UNIVERSITY OF LEICESTER

Graduate School

Studies on the Smooth and Rough colony morphotypes of *Mycobacterium abscessus* and their relevance to infection transmission

Thesis submitted for the degree of

Doctor **o**f **P**hilosophy

at the **U**niversity of **L**eicester

By

Enass Abdul Kadhum Abed Al-Mkhadhree, MSc

Department of Infection, Immunity and Inflammation

University of Leicester
Leicester
UK

2020

Abstract

Mycobacterium abscessus **(*Mab*) is an opportunistic pathogen causing severe infections in patients with pre-existing lung disease, notably cystic fibrosis; it was first described in 1953 after isolation from a knee abscess (type strain *Mab* ATCC 19977). Despite its presumed environmental reservoir, recent evidence indicates patient to patient transmission.**

Mab **is divided into two colony morphotypes, Smooth (S) and Rough (R); the former are associated with high production of glycopeptidolipid (GPL) while the latter have less or no GPL production. S to R transition is based on mutations that eliminate GPL expression. It was hypothesized that Smooth and Rough (R) colony morphotypes reflect differences in cellular hydrophobicity, which in turn affect capacity for aerosol transmission.**

The properties of type and clinical S and R variants were compared. R strains grew differently in broth and biofilm assays, were consistently more hydrophobic and showed no demonstrable GPLs. However, no differences between S and R were observed in Goldberg drum aerosol or desiccation survival experiments and the ATCC S strain was clearly more resistant to UV exposure. Preparation of *ffluc* (luciferase)-expressing S and R strains facilitated the desiccation and UV assessments and enabled analysis of hydrophobicity attributable to the variants in mixed cultures. In the latter case, the S strain appeared to be as hydrophobic as the R when the two strains we mixed. It was speculated that redistribution of GPLs might be responsible for this and that this might explain the lack of difference between S and R observed in the aerosol survival experiments. Finally, genome and transcriptome analyses revealed several new and previously recognised polymorphisms associated with R transition.

Overall, little support for R strains being adapted to airborne transmission was found and a wider range of genomic targets than previously recognised appeared to underpin S to R transition.

Acknowledgement

"In the name of ALLAH, most gracious and most merciful"

I would like to express my sincere gratitude to my supervisor, Prof Mike Barer for his constant support and motivation. It would never have been possible for me to take this work to completion without his incredible support and encouragement. I would also like to thank my second supervisor Dr Helen O'Hare, with her assistance I have achieved many of the objectives I set out to do at the beginning of my journey. I want to thank the Ministry of Higher Education and Scientific Research (MOHESR) for offering the fund for this study and without their financial support, this research would not have been possible. My sincere thanks goes to my progress review panel members Prof Martha Clokie and Dr Julie Morrisey, Dr Natalie Garton and Dr Richard Haigh, for providing me with critical feedback and assisting me throughout my PhD. Furthermore, I am very grateful to have been a part of lab136, and give thanks to all past and present members. Especially many thanks to Dr Heyam Aloady, Dr Robeena Farzand, Dr Hanaa Galeb, Dr Amal Hamlan, Dr Mutlaq Alshammri, Valeria Quimper, Shaikha Albatli, Malikah Sadik, Irina Elliott, Anika Wisniewska, and Malgorzata Grazyna for working together, for stimulating discussions, and for all the fun we have in my first two years of PhD. My sincere thanks also go to Dr Caroline Williams.

I would like to express my deepest thanks to my parents especially my father Abdul Kadhum Abed and my mother Soad Yasser, for all the inspiration, support and good education. My thanks also go especially my close friends and brothers Firas, Hassan, Hyder and Monthather for their belief in me and care during every stage of my life. Special thanks to my lovely sister Sarah and her family.

List of Abbreviations

ADC	Albumin-dextrose-catalase
AG	Arabinogalactan
AGI	All-glass impinger
ATP	Adenosine Triphosphate
BD	Before Desiccation
Bp	Base pair
CFU	Colony-forming unit
CL	Containment level
CSH	Cell surface hydrophobicity
DNA	Deoxyribonucleic acid
DW	distilled water
ECM	extracellular matrix
ffluc	Firefly Luciferase
GPLs	Glycopeptidolipids
GTC	Guanidinium thiocyanate
HEPA	High Efficiency Particulate Air
II	inhibition index
INF	Interferon
KDa	Kilo Daltons
Kan	Kanamycin
LAM	Lipoarabinomannan
LM	Lipomannan
LOSs	Lipooligosaccharides
LPS	Lipopolysaccharides
LTBI	Latent tuberculosis infection
Mab	*Mycobacterium abscessus*
mAGP	Mycolyl-arabinogalactan-peptidoglycan
Mas	Mycolic acids
MATH	Microbial adhesion to hydrocarbons

Mb	Million base
Med	Median
MIM	mycobacterial inner membrane
MOM	mycobacterial outer membrane
MPN	Most probable number
Mtb	*Mycobacterium tuberculosis*
NS(ns)	Non significant
NTM	Non-Tuberculous Mycobacteria
OD	Optical density
P1	Patient 1
P2	Patient 2
P3	Patient 3
P4	Patient 4
PBS	Phosphate-buffered Saline
PCR	Polymerase chain reaction
PDIMs	phthiocerol dimycocerosates
PG	Peptidoglycan
PI	Phosphatidylinositol
PMA	Propidium monoazide
PUM	Phosphate-urea-magnesium
Q	Quantile
R	Rough
R1	Rough (Patient 1)
R2	Rough (Patient 2)
R3	Rough (Patient 3)
R4	Rough (Patient 4)
R10	Rough (Patient 10)
R11	Rough (Patient 11)
RCBA	Relative Congo Red Binding Assay
RH	Relative humidity
RT	Room temperature

RIN	RNA integrity number
RNA	Ribonucleic acid
RNA-seq	RNA sequencing
RNIs	Reactive nitrogen intermediates
ROS	Reactive oxygen species
RPKM	Reads per Kilobase per Million mapped reads
RT	Room temperature
RT-qPCR	Reverse-transcription quantitative PCR
S	Smooth
S1	Smooth (Patient 1)
S2	Smooth (Patient 2)
S3	Smooth (Patient 3)
S4	Smooth (Patient 4)
S5	Smooth (Patient 5)
S6	Smooth (Patient6)
S7	Smooth (Patient 7)
S8	Smooth (Patient 8)
S9	Smooth (Patient 9)
SAT	Salt aggregation test
S-CFU	Colony-forming unit, Spread plate
SN	Supernatant
TB	Tuberculosis
TC	Total count
TF	transcription factor
TMM	Trimmed Mean of M values
TPM	Transcript per Million
UQ	Upper quartile
WGS	Whole genome sequencing
WT	Wild type

Units of Measurement

^{o}C	Centigrade
CFU	colony forming unit
g	gram
hrs	hour
µg	microgram
µl	microliter
µM	micromolar
mg	milligram
mg/ml	milligram per millilitre
ml	millilitre
mm	millimetre
mM	millimolar
min	minute
RLU	relative light unit
rpm	revolutions per minute
s	seconds
w/v	weight per volume

List of Figures

List of Tables

Table of Contents

Chapter 1

Introduction

1.1 Introduction

Mycobacterium abscessus is a rapid growing mycobacterium that is recognised as an opportunistic pathogen causing life threatening infections in patients with pre-existing lung damage. Although non-tuberculous mycobacteria are generally considered to be transmitted from environmental sources, recent studies have provided evidence that person to person transmission may occur, particularly amongst patients with cystic fibrosis. The mechanisms underpinning this likely airborne transmission are not understood. An investigation into some of the features of *M. abscessus* that potentially underpin its transmission is described in this thesis.

1.2 Mycobacteria

Mycobacteria are aerobic, non-motile, and acid-fast bacilli (0.3 to 0.5 µm in diameter)(Wayne and Kubica 1986). They are categorised as acid-fast because they resist decolourisation from acid alcohol after staining with carbol-fuchsin (Allen 1992). The acid-fastness is linked to the high lipid content specifically the mycolic acid in the cell envelope (Watanabe *et al.* 2001). Mycobacteria comprise more than 100 species ranging from harmless saprophytes to serious pathogens; they are classified into two major groups based on slow and fast growth rates *in vitro* (Figure 1.1). Slowly growing mycobacteria (SGM) include the major pathogenic species *Mycobacterium tuberculosis*, *Mycobacterium leprae*, and *Mycobacterium ulcerans*. However, the majority are rapid growing mycobacteria (RGM) and are found in environmental habitats including soil and water (Gonzalez-Santiago and Drage 2015).

M. tuberculosis (*Mtb*) is the causative agent of tuberculosis (TB), one of the ancient diseases that is still a threat to public health globally (Sandhu 2011). TB infection develops when *Mtb*-containing aerosols enter the lung, where phagocytosis by alveolar macrophages is thought to be a key first step enabling initial intracellular replication and immune sensitisation. At this initial interaction, macrophages may clear much of the pathogen but many bacilli are adapted to persist (Bermudez and Goodman 1996).

In due course sites of infection are partitioned off by the establishment of granulomas (Saunders 2000). The persistence of *Mtb* in granulomas is not fully understood but remains an important topic for further study (Peyron *et al.* 2008). One mechanism contributing to long term survival may be storage of lipids, particularly triacylglycerol (TAG), which occurs after exposure to multiple stresses (Deb *et al.* 2009).

Environmental or non-tuberculous mycobacteria (NTM) have received less attention due to their relatively infrequent isolation from human infections in the past (Wagner and Young 2004). NTM include what are generally considered opportunistic agents of human infection such as *Mycobacterium avium, Mycobacterium genavense*, *Mycobacterium abscessus* (hereafter *Mab*), *Mycobacterium marium*, and *Mycobacterium intracellulare*. *Mab* is one of the major human opportunistic pathogens causing several severe infections (Falkinham 3rd 1996).

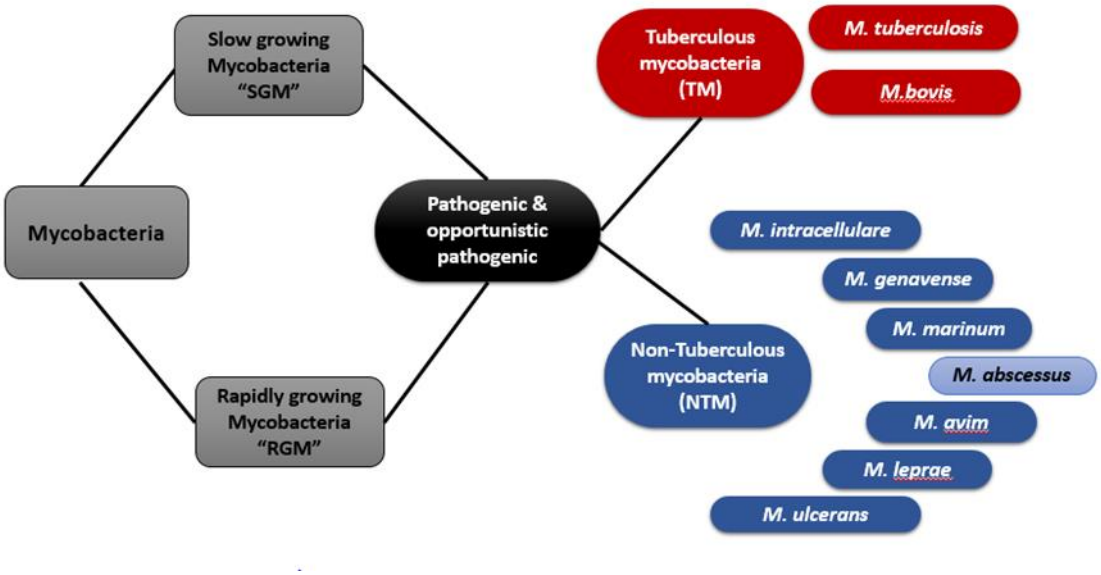


Figure 1.1: Classification of Mycobacteria in to two major groups based on their growth characteristics.

1.3 Mycobacterium abscessus

Mab was first described by Moore and Frerichs in 1953 after its isolation from a knee abscess (type strain *Mab* ATCC 19977)(Moore and Frerichs 1953). Nevertheless, *Mab* has a significant homology both genetically and phenotypically, with slow growing mycobacterial pathogens, for example, with the formation of granulomas (Rottman *et*

al. 2007). Moreover, *Mab* is closely related to *Mycobacterium chelonae* in the *M. chelonae*/*abscessus* group (Baron 1999; Brown-Elliott and Wallace 2002). This grouping reflected their similar biochemical profiles, however, DNA-DNA hybridization identified significant non-homology and separate species status was assigned (Kusunoki and Ezaki 1992; Brown-Elliott and Wallace 2002).

Mab has been further classified into three sub-species; *abscessus*, *massiliense,* and *bollettii* (Adékambi *et al.* 2004; Adékambi *et al.* 2006). Natural sources are soil and water and including municipal tap water (Carson *et al.* 1988). Members of this abscessus group can also grow in various disinfectants such as chlorine and glutaraldehyde (Steingrube *et al.* 1991; Brown-Elliott and Wallace 2002). *Mab* can be easily grown in laboratory on Middlebrook's 7H10 agar, visible growth appearing within 3-7 days (Metchock *et al.* 1999). Although they are non-flagellated and non-piliated bacteria, some members show gliding motility on solid media (Howard *et al*. 2006).

Macroscopically *Mab* are divided into two morphotypes (Figure1.2); non-cord forming with S colony morphology (designated S hereafter) and cord forming with R colony morphology (designated R) (Sanguinetti *et al.* 2001; S. Howard *et al.* 2006). The former is associated with high production of glycopeptidolipids (GPLs) while the latter have less or no GPL production (S. Howard *et al.* 2006; Pawlik *et al.* 2013). One of the major similarities between *Mab* and *Mtb* is induction of granulomas and persistence inside granulomas (Medjahed *et al.* 2010; Llorens-Fons *et al.* 2017). R morphotypes are associated more with severe illness in humans compared to S (Sanguinetti *et al.* 2001; Catherinot *et al.* 2007; Jönsson *et al.* 2013). A recent study on sequential isolates of *Mab* from nine patients over 8 years revealed that R isolates dominated at the later stage of disease, while at early stages the majority showed S morphology (Park *et al.* 2015; Llorens-Fons *et al.* 2017).

NTM can colonise humans from various environmental sources with or without any signs of infection. The significant increase in frequency of both NTM colonisation and infection have been reported in Taiwan (Lai *et al.* 2011). Furthermore, mixed morphotypes (S/R) of *Mab* have been detected, significantly associated with cystic fibrosis patients with

lung infections or chronic colonization (Sapriel *et al.* 2016). However, no association was detected between *Mab* drug resistance and its colonization (Sapriel *et al.* 2016; Khamngam *et al.* 2019). Increasing clinical recognition of *Mab,* particularly in patients with pre-existing lung disease, has stimulated investigation of disease pathogenesis of this species using animal models including mice and zebra fish embryos (Bernut *et al.* 2017). R morphotype infections develop severe disease and high mortalities in both mice and zebra fish embryos, while S morphotypes are unable to produce any signs of infection (Byrd and Lyons 1999).

Mab are reported to switch between S and R morphotypes (Figure 1.2) permitting them to adjust between a colonizing and a virulent phenotype (Falkinham *et al.* 2001; Kuo *et al.* 2011).

Figure 1.2: Smooth (S) and Rough (R) morphotypes of *Mab*.
The figure shows magnified colonies after 6 days of incubation at 37°C on 7H10 agar**.**

1.2.1 *Mab* infections: Laboratory studies

Mab is currently considered as an emerging opportunistic pathogen (Jönsson *et al.* 2013) causing disease both in immunocompetent and immunosuppressed patients (Rhoades *et al.* 2009). Beside several pulmonary diseases, this pathogen also causes dermatologic and soft tissue infections, which usually follow surgical interventions (Bechara *et al.* 2010). The virulence of this bacterium has been linked to cell wall composition specifically presence or absence of GPL (Ripoll *et al.* 2007). In 1999, Byrd and Lyons demonstrated that S strains with GPL in their cell wall are less virulent than the R variants

lacking GPL (Byrd and Lyons 1999). Additionally, *Mab* lacking GPL are capable of replication in macrophages and induce cytokine production. The major differentiating feature between the disease and non-disease causing mycobacteria is their ability to survive and replicate within the macrophages and dendritic cells and to prevent phagosome fusion with lysosomes (Roux *et al.* 2016). This suggested that the R morphotype is more virulent (Catherinot *et al.* 2009; Davidson *et al.* 2011).

Mab establishes chronic disease that is linked to the granuloma formation and caseous lesion production (Cullen *et al.* 2000; Catherinot *et al.* 2007; Catherinot *et al.* 2009) and the organism is considered as a pathogenesis model enabling investigation of the mechanisms underpinning mycobacterial infections including TB (Medjahed *et al.* 2010). However, the immunopathological mechanisms of *Mab* pathogenesis are yet to be understood.

One of the major challenges in understanding of *Mab* pathogenesis, is development of a suitable infection model to study both acute and chronic stages of the infection (Bernut *et al.* 2017). Currently, the zebrafish embryo is considered particularly useful. Zebrafish embryos show many similarities to the human immune system and importantly offer optical transparency that allows a sophisticated sequential visualization of the infection process (Bernut *et al.* 2017).

R and S *Mab* morphotypes can grow in macrophages, however, they produce different phenotypes during infection. The former rapidly infect macrophage phagosomes with up to 30 % of cells loaded with aggregates of more than 5 bacterial cells at 3 hours leading on to early host cell lysis, while the S morphotype bacilli remain within macrophages (Byrd and Lyons 1999; Roux *et al.* 2016). This unique feature of R morphotypes reflects the tendency of cells to grow in close apposition leaving no space and resulting in large clumps (Sánchez-Chardi *et al.* 2011). Macrophages infected with R morphotypes burst causing a release of bacilli out of the cells, while macrophages with S morphotype infection remain intact.

With the use of zebra fish embryos, it is now possible to see the bacilli growing in large numbers and forming cords in a manner unachievable with macrophage or neutrophil infection (Bernut *et al*., 2016). In the zebra fish model of infection, the S morphotype was not infectious and the R morphotype did not produce cords suggesting impaired replication (Halloum *et al.* 2016). Patients infected with R strains remain asymptomatic until inflammation supervenes resulting in clinical symptoms (Kreutzfeldt *et al.* 2013). Moreover, infections caused by this bacterial morphotype are particularly problematic due to its resistance to several antibiotics (Jeon *et al.* 2009). Some major infections caused by *Mab* include community acquired pulmonary diseases, disseminated, localised, and nosocomial infections, and also infections in cystic fibrosis patients.

1.2.1 *Mab* infections: Clinical infections

Mab infections generally progress very slowly. Risk factors include bronchiectasis, cystic fibrosis, tuberculosis and sarcoidosis. Females who are white, non-smokers and older than 60 years showing no previous history of lung disease also constitute significant patient groups (Brown-Elliott and Wallace 2002; Koh *et al.* 2007).

Disseminated infections involve multiple nodular skin lesions and positive blood cultures (Wallace *et al.* 1983; Wagner and Young 2004). In contrast, infection following skin trauma and contact with a contaminated objects (e.g. soil, water, clinical instruments) produce localised infections such as soft tissue abscesses or bone and joint infections (Moore and Frerichs 1953; Wallace *et al.* 1983; Wagner and Young 2004).

Cystic fibrosis (CF) is an autosomal genetic disorder in which mutation in a protein named cystic fibrosis transmembrane conductance regulator (CFTR) has occurred. This gene encodes a chloride channel which is situated in the membrane of the epithelial cells in exocrine glands. These channels are responsible for the transport of ions and fluids both in and out of the epithelial surfaces (Amaral 2005). Most CF patients have chronic lung diseases due to disturbance in mucociliary clearance of their airways and this predominantly accounts for their morbidity and mortality (Cullen *et al.* 2000; Stern and Rosenstein 2000). The problematic pathogens in these patient groups include: *Staphylococcus aureus*, *Pseudomonas aeruginosa*, *Haemophilus influenzae* and

environmental mycobacteria (Cullen *et al.* 2000). Before 1990, infections of CF patients with environmental mycobacteria were uncommon. Recently, in CF patients prevalences of 4-20 % of environmental mycobacterial infections have been reported , with up to 18 % caused by *Mab* (Griffith *et al.* 2007). Such patients show a wide range of presentations, ranging from asymptomatic colonisation to severe disease with reduced lung function (Sanguinetti *et al.* 2001; Griffith 2003; Griffith *et al.* 2007).

Presence of *Mab* is problematic, as this organism is naturally highly resistant to multiple antibiotics and difficult to treat. In many cases antibiotics only reduce the bacterial load and eradication is not achieved (Cullen *et al.* 2000; Griffith *et al.* 2007). The intrinsic and acquired resistance of *Mab* to commonly used antibiotics limits the chemotherapeutic options for clinicians (Nessar *et al.* 2012). Intrinsic resistance is attributed to a combination of the permeability barrier of the complex multilayer cell envelope, drug export systems, antibiotic targets with low affinity and enzymes that neutralize antibiotics in the cytoplasm (Nessar *et al.* 2012). However, the acquired resistance has only been observed for aminoglycosides and macrolides, which is caused by mutation of genes encoding the antibiotic targets (Ramaswamy and Musser 1998; Ramaswamy *et al.* 2003; Louw *et al.* 2009).

1.3 *Mab* transmission

Slow growing NTM are transmitted from a variety of different habitats. *M. avium* and *M. intracellulare* aerosolise from water to air by droplet formation. Bacilli at the water surface are ejected in droplets formed by turbulence (Parker *et al*. 1983). *M. avium* pulmonary infection has been associated with exposure to aerosols generated by showers (Falkinham 2003) and several studies indicate mycobacterial transmission associated with aerosols generated in hot tubs and spas (Embil *et al.* 1997; Kahana *et al.* 1997; Mangione *et al.* 2001; Rickman *et al.* 2002).

The actual mode of transmission of *Mab* in to the lungs of new hosts remains unclear, however, a number of clinical cases indicate droplet inhalation (Falkinham 2003; De Groote *et al.* 2006). Exposure to aerosols containing mycobacterial cells appears to be a prominent feature of outbreaks of respiratory disease (Falkinham 2003).

