISWIRTH, B. N., WHITTAM, T. S. & MUSSER, J. M. 1997. Restricted structural gene polymorphism in the Mycobacterium tuberculosis complex indicates evolutionarily recent global dissemination. *Proc Natl Acad Sci U S A,* 94**,** 9869-74.

SRIKANTH, P., SUDHARSANAM, S. & STEINBERG, R. 2008. Bio-aerosols in indoor environment: composition, health effects and analysis. *Indian J Med Microbiol,* 26**,** 302-12.

STALLINGS, C. L. & GLICKMAN, M. S. 2010. Is Mycobacterium tuberculosis stressed out? A critical assessment of the genetic evidence. *Microbes Infect,* 12**,** 1091-101.

STALLINGS, C. L., STEPHANOU, N. C., CHU, L., HOCHSCHILD, A., NICKELS, B. E. & GLICKMAN, M. S. 2009. CarD is an essential regulator of rRNA transcription required for Mycobacterium tuberculosis persistence. *Cell,* 138**,** 146-59.

STEAD, W. W., SENNER, J. W., REDDICK, W. T. & LOFGREN, J. P. 1990. Racial differences in susceptibility to infection by Mycobacterium tuberculosis. *N Engl J Med,* 322**,** 422-7.

STEWART, G. R., ROBERTSON, B. D. & YOUNG, D. B. 2004. Analysis of the function of mycobacterial DnaJ proteins by overexpression and microarray profiling. *Tuberculosis (Edinb),* 84**,** 180-7.

STEWART, G. R., WERNISCH, L., STABLER, R., MANGAN, J. A., HINDS, J., LAING, K. G., YOUNG, D. B. & BUTCHER, P. D. 2002. Dissection of the heat-shock response in Mycobacterium tuberculosis using mutants and microarrays. *Microbiology,* 148**,** 3129-38.

STOKES, R. W. & DOXSEE, D. 1999. The receptor-mediated uptake, survival, replication, and drug sensitivity of Mycobacterium tuberculosis within the macrophage-like cell line THP-1: a comparison with human monocyte-derived macrophages. *Cell Immunol,* 197**,** 1-9.

SULTAN, L., NYKA, W., MILLS, C., O'GRADY, F., WELLS, W. & RILEY, R. L. 1960. Tuberculosis disseminators. A study of the variability of aerial infectivity of tuberculous patients. *Am Rev Respir Dis,* 82**,** 358-69.

SUNDARAMURTHY, V. & PIETERS, J. 2007. Interactions of pathogenic mycobacteria with host macrophages. *Microbes Infect,* 9**,** 1671-9.

SUZUKI, R. & SHIMODAIRA, H. 2006. Pvclust: an R package for assessing the uncertainty in hierarchical clustering. *Bioinformatics,* 22**,** 1540-2.

SUZUKI, T., HIGGINS, P. J. & CRAWFORD, D. R. 2000. Control selection for RNA quantitation. *Biotechniques,* 29**,** 332-7.

TAKAYAMA, K., WANG, C. & BESRA, G. S. 2005. Pathway to synthesis and processing of mycolic acids in Mycobacterium tuberculosis. *Clin Microbiol Rev,* 18**,** 81-101.

TANEJA, N. K., DHINGRA, S., MITTAL, A., NARESH, M. & TYAGI, J. S. 2010. Mycobacterium tuberculosis transcriptional adaptation, growth arrest and dormancy phenotype development is triggered by vitamin C. *PLoS ONE,* 5**,** e10860.

TELLIER, R. 2006. Review of aerosol transmission of influenza A virus. *Emerg Infect Dis,* 12**,** 1657-62.

THELLIN, O., ZORZI, W., LAKAYE, B., DE BORMAN, B., COUMANS, B., HENNEN, G., GRISAR, T., IGOUT, A. & HEINEN, E. 1999. Housekeeping genes as internal standards: use and limits. *J Biotechnol,* 75**,** 291-5.