Multiple outbreaks caused by *Mab* have been reported associated with contaminated objects. One major outbreak was reported in the USA and involved 87 people with abscesses due to an unlicensed injectable medicine (Galil *et al.* 1999). Several other documented outbreaks have been related to contaminated bronchoscopes, contamination of samples in mycobacteriological laboratories and by contaminated hospital water systems (Zhang *et al.* 1997; Lai *et al.* 1998; Wallace Jr *et al.* 1998a; Blossom *et al.* 2008).

Aerosols and fomites are both linked to the transmission of outbreak strains (Bryant *et al.* 2016). Inhalation of bacteria into the lung facilitates colonisation in the lung airways leading on to invasive lung infection (Thomson *et al*., 2007).

Transmission of *Mab* occurs indirectly by the contact between an infected person and the environment, e.g. contaminated objects. Dissemination of *Mab* through the contaminated injections that led to intramuscular infections (Khermosh *et al.* 1979; Teenaged 1996). Moreover, several other studies demonstrated that infections of soft tissues occurs by cutaneous implantation of contaminated contraceptives and injections (Alfa *et al.* 1995; Fox *et al.* 2004). Prospectively *Mab* infection was observed in an immunocompetent patient after micrographic surgery, which was also linked to the contaminated surgical tools (Fisher and Gloster 2005).

The reasons for increased frequency of *Mab* infection in cystic fibrosis patients are not established. However, infection has recently been shown to be enhanced in a CFTR-depleted zebrafish model (Bernut *et al.* 2019). It has also been suggested that exposure to *Mab* from biofilms in showerheads or other aerosol sources, has contributed to infections in immunocompromised patients (Feazel *et al.* 2009; Renna *et al.* 2011).

1.3.1 Aerosol transmission of *Mab*

Presence and survival of mycobacteria in aerosol are also dependent on several physiochemical conditions, for example, salts or detergents may reduce the rate of transmission of mycobacteria from water to air (Parker *et al.* 1983). *Mtb* can be transmitted from person to person via aerosol, however, very little is known about the

aerosolisation of other mycobacteria such as *M. avium*, *M. intracellulare,* and *Mab*. Previously, it was thought that environmental opportunistic mycobacteria could not be aerosolised and transmitted this way from person to person (Wolinsky 1979). However, their transmission was linked to aerosols from aqueous suspension in certain cases (Wendt *et al.* 1980; Young 1993).

Recently, two very important studies have been published concerning the global transmission of *Mab* (Bryant *et al.* 2013; Bryant *et al.* 2016). Bryant *et al.* (2013) used whole genome sequencing to demonstrate that some strains were frequently transmitted between cystic fibrosis patients, while others were not. This gave rise to suspicion that airborne transmission was involved. Falkinham also reported that airborne transmission of NTM is possible (Falkinham 2003), a feature supported by the demonstration of exhaled NTM by our group (Williams *et al.* 2014). Bryant *et al*. (2016) stated that most *Mab* infections were likely acquired through fomites or aerosols.

Person to person transmission of *Mab* is yet to be directly demonstrated. It is important to define the extent of spread of *Mab* among individuals as person to person transmission of other cystic fibrosis pathogens such as *Pseudomonas aeruginosa* has been clearly defined (Bange *et al.* 2001). Cystic fibrosis patients were found to have shared identical type of *Mab* suggesting clear possibility of *Mab* transfer from one individual to another, or a common source exposure (Aitken *et al.* 2012). Another recent study by Harris *et al*. reported that the person to person transmission of *Mab* was detected in a sibling pair (Harris *et al.* 2015).

Mab has been found in domestic water and aerosols generated by showers in the homes of patients with pulmonary infections. This further supports the theory of transmission of these bacteria via aerosols (Thomson *et al.* 2013). Importantly, direct person-to-person transmission of *Mab* by aerosol inhalation has been considered important in recent world-wide surveys (Bryant *et al*, 2013; Bryant *et al*. 2016). Given the greater *in vitro* virulence of the R morphotype and suggestions that this phenotype may aerosolise preferentially (Jankute *et al.* 2017), the significance of S to R transition for aerosol transmission has been considered further here (See section 1.3.1).

1.4 The features of Mycobacterial cell wall

Mycobacteria have a unique lipophilic cell envelope with low permeability to aqueous solutes (Draper, 1998); it consists of three layers: a typical inner plasma membrane, a complex cell wall and an outer membrane (Daffé and Draper 1997). Figure 1.3 shows the organization of the mycobacterial cell wall. It depicts the peptidoglycan layer, and the core region that is attached to the branched polysaccharide (arabinogalactan).

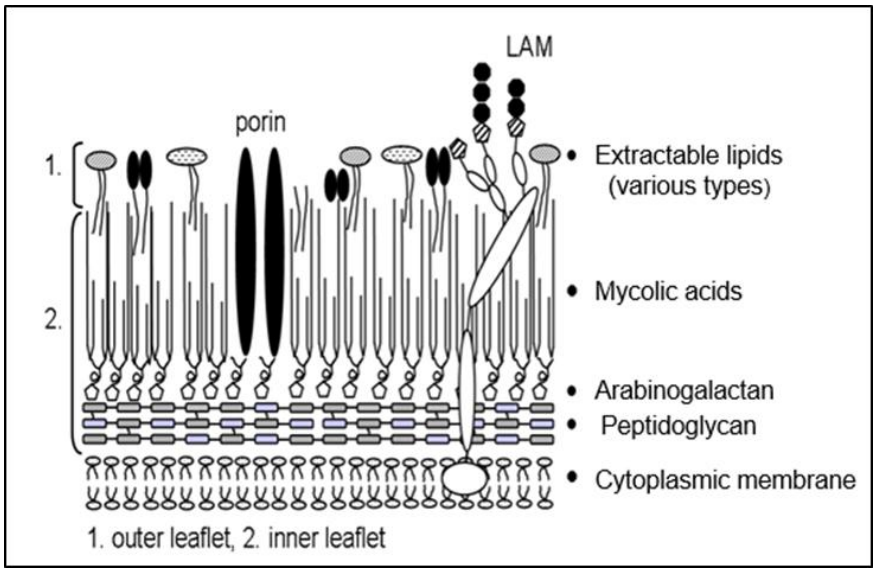


Figure 1.3: Schematic of the mycobacterial cell wall.
The figure is adapted from references (Draper 1998; Brennan 2003; Hong and Hopfinger 2004). Mycolic acids are high molecular weight (C_{60}–C_{90}) α-alkyl, β-hydroxy fatty acids, and are a major component of the cell wall of different species of mycobacteria. They constitute about 40-60 % of the cell envelope's dry weight. They are arranged in a bilayer, which allows the cell wall to be arranged in tightly packed layers (Minnikin *et al.* 1982). This establishes a permeability barrier of low fluidity (Liu, Rosenberg, and Nikaido 1995). Mycolic acids are branched fatty acids; one branch is composed of 56 carbon atoms while the other consists of 20-24 carbon atoms. Different models have been proposed for its structure and arrangement. According to the current model, the inner part is made up of highly structured mycolic acids arranged perpendicular to the cell wall and covalently linked to arabinogalactan, while the outer part is composed of other lipids (Nikaido *et al.* 1993; Daffé and Draper 1997). The outer aspect of the outer membrane displays extractable lipids and proteins that are not covalently linked to the core as well as and lipoarabinomannan (LAM) (Brennan 2003; Hong and Hopfinger 2004).

Among the above-mentioned extractable lipids, the current project will focus on GPLs, which are also known as c-mycosides. These contain a tripeptide amino-alcohol core with two monosaccharide residues linked to long chain fatty acids (Nikaido *et al.* 1993). The complex lipids that interact with other features in the outer layer, resulting in the

formation of a definite and stable surface (Daffé *et al.* 2014; Jankute *et al.* 2017). Any disruption in the components of the cell wall, particularly lipids potentially alter features, such as cording and cell surface hydrophobicity (Bloch *et al.* 1953).

1.4.1 Glycopeptidolipids (GPLs)

GPLs are a class of glycolipids produced by rapidly growing NTM such as *Mab* and *M. chelonae* (Ripoll *et al.* 2007). GPLs are composed of a lipopeptide core structure containing a 3-hydroxy or a 3-methoxy C26-C33 fatty acyl chain N linked to a tripeptide-amino-alcohol core generally made up of D-phenylalanine-D-allo-threonine-D-alanine-L-alaninol (Figure 1.4 B) (Schorey and Sweet 2008). This lipopeptide core is glycosylated with the allo-threonine glycosidically linked to a 6-deoxy-α-L-talose (6-deoxytalose) and the alaninol glycosidically linked to an α-L-rhamnose (rhamnose). GPLs are classified into alkali-stable C-types and alkali-labile serine-containing GPLs. C-types are found in saprophytic mycobacteria including *M. smegmatis*, and also in opportunistic pathogens such as *M. avium*, *M. chelonae*, and *Mab* (Schorey and Sweet 2008). On the other hand, alkali-labile serine-containing GPLs are found in *Mycobacterium xenopi* (Besra *et al.* 1993).

The nonspecific GPLs (nsGPLs) may also be O-acetylated at various locations, depending on the strain. *M. smegmatis, Mab* and *M. chelonae* produce nsGPLs that contain a 6-deoxytalose, that is 3,4-di-O-acetylated, and a rhamnose, that is 2,3,4-triO-methylated or 3,4-di-O-methylated (Patterson *et al.* 2000; Villeneuve *et al.* 2003; Ripoll *et al.* 2007).

Many clinically important NTM produce GPLs and exhibit a cording phenotype. However, only a few are capable of transitioning between S and R morphotypes. Yet, the correlation between the presence or absence of GPLs and invasive infection is clear. Brennan *et al*. reported that the *Mycobacterium avium complex* (MAC) expresses strain-specific GPLs distinct from core GPLs produced by *Mab*; the latter contains modifications by the addition of oligosaccharides (Brennan *et al.* 1981). Similarly, *Mycobacterium kansasii* also produces lipids (lipooligosaccharides) distinctive from GPLs; it forms S colonies on solid media and microscopic bacterial aggregates in broth culture, but does not display cording as found in *Mtb* and *Mab* (Tu *et al.* 2003; Nataraj *et al.* 2015). It is

worth noting that some mycobacteria, such as *M. marinum*, exhibit cording but are incapable of expressing GPLs (Hall-Stoodley *et al.* 2006). GPLs prevent the interaction of underlying surface molecules such as trehalose polyphleates (TTP) between bacteria that might play a role in cording exhibited by R variants (Llorens-Fons *et al.* 2017). Cell wall localization and chemical structure of the GPLs in the *Mab* cell envelope (Figure 1.4) was explained by Gutiérrez (Gutiérrez *et al.* 2018).

Mab GPLs are non-immunogenic while those of MAC are linked with virulence and initiating immune responses (Schorey and Sweet 2008). GPL-defective MAC mutants served to elucidate the role of GPLs in immune responses (Irani *et al.* 2004; Krzywinska *et al.* 2005; Bhatnagar and Schorey 2006). Several studies examined whether GPLs could work in the modulation of a T-helper-1 (Th1) response and have shown that GPLs can down-regulate the Th1 responses and protect the pathogen (Pourshafie *et al.* 1993; Horgen *et al.* 2000). In contrast, other studies indicate that intact GPLs were not inhibitory (Barrow *et al.* 1993) and only GPLs in which oligosaccharides were removed from the *allo*-threonine by β-elimination were capable of downregulating a Th1-type response (Tassell *et al.* 1992; Rastogi and Barrow 1994).

Additional studies have examined how GPLs direct or modulate a proinflammatory response, such as release of prostaglandins, leukotrienes, IL-1, IL-6, and TNF-α (Barrow *et al.* 1993; Pourshafie *et al.* 1993; Barrow *et al.* 1995; Horgen *et al.* 2000; Sweet and Schorey 2006). The ability of serovar-specific GPLs (ssGPLs) to stimulate the release of proinflammatory mediators appears to be structure specific, as certain ssGPLs are proinflammatory while others are not (Barrow *et al.* 1995; Sweet and Schorey 2006). This indicates that slight structural modifications can alter the way in which the GPL interacts with host-cell receptors.

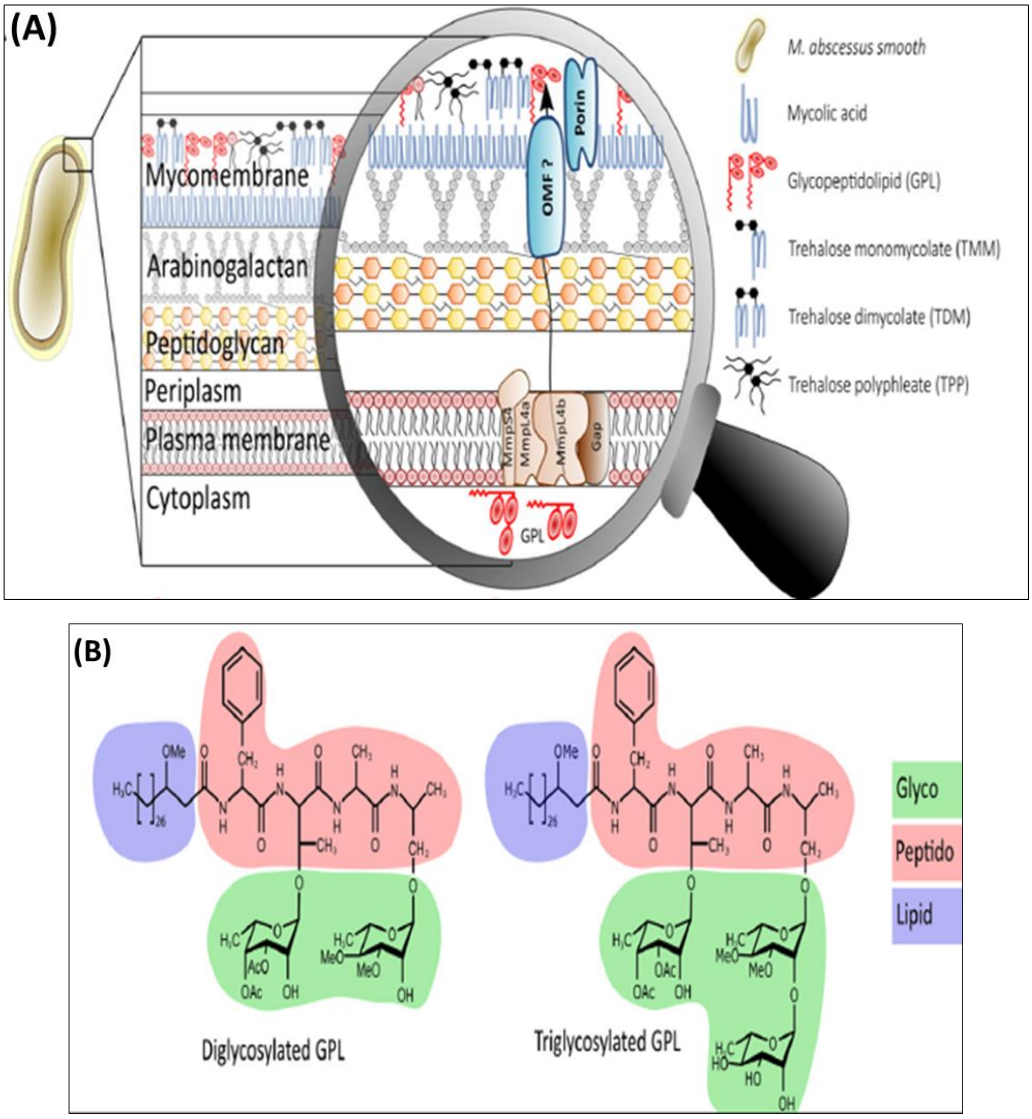


Figure 1.4: Cell wall localization and chemical structure of the GPLs in the *Mab* cell envelope.

(A) Schematic representation of the plasma membrane proteins contributing to the transport of GPL and illustrating different extractible lipids (such as GPLs) of the outer leaflets of the mycomembrane. (B) Chemical structure of the diglycosylated (apolar) and triglycosylated (polar) GPL. GPL has variations in chemical structure; only one structure is represented here. Variations occur in the lipid chain length or in the hydroxylation/O-methylation status of the various monosaccharides (Gutiérrez *et al.* 2018).

1.4.2 GPL synthesis

GPL is the most abundant glycolipid in the wild type of *Mab* cell wall and its production is controlled by the GPL encoding operon (Ripoll *et al.* 2007) (Figure 1.5). To understand the role of GPLs in mycobacterial virulence, it is important to unravel the mechanism of GPL synthesis. In 1999, Belisle *et al*. reported, for the first time, the gene cluster

responsible for the synthesis of GPLs in *M. avium*. It was described that the gene cluster was about 22 to 27 kb. This region was designated as the ser2 cluster (Belisle *et al.* 1991). Subsequent studies on individual genes in the cluster addressed the roles in the production of the lipopeptide core (Mills *et al.* 1994).

In 1999, Billman-Jacobe and colleagues isolated a mycobacterial peptide synthetase (mps) encoding gene for the first time. The mps is required for the genesis of the lipopeptide core of the GPLs in *M. smegmatis* (Billman-Jacobe, Haites, and Coppel 1999). Additionally, a gene encoding a polyketide synthase (*pks*) together with the gene *gap*, are important for the formation of the lipopeptide core and the transport of GPLs to the cell surface of *M. smegmatis* (Sondén *et al.* 2005).

In 2006, Freeman *et al*. demonstrated that pstA and pstB were similar to mps found in *M. smegmatis*, with both having a role in lipopeptide core formation (Freeman *et al.* 2006). Furthermore, a small integral membrane protein named GPL addressing protein (Gap), encoded by *gap*, is specifically required for the transport of the GPLs to the cell surface. Gap proteins provide a new model for the transport of molecules across the mycobacterial envelope, and play a critical role in mycobacterial virulence (Sondén *et al.* 2005).

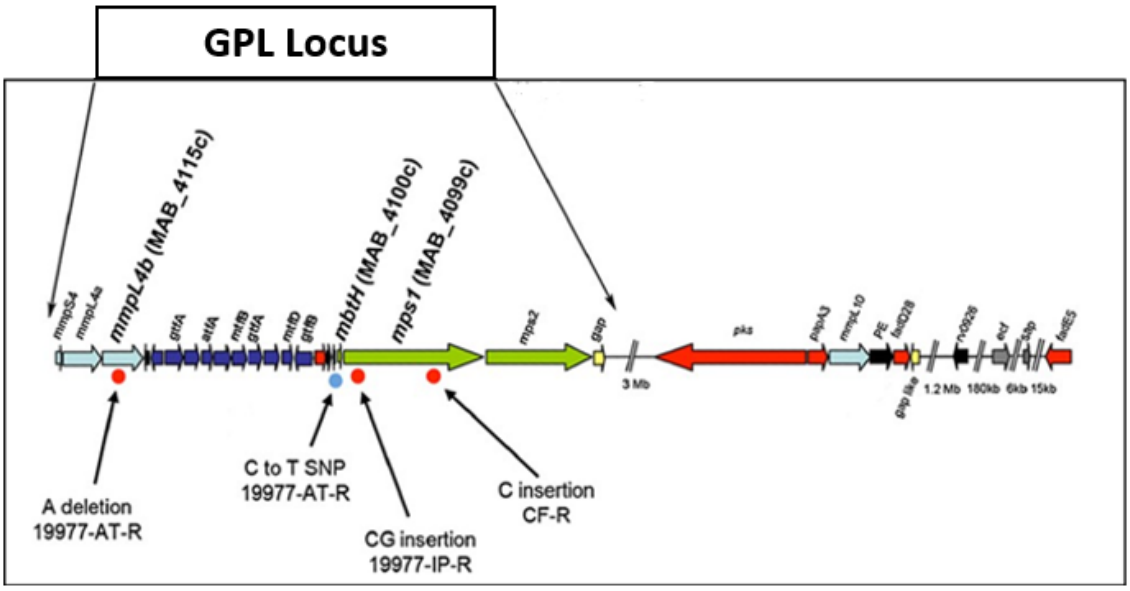


Figure 1.5: Schematic of sequence polymorphisms among three isogenic S/ R pairs of *Mab*.

This study revealed different small insertions, deletions (indels) or single nucleotide polymorphisms within the non-ribosomal peptide synthase gene cluster *mps1-mps2-gap* or *mmpl4b* in the three R variants. This is consistent with the transcriptional differences identified within this genomic locus that is implicated in the synthesis and transport of GPLs (Pawlik *et al.* 2013).

To understand the genetics of the GPL locus, S and R morphotypes were compared with isogenic S and R variants. This revealed the presence of insertions, deletions and single nucleotide polymorphisms with in the GPL locus (Pawlik *et al.* 2013). A single nucleotide deletion in *mmpL4b* and nucleotide insertions in *mps1* are reported in three R variants, compared to their isogenic S variants. Furthermore, considerable levels of difference have been observed at the transcriptomic level between S and R isogenic strains. Insertion polymorphisms in the *mps1* sequence abolish the transcription of *mps1, mps2* and *gap,* and GPL production in the R morphotypes (Pawlik *et al*. 2013). Disease stages can also affect the mutation rate in the GPL locus, as mutations in *mps2, mmpL4a and mmpS4* are found in the R morphotypes isolated from chronic stages of infection (Park *et al.* 2015). These mutations within the GPL encoding gene cluster abolish GPL production, resulting in the transition of S to the R morphotypes (Medjahed and Reyrat 2009; Nessar *et al.* 2011a).

A study by Medjahed *et al*. in 2009 demonstrated that mutations within the 22 kb region of the GPL locus affect the production of GPLs on the cell surface. Complete deletion of *mmpl4b* abolishes GPL production, leading to the production of R colonies (Medjahed and Reyrat 2009). Point mutations in MmpL4a at Tyr842 or MmpL4b at Tyr854, presumably are responsible for the proton-motive force of the MmpL proteins, and also affect GPL production promoting the transition from the S to R morphotype (Bryant *et al.* 2016). Apart from GPLs, mycobacteria contain a large variety of other unusual lipids, including trehalose polyphleates (TPP), which play an important role in the organization of the cell wall as well as contributing to pathogenesis (Burbaud *et al.* 2016).