THOMPSON, R. & SMYTH, G. 2010. *Walter+Eliza Hall Bioinformatics: Compare Groups of Growth Curves* [Online]. Melbourne. Available: http://bioinf.wehi.edu.au/software/compareCurves/ [Accessed 10 August 2011].

THORREZ, L., VAN DEUN, K., TRANCHEVENT, L. C., VAN LOMMEL, L., ENGELEN, K., MARCHAL, K., MOREAU, Y., VAN MECHELEN, I. & SCHUIT, F. 2008. Using ribosomal protein genes as reference: a tale of caution. *PLoS One,* 3**,** e1854.

TIMM, J., POST, F. A., BEKKER, L. G., WALTHER, G. B., WAINWRIGHT, H. C., MANGANELLI, R., CHAN, W. T., TSENOVA, L., GOLD, B., SMITH, I., KAPLAN, G. & MCKINNEY, J. D. 2003. Differential expression of iron-, carbon-, and oxygen-responsive mycobacterial genes in the lungs of chronically infected mice and tuberculosis patients. *Proc Natl Acad Sci U S A,* 100**,** 14321-6.

TONDEUR, S., AGBULUT, O., MENOT, M. L., LARGHERO, J., PAULIN, D., MENASCHE, P., SAMUEL, J. L., CHOMIENNE, C. & CASSINAT, B. 2004. Overcoming bacterial DNA contamination in real-time PCR and RT-PCR reactions for LacZ detection in cell therapy monitoring. *Mol Cell Probes,* 18**,** 437-41.

TRICARICO, C., PINZANI, P., BIANCHI, S., PAGLIERANI, M., DISTANTE, V., PAZZAGLI, M., BUSTIN, S. A. & ORLANDO, C. 2002. Quantitative real-time reverse transcription polymerase chain reaction: normalization to rRNA or single housekeeping genes is inappropriate for human tissue biopsies. *Anal Biochem,* 309**,** 293-300.

TSUCHIYA, S., YAMABE, M., YAMAGUCHI, Y., KOBAYASHI, Y., KONNO, T. & TADA, K. 1980. Establishment and characterization of a human acute monocytic leukemia cell line (THP-1). *Int J Cancer,* 26**,** 171-6.

TUFARIELLO, J. M., CHAN, J. & FLYNN, J. L. 2003. Latent tuberculosis: mechanisms of host and bacillus that contribute to persistent infection. *Lancet Infect Dis,* 3**,** 578-90.

TUFARIELLO, J. M., JACOBS, W. R., JR. & CHAN, J. 2004. Individual Mycobacterium tuberculosis resuscitation-promoting factor homologues are dispensable for growth in vitro and in vivo. *Infect Immun,* 72**,** 515-26.

TUFARIELLO, J. M., MI, K., XU, J., MANABE, Y. C., KESAVAN, A. K., DRUMM, J., TANAKA, K., JACOBS, W. R., JR. & CHAN, J. 2006. Deletion of the Mycobacterium tuberculosis resuscitation-promoting factor Rv1009 gene results in delayed reactivation from chronic tuberculosis. *Infect Immun,* 74**,** 2985-95.

ULRICHS, T., KOSMIADI, G. A., JORG, S., PRADL, L., TITUKHINA, M., MISHENKO, V., GUSHINA, N. & KAUFMANN, S. H. 2005. Differential organization of the local immune response in patients with active cavitary tuberculosis or with nonprogressive tuberculoma. *J Infect Dis,* 192**,** 89-97.

UNIVERSITY OF THE BASQUE COUNTRY. 2010. *In silico simulation of molecular biology experiments* [Online]. Bilbao, Spain. Available: http://insilico.ehu.es/ [Accessed 14/01 2010].

UNNIRAMAN, S., CHATTERJI, M. & NAGARAJA, V. 2002. DNA gyrase genes in Mycobacterium tuberculosis: a single operon driven by multiple promoters. *J Bacteriol,* 184**,** 5449-56.