1.4.3 Trehalose polyphleates (TPP)

TPPs are a family of trehalose-based lipids that were first described and characterized in *Mycobacterium phlei* as trehalose, which were acylated with long polyunsaturated fatty

acids (called phleic acids because they were first described in *Mycobacterium phlei*) (Asselineau *et al.* 1972; Asselineau and Montrozier 1976). A polyketide synthase gene is one of the genes responsible for producing TPP (Burbaud *et al*. 2016). Recently, TPPs have been detected in *Mab*, *M. smegmatis*, *M. avium*, and other NTMs; moreover, those from *M. smegmatis* have been structurally characterized (Burbaud *et al.* 2016)**.** TPPs are associated with formation of cords and clumps in *Mab* R strains (Llorens-Fons *et al.* 2017). It is also proposed that TPPs contribute both to the structural organization of the cell envelope and virulence.

1.5 Hydrophobicity

It is widely recognised in bacteriology that S and R colony types are associated, respectively, with lesser and greater cellular hydrophobicity (Minnikin *et al.* 2015) and this feature has been addressed systematically with *Mab* here.

Hydrophobicity has been defined from both a thermodynamic and phenomenological view. When polar compounds are shifted to aqueous media, a large positive change in heat capacity occurs (Edsall 1992). Some amino acid side chains of proteins are analogous to polar molecules in water (Kauzmann 1959). Breslow states that 'the hydrophobic effect is the tendency of nonpolar species to aggregate in water solution so as to decrease the hydrocarbon-water interfacial area'. The hydrophobic bonds form from the tendency for the association between two apolar molecules which is more favourable compared with water (Breslow 1991). Bacteria and other microorganisms use hydrophobicity to stick to abiotic surfaces. However, the contribution of hydrophobic interactions in pathogens' initial adhesion to tissues needs to be investigated in more detail (Doyle 2000).

1.5.1 Mycobacterial cell surface hydrophobicity

The cell envelope is critical for the mycobacterial physiology, primarily because it is linked to many essential processes. These include the protection of the bacterial cell from hostile environments, mechanical resistance of the cells, transport of solutes and proteins, adhesion to receptors. The hallmark of mycobacteria is their abundance in lipid, constituting up to 40 % of the dry weight of the tubercle bacillus (Brennan and

Goren 1979; Daffé and Draper 1997). The mycobacterial cell wall contains up to 60 % of lipids, as compared with some 20% for the lipid-rich cell walls of Gram-negative microorganisms (Brennan and Goren 1979). These lipids include the unusually long-chain fatty acids (mycolic acids, MA) covalently linked to the cell wall polysaccharide arabinogalactan (AG) and also esterifying trehalose, as well as the numerous classes of exotic compounds typifying the *Mycobacterium* genus (Chiaradia *et al.* 2017). To these lipids have been attributed GPL many of the biological properties of mycobacteria bacillus (Brennan and Goren 1979; Daffé and Draper 1997). Additionally, variabilities in cell surface lipids have been reported in different mycobacterial species. The notable examples are the presence of GPL in cell surface of S morphotype of *Mab* and its absence in *Mtb*, while phenolic glycolipid is only present in *Mtb* strains (Minnikin *et al.* 2015; Gutiérrez *et al.* 2018).

A previous study has reported that *Mtb* has greater cell surface hydrophobicity than many other mycobacterial species (Minnikin *et al.* 2015). The slow growth of *Mtb* may reflect the presence of its rich, waxy outer layer, which reduces permeability to the cell wall (Brennan and Nikaido 1995). This lipid-rich outer layer also allows bacteria to adhere to solid surfaces. This property might help mycobacteria in terms of persistence, and resistance against removal from their niche (Mangion *et al.* 2001). It was noted that the R variant of *Mab*, which does not produce polar glycopeptidolipids, is likely to be more hydrophobic than the S strain (Jönsson *et al.* 2007; Minnikin *et al.* 2015). Moreover, another study reported that the R colony *Mab* exhibited more cell surface hydrophobicity than the S morphotype when hexadecane partitioning assays were applied (Viljoen et al., 2018).

There is a link between cell surface hydrophobicity and the aerosolisation of environmental mycobacteria, as has been previously indicated with *M. intracellulare* by Parker (Parker *et al*. 1980). In this study, a higher number of *M. intracellulare* cells were found than *Mycobacterium scrofulaceum* in aerosols from laboratory experiments and natural seawater. In addition, via microscopy analysis, they noticed that while *M. intracellulare* formed aggregates, *M. scrofulaceum* did not. The hydrophobic nature of *M. intracellulare* was the reason attributed to this aggregation (Parker et al. 1983).

Jankute and colleagues demonstrated that the enhanced cell surface hydrophobicity of *Mtb* is due to changes in the composition of the lipids in its cell envelope, which may have led to successful aerosol-transmission (Jankute *et al.* 2017). The *Mab* R morphotype has previously been reported as more hydrophobic than the S, which might facilitate the propensity for aerosol transmission (Jankute *et al.* 2017).

The current project explored possible links between cell surface hydrophobicity and transmission of *Mab* via aerosols, and investigated the roles of two morphological variants, the R (GPL-) and S (GPL+).

1.6 Mycobacterial lipid bodies

Burdon first observed lipid bodies (LBs) in mycobacteria in a 1946 study, by using Sudan black to stain lipid droplets (now referred to as LBs) (Burdon 1946). LBs are intracellular, spherical, neutral lipid-filled inclusions containing triacylglycerol (TAG) and wax ester (Binjomah 2014). LBs are cytoplasmic organelles involved in the storage and processing of lipids and are present in most cell types and organisms. Very little is known about triacylglycerol (TAG) accumulation in NTM; only a few studies have focused on the model of rapid-growing species *M. smegmatis* (Garton *et al.* 2002; Dhouib *et al.* 2011).

Neutral lipids are uncharged lipids such as TAG and wax esters. They accumulate to make lipid bodies (LBs) when the growth of a bacterium is limited and there is an excess amount of carbon. A study performed on *Rhodococcus* spp. showed the critical role of TAGs using a TAG lipase degradation inhibitor. When cells were able to degrade TAG, their survival during stressful conditions (e.g. desciccation or UV exposure) was not affected (Urbano *et al.* 2013). On the other hand, cell survival dropped dramatically when the degradation of TAG was inhibited. Similarly, the inhibition of TAG degradation caused a decrease of *Rhodococcus* cell survival in the presence of H_2O_2 under carbon-starvation conditions (Bequer Urbano *et al*. 2013).

LBs are also thought to play a role in aerosolisation of bacterial cells, where their contribution to buoyancy is also potentially significant. Intracellular lipophilic inclusions (ILIs; now known as LBs) can usually be detected in sputum samples containing tubercle

bacilli (Garton *et al.* 2002; Garton *et al.* 2008). Accumulated lipids facilitate the state of dormancy that allows the organism to survive and recommence replication when the host's immune system is weakened also the active infection leads to the spreading of the organism (Daniel *et al.* 2004).

1.7 Biofilm

Biofilms have been defined as “matrix-enclosed bacterial population’s adherent to each other and/or to surface or interfaces” (Costerton and Lewandowski 1997). Approximately 99.9 % of bacteria present in natural environments grow in biofilms on a wide range of surfaces (Costerton *et al.* 1978) . Biofilm formation is a multistep process, and bacteria start forming it by adhering to a solid surface. Once the bacteria form a stable association with the surface, they begin to excrete a slimy and sticky substance that allows bacteria to attach to surfaces and establish a three-dimensional biofilm (Donlan, Skog, and Byrne 2012; Watnick and Kolter 2000). Upon maturity, the biofilm becomes flat or mushroom-shaped (Stoodley *et al*. 2002). These three developmental stages of biofilm were found in motile bacteria such as *Escherichia coli* and *Vibrio cholerae*, as well as non-motile species such as *S. aureus* and Mycobacterium spp. (Hall-Stoodley and Lappin-Scott 1998).

1.7.1 Mycobacterial Biofilm

Mycobacterial biofilm studies report a little different structure which include cells that adhere to the hydrophobic solid surfaces (attached) and floating mats (pellicles) on the surface of a media of liquid culture, and also to the (planktonic) cells in the liquid (Binjomah 2014) (Figure1.6).

A number of non-tuberculous mycobacteria such as *M. smegmetis*, *M. avium* or *M. chelonae*, form biofilms which allow them to survive during starvation and resist antibiotics (Ojha *et al.* 2005). As explained earlier, biofilm formation is a multi-step process involving an environmental surface and polymers secreted by bacteria (Srey *et al.* 2013). The mechanism is started by attachment cell-to-cell communication on a surface, extracellular matrix (ECM) synthesis that then surrounds fundamental cells and

ends by cellular detachment (Stoodley *et al*. 2002). In the past, mycobacterial biofilms where thought to differ structurally from other bacterial biofilms. However, it has since then been found that mycobacterial cells attach to hydrophobic surfaces and that cells were found to be surfing in the liquid culture media (Pellicles).

Biofilms might facilitate the transmission of RGM and when specific bacteria are aerosolised and inhaled, they can initiate infections. (Stoodley *et al.* 2002). Clary *et al*. showed that *Mab* can form biofilm and that these have characteristics that could promote *Mab* survival and persistence in infection (Clary *et al.* 2018).

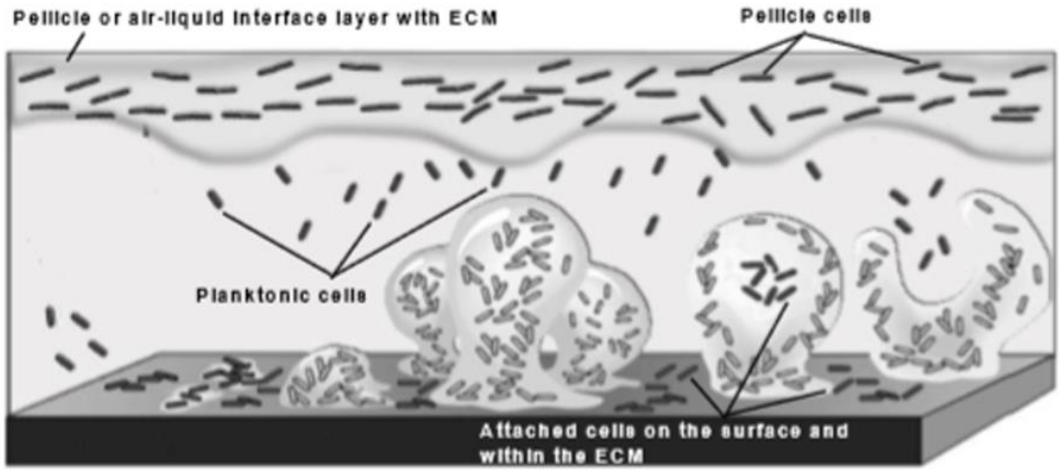


Figure 1.6: Schematic of a typical mycobacterial biofilm.
Attached are planktonic cells and the pellicle layer during the development cycle. Figure adapted from (Binjomah 2014).

The significant role of GPLs in *M. smegmatis* biofilm formation has been demonstrated (Recht *et al*. 2000). GPLs are also important during multicellular growth of *M. avium*, clarifying the important role in biofilm formation (Yamazaki *et al.* 2006). Additional studies have shown that GPL-producing strains of *Mab* form biofilm, whereas the R phenotype lacked this function (Howard *et al*. 2006b).

1.8 Stress in Mycobacteria

Stress is defined as harmful and unstable conditions that cause damage to a cell. Mycobacteria have been shown to tolerate many of the stresses faced during transmission, including starvation, osmotic pressure, exposure to ultraviolet light and

dehydration (Saviola 2010b). Mycobacteria are resistant to different stresses partly due to their thick, waxy cell wall. They resist various environmental stressors such as exposure to antibiotics (Nguyen et al. 2005).

The proper structure of the cell wall of the organism is critical for the conservation of cell shapes and for mycobacterial ultrastructure such as colony morphology (Saviola 2010a). Trehalose-dimycolate (TDM) is responsible for *Mtb* self-association into structures that appear as cords, which likely provides a physical barrier to stresses and protects from cell damage. In addition to innate resistance to stress due to a robust cell wall, mycobacteria also respond to the environmental stressors via specific gene induction (Harland *et al.* 2008). Thus, in the presence of a variety of stresses including heat shock, low pH, and low oxygen tension, bacteria will upregulate various genes needed to withstand environmental stressors. Bacteria ramp up certain stress induced pathways that are not needed or perhaps are detrimental during growth under non stress conditions (Saviola 2010a).

Clearly, mycobacteria use numerous mechanisms that can respond to environmental stressors; these systems likely developed due to selective pressure from the host's immune system (Saviola 2010). Understanding stress response systems can hopefully help to improve treatments that will prevent mycobacterial attempts to resist the host and promote the removal of mycobacteria from the host (Saviola, 2010). The mechanisms used by bacteria to tolerate stress are not yet clear. For example, transmission of *Mtb* is important for its long-term survival. Due to this, the pathogen is required to gain essential properties that are useful for transmission (Saviola 2010). It is believed that non-replicating, LB-positive cells might have the essential requirements that enhance the pathogen's survival when faced with environmental stresses that may be present during transmission when establishing a new infection (Barer and Garton, 2010).

One of the major hurdles for the bacteria during prolonged survival in the host is oxidative and nitric oxide stress. Voskuil *et al*. (2011) undertook a comprehensive study to profile whole genome expression in response to oxidative (H_2O_2) and nitric oxide (NO)

stresses (Voskuil *et al.* 2011). Total RNA isolated from *Mtb* exposed to a defined, but wide range of concentrations of H_2O_2 or NO donors was compared using oligonucleotide microarrays representing all genes of *Mtb*. Expression of a majority of the genes in the dormancy regulon was enhanced at low concentrations of NO, but without affecting growth of *Mtb*. At higher concentrations, numerous genes with functions related to oxidative stress were also induced concurrently with the dormancy regulon (Voskuil *et al.* 2011). The mechanisms that used by mycobacterium for tolerance of stress are yet to be determined.

1.8.1 Mycobacteria and Desiccation

Desiccation is the condition of the absence of water (dryness) to a level where a cell can tolerate DNA and protein damage. This type of environmental stress poses a great challenge for many bacterial cells. Loss of water leads to a decrease in turgor pressure and changes the biochemical proprieties of the cell membrane, resulting in damage to the cell envelope. The drying of a cell can also lead to conformational changes to DNA and the denaturation of intracellular proteins (Potts 1994; Wolfe and Bryant 1999).

Dehydration tolerance is the capability of cells to experience almost complete dehydration by air-drying without losing viability (Billi and Potts 2002). The greatest life-threatening abiotic stress is low water potential, which adversely affects all biological functions (Krisko and Radman 2010).(Section 4.1.5).

In an individual with active pulmonary disease, *Mtb* is expelled into the airways during the course of infection. After several generations within the human body, *Mtb* eventually arrives inside microdroplets to the outside environment, droplets then evaporate to create droplet nuclei. Therefore, it is vital for the infectivity of mycobacteria that it can resist desiccation from several hours. Mycobacteria which can persist for longer periods of time may eventually infect a larger number of hosts (Saviola 2010a).

The role of TDM was examined in model mycobacterial membranes and was found to contribute to their resistance to desiccation (Harland et al. 2008). During a transition

from a fairly wet environment to an aerosol where desiccation predominates, this desiccation resistance may theoretically benefit the bacilli. Interestingly, synthetic trehalose glycolipids also convey desiccation resistance to lipid membranes (Harland et al. 2009). These compounds can stabilise membranes that are often not well suited under conditions where drying is a concern. Therefore, TDM could give mycobacteria a competitive advantage by allowing their membranes to resist desiccation and thus permit the bacilli to survive to establish a productive infection in another host inhaling the droplet nuclei (Saviola 2010a). Some bacteria have developed adaptive mechanisms in response to desiccation such as the production of altered forms such as endospores or cysts; these more resistant forms require significant time and resources to develop (Berleman and Bauer 2004; Lebre *et al.* 2017). Moreover, desiccation stress is also associated with osmotic stress (as extracellular components in bacterial suspensions increase in concentration during drying). To counter osmotic stress, some bacteria produce small hydrophilic molecules, known as compatible solutes, in the cytoplasm. Accumulation of compatible solutes in the cytoplasm reduces the difference in osmolarity inside and outside the cells.

1.9 Genetic analyses

Genotyping and molecular typing has helped to understand the epidemiology of infectious diseases. Genetic tools have assisted our comprehension of the biology and pathogenesis of mycobacteria for more than a decade. Molecular techniques have been used to discover mycobacterial gene expression required for adaptation and sustainability during infection and dormant states. Gene over-expression studies help us to understand the role of essential genes in mycobacteria (Ehrt *et al.* 2005).

The emergence of NTM as important environmental pathogens has stimulated the search for molecular markers to identify their sources and determine their virulence mechanisms (Behr and Falkinham Iii 2009). The use of genome sequencing for a various number of NTMs has allowed for these organisms to be classified (Behr and Falkinham Iii 2009).

Multilocus sequence typing (MLST) typically lacks the level of discrimination necessary for source tracking; however, it produces unambiguous data on strains that can effectively help rule out an epidemiologic link between different patterns. Additionally, it provides important information on population structure that cannot be obtained using classical tracking methods, such as pulse field gel electrophoresis and restriction fragment length polymorphism. For example, MLST demonstrated the diversity amongst isolates of *M. avium subsp.* compared with *M. avium* subsp. hominissuis and *M. intracellulare* (Turenne *et al.* 2008).

16S rRNA gene sequencing has been suggested for the identification of NTM but has shown limitations in terms of distinguishing among some closely related species (Cloud *et al.* 2002). Based on differences, other target genes have been proposed for the identification of closely related mycobacteria (Soini and Viljanen 1997; Blackwood *et al.* 2000). Importantly, *M. abscessus*, *M. massiliense* and *M. bolletii* (have an identical 16S rRNA gene sequence) are still difficult to differentiate with sequencing of other gene targets (such as partial sequencing of *rpoB*) (Simmon *et al.* 2007). There is no clear understanding as to how genomic composition determines bacterial pathogenicity, virulence, and response to treatment (Macheras *et al.* 2009).

The genome of *Mab*, type strain ATCC 19977, was first sequenced in 2009 (Ripoll *et al.* 2009) and will be discussed further in chapter 5. This information has assisted use of *Mab* as a model for studying pathogenic mechanisms of mycobacteria and, specifically, non-tuberculosis infection (Medjahed *et al.* 2010). However, further research in the molecular microbiology and pathophysiology of *Mab* using genetic tools and knock-out mutants could provide a deeper insight into the phenotype of both the S and R morphotypes (Medjahed and Reyrat 2009).

1.9.1 Gene expression

Gene expression is the process of translating the instructions in DNA into a functional product, such as a protein. Serially, first mRNA is transcribed using the genes encoded on a DNA, this transcribed genetic information is used to produce proteins. Genes that regulate the expression of bacterial genes during an infection are important markers of

the adaptive processes in mycobacterial infections (Ehrt *et al.* 2005). Inducible gene expression systems are important methods for studying the virulence gene function and for validating drug targets (Mnaimneh *et al.* 2004).

The term "transcriptome" was first introduced in the 1990s (Piétu et al. 1999, Velculescu *et al*. 1997) and means the sum total of all messenger RNA molecules expressed from the genes of an organism at any particular moment. Previously, various technologies were developed to understand and quantify the transcriptome, which was followed by the first attempt of transcriptomic analysis on a human brain (Adams et al. 1991). This technology has been developed significantly, resulting in its widespread use and application within the biological sciences.

1.9.2 The bacterial transcriptome

Crick introduced the central dogma of molecular biology with information flowing from DNA to RNA by transcription and then to protein via translation (Crick 1970). The pattern of gene expression can change in response to environmental factors and alter organism phenotype. Thus the specific mRNA molecules present reflect both the cell's identity and its biological activities (Kukurba and Montgomery 2015). Collectively defined as the transcriptome, this profile of transcripts supplies the information needed to produce proteins required under specific conditions. Transcriptome analyses can be used to detect genes that are critical for survival under specific conditions (Brown 2016).

In this study, transcriptome analysis was performed for the total RNA isolated from *Mab* from liquid culture to investigate the expression profile of different genes involve in GPL production.

1.9.3 Gene expression analysis

Gene expression can be analysed by the isolation of transcribed RNA followed by its amplification, detection and the quantitation of the initial isolated RNA. In the last decade, the study of gene expression has been restricted to quantitative PCR (qPCR) analyses of candidate genes, or has relied on cross-species hybridization of microarrays. Nowadays, gene expression analysis is also achieved through next generation sequencing technologies (NGS) of mRNA (Metzker 2010). An example of NGS technology

is high throughput RNA sequencing (RNA-seq), also called whole transcriptome shotgun sequencing or high-throughput RNA sequencing (Morin *et al*. 2008). RNA-Seq technology has transformed the method of transcriptome profiling, allowing for the precise measurement of the quantity of RNA in a biological sample at a specific time (L. Wang *et al.* 2009; Chu and Corey 2012). Deep-sequencing technologies are used to measure the complete set of transcripts in a cell, and their quantity, to determine a specific developmental stage or physiological condition.

When determined by NGS, RNA-seq transcriptome information includes transcriptional start sites, transcript content, antisense RNAs and mRNA abundance (Filiatrault *et al.* 2011). Other technologies that can be used include Northern blotting, Real time Reverse transcription-PCR, Sanger sequencing and DNA microarray.

1.9.4 Real time Reverse transcription-PCR

Also known as RT-qPCR, this method provides unmatched sensitivity and speed when determining the levels of specific transcripts (Wong and Medrano, 2005). In this method, RNA is extracted and reverse-transcribed by using set of primers to form cDNA. Then, the segment amplification of the cDNA of interest is achieved by using target gene-specific PCR primers. The reaction is followed in RT-qPCR by detecting the characterisation of intensity fluorescence by a value called the threshold cycle (C_t), which is the time when the intensity of fluorescence is greater than the fluorescence of the background. When the initial transcription numbers of the target gene increase, the C_t value decreases (Wagner, 2013). The features of the RT-qPCR technique are used to increase the throughput of experiments and reduce the quantity of input RNA necessary; however, the ability to request only hundreds of known transcripts is limited at one time and does not cover extensive-scale transcriptomes (VanGuilder *et al.* 2008).

1.9.5 Sanger sequencing

Sanger sequencing is another technique which, in contrast to RT-qPCR, uses one primer instead of two. The amplification process creates copies by using one strand of the DNA as a template to amplify the region of interest. Each cycle of amplification forms one

copy in the same direction as the primer, but in the following cycles, it cannot be used as a template. In the sample, the original template of DNA is the source of all amplification in the reaction. Thus, enough copies of the original the DNA template must be included at the start of the process, in order for the successful amplification of the target sequence, which can then be visualised by the automated sequencing equipment (Heather and Chain 2016).

1.10 Aim

The major aim of this study was to investigate the possible relationships of the S and R morphotypes of *Mab* to its aerosol transmission and to better understand the genetic polymorphisms associated with S to R transition. Two hypotheses were proposed:

Hypothesis 1: Bacterial surface hydrophobicity is a key factor in determining the propensity of cells to enter aerosols and alterations in this property may influence physicochemical and biological factors that impact the transmission of *Mab*.

Hypothesis 2: Hydrophobic *Mab* cells are more resistant to stresses encountered during aerosol transmission.

1.11 Objectives

1. To characterise phenotypic features of S and R morphotypes of *Mab*, including growth kinetics, hydrophobicity, lipid and LB content and biofilm formation. (Chapter 3).
2. To determine whether S to R morphotype switching can be enhanced by exposure to antimicrobials and culture on detergent containing medium. (Chapter 3).
3. To compare survival of S and R morphotypes in aerosols and through exposure to desiccation and to UV radiation. (Chapter 4).

4. To investigate the genetic basis of S to R switching by WGS and transcriptional analyses. (Chapter 5).

Chapter 2

Materials and Methods

2.1 Mycobacterial strains

Bacterial strains used in this work are summarised in Table 2.1. *Mycobacterium abscessus* ATCC19977 as (lab control strains /S and R)) were obtained from the Belgian Coordinated Collections of Microorganisms/Institute of Tropical Medicine, Antwerp (BCCM/ITM).

Table 2.1: Bacterial strains used in this study

Strain	Subspecies	Description	Source
Mycobacterium abscessus ATCC19977 Lab control strains (S and R)	*Mycobacterium abscessus* subspecies *abscessus*	Rapid-growing/pathogenic mycobacterium	C04975 BCCM/ITM
Mycobacterium abscessus	*Mycobacterium abscessus* subspecies *abscessus*	Eleven clinical isolates were obtained from CF patients (For four clinical isolates, Smooth and Rough morphotypes were purified from the original mixed colony profile, two isolates showed pure Rough colony morphotype, while five isolates showed pure Smooth colony morphotype)	Leicester Royal Infirmary 2016-2017
Mycobacterium smegmatis	*Mycobacterium smegmatis* mc^2 155	Rapid-growing/ non-pathogenic mycobacterium	Lab 136 archive stocks

2.2 Culture media and reagents

All chemicals and media used were obtained from Sigma-Aldrich (Poole, Dorset, UK) or Fisher Scientific (Loughborough, Leicestershire, UK), and the media for bacterial culture were obtained from Becton Dickinson Biosciences (Oxford, UK). Polypropylene centrifuge tubes (Falcon 15 and 50 ml) were from Greiner Bio-One (Stonehouse, UK), universal tubes, 1.5 ml and 2 ml microfuge tubes were from StarLab (Milton Keynes, UK). Sterile 2-3mm glass beads were added to the cultures for both phenotypes S and R to avoid clumps. The sterilisation of media and reagents was achieved by autoclaving at 121°C at 15 psi for 15 min unless otherwise stated.

2.2.1 Growth media

2.2.1.1 Middlebrook 7H9-ADC-Tween 80 broth

Middlebrook 7H9 medium was prepared by adding 2.35 g of broth powder to 450 ml of deionised water containing 1.25 g glycerol. After sterilisation via autoclaving, the suspension was stored in the dark at room temperature (RT). At the time of use, the broth was supplemented with 10 % (v/v) Albumin-dextrose-catalase (ADC) and 10 % (w/v) Tween 80 to a final concentration of 0.05 % (w/v).

2.2.1.2 Middlebrook 7H10-OADC agar

7H10 agar was prepared by dissolving 7.6 g of agar powder in 360 ml of de-ionised water containing 2.5 g glycerol. After sterilisation via autoclaving, the prepared media was stored at room temperature in the dark. The agar was always supplemented with 10 % (v/v) Oleic acid-albumin-dextrose-catalase (OADC).

2.2.1.3 Mueller Hinton Broth

Mueller Hinton Broth media was prepared by adding 21 g of broth powder in 1000 ml of deionised water. After sterilisation via autoclaving, the suspension was stored in the dark at room temperature (RT).

2.2.1.4 Sauton's Medium

The following components were mixed and made up to a volume of 1 litre to produce Sauton's Media, the pH was adjusted to 7.0.

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

2.2.1.5 Luria-Bertani (LB) agar

The solution was made up to 400 ml with distilled water and the pH was adjusted to 7.4 and autoclaved then stored in the dark at RT.

Bacto-tryptone	4 g
Bacto-yeast extract	2 g
NaCl	2 g
Agar-powder	1.5 % (w/v)

2.2.2 Reagents

2.2.2.1 Albumin-Dextrose-Catalase growth supplement (ADC):

Amounts to make 250 ml are described. All the chemicals were first dissolved in 150ml of de-ionised water then made up to 250 ml after dissolution was complete.

Bovine Serum Albumin fraction V	12.50 g
Sodium Chloride	2.125 g
D-Glucose	5.00 g
Catalase	10.0 mg

The solution was filter sterilised (0.2 µm filter, Nalgene, Hereford, UK) and stored at 4°C.

2.2.2.2 Oleic Albumin-Dextrose-Catalase growth supplement (OADC)

The OADC was modified from ADC supplement with addition of 13.9 ml Oleic Acid solution (1 % w/v in 0.2 M sodium hydroxide), followed by filter sterilisation and storage at 4°C.

2.2.2.3 SDS solution 10 % w/v

10 g of SDS in 100 ml distilled water was dissolved by heating at 65 °C for 20 min.

2.2.2.4 Auramine O

To prepare Auramine O, solution A (0.1 g of Auramine O powder suspended in 10 ml of 95 % (v/v) ethanol) and solution B (3.0 g of phenol crystals dissolved in 87 ml de-ionised water) were mixed together. This stain was then stored in an amber bottle at room temperature. To avoid degradation of the stain, the bottle was stored in a dark and cool place.

2.2.2.5 Nile Red

A concentrated stock solution (0.5 mg/ml) of Nile red (Sigma Aldrich) was prepared by dissolving Nile red crystals in absolute ethanol. The solution was vortexed vigorously to dissolve the Nile red crystals. The prepared solution was filtered and stored in the dark at -20°C. A fresh concentrated stock was made every 4 weeks. Prior to use, a working stock of Nile red was prepared to a final concentration of 10 µg/ml in PBS.

2.2.2.6 0.5 % (v/v) Acid alcohol

To prepare 0.5 % v/v Acid alcohol solution 0.5ml of 1M hydrochloric acid (HCl) was added to 100 ml of ethanol (70 % v/v).

2.2.2.7 10 % (w/v) Tween 80

Tween 80 is used to reduce the clumping of the mycobacteria species during growth in liquid culture. Tween 80 was prepared by dissolving 10 g of Tween 80 in distilled water to a final volume of 100 ml and final concentration of 10 % (w/v). The solution was sterilised by filtration through a 0.2 µm filter unit and stored at 4°C.

2.2.2.8 50 % (w/v) glycerol solution

50 % (w/v) glycerol solution was prepared in 50 ml of de-ionised water. The solution was sterilised by autoclaving and was stored at room temperature.

2.2.2.9 Phenol: chloroform: isoamyl alcohol 25:24:1 (v/v/v)

A commercially available mixture of phenol; chloroform; isoamyl alcohol in the ratio of 25:24:1 (v/v/v); saturated with 100 mM TRIS pH 8.0; contains ~0.1 % 8-hydroxyquinoline was used in this study.

2.2.2.10 Chloroform: isoamyl alcohol 24:1 (v/v)

A commercially available mixture of Chloroform ($CHCl_3$): isoamyl alcohol 24:1 v/v was used in the current study.

2.2.3 Buffers

2.2.3.1 Phosphate Urea Magnesium Sulphate buffer (PUM)

The phosphate urea magnesium sulphate buffer was prepared as explained previously (Stokes *et al*. 2004). Briefly, the mentioned components were dissolved in 900 ml of de-ionised water and stirred for 30 minutes.

Potassium phosphate dibasic trihydrate ($K_2HPO_4.3H_2O$)	22.2 g
Potassium phosphate monobasic anhydrous salt (KH_2PO_4)	7.26 g
Urea	8 g
Magnesium sulphate heptahydrate ($MgSO_4.7H_2O$)	0.2 g

The final pH was adjusted to 7.1 and finally the buffer was sterilised by autoclaving.

2.2.3.2 Phosphate-buffered Saline (PBS)

To prepare PBS buffer one tablet of PBS (Sigma-Aldrich, cat no: P3813) was dissolved in 200 ml of de-ionised water to give a solution with final concentration of 0.01 M phosphate buffer, 0.002M potassium chloride and 0.137 M sodium chloride at pH of 7.4. The solution was autoclaved at 121°C for 20 min.

2.2.3.3 Lysis buffer

A solution of 50 mM Tris.HCl pH8.0, 10mM EDTA, 100 mM NaCl was prepared and sterilised by autoclaving. RNAse (DNAse free) was added to a final concentration of 200 µg/ml immediately before use.

2.2.3.4 TE (Tris / EDTA) buffer:

A solution of 10 mMTris.HCl (pH8.0) and 1 mM EDTA was prepared and sterilized by autoclaving.

2.3 General methods

2.3.1 Optical density (OD)

The OD of bacterial suspensions were determined using a Sanyo SP75 UV/Vis spectrophotometer (Watford) using a wavelength of 600 nm. To ensure accuracy, 7H9 broth or PUM buffer (Phosphate-urea-magnesium) was used as a blank. The optical density (OD) of bacterial cultures were measured by transferring 1 ml of the bacterial suspension to a 1.5 ml cuvette. Next, the cuvette was sealed with autoclave tape and laboratory film and measured using the spectrophotometer. Cultures with an OD greater than 1 were diluted 10-fold prior to measuring.

2.3.2 Colony Forming Unit (CFU) Count Method

Drop, spot and spread plate methods were used to perform colony counts, which are modified methods from Miles and Misra, as surface viable cell counting methods (Miles *et al.* 1938) . Ten-fold serial dilutions of cell suspension were performed by adding 50 µl of the cell suspension to 450 µl of 7H9-ADC-Tween 80 media in 1.5 ml microfuge tubes.

Three 20 µl drops from each dilution were plated onto 7H10-OADC agar plates. Each agar plate was separated into 4 sectors and each sector was designated to one of the dilutions. Once the drops were dry, the plates were sealed with laboratory sealing film, inverted, and incubated at 37 °C until isolated colonies were visible. For the spread plate method, 100 µl was plated. The plates were allowed to dry and were then sealed with laboratory sealing film and incubated at 37 °C.

Final counts of 10-100 colonies (averaged over the three replicates of 20 µl spots) were used for the final calculation of colony forming units (CFUs), using the following equation:

$$CFU / ml = A \times D \times 50$$

A = average colony count per 20 µl spot

D = Dilution factor

2.3.3 Bacterial stock cultures for long-term storage

To prepare stock cultures, bacteria were grown to the exponential phase. The actively growing bacteria were then mixed equally at a 1:1 ratio with 50 % (w/v) glycerol solution and aliquotted in to 1.5 micro centrifuge tubes before being stored at -80 °C for future use.

2.3.3.1 The treatment of bacterial suspensions with syringe to homogenise the suspensions

To avoid and break up bacterial clumps in liquid suspensions, a blunt 25G syringe needle (Henke Sass Wolf, Germany) was used. The cell suspension was passed through the needle 5-9 times to break up the clumps.

2.3.4 Cultivation of *Mab*

Mab S and R variants are hazard group 2 microorganisms and must be handled with the appropriate safety measures to avoid any possible aerosol generation and exposure. Therefore, in this study, all experiments with this strain were carried out in Class II microbiological safety cabinets. Freeze-dried samples of *Mab* strain ATCC 19977 (wild type) S, R variants along with clinical isolates, 250 ul of culture (OD 0.8-1) were used and

mixed gently. The cultures were inoculated on 7H10 agar plates as well as in 7H9 broth at 37 °C for three days with shaking at 130 rpm overnight. Then, this was used to inoculate universal tubes containing 5 ml of 7H9-ADC-Tween-80 broth. Sterile 30 (2-3 mm glass) beads were added, and the cell suspension was vortexed vigorously until no obvious clumps were detected. This was then incubated at 37°C with shaking at 130 rpm overnight. The cultures were used to inoculate 3 x 125 ml conical glass flasks, each containing 25 ml of 7H9-ADC-Tween 80 broth to an OD_{600} of 0.005. Sterile 30 (2-3 mm glass) beads were also added to these, after which they were incubated as above until grown to the exponential or stationary phase, depending on the purpose of experiments.

2.4 Fluorescence microscopy

2.4.1 Immobilisation of *Mab* onto slides for microscopy

Mab cells were immobilised onto glass slides using the Bellco slide micro-chamber system described by Walker (Walker *et al.* 1994). Figure 2.1 illustrates the system. To prepare the slides, 50 µl of cell suspension was dispensed into the wells of the chambers, followed by centrifugation at 1000 × *g* for 10 minutes in an IEC Centra-4 X centrifuge (International Equipment Company, Dunstable, and Bedfordshire, UK). The supernatant was carefully removed using a pipette. The coverslip was left to dry in a dark place. The smears were mounted in 10 % of glycerol (w/v) in PBS and sealed with transparent nail varnish.

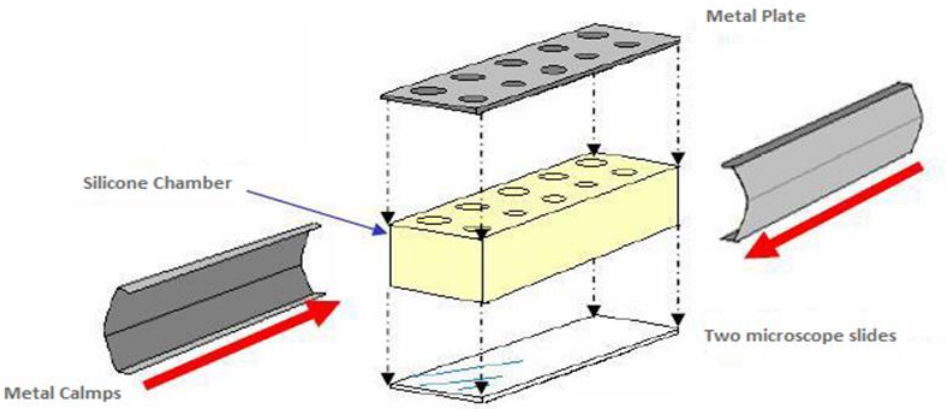


Figure 2.1: Schematic representation of the assembly of the Belco micro-chamber.

For producing bacterial monolayers for microscopy. Figure is adapted from Sherratt (Sherratt 2008).

2.4.2 Labelling of *Mab* with Nile red

Nile red staining was used to stain the immobilised smears of bacterial cells. The immobilised cells were covered in stain solution (10 µg/ml) for 15 minutes, followed by a water rinse. The slides were mounted using 10 % glycerol (v/v in PBS) and then, the slides were air-dried in the dark to avoid Nile red degradation from light, after which glass coverslips were applied and sealed with colourless nail varnish.

2.4.3 Image recording

All stained slides were prepared and imaged on the same day using an inverted microscope (Nikon Eclipse Ti) with the X100 objective and immersion oil. Images were recorded with a 12/10 bit, high speed DS-U3 CCD camera Build 831 (Nikon Corporation, Japan) using the Nikon NIS Elements Imaging Software. The two sets of filters which were used for the epifluorescence microscopy are shown in Table 2.2.

Table 2.2: Filters sets which were used for epifluorescence microscopy.

Filter Block (Chroma)	Excitation	Dichroic mirror	Emission
Auramine (31015)	460± 25nm	500DCLP	550 ± 25nm
ET-Texas Red/mCherry (49008)	560 ±20nm	585LP	630 ± 37.5nm

2.4.4 Image analysis

The image analysis was carried out by using the method established by Andrew Bell at the University of Leicester (BELL 2013). The method calculates fluorescence intensity per unit all area data using the ImageJ-based software (National Institutes of Health, Bethesda, Maryland). To measure the intensity of fluorescence from the region of interest, phase images were identified by ImageJ through thresholding. The percentage of lipid bodies and acid fastness of the cells was measured by calculating the change in

fluorescence intensity of each cell per unit area, along the spline of the cell. The ratio of peak to trough intensity values were automatically determined by the ImageJ software. 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 signals (BELL 2013). Finally, the data was exported to Graph Pad Prism (version 7.0) to generate graphs and for statistical analysis.

2.5 Graphical representation and Statistical analyses of significance

The statistical analyses and significance tests in this study were performed using Prism 7 (Graph Pad Software, Inc.) Statistical software. The significance symbols used throughout this research is presented graphically in Table 2.3.

Table 2.3 Display of significance used in this thesis.

Significance	Denoted by
P ≤ 0.05	*
P ≤ 0.01	**
P ≤ 0.001	***
P ≤ 0.0001	****
P > 0.05	ns

Chapter 3

Phenotypic characterisation of *Mycobacterium abscessus*

3.1 Introduction

It has been well established that S and Rough (S and R) colonies of *Mab* are, respectively, moist and dry, with the latter show cording similar to *M. tuberculosis* (Byrd and Lyons 1999). *Mab* is fast growing and easily grown *in vitro* in many synthetic or complex laboratory media. The colonies are usually visible in a few days compared to several weeks for slow-growing mycobacterium (Jankute *et al.* 2017).

Hydrophobicity is the tendency of nonpolar lipid compounds to aggregate in a water solution by decreasing the total area of contact between hydrocarbon and water. (Belisle *et al.* 1991). The mycobacterial cell envelope has several polar and non-polar lipids (detailed in section 1.4) determining cell surface hydrophobicity (see details in section 1.5) (Minnikin *et al*. 2015). There are several techniques to measure cell surface hydrophobicity (CSH) of microbes including microbial adhesion to hydrocarbons (MATH), contact angle measurement, the salt aggregation test, microsphere adhesion to cells and hydrophobic interaction chromatography (Rosenberg *et al.* 1980). Microbial adhesion to hydrocarbons is widely deployed in this regard and has been used in many studies such as those directed to, bioremediation (Baldi *et al*., 1999), pathogenesis (Lee and Yii, 1996) and food contamination (Lee *et al*., 2017). It is worth noting that very little is known about the cell surface hydrophobicities of R and S variants of *Mab* (Costerton *et al.* 1995). However, the hexadecane-based hydrophobicity examinations for other mycobacterial species showed that rough mycobacterial phenotypes are generally more hydrophobic than S (Minnikin *et al*. 2015).

In the current study, the Relative Congo Red Binding (RCRB) assay was used as a second method to measure CSH; Congo Red is an amphiphilic dye that interacts with the lipophilic regions of the cell envelope (Cangelosi *et al.* 1999; Etienne *et al.* 2002; Etienne *et al.* 2009). Previously, Congo Red has been used to differentiate between pathogenic and non-pathogenic strains of bacterial species (Kay *et al*.1985; Qadri *et al*.1988). Additionally, Congo Red has also been used to evaluate the presence of hydrophobic cell surface proteins in enterovirulent *Shigella* spp. (Ambalam *et al.* 2012).

Among various non polar lipids, GPLs are produced by most of the fast growing NTM and their presence affects important phenotypes such as cell envelope permeability (Ripoll *et al.* 2009). Previous studies indicate that presence or absence of GPLs is central respectively to the S and R morphologies of *Mab* (Schorey and Sweet 2008). Moreover, GPLs also have a key role in cell surface hydrophobicity. Apart from GPLs, TAGs that are found in the cell envelope in addition to LBs (see section 1.6) (Garton *et al.* 2002) may contribute to CSH. However, whether *Mab*, like *Mtb,* has the mechanism that needed to obtain and store lipids as TAG, it remains untreated (Viljoen *et al.* 2016) . There is a gap of information about the LB accumulation and their variations in the S and R variants and this aspect has been addressed here.

Biofilm formation is one of the natural and important features of bacteria and the great majority of bacteria including the NTM *M. smegmatis*, *M. avium* and *M. chelonae*, are capable of this growth pattern (Costerton *et al.* 1995; Ojha *et al.* 2005). Biofilm formation appears to contribute to the persistence of pathogenic bacteria in chronic pulmonary diseases such as cystic fibrosis (Clary *et al.* 2018), where the presence of slime-encased bacterial biofilms on the surface of tissues and implanted devices is observed (Palmer and White 1997). Thus, biofilm growth experiments on R and S variants of the type strain and clinical isolates of *Mab* from cystic fibrosis patients are reported here.

Mab is not only pathogenic but also resistant to many antimicrobials (Maurer *et al.* 2014). While this may result from poor cell envelope permeability, many potential resistance-encoding genes have been recognised in the *Mab* (Ripoll *et al.* 2007). The major threat posed by this pathogen is due to its high natural resistance to a broad range of antibiotics, which is of particular concern in public health in situations (Nessar *et al.* 2012; Bryant *et al.* 2013). Antibiotic treatment of *Mab* infection is guided by *in vitro* antimicrobial susceptibility testing (AST) and usually includes a macrolide, such as Clarithromycin (CLR) Orazithromycin (AZM), Amikacin (AMK), and various combinations of Moxifloxacin (MOX), Linezolid (LIN), Tigecycline (TIG), Imipenem (IMI), or Cefoxitin (CEF) (Griffith *et al.* 2007; Brown-Elliott *et al.* 2012) . Clarithromycin is an antibiotic used to treat several bacterial infections which includes pneumonia, strep. throat, and skin

infections (Langtry and Brogden 1997). Clarithromycin produces its antimicrobial effect by inhibition of intracellular protein synthesis and also causes significant morphological changes in the cell wall of *M. avium* complex (Peters and Clissold 1992). It became the drug of choice for *Mab* infections, and therapeutic successes were reported (Griffith *et al.* 1993; Maxson *et al.* 1994). Ciprofloxacin is also an antibiotic that is used to treat bacterial infections and functions by inhibiting DNA gyrase (Drlica and Zhao 1997). Reports showed that R morphotypes are predominant in chronic infections (Nessar *et al.* 2012; Sapriel *et al.* 2016; Kham-ngam *et al.* 2019) indicating their better ability to withstand the selective pressure of host immune system and also selective pressure of antibiotic treatments. Therefore, in this study antibiotics were used to confirm their effects on S to R transition.