VAN BEEK, S. C., NHUNG, N. V., SY, D. N., STERK, P. J., TIEMERSMA, E. W. & COBELENS, F. G. 2011. Measurement of exhaled nitric oxide as a potential screening tool for pulmonary tuberculosis. *Int J Tuberc Lung Dis,* 15**,** 185-92.

VAN DER GEIZE, R., YAM, K., HEUSER, T., WILBRINK, M. H., HARA, H., ANDERTON, M. C., SIM, E., DIJKHUIZEN, L., DAVIES, J. E., MOHN, W. W. & ELTIS, L. D. 2007. A gene cluster encoding cholesterol catabolism in a soil actinomycete provides insight into Mycobacterium tuberculosis survival in macrophages. *Proc Natl Acad Sci U S A,* 104**,** 1947-52.

VAN RUTH, S. M., GROSSMANN, I., GEARY, M. & DELAHUNTY, C. M. 2001. Interactions between artificial saliva and 20 aroma compounds in water and oil model systems. *J Agric Food Chem,* 49**,** 2409-13.

VANDESOMPELE, J., DE PRETER, K., PATTYN, F., POPPE, B., VAN ROY, N., DE PAEPE, A. & SPELEMAN, F. 2002. Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes. *Genome Biol,* 3**,** RESEARCH0034.

VANDESOMPELE, J., KUBISTA, M. & PFAFFL, M. W. 2009. Reference Gene Validation Software for Improved Normalization. *In:* LOGAN, J., EDWARDS, K. & SAUNDER, N. (eds.) *Real-Time PCR: Current Technology and Applications*. Norfolk, UK: Caister Academic Press.

VAZQUEZ, M. & HANSLMEIER, A. 2006. *Ultraviolet Radiation in the Solar System,* Dordrecht, Netherlands, Springer.

VOSKUIL, M. I. 2004. Mycobacterium tuberculosis gene expression during environmental conditions associated with latency. *Tuberculosis (Edinb),* 84**,** 138-43.

VOSKUIL, M. I., BARTEK, I. L., VISCONTI, K. & SCHOOLNIK, G. K. 2011. The response of mycobacterium tuberculosis to reactive oxygen and nitrogen species. *Front Microbiol,* 2**,** 105.

VOSKUIL, M. I., SCHNAPPINGER, D., RUTHERFORD, R., LIU, Y. & SCHOOLNIK, G. K. 2004. Regulation of the Mycobacterium tuberculosis PE/PPE genes. *Tuberculosis (Edinb),* 84**,** 256-62.

VOSKUIL, M. I., SCHNAPPINGER, D., VISCONTI, K. C., HARRELL, M. I., DOLGANOV, G. M., SHERMAN, D. R. & SCHOOLNIK, G. K. 2003. Inhibition of respiration by nitric oxide induces a Mycobacterium tuberculosis dormancy program. *J Exp Med,* 198**,** 705-13.

VOTYAKOVA, T. V., KAPRELYANTS, A. S. & KELL, D. B. 1994. Influence of Viable Cells on the Resuscitation of Dormant Cells in Micrococcus luteus Cultures Held in an Extended Stationary Phase: the Population Effect. *Appl Environ Microbiol,* 60**,** 3284-3291.

VOYNOW, J. A. & RUBIN, B. K. 2009. Mucins, mucus, and sputum. *Chest,* 135**,** 505-12.

WALL, D. M., DUFFY, P. S., DUPONT, C., PRESCOTT, J. F. & MEIJER, W. G. 2005. Isocitrate lyase activity is required for virulence of the intracellular pathogen Rhodococcus equi. *Infect Immun,* 73**,** 6736-41.

WALTERMANN, M. & STEINBUCHEL, A. 2005. Neutral lipid bodies in prokaryotes: recent insights into structure, formation, and relationship to eukaryotic lipid depots. *J Bacteriol,* 187**,** 3607-19.