Sodium dodecyl sulfate (SDS) is an ionic detergent that is suitable for the rapid destroyed of biological membranes. It is considered a key component of several reagents that used to purify nucleic acids because SDS abilities to quickly disrupt the tissue and to inhibit both deoxy ribonuclease (DNase) activity and RNase (Farrell Jr 2006; Farrell 2010). The toxic influences of SDS on bacteria have been informed (Kumar *et al.* 2014). The increased quantity of SDS in the cytoplasm led to misfolding of denatured proteins which could be cytotoxic (Rajagopal *et al.* 2002). The SDS impact different cell organelles and affected genes regulation in *Saccharomyces cerevisiae* (Sirisattha *et al*. 2004).

In mycobacteria, changes in the lipid content of the cell envelope have been associated with differing susceptibility to the action of SDS. For example, absence of phthiocerol dimycocerosates makes *Mtb* more susceptible to SDS compared to parent cells (Siméone *et al.* 2007). This could be due changes in the permeability of cell envelope of *Mtb*. Therefore, we used SDS in this study to figure out the effects of SDS on the genome of *Mab* and its transition from S to R morphotype.

In view of the reported tendency for the *Mab* S morphotype to switch to the R form in chronic infection (Kreutzfeldt *et al.* 2013), an initial investigation into the influence of antimicrobials and detergent on S to R switching were performed here. Clarithromycin, Ciprofloxacin and SDS were selected as a suitable stress for selection of mutants having undergone S to R transition.

3.1.1 Overall aim:

The work of this chapter focused on the phenotypic characterisation of *M. abscessus* strain ATCC 19977 S and R morphotypes obtained from BCCM/ITM. This strain was archived and used for all other experiments. Clinical isolates were obtained from four patients with cystic fibrosis.

3.1.2 Objectives

1. To study the growth kinetics of both S and R morphotypes of *M. abscessus* strain ATCC 19977.
2. To investigate the cell surface hydrophobicity using hexadecane partitioning and Congo Red binding assays.
3. To study the lipid body profile and GPL production in S and R morphotypes of the lab strain and clinical isolates.
4. To examine the biofilm formation ability of *Mab* S and R morphotypes.
5. To investigate S to R switching frequencies and the influence of antimicrobials and the detergent sodium dodecyl sulphate thereon.

3.2 Methods

3.2.1 Measurement of growth of *Mab*

For optical densities**,** a single colony from both S and R of the lab control and also clinical isolates were inoculated into 5ml 7H9 and grown (37 ^{o}C at 130 rpm) until late exponential phase (OD_{600} 0.9). The OD_{600} of culture was then adjusted to 0.08 in 7H9. 150 µl aliquots added to Bioscreen plate wells and the cultures were incubated in the Bioscreen plate reader (Labsystems, Cambridge UK) at the required temperature 37 ^{o}C for 3 days with constant shaking. The Optical of density readings were taken every 3hr.

3.2.2 CSH measurement of *Mab* using MATH technique

The hydrophobicity assay was modified from Stokes *et al.* (2004). Briefly, both S and R variants of *Mab* and clinical isolates were cultured from frozen stock (-80 °C) on 7H10 then a single colony from each sample was inoculated in the 5 ml of 7H9 (3 days). The culture was then normalised to OD_{600} 0.003 and allowed to grow to the early stationary phase OD_{600} 1.0 after one day following single colony incubation. At the respective time points, 10 ml of bacterial suspension were harvested and pelleted at 2000 x *g* for 20 min, and re-suspended in equal volume of PUM buffer (Phosphate-urea-magnesium) and the optical density was adjusted to OD_{600} 0.5. From the adjusted suspension, 1.5 ml of S, R and clinical isolates were transferred to borosilicate glass tubes (dimensions: 11 by 120 mm) in which they were incubated with 1.2 ml of hexadecane for 8 minutes at 37 °C. The suspension was vortexed (8 sec) to fully mix the culture and hexadecane then incubated for 15 min at room temperature. Finally, the lower aqueous layer was aspirated using a borosilicate glass Pasteur pipette and OD_{600} (Figure 3.4). The percentage of bacteria in the aqueous phase was calculated using the following formula: % of hexadecane adherence = 100 – (100 x (OD600 after / OD600 before hexadecane).

3.2.3 Relative Congo Red binding (RCRB) assay

The Congo Red (CR) binding assay was used for quantitative assessment of relative *Mab* hydrophobicity (Cangelosi *et al*. 1999). Similar to section 3.2.1, both S and R variants of *Mab* and clinical isolates were cultured from frozen stock to 7H10 and a single colony from each sample was inoculated in the 5ml of 7H9 and incubated at 37 °C for 2 days.

Each sample was then sub-cultured onto fresh 7H10 agar from an adjusted suspension in 7H9 at OD_{600} of 0.05. Serial dilutions were made when cultures reached late exponential phase (0.9-1.0 OD_{600}) then 100 µl from each dilution was plated onto 7H10-OADC- plate supplemented with CR dye at a final concentration of 100 µg/ml and incubated at 37 °C until colonies became visible (5 to 9) days. 9 -15 colonies were scraped off the plates and suspended in 3.5 ml sterile deionized water and OD_{600} was adjusted to OD_{600} 0.5. The suspension was centrifuged at 2000 x *g* for 10 min and the pellet was then resuspended in 200 µl acetone, vigorously vortexed, then incubated for 2 hrs at room temperature. Cells were then pelleted 4,000 x *g* for 5min. The A_{490nm} of the resultant supernatants were measured (Microplate reader model: LE808, BioTek, Dorest, UK) with acetone as a blank. The RCRB value was defined as the OD_{490} of the acetone extracts divided by the OD_{600} of the original cell suspension (Cangelosi *et al.* 1999).

3.2.4 Lipid body profile

Both S and R variants of *Mab* and the clinical isolates were grown for early visible growth on 7H10 plate. The bacterial lawn was scraped off the plate and suspended into 10 ml de-ionized water and the OD_{600} adjusted to 1.0. The cells were first immobilised onto glass slides using the Bellco slide microchamber system described by Walker *et al.* (1994) (Figure 3.6). Briefly, 50 µl of cell suspensions from each sample were dispensed into the wells of chamber followed by centrifugation at 1000 × *g* for 10 minutes in an IEC Centra-4X centrifuge (International Equipment Company, Dunstable, and Bedfordshire, UK).The supernatant was carefully removed by micropipette. The slides were kept at room temperature to air dry for 30 minutes.

3.2.5 Lipid Extraction

The method of total lipid extraction was adapted from (Burbaud *et al.* 2016). Briefly, Equivalent masses of cell pellets (0.2 g) were taken in 15 ml centrifuge tubes. An equal volume of (1:2 v/v) chloroform ($CHCl_3$): methanol (MeOH) (2ml) was added to the cell pellet and mixed using a rotating mixer for 1hr. The mixture was centrifuged 3000 × *g* for 10 minutes to pellet cells then the supernatant was carefully removed to a fresh 15ml

centrifuge tube. The extraction was repeated and the extract removed and pooled with the first chloroform ($CHCl_3$): methanol (MeOH) 1:2 (v/v) extract in the 15 ml centrifuge tube. To the combined extract was added half the total volume of $CHCl_3$ to bring the $CHCl_3$: MeOH to 2:1 (v/v). To the combined extract was added 20 % (v/v) of the total volume of water to wash the organic extracts (1.2 ml) and mixed on a rotating mixer for 5min. The two phases (organic and aqueous) were separated by centrifugation at 3000 × *g* for 10 minutes and the lower organic phase were carefully removed to a glass vial and dried under a stream of Nitrogen. The dry lipid extract was stored at -20°C. For Thin layer Chromatography (TLC) analysis, the lipid extracts were dissolved in 100 µl. $CHCl_3$: MeOH 2:1 (v/v) and 2 µl of each extract from each strain were applied to a TLC plate (Merck silica gel 60 F254, 20 x20 cm aluminum sheet) the desired solvent system and run in $CHCl_3/CH_3OH$ (90:10, v/v). The lipid spots were visualized by spraying the plates with a 0.2 % anthrone solution (w/v) in concentrated H_2SO_4, followed by heating.

3.2.6 Biofilm formation

Biofilm assay was performed using a modified 12 well tissue culture-based method (Ojha *et al.* 2008). The culture was prepared by inoculating a single colony from freshly grown bacteria on a 7H10 plate. Both S and R variants of *Mab* and the clinical isolates were grown in 12 well polystyrene tissue culture plates (Corning®Costar® Cell Culture Plates) by inoculating 3 ml of Sauton's medium with exponentially growing bacteria from 7H9 diluted 1:100. The plates were sealed with PetriSEAL™ tape properly then incubated statically at 37 °C from 9 days to 14 days. The biofilm development was visually assessed at different time points.

3.2.7 Minimum inhibitory concentration (MIC) by broth dilution method

The antibiotic susceptibility test was performed by broth micro-dilution (Lee *et al.*, 2007). Briefly, the inoculum suspensions were prepared by diluting broth culture in sterile water to the density of 0.5 McFarland standards or OD_{600} (0.08 to 0.1). The final concentration of *Mab* was 1-5 × 10^5 CFU/ml. The suspension were vortexed then dispensed into the required wells of a 96-well polystyrene (NUNC) microtiter plates containing broth (Mueller-Hinton; MH) medium and ciprofloxacin CIP (Sigma) at serial

diluted concentration of 8 µg/ml to 0.016 µg/ml. Final inoculum were between 1×10^4 to 5×10^4 CFU per well in a volume of 0.2 ml. Inoculated plates were covered and incubated at 37 °C for 5 days. MIC values were defined as the lowest concentration of drug preventing visible growth.

Clarithromycin was used at serial two-fold dilutions from 8 µg/ml to 0.03 µg/ml. The final inoculum was between 1×10^4 to 5×10^4 CFU/well. Plates were incubated at 37 °C for 5 days. All measurements were performed in triplicate at two independent occasions. For the effect of sub-inhibitory MIC, the bacteria were grown in MH broth supplemented with sub-inhibitory MIC of clarithromycin until exponential phase. To observe the effects of antibiotic on mutation, culture was diluted (10^{-5}) and 0.5 ml was plated in large bio assay dish (Fisher Scientific, Loughborough, Leicestershire, UK).

3.2.8 Investigation of selective concentration for R strain using SDS:

SDS was used as selective reagent to find a concentration at which the S strain shows susceptibility and R strains did not. Both strains were grown on medium containing SDS (0.006 to 5 % (w/v)). Both phenotypes showed growth on medium containing 0.1 or below 0.1 % (w/v) of SDS, whereas no growth was observed on medium containing more than 0.1 % of SDS. Interestingly, mutation transition of S to R was observed using 0.1 % (w/v) SDS. The mutant was archived and was confirmed for phenotypic features including cording appearance and hydrophobic nature. It was assumed that due to loss of GPL R strain would be less permeable to SDS and would show resistance.

3.3 Results

Mab strain ATCC 19977 was used as a standard reference and clinical isolates from four CF patients were compared against this.

3.3.1 Culture of *Mab* ATCC 19977 and its growth characteristics

Freeze-dried cells of *Mab* strain ATCC 19977 were obtained from BCCM/ITM. The received material was suspended in 7H9 then inoculated into both liquid and solid media (7H9 and 7H10, respectively). The agar cultures revealed a mixture of R and S colony morphologies with a majority of S. Pure growth of the separate morphologies was obtained (Figure 3.1) and archived for all subsequent experiments. The S variant grew as creamy, white, raised and shiny colonies, while the R variant colonies were flat and dry.

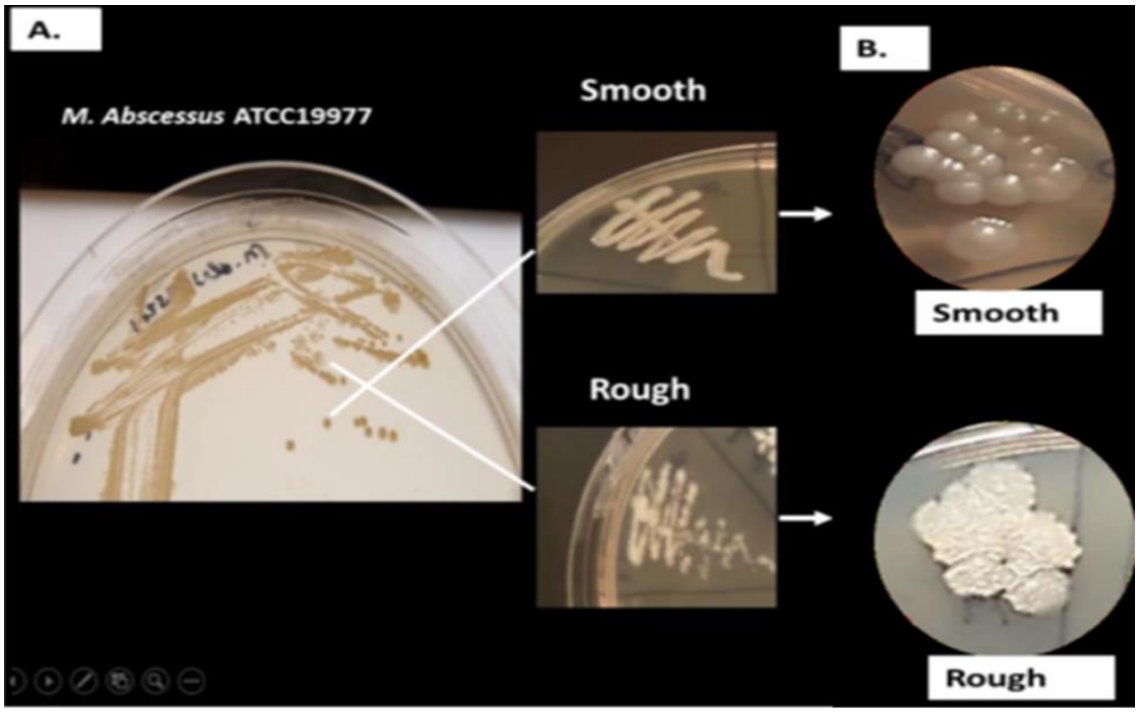


Figure 3.1: Separation of S and R colony variants obtained from *Mab* ATCC1997.

Mab usually takes 3 to 5 days at 37◦C to obtain viable growth on solid medium, however the doubling time is highly dependent on the inoculum size and medium used in broth medium (Cortes *et al.* 2010). The growth kinetics of S and R strains were investigated

first by CFU counts in the control ATCC strain, then by OD measurement in all five S and R pairs. The results are shown in Figures 3.2 and 3.3.

Control strain CFU counts showed slow initiation of growth of the R (control) whereas the S (control) showed an increase of 2 log in the first 20 hrs. S (control) viable counts increased by 8 fold from 20 hrs until 30 hrs of incubation. However, the R (control) grows smoothly and rapidly after 20hrs of lag phase and the CFU counts were increased by two logs between 20 and 30 hrs. This suggests that after 20 hrs, the S morphotype of control strain growth rate is slower than that of the R.

Similar differences were apparent in the OD readings for the control S and R strains (Figure 3.2 A). After the lag phase the control S strain showed fast growth whereas, the R strain grew significantly slower for 30 hrs and then there was rapid growth for 10 hrs followed by slower growth for the remaining ten hr of observation; stationary phase was not reached in this time (readings for optical densities were taken at the similar times as CFU determination). It is worth noting that R morphotype was clumped while growing in 7H9 broth medium that might have resulted in the lower OD and also for the greater variation between the measurements of replicates compared to higher ODs of S (control) strain. Therefore, two different techniques were used to have clear evidence of differences in the growth phenotype of the S and R morphotypes.

The growth curves of clinical isolates (Patients 1 to 4), as measured by OD, also showed significant and similar differences between the S and R morphotypes. All clinical S isolates showed extended lag phase ranging from 15 to 18 hr followed by fast growth in exponential phase. However, R isolates showed a longer lag phase and multiple growth phases by OD measurement. R1, R2, were observed to have 24 hr lag phases, the lag phase of R3 was shorter, whereas R4 showed a lag phase of 42 hrs followed by rapid growth for 30 hrs. One of the obvious differences between the growth of S and R strains was that R stains have longer lag phases, with slower growth than corresponding S morphotypes in first exponential phase and a second brief period of more rapid growth.

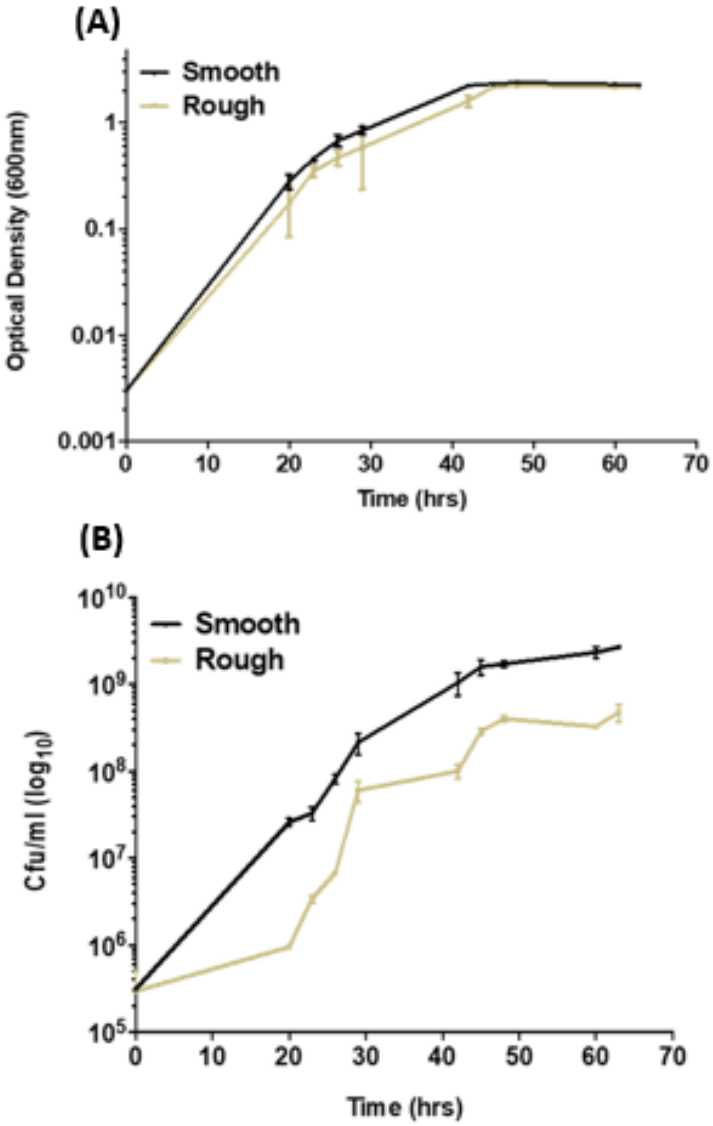


Figure 3.2 : Growth dynamics of *Mab* S and R morphotypes.
Measured over the course of 3 days with manual OD_{600} readings **(A)** and CFU/ml **(B).** The experiment was performed on three independent occasions and each time run with three technical replicates. The OD was measured from culture grown in flasks and CFU counts were performed from the same batch culture. The plots are the representative of combined data from two independent occasions and the error bars represent standard deviation of the six replicates.

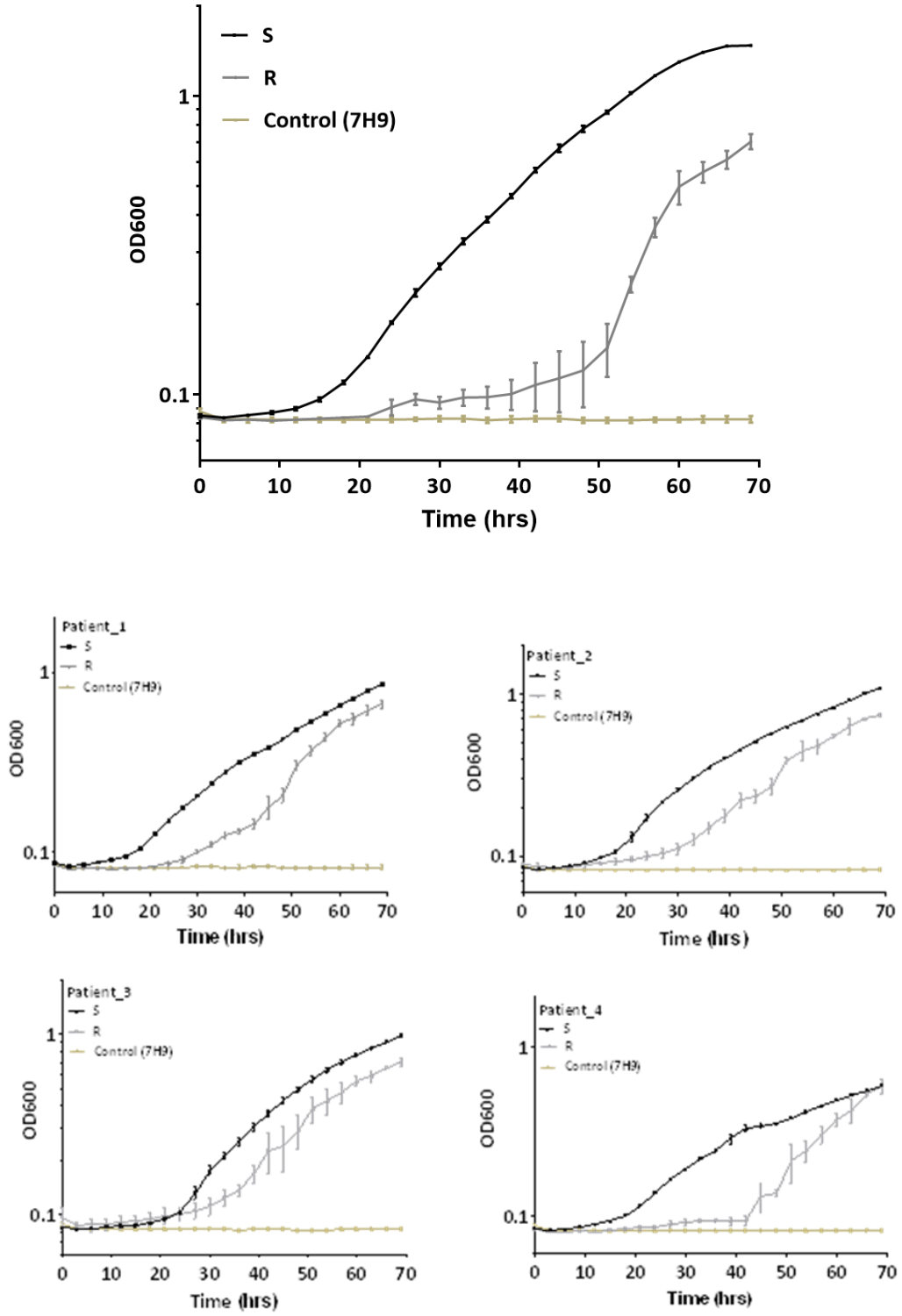


Figure 3.3 : Growth characteristics of *Mab* lab control strain (S and R morphotypes) and the clinical isolate (S and R morphotypes) of four CF patients.

Growth dynamics of the control ATTC strain and clinical isolates were tested by growing bacterial cells in 100 wells plate with incubation in a Bioscreen with measurement of optical densities every three hrs for 70 hrs. Error bars represent SD of triplicate wells from two biological, the experiment was performed two independent times with three replicates for each. The Control (7H9) represents non-inoculated wells.

3.3.2 CSH measurement using MATH and RCRB assay of *Mab*

Two complementary methods were employed to determine the relative CSH of *Mab* lab control strain S and R morphotypes and those of the clinical isolates. Firstly, hexadecane partitioning of cells between the organic hexadecane phase and an aqueous buffer was used to study cell surface hydrophobicity. The R morphotype of all study strains were significantly more hydrophobic than the S morphotype (Figure 3.4A, B), demonstrating the hydrophobic nature of R cells. In addition, pure isolates of S (5 patients), pure R (2 patients), the S-SDS and R-SDS isolates were also investigated and the results reflected those found with the lab control and other clinical isolates (Appendix I , Figure 1).

Furthermore, a RCRB assay was also used evaluate the CSH of the *Mab* control strain and clinical isolates of *Mab*. CR is an amphiphilic dye that interacts with the lipophilic region of the cell surface (Cangelosi *et al.* 1999) assessing overall hydrophobicity and was utilised as an alternate method to determine the relative hydrophobic nature of the control strain morphotypes and those of the clinical isolates. In parallel, the strains were cultivated on medium containing CR (100 μg/ml). Cultivation in the presence of CR showed uniform uptake of dye by R colonies while S colonies remained un- or poorly stained with the CR (Figure 3.4C). Absorbance measurements of acetone extracts of these reflected this also (Figure 3.4D). Results from both assays show that S morphotypes have less cell surface hydrophobicity and therefore tend to stay in aqueous phase in MATH assay and show less interaction with CR in RCRB assay. This experiment indicated a clear difference in hydrophobicity between two variants of same species.

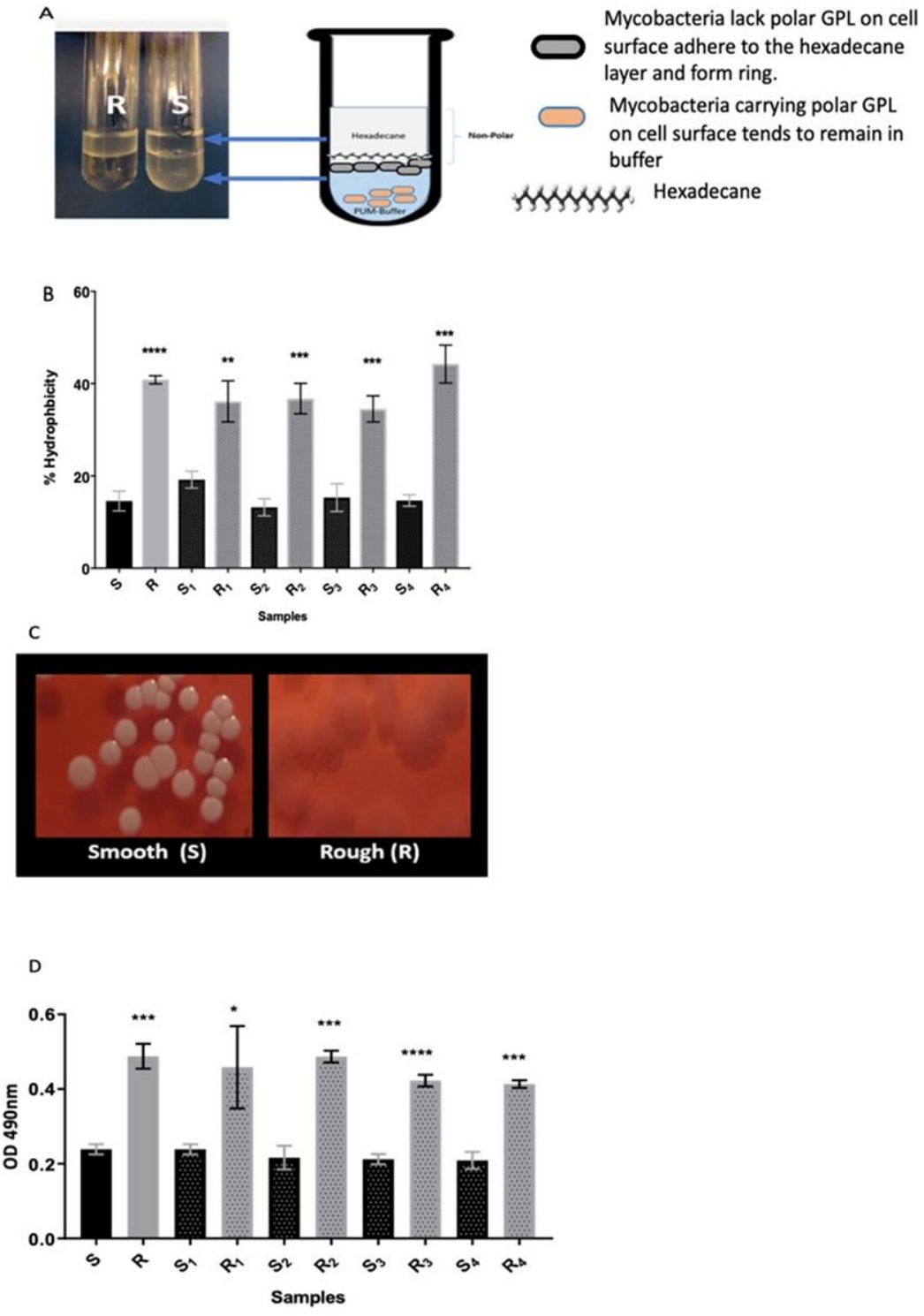


Figure 3.4 : Measurement of CSH of lab control and clinical isolates of *Mab* using MATH and RCRB assay.

(A) Hexadecane treatment of *Mab* suspended in with PUM buffer. The organic hexadecane layer shows greater turbidity in tubes containing R (left hand side) while the turbidity is greater in the aqueous PUM phase with the S morphotype. **(B)** MATH test for *Mab* lab control strains S and R

and clinical isolates (S_1, R1; S_2, R2; S_3, R3; S_4, R4) respectively. *Mab* strains were suspended in an aqueous buffer mixed with hexadecane. Following separation of the two phases, the OD_{600} of the aqueous phase was measured. Error bars signify the standard deviation of three independent experiments performed in triplicate. Statistical analysis was performed with multiple t-test using Sidak's method. **(C)** Staining of the *Mab* control strain S and R morphotypes on solid medium. **(D)** Absorbance of CR in acetone extract of colonies grown on solid media of clinical isolates and control strains of *Mab*. Error bars indicate the standard deviation of three independent experiments performed with triplicates. Statistical significance was determined by multiple t-test ($p \leq 0.05$*) ($P \leq 0.01$**) ($P \leq 0.001$***) ($P \leq 0.0001$****).

3.3.3 LB formation by *Mab*

As LB content may also be related to CSH, this was also determined in R and S morphotypes, following growth on 7H10. Nile red is a neutral lipid stain strongly soluble in lipid and highly fluorescent in a hydrophobic environment (Greenspan *et al.* 1985). Additionally, Nile red is a commonly used lipophilic stain for intracellular TAG detection in *Mycobacteria* (Garton et al. 2002). Nile Red fluorescence intensity reflects the amount of lipid within a stained cell. Nile red short excitation and yellow/gold emission (≤580 nm) wavelengths favour the detection of highly hydrophobic environments of neutral apolar lipids such as TAG in LBs, whereas longer excitation red wavelength emission (≥ 590 nm) detects all lipids, including polar lipids resulting from interactions with intracellular membrane phospholipids (Greenspan and Fowler 1985). Nile Red was used to stain lipids to perform fast screening of lipid contents in S and R variants of *Mab*. A comparison between two phenotypes of *Mab* revealed that Nile Red fluorescence intensity was similar in both control strain S and R morphotypes. However, significant variability was observed between S and R morphotypes of clinical strains (Figure 3.5A) with a tendency for higher fluorescence per unit cell area in S strains S2-4, but the converse in S1. The relative differences seen in the yellow-green fluorescence intensity values (S vs R) show R1 again being different from the other three clinical pairs (Figure 3.5B).

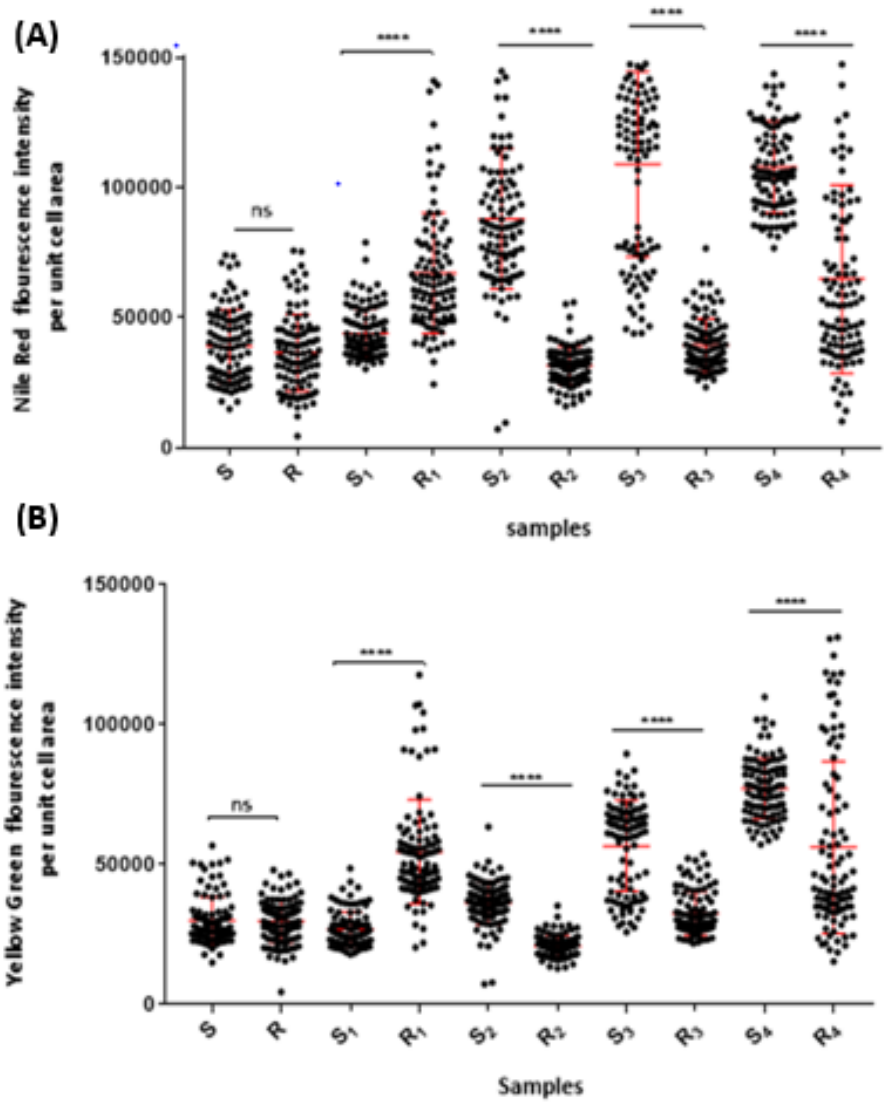


Figure 3.5: Fluorescence intensity of *Mab* stained with Nile red.
Comparison of fluorescence intensity per unit cell area between S and R variants of *Mab* control strains and the clinical isolates. **(A)** Red fluorescence emission (Ex560 / Em630) and **(B)** Yellow fluorescence emission (Ex460 / Em550). Results assessed from triplicate slides at each time point (at least 100 cells per slide were imaged) and were expressed as Nile red fluorescence intensity per unit cell area. Each individual cell is represented by one point; median and standard deviation of the measurements are indicated with red bar. Multiple t-test was used to measure statistical significance. Multiple t-test. (ns $P > 0.05$) ($P < 0.0001$****).

Fluorescence microscopy images for Nile-red stained preparation of the S and R variants of the control *Mab* strain cultured on 7H10 are shown in the Figure 3.6. Preparations were examined for both yellow-green and red flourescence emission for a better understanding of the distribution of both polar and a polar-lipids. LBs were observed in both S and R morphotypes. LB positivity as a percentage of the total cell population with LBs (%LB) in the *Mab* control strain and clinical isolates is presented in Table 1 in Appendix II.

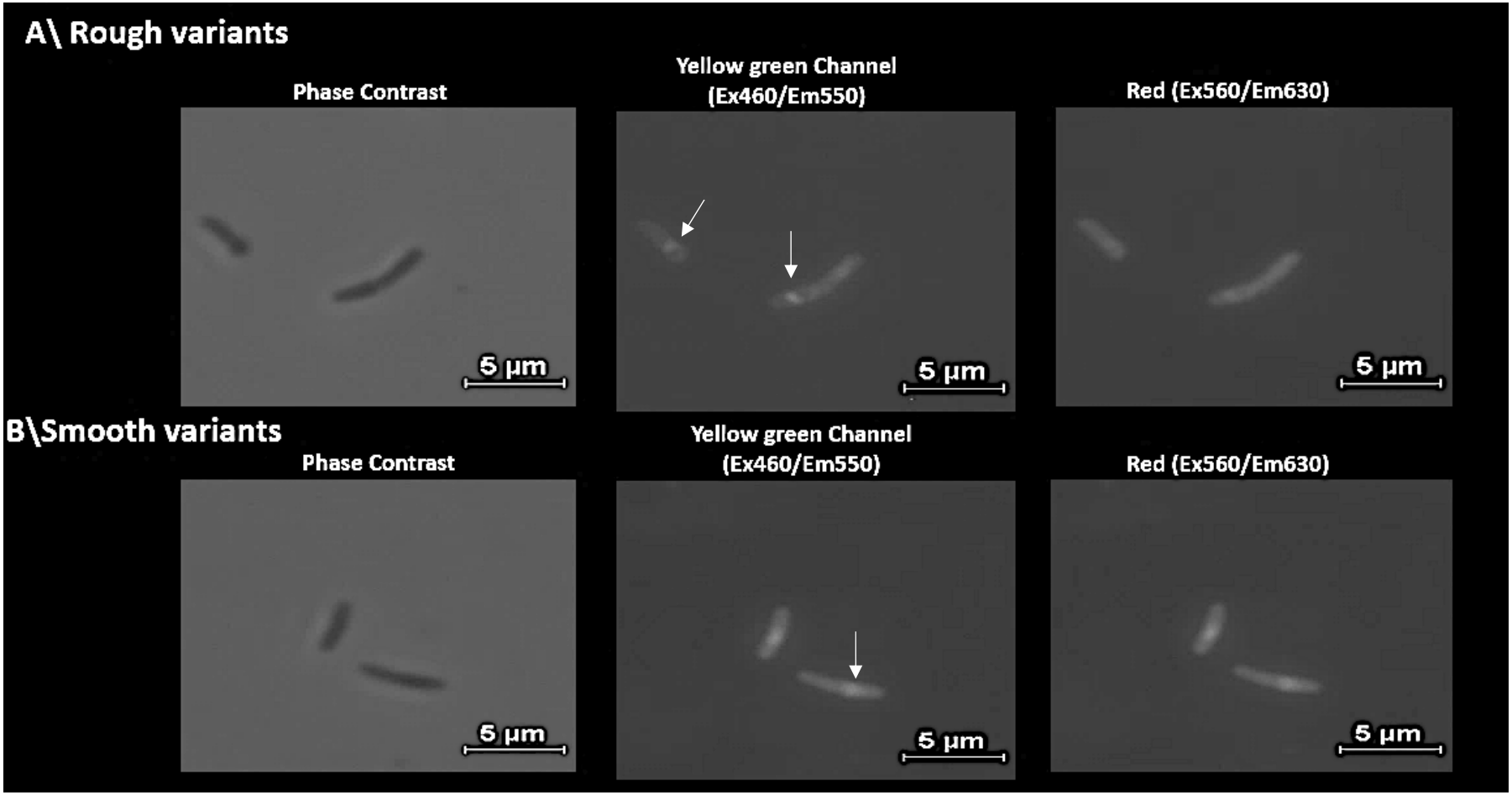


Figure 3.6: Fluorescence microscopy images for slides prepared from S and R variants of *Mab* control strains cultured on 7H10.
(A) Rough (R) variant. **(B)** Smooth (S) variant. Images from left to right are; Phase contrast, Yellow green fluorescence at Ex460/Em550 (expected to highlight more non-polar lipid such as TAG) and Red fluorescence at Ex560/Em630 (all lipids).. Arrows highlight LBs within the cells.

3.3.4 Thin Liquid Chromatography of lipid extracts to profile extractable lipids of *Mab*

Since S and R colony morphologies have been associated with the presence and absence of GPL, respectively, GPL expression was examined in both morphotypes using TLC to profile lipids in extracts. *Mab* control strain S and R morphotypes and those of four clinical isolates were tested for GPL production. All of the *Mab* S variants of the control and 4 clinical isolates expressed GPL, as expected (Burbaud *et al*. 2016), whereas R morphotypes and the clinical R isolates were deficient of GPL production (Figure 3.7). Notably, both S and R strains synthesized a lipid component putatively identified as TPP (Burbaud *et al.* 2016). However, R strains produced large amounts of TPP compared to S strains, except S4.

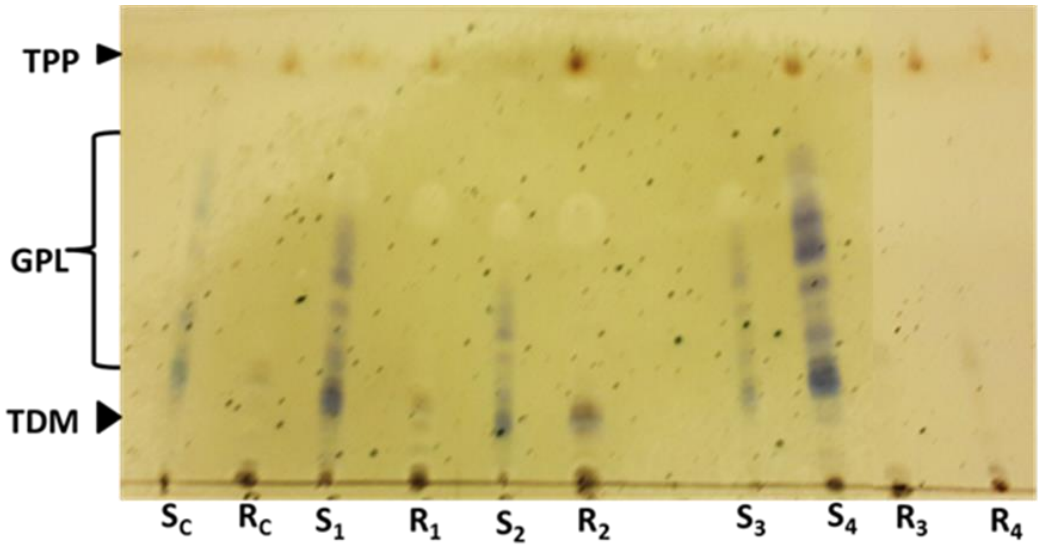


Figure 3.7: Thin layer chromatography profile (TLC) of lipid extracts of *Mab*.
To control the amount of extracts added, 0.2 g of biomass was used to extract the lipids and equivalent volumes of extract were run in $CHCl_3/CH_3OH$ (9:1, v/v). The spots were visualized by spraying the plates with a 0.2% anthrone solution (w/v) in concentrated H_2SO_4, followed by heating. 0.2 % w/v anthrone in H_2SO_4 (Burbaud *et al.* 2016). TPP – trehalose polyphelates, GPL – glycopeptidolipids, TDM –trehalose dimycolate. S; R morphotype; 1, 2, 3, 4 clinical isolates (S and R morphotypes were purified from original mixed colony profile).

3.3.5 Biofilm formation by *Mab*

Biofilm formation by the S and R phenotypes of *Mab* was studied. In conditions of static culture, both S and R variants formed of biofilm-like structures (Figure 3.8). R was able to produce more biofilm after 9 days of static incubation compared to the S strain. However, both strains had produced biofilm with three distinguishable layers after fourteen days of incubation.