WANG, X., WANG, H. & XIE, J. 2011. Genes and regulatory networks involved in persistence of Mycobacterium tuberculosis. *Sci China Life Sci,* 54**,** 300-10.

WARNER, D. F. & MIZRAHI, V. 2006. Tuberculosis chemotherapy: the influence of bacillary stress and damage response pathways on drug efficacy. *Clin Microbiol Rev,* 19**,** 558-70.

WAYNE, L. G. & HAYES, L. G. 1996. An in vitro model for sequential study of shiftdown of Mycobacterium tuberculosis through two stages of nonreplicating persistence. *Infect Immun,* 64**,** 2062-9.

WAYNE, L. G. & KUBICA, G. P. 1986. The Mycobacteria. *In:* SNEATH, P. H. A., MAIR, N. S., SHARPE, M. E. & HOLT, J. G. (eds.) *Bergey's Manual of Systematic Bacteriology*. 1st ed. Baltimore: Williams and Wilkins.

WAYNE, L. G. & SOHASKEY, C. D. 2001. Nonreplicating persistence of mycobacterium tuberculosis. *Annu Rev Microbiol,* 55**,** 139-63.

WHEELER, P. R. & BLANCHARD, J. S. 2005. General Metabolism and Biochemical Pathways of Tubercle Bacilli. *In:* COLE, S. T. (ed.) *Tuberculosis and the Tubercle Bacillus*. Washington, D.C.: ASM Press.

WILLIAMS, A., JAMES, B. W., BACON, J., HATCH, K. A., HATCH, G. J., HALL, G. A. & MARSH, P. D. 2005. An assay to compare the infectivity of Mycobacterium tuberculosis isolates based on aerosol infection of guinea pigs and assessment of bacteriology. *Tuberculosis (Edinb),* 85**,** 177-84.

WONG, M. L. & MEDRANO, J. F. 2005. Real-time PCR for mRNA quantitation. *Biotechniques,* 39**,** 75-85.

WORLD HEALTH ORGANIZATION 2011. WHO Global Tuberculosis Control Report 2011. Geneva: World Health Organization.

WU, S., HOWARD, S. T., LAKEY, D. L., KIPNIS, A., SAMTEN, B., SAFI, H., GRUPPO, V., WIZEL, B., SHAMS, H., BASARABA, R. J., ORME, I. M. & BARNES, P. F. 2004. The principal sigma factor sigA mediates enhanced growth of Mycobacterium tuberculosis in vivo. *Mol Microbiol,* 51**,** 1551-62.

YANG, X., GAO, J., SMITH, I., DUBNAU, E. & SAMPSON, N. S. 2011. Cholesterol is not an essential source of nutrition for Mycobacterium tuberculosis during infection. *J Bacteriol,* 193**,** 1473-6.

YOUNG, D. B. & DUNCAN, K. 1995. Prospects for new interventions in the treatment and prevention of mycobacterial disease. *Annu Rev Microbiol,* 49**,** 641-73.

YUAN, Y., CRANE, D. D. & BARRY, C. E., 3RD 1996. Stationary phase-associated protein expression in Mycobacterium tuberculosis: function of the mycobacterial alpha-crystallin homolog. *J Bacteriol,* 178**,** 4484-92.

YUAN, Y., CRANE, D. D., SIMPSON, R. M., ZHU, Y. Q., HICKEY, M. J., SHERMAN, D. R. & BARRY, C. E., 3RD 1998. The 16-kDa alpha-crystallin (Acr) protein of Mycobacterium tuberculosis is required for growth in macrophages. *Proc Natl Acad Sci U S A,* 95**,** 9578-83.

ZHANG, Y. & AMZEL, L. M. 2002. Tuberculosis drug targets. *Curr Drug Targets,* 3**,** 131-54.

ZUMLA, A., MWABA, P., SQUIRE, S. B. & GRANGE, J. M. 1999. The tuberculosis pandemic--which way now? *J Infect,* 38**,** 74-9.