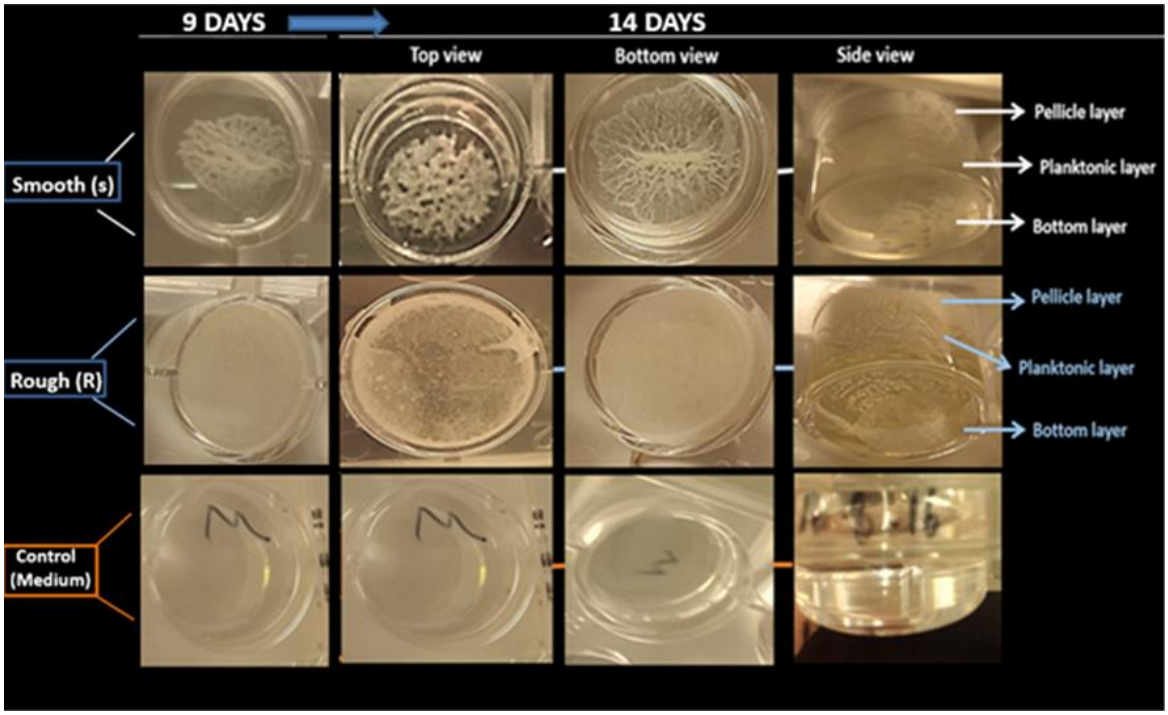


Figure 3.8: Biofilm formation.

Top view of wells from 12 well plate after 14 days *Mab* growth, S and R are showing biofilm. In the side view of the well there are clearly visible three layers, bottom, pellicle and planktonic phase formed by both variants.

3.3.6 Effects of Sub-inhibitory concentration (SIC) of antimicrobials on phenotype transition of *Mab*

One of the aims of this project is to investigate the effects of various antibiotics on the mutation that affects the colony. This project started working with the measurement of MIC of Ciprofloxacin for both S and R phenotype of *Mab*. The MIC for S was 4 µg/ml while 2 µg/ml for the R phenotype (Figure 3.9 A). Based on these findings 2 µg/ml and 1µg/ml concentration were choose as SIC of Ciprofloxacin for S and R, respectively while the measurement of MIC of Clarithromycin for both S and R phenotype of *Mab* were 0.5 µg/ml (Figure 3.9 B).

In order to determine whether exposure to antibiotics used in *Mab* treatment might influence S to R switching, sub-inhibitory concentration for S and R phenotypes of 2 and 1 µg/ml for Ciprofloxacin were selected, whereas, an SIC for Clarithromycin of 0.25 µg /ml was selected for both morphotypes (Figure 3.9). The test strains were grown in broth medium supplemented with the chosen sub-inhibitory concentrations then plated in serial dilutions onto agar medium in in large bio-assay dish. In excess of 3,000 well separated colonies were screened for each test exposure. No transitions, S to R or R to S were observed.

SDS was also used as selective reagent to find a concentration to which S strain showed susceptibility, but not the R. Both S and R strains were grown on medium containing SDS (0.006 to 5 % (w/v)). Data demonstrated that both phenotypes were able to grow on medium containing 0.1 % (w/v) or below 0.1 % (w/v) of SDS, whereas no growth was observed on medium containing more than 0.1 % (w/v) of SDS. Interestingly, a single colony of the R (R –SDS) morphotype was obtained from an S inoculum using 0.1 % (w/v) SDS. This was archived and subjected to assessment in genotypic and phenotypic analyses.

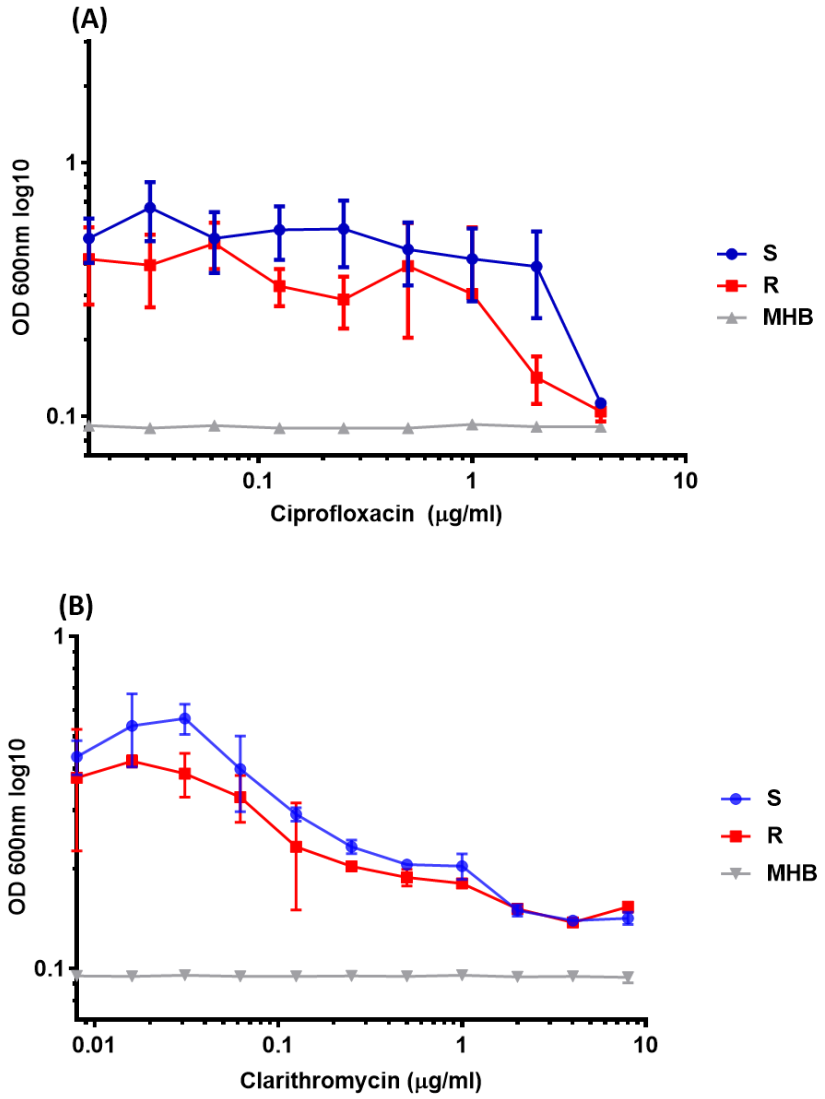


Figure 3.9: Choosing sub-inhibitory concentrations of Ciprofloxacin and Clarithromycin for *Mab* S and R lab control strains. **(A)** Sub-inhibitory concentrations for S and R phenotypes were respectively 2 and 1 µg/ml of Ciprofloxacin. **(B)** Sub-inhibitory concentrations of Clarithromycin was 0.25 µg/ml for both morphotypes. MIC was performed using broth dilution method and MHB was Mueller Hinton broth control.

3.4 Discussion

3.4.1 R morphotypes have a long lag phase

The importance of understanding the biology of the two *Mab* morphotypes (S and R) is the growth curve of the control and clinical strains isolated from CF patients. *Mab* S and R colonies exhibited a shiny, moist appearance whereas, R colonies were textured and had an irregular dry surface, with many crinkles and ridges. Many studies have correlated the R morphotype with the loss of ability of GPL production, affecting the surface phenotype (S.T. Howard *et al.* 2006; Nessar *et al.* 2011b; Howard 2013). The population doubling time of *Mab* is highly dependent on the *in vitro* conditions such as type of nutrient media (Cortes *et al.* 2010).

In this study, in the first 20h of culture the S morphotype increased CFU by 2 logs, whereas the R morphotype increased CFU < 1 log. It may be that S had a shorter lag, than R, however further detail investigations at early time points are needed to confirm extended lag phase phenotype of R strain. The Bioscreen OD measurements show that the lag of the R morphotypes was greater than the corresponding S morphotypes for the clinical strains (Figure 3.3).

From 20h-30h, the S morphotype increased CFU just over 1 log (Figure 3.2). The OD measurements show, that once growth initiates, the R morphotypes show a period of more rapid growth than the S morphotypes. The control R morphotype also shows this multiphasic pattern of growth (for CFU and OD) (Figure 3.2 and 3.3). This might be due to metabolic modification or fitness cost of R compared to S variants. Cortes (2010) reported that the culture of *Mab* in 7H9 to mid logarithmic phase requires 24 to 36 hr, and it will typically take 3 to 5 days on 7H10 agar plates at 37◦C to obtain viable transformants (Cortes *et al.* 2010).

3.4.2 *Mab* R morphotypes show relatively higher CSH

Microbial surface hydrophobicity was assessed with the MATH method as described by Rosenberg *et al*. (1980). The assay explores the relationship between the degrees of cellular interaction with liquid hydrocarbon (Rosenberg *et al*. 1980). The technique of MATH was chosen here because it is simple, rapid and low cost. A strong relationship

between CSH and the propensity of NTM for aerosolisation has been reported previously (Falkinham 2003). Here, R variants of the control and clinical isolates were found to be significantly more hydrophobic compared to the S.

In addition to assessing CSH, the CR method results have also been associated with virulence (Qadri *et al.* 1988; Cangelosi *et al.* 1999). Etienne and co-authors showed that a GPL deficient mutant of *M. smegmatis* becomes CR positive compared to wild type (Etienne *et al.* 2002). Minnikin *et al*. (2015b) demonstrated that WT *M. kansasii* and *M. canettii* (smooth colony morphologies) stained weakly by CR; in contrast, the dye was strongly taken up by the R strains of *Mtb*. Here, differences in control and clinical isolate cell surface hydrophobicity were demonstrated by growing these on CR containing agar plates.

RCRB assay was used with both phenotypes, S and R, of *Mab* control strain and clinical isolates for comparison to MATH. Findings suggested that R *Mab* variants can stain strongly and uptake the Congo Red dye and this difference in hydrophobicity correlated well with the MATH results. Therefore, this indicates that both hexadecane partitioning and CR staining investigations could demonstrate a clear difference in hydrophobicity between two variants of the same species. *Mab* strains were also used recently to demonstrate a positive correlation between GPL production and hydrophobicity (Viljoen *et al.* 2018). It was demonstrated that a *Mab mmpL4a* GPL-null mutant with an R morphotype, showed a similar, more hydrophobic phenotype when compared to the wild type (S morphotype). A *Mab mmpS4* mutant, that produced low levels of GPL, was found to have and intermediate hydrophobicity between the wild type and the *mmpL4a* mutant. Results of this study combined with the previously published data, suggest that the differences in cell surface hydrophobicity is a result of altered GPL expression.

3.4.3 GPL and LB formation by *Mab*

Most of the fast growing NTMs have the ability to produce GPL, which is one of the fundamental lipid components that play a major role in the bacterial phenotype such as cell wall permeability (Ripoll *et al.* 2009). To further investigate this phenomenon, the GPL production was examined in *Mab* variants. Here, variabilities were observed not

only between two variants of the *Mab* control strain, but also among clinical isolates for four patients, in which same patient showed mix colonies. The results showed that the S morphotype control strain expresses GPL whereas, the R variant did not. The S and R variants of the four clinical isolates showed a similar result. Similar variation in GPL reduction in mutants with rough morphotype was observed by Viljoen *et al.* supporting the findings of this study (Viljoen *et al.* 2018).

In order to understand the LB content of the S and R variants of *Mab*, the LB profile was analysed to determine the level of lipid accumulation differed in these. Increases in lipid production and accumulation of TAG were reported in later *Mab* clinical isolates obtained from chronically infected CF patients compared to the acute infections (Park *et al.* 2015). This observation suggested that TAG accumulation might be linked to chronic infections of *Mab*. Similar adaptive changes and enhanced lipid accumulation were observed in *Mtb* in conditions resulting in dormancy (Daniel *et al.* 2004). Thus, it has been suggested that *Mab* can gain selective advantages for storing fatty acids as TAG, instead of using them for the synthesis of complex lipids in the human lung environment (Park *et al.* 2015). In addition to metabolic adaptations, this also indicates that *Mab* might acquire genetic modifications during its long-standing infection period that as a consequence, may lead to increase lipid production (Viljoen *et al.* 2016). To further investigate this phenomenon, LB production and accumulation was examined in *Mab* variants. Variability was observed in lipid accumulation (revealed as differences in Nile Red fluorescence intensity per unit cell area) not only between two variants of *Mab,* but also among clinical isolates. In addition, the % LB content also showed the variability in lipid accumulation among S and R morphotypes. Thus, findings of indirect measurement of Nile Red fluorescence per unit area, combined with % LB suggested that lipid accumulation could be strain specific. This might also be due to the strains have been isolated from patients at different stages of clinical infection and these strains might represent their different metabolic adaptations.

TPP is external lipid and is thought to be exposed on the cell surfaces of *Mab*, Specifically in R variants it plays an important role in the pathogenesis and has a function in the formation of clumps and cords in *Mab* (Llorens-Fons *et al.* 2017). To confirm the

presence of TPP, lipid profile analysis was used and it was observed that R stains produce more TPP. Burbaud (2016) demonstrated that both stains S and R able to produce TPP (Burbaud *et al.* 2016).

3.4.4 *Mab* forms a biofilm-like structure

Bacterial biofilms are microbial aggregates, in which cells adhere to each other and onto biotic or abiotic surfaces. Biofilm formation begins with the attachment of free-floating microorganisms to a surface followed by cell-cell attachment (Hori and Matsumoto 2010). Byrd and Lyons have stated that the *Mab* R phenotype fails to produce biofilms (Byrd and Lyons 1999). However, this phenomenon was recently reinvestigated and it was reported that R morphotypes can also grow in biofilm-like structures (Clary *et al.* 2018). This supports findings of our investigation. In the context of cell surface characteristics, hydrophilicity and hydrophobicity can also influence the level of bacterial surface adhesion and the nature of biofilm formation (Krasowska and Sigler 2014) in *M. smegmatis* (Recht *et al.* 2000), *Mab* (S. Howard *et al.* 2006) and *M. bolletii* (Bernut *et al*. 2016b). Hence, the presence of GPL on the cell surface of S variants, can facilitate biofilm formation in liquid medium, whereas lack of GPL can enhance bacterial aggregation and cording.

This study, demonstrated a particular biofilm-like structure, formed by *Mab*, including cells that were attached to hydrophobic solid surfaces and floating mats (pellicles) on the surface of a liquid culture medium, together with planktonic cells in the liquid. This has previously been observed for *Mtb* (Binjomah, 2014). The present study has also shown that both variants of *Mab* have the ability to form biofilm, but at different rates. It was observed that the R phenotype formed pellicle layers more quickly, that might be due to its cording phenotype. Moreover, S required a longer period of incubation to form biofilm. These results could be supported by the phenomenon, in which hydrophobic bacteria have reduced repulsion from the extracellular matrix (Donlan 2002). Here, data suggested that biofilm-like structures of the R phenotype might give rise to the tendency of more frequent cell interactions compared to S variants.

3.4.5 The effects of antibiotics and detergent on the mutagenesis and expression of R morphotype

Selective pressure of antibiotic exposure affects several aspects such as mutations, which can cause antibiotic resistance in mycobacteria (Martinez and Baquero 2000; Brimacombe *et al.* 2007). Our preliminary results indicated that sub-inhibitory concentrations of both ciprofloxacin and clarithromycin did not show a significant effect on transition between *Mab* S and R phenotypes. Unfortunately, mutation/transition from S to R in the presence of antibiotic pressure was not detected using current parameters. It might have been possible if the exposure time to antimicrobial was increased. Another reason could be a very low mutation frequency that is low below the level of detection.

Fortunately, one event of S to R transition was detected when lab control strain of S morphotype was exposed to 0.1 % (w/v) SDS. The experiment was repeated several times, but no other transition of S to R was detected, suggesting a very low frequency of mutation affecting the GPL production. SDS showed an effect of promoting genetic mutation (Section 5.3.2) that supports the previous reports of toxic effects of SDS on cell envelope integrity, protein denaturation and gene regulation (Rajagopal *et al.* 2002; Sirisattha *et al.* 2004).

3.5 Conclusions

1. Growth is clearly different between S and R morphotypes. The most obvious deference is a difference in lag phases, but there also appears to be faster and slower periods of growth R strains as measured by optical density.
2. R morphotypes were consistently more hydrophobic than their S counterparts.
3. Nile red fluorescence intensity analysis showed no significant difference for the S and R morphotypes of control strain, but inconsistent differences were observed for the different morphotypes of the clinical isolates. No consistent pattern of LB content with morphotype was observed.
4. The R strain produced biofilm earlier, and with different structure, compared to the S strain.
5. Lipid profile analysis showed a clear loss of GPL in R strains and possible enhancement in TPP levels.
6. Sub-inhibitory levels of ciprofloxacin and clarithromycin did not yield R variants from S, or vice versa, while a single R variant was isolated on 0.1% (w/v) SDS.

Chapter 4

Comparison of Survival of *Mab* S and R in environmental conditions associated with airborne transmission.

4.1 Introduction

4.1.1 Investigating the survival of *Mab* in aerosols using Goldberg Drum technique.

Non-tuberculous mycobacteria (NTM) are increasingly found in the sputum of people with cystic fibrosis (Floto *et al.* 2016). Specifically, *Mab* has emerged as a potentially important pathogen, with evidence of accelerated lung function decline (Esther *et al.* 2010; Esther Jr *et al.* 2010). *Mab* has been isolated from household water and has been previously isolated from the shower aerosols of people with pulmonary NTM disease (Clifton *et al.* 2008; Thomson *et al.* 2013).

Slow growing NTM are readily transmitted between different habitats. *M. avium* and *M. intracellulare* aerosolise from water to air by droplet formation, and the cells attached to these droplets can remain bound to the surface of the water (Falkinham 2003). The bursting of droplets can result in the ejection of droplet materials to a height of approximately 10 cm (Parker *et al.* 1983; Falkinham 2003). Loudon and his colleagues showed that mycobacteria, including *Mtb*, show survival rates from 13 to 56% in aerosols 6 hours after nebulisation (Loudon *et al*., 1969) while, more recently, a survival half-life life of 30 minutes has been reported for *Mtb* (Lever *et al*., 2000). Bryant *et al*. (2016) showed that the majority of *Mab* infections were acquired potentially through fomites or aerosols. The actual mode of transmission of *Mab* into the lungs of the host is still not fully understood, however, Harris (2015) detected person-to-person transmission of *Mab* between a sibling pair with cystic fibrosis. It was suggested that other parameters might play a role, the environment in particular, which need to be investigated further (Harris *et al.* 2015).

The ability of pathogenic *Mab* to transmit between patients is the biggest concern at this moment in time. It would be beneficial to understand *Mab*'s mechanisms of transmission and discover the key to its survival

4.1.2 Goldberg Drum

The Goldberg Drum (2 feet wide and 6 feet in diameter) was described in 1958 (Goldberg *et al.* 1958). The apparatus is made of aluminium with an internal capacity of 70L. It was firstly designed to study the survival of microorganisms within aerosols (size 1-6 µm) under controlled RH and temperature. The Goldberg Drum is connected such that conditions of the humidity of the air within the system are at a set point. The Henderson apparatus has two main functions, delivering the generated aerosols to the drum and controlling the RH inside the drum. It consists of a chamber and a spray tube. The drum contains a series of baffles that help maintain aerosols though its rotation at 3.5 rpm. In addition, the apparatus has a temperature and humidity data logger that allows a continuous monitoring of environmental conditions (Thompson *et al.* 2011) .

4.1.3 Henderson apparatus

The Henderson device is used in conjunction with the Goldberg drum to control the relative humidity inside; it also aids delivery of the generated aerosols. The apparatus consists of a chamber and two spray tubes. The chamber consists of three parts: a pump that generates air, a tank of silica gel (provides dry air) and a tank of water (provides wet air). The generated air travels through these tanks before reaching the first spray tube. Depending on the required Relative humidity (RH), the volume of the air that passes through each tank is controlled by the piccolo device. The piccolo apparatus permits the equilibration of aerosols with the main air flow inside the second spray tube before reaching the drum.

4.1.4 Ultrasonic NE-U780 Omron nebuliser

The ultrasonic NE–U780 Omron nebuliser apparatus (https://www.omron-healthcare.com/en) has a vibrator/sonicator positioned under a water bath. The vibrations are transferred through the water to the medication cup that contains the bacterial suspension. This results in the production of bacterial aerosols (Figure 4.1). A 3-jet Collison nebuliser was connected up-stream from the Omron nebuliser. This was

filled only with DW and was necessary to provide the pressure to carry the Omron aerosols into the Drum (Figure 4. 2).

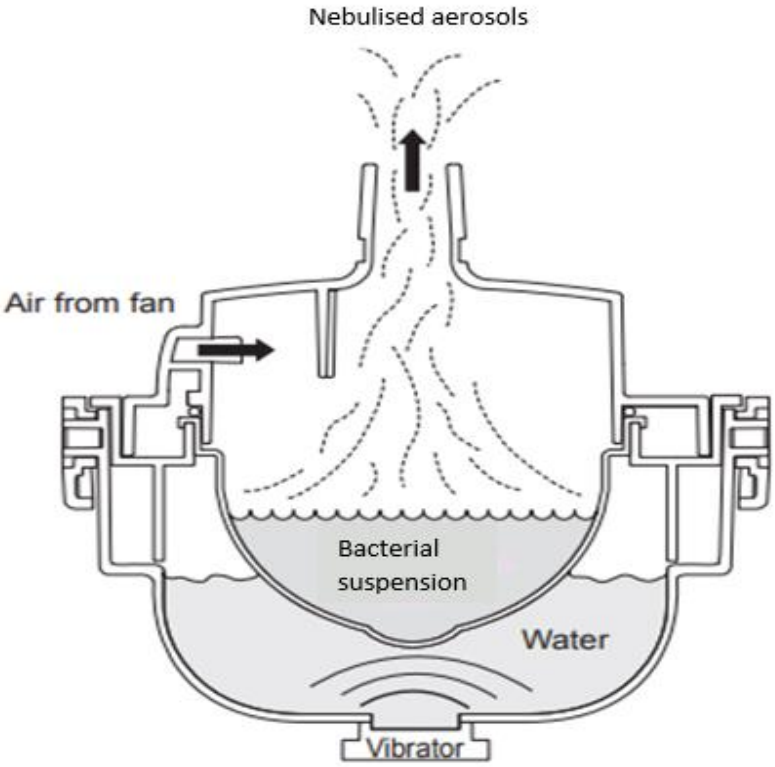


Figure 4.1: Ultrasonic NE-780 Omron nebuliser.
The bacterial suspension is located inside the medication cup. (Image is modified from Omron instruction manual).

Figure 4.2: Modifications were applied when Omron nebuliser was used in the experiment to maintain similar internal air flow to Collison nebuliser.

4.1.5 Desiccation in Mycobacteria

When water is absent, the microbial cell can experience protein and DNA damage (Potts, 1994). A unique ability to tolerate desiccation had been found in mycobacteria and has been correlated to the distinctive cell envelope structure. The mycobacterial outer lipid layer associates with mycolic acids and includes glycolipids such as trehalose 6,6'-dimycolate (TDM) (Harland *et al.* 2008). It was found that TDM could be dried and desiccated without any significant damage to the cell membrane, suggesting its preservative role during desiccation (Harland et al. 2008). Moreover, findings on the impact of LBs, suggest that these could play an essential role in the survival of stress have been reported. *Mtb* growth was reported to be LB positive under anaerobic conditions and more infectious by 10-fold, when compared with aerobically grown controls in a guinea pig aerosol model of infection (Bacon *et al.* 2004) .

In addition, triacylglycerols (TAGs) that are prominent in LBs, were found to have an important role in the metabolism of *Rhodococcus opacus*, as the TAGs were oxidised to produce energy during water stressful conditions (Alvarez et al. 2004). Furthermore, it was found that inhibition of lipase breakdown of TAGs, or β-oxidation of released fatty acids, adversely affects the survival of *Rhodococcus* spp. during desiccation, carbon starvation and UV-irradiation (Alvarez et al., 2004; Urbano et al., 2013). LBs play an important role in the survival of *Rhodococcus* spp. and could also have a similar role in mycobacteria. It has been observed that during desiccation, mycobacterial cells are exposed to more physiological challenges than in osmotic stress situations (Billi and Potts 2002). Bacterial viability is usually defined as the ability to form progeny (Cangelosi and Meschke, 2014). It refers to any method that identifies cellular metabolism and functional cell membranes or successful DNA replication, which are considered as indirect methods of inferring viability.

It is worth noting, that the viability depends on cellular metabolism, which in itself is an indicator of bacterial viability. There are techniques which detect the metabolic activity in cells by assessing ATP production (Thore *et al.* 1975; KODAKA *et al.* 1996; Venkateswaran *et al.* 2003). When the luciferase enzyme and luciferase substrate are

present, the catalysed oxidation reaction leads to the production of light (Figure 4.3). Relative light units (RLU) are used as measure of light intensity and RLU readings can be used as an indirect measure of ATP concentration (Stanley 1989; Selan *et al.* 1992; Siragusa *et al.* 1996). This method can provide a plausible estimate of viability activity and was used to assess the survival of *Mab* during desiccation and UV tolerance assays performed here.

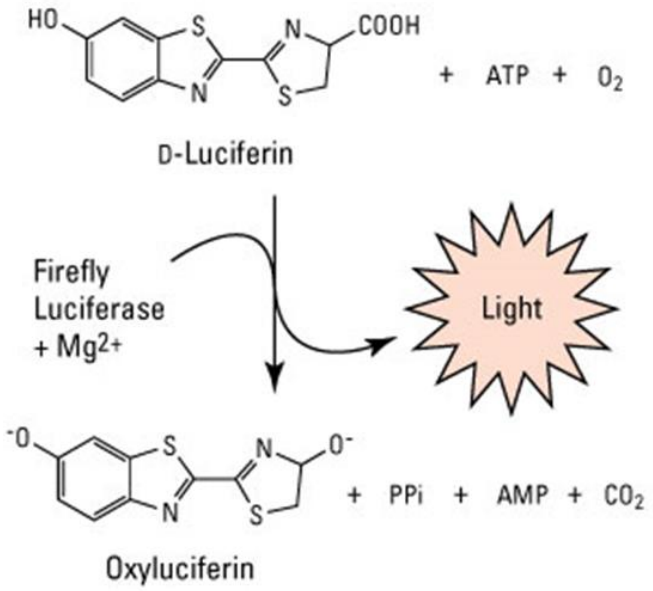


Figure 4.3: **Diagram of the reaction of firefly luciferase**. Image taken from Thermo Fisher Scientific "Firefly Luciferase Assays & Vectors".

The work in this chapter focuses on designing experiments to examine differences in the survival between S and R phenotypes of *Mab*, in terms of stability in aerosols. This was done using the Goldberg drum system with nebulisation using the Ultrasonic NE-U780 Omron nebuliser. Moreover, the features which influence the bacterial propensities for aerosolisation and stability in aerosol, were investigated and determined. Previous experiments on aerosol survival have used many types of nebulizers; however, none have used the Ultrasonic NE-U780 Omron nebulizer. Once aerosolized, various stressors including desiccation, UV light exposure and osmotic stress are faced by *Mtb* (Saviola, 2010). The effect of these environmental stressors, specifically desiccation and UV, on the survival of *Mab* during transmission is the focus of this chapter.

In previous studies, it has been suggested that the *Mab* R phenotype would be hydrophobic, more virulent and aerosol transmissible when compared with the S phenotype (Jankute et al. 2017; Minnikin *et al*. 2015; Jönsson *et al.* 2007). This chapter focuses on the investigation of the *Mab* survival for both phenotypes to determine whether R morphotypes would survive environmental stresses associated with aerosol transmission, better than morphotypes in aerosol experiment by using a Goldberg drum in accompanied with Henderson device and then desiccation.

4.1.6 Aim:

In this chapter, the aim was to investigate any differences in the survival rate of *Mab between* the two phenotypes S and R in aerosols over a period of two hours and to compare the survival rate during stress tolerance assays.

4.1.7 Objectives:

1. To investigate whether the process of nebulisation has an effect on *Mab* cell integrity.
2. To investigate the stability pattern of *Mab* in aerosols following aerosolisation with an Ultrasonic NE-U780 Omron nebuliser.
3. To examine the impact of a freeze-thaw cycle on the two morphotypes S and R of *Mab* harvested from the aerosol phase.
4. To prepare luciferase reporter strains of S and R morphotypes for comparison in survival assays.
5. To compare the survival S and R morphotypes of *Mab* after desiccation.
6. To determine differences in survival rates S and R morphotypes of *Mab* after exposure to UV irradiation.

4.2 Materials

Kanamycin

108 mg of kanamycin was dissolved in 4.32 ml of distilled water (final concentration of 25 mg/ml), vortexed and sterilized using a 0.2 µm filter with a 5 ml syringe and a 0.8 mm x 50 mm needle. 500 µl aliquots of the kanamycin stock was stored in Eppendorf tubes at -20°C.

Luciferin

50 mg of D-Luciferin potassium salt (Promega) was mixed with 1.67 ml of distilled water (final concentration of 30 mg/ml). The solution was separated into 140 µl aliquots and stored at -80°C until use.

4.3 Methods

4.3.1 Preparation of growth conditions

4.3.1.1 *Mab* survival in aerosol

Bacterial growth conditions were based on a previous study by Lever and colleagues which investigated the survival of *Mtb* in aerosols (Lever *et al.* 2000). Firstly, 100 µl of a frozen stock of *Mab* for both S and R phenotypes (OD = 0.9-1.0) were used as an inoculum for 7H10 agar (supplemented with OADC). The inoculum was horizontally over the plates by using a cotton tipped swab (Alpha laboratories; Code 42141500) to form a lawn of bacterial growth all over the plate, three to five days post-incubation at 37°C. The cells of both S and R morphotypes were scraped from 3 plates of each morphotype using a scraper (250mm, Fisher Scientific), and were re-suspended into 25 ml of de-ionised water. To remove clumps in the suspension, glass beads (2-3 mm) were added and the suspension was vortexed vigorously until no clumps were seen. Then, the suspensions were adjusted to an optical density (OD_{600}) of 2.0.

4.3.1.2 Preparation of the *Mab* inoculum for the desiccation assay

A 100µl of suspension of both the S and R morphotype cultures at an OD600 of 1.5 were used to inoculate 7H10 agar plates and spread using a sterile cotton swab and incubated as described above. The lawn growth at 5 days was harvested as previously described and cells were re-suspended in PBS, treated with glass beads as described above. The suspension was then adjusted to an OD600 of 1.0.

4.3.2 The generation and sampling of aerosols

These aerosol experiment were conducted at PHE Porton Down. The sampling and aerosol generation of *Mab* were performed in an isolator at contaminant level 2, as *Mab* is classed as a hazard group two pathogen. An ultrasonic NE – U780 Omeron nebuliser was used to generate aerosols, using 10 ml of each the S and R bacterial suspensions which were mixed together after each was individually normalized to an OD_{600} of 2.0 and vortexed with glass beads to make a homologous suspension. The nebulizer ran at 26–28 psi and distributed aerosols into the Goldberg Drum (70L). The aerosol was maintained at a mean RH of 85 ± 10 % and a mean RT of 22 ± 2°C.

All-glass impingers (AGI-30), as described by Wolf (1959) were used to collect all aerosol samples into water (Wolf *et al*. 1959). These samplers were used at a flow rate of 11.5 ml/min for 5 min.

4.3.3 Test procedure

The Goldberg drum and Henderson device were linked to each other (Figure 4.4). In order to achieve the desired RH, the drum was pumped with conditioned air for 30 minutes prior to each experiment. Before the start of the nebulisation, a negative control aerosol sample was removed from the drum. This was done to guarantee that no cross contamination would occur between experiments. After this, the *Mab* aerosols were sprayed into the Goldberg drum for a duration of 5 minutes. Aerosolised samples were collected into All-glass impingers (AGIs) that contained 10 ml of deionised water (DW) for sampling at 0, 15, 30, 60, 90, and 120 minutes after nebulisation, sample acquisition time was one minute. The air extracted from the drum was replaced with new air throughout the sampling time. The new conditioned air originated from the

Henderson apparatus (preconditioned at the appropriate RH). At each sample time point, samples were extracted and the aerosols were diluted within the drum. Following this, glycerol was added to a final concentration of 20 % v/v to each sample. The influence of the nebulisation process on the spray suspension was evaluated by collecting two samples from the nebuliser reservoir at time pre- and post-nebulisation. Glycerol was added to these samples as previously described. The Goldberg drum system and a schematic depicting key features are shown in Figure 4.4 and Figure 4.5.

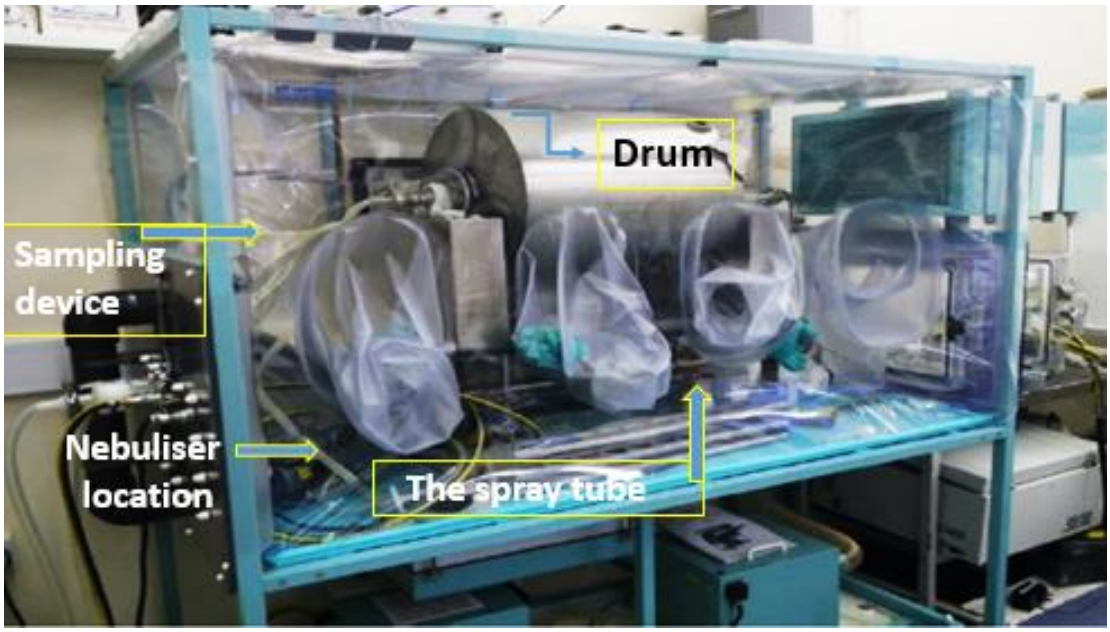


Figure 4.4**: Image of the Goldberg drum**
The nebuliser is positioned at the drum's left side connected to the spray tube. Organisms are sprayed into the drum at the other end and kept in an aerosol by rotation (3.5 rpm) and a series of internal baffles. Samples are extracted from the drum via the sampling apparatus located at near end.

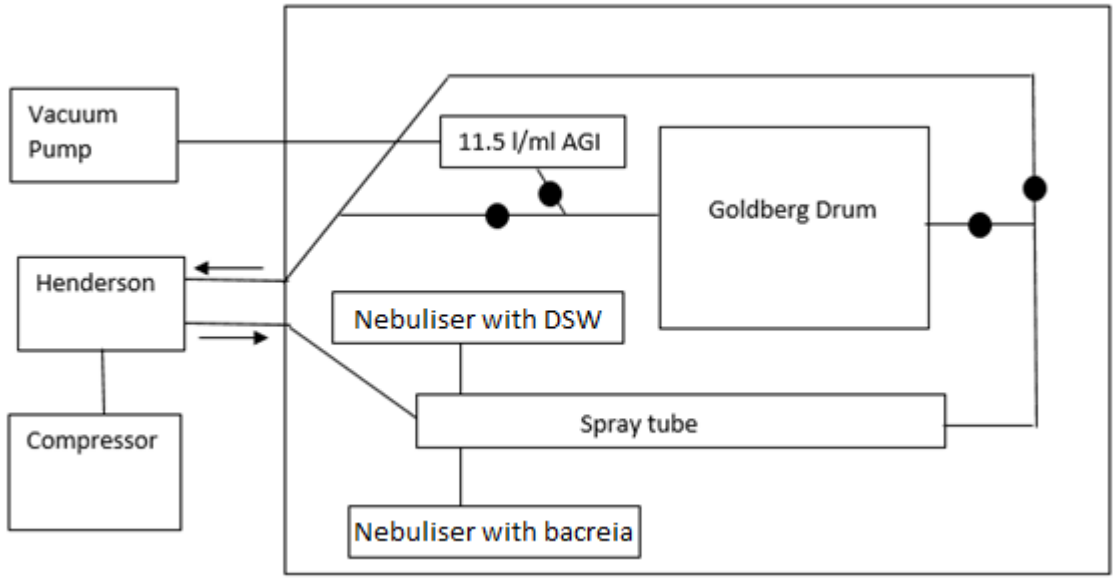


Figure 4.5**: A Schematic diagram of the Goldberg drum system.**
The diagram shows a spray tube connected to two nebulisers inside the isolator, one contains the bacterial suspension and the other one contains deionized water to help with mixing the generated aerosols inside the spray tube. Tubing is used to link the spray tube to the Goldberg drum. The drum is connected to the AGI sampler. Clamps (⬤) are opened and closed during sampling (Alshammri 2019).

4.3.4 Processing samples

The start of the experiment was performed at Porton Down, with each experiment containing two nebulises reservoir samples, taken pre and post-nebulisation. 6 aerosol samples were collected and each sample was serially diluted to calculate the colony forming units by the spread plate method (S-CFU) to determine counts of both morphotypes of *Mab* on Middlebrook 7H10-OADC. Aliquots of the neat samples were also stored at -80°C and then sent to the University of Leicester for replicate culture studies after thawing.

S-CFUs were determined with 100 µl aliquots of serial ten-fold dilutions in DW, added to agar and spread evenly using cotton-tipped swabs. Each plate was sealed with laboratory sealing film and incubated at 37 °C until the colonies became visible. The final calculation of colony forming units (CFU) was performed by counting 10-330 colonies in single plate, using the following equation on three independent occasions.

$$S\text{-}CFU/ml = C \times D \times 10$$

C= count per plate

D= Dilution factor

4.3.5 Preparation of S and R luciferase reporter strains of *Mab*

These strains were prepared in order to enable rapid viability assessments in survival assays. While it had been hoped to achieve this in the Goldberg drum experiments, permission to handle the genetically modified strains at Porton Down was not obtained in time to achieve this, so the strains were only used in desiccation and UV exposure experiments.

4.3.5.1 *ffluc* Plasmid Preparation

The plasmid pMV306G13 *ffluc* (Andreu *et al.* 2013) was purchased from Addgene (plasmid # 49997) and was received as an *Escherichia coli* agar stab. The *E. coli* was streaked on a Luria agar plate with 50 μg/ml kanamycin. Single colonies were taken and used to inoculate 5ml Luria broth also containing 50 μg/ml kanamycin. Following overnight incubation (37°C, 200 rpm), plasmid was prepared using a Plasmid Mini Kit (Bioline) and the concentration measured using a Nano Drop. Luciferase *Mab* S and R strains were transformed with this pMV306G13 *ffluc* plasmid. Cells of *Mab* S and R were streaked onto different concentrations of kanamycin along with the WT as a control, using parameters found in Viljoen (Viljoen *et al.* 2018). 50 μg/ml kanamycin was chosen for selection of *Mab* transformants.

4.3.5.2 Preparation of electrocompetent *Mab*

The method of preparation of mycobacteria for electroporation was modified from Parish and Stoker (Parish and Stoker 1998). Fresh 7H10 agar plates were first streaked with wild type *Mab* S and R phenotypes and incubated at 37 °C for 5 days. Then, a single colony from each sample was inoculated in 5 ml of 7H9 broth. The following day, the culture was normalized to an OD_{600} of 0.003 and incubated. When the cultures reached an OD_{600} of 0.8, they were harvested by centrifugation at 2,000 x *g* for 10 minutes. The

supernatant was removed and cells were washed twice with 20 ml of ice-cold 10 % (v/v) glycerol in distilled water, and centrifuged at 2,000 x *g* for 10 minutes (this step was then repeated once more). The remaining cell pellet was resuspended in 2.5 ml of ice cold 10 % (w/v) glycerol in distilled water, and 400 µl aliquots were snap frozen by submerging the microfuge tubes in a dry ice and Industrial Methylated Spirit (IMS) slurry, and stored at –80 °C.

4.3.5.3 Transformation of *Mab*

Aliquots (400 µl) of electrocompetent S and R *Mab* cells were thawed on ice and 5 µl of plasmid DNA was added. These were mixed and held on ice for 10 minutes. The preparations were then transferred to a pre-chilled electroporation cuvette (0.2 cm gap) and placed between terminals of a Gene Pulser instrument (Biorad). Electroporation was performed at 2.5 kV, 25 µF, 1000 Ω (Kendall & Frita, 2008), followed immediately by the addition of 1ml of 7H9 complete broth, after which they were transferred to 30 ml universal tubes containing 5ml of 7H9 complete media. Transformed *Mab* cells were incubated at 37 °C in a shaking incubator at 130 rpm for 120 minutes to allow for cell recovery and expression of the antibiotic resistance genes.

4.3.5.4 Selection of transformants

Cells of both strains were plated on 7H10 agar plates containing 50 µ g/ml kanamycin and incubated at 37 °C until transformants were observed (about 3 days). Single colonies were picked and streaked on patch plates (to maintain particular clone) and incubated until growth was detected. Starter cultures (5 ml of 7H9 complete media) were inoculated using the patch plates and luminescence was tested (luciferase assay on culture samples) once an OD_{600} of 0.8 was reached. Frozen stocks were made. Luminescence was also tested on subculture in order to detect any drop in luminescence.

4.3.6 Luciferase Assay

The luciferase assay was based on methods described by Andreu (Andreu *et al* 2010). The substrate, D-luciferin potassium salt (Promega), was stored and dissolved in distilled

water in single-use aliquots at -80 °C at a concentration of 30 mg/ml. Before use, this was thawed, vortexed and diluted in PBS to a standard working concentration of 2.16 mM. In a 96-well plate, 50 µl aliquots of the sample and a blank (medium only) were placed in individual wells in triplicate. A Varioskan Flash Multimode Reader was set to dispense 50 µl of luciferin solution into each well, shake for 1 minute at 600 rpm, and to detect any emitted light for 10 seconds. Detected light was reported as RLUs. If blank wells returned positive values due to background signals, they were averaged out and subtracted from each sample value. An ATP determination kit (Thermo Fisher) was also used as a standard to compare values between runs (refer to 'ATP Determination Kit for use as a standard' in Results). The limit of detection was based on the average reading of the blanks.

4.3.7 Desiccation Assay (plate method, devised by Valeria Quimper of this lab)

The plate method of a mycobacterial desiccation assay was performed by pipetting five 2 µl drops of the prepared *Mab* inoculum (OD 1.0 in PBS) in triplicate wells (only wells located at each corner of the plate were used) of a 24-well white polystyrene plate. Plates were left to desiccate inside a class 2 biosafety cabinet for the allotted times. Cells were then resuspended in 490 µl of complete 7H9 broth, thereafter 500 µl of complete 7H9 was pipetted in triplicate manner on the remaining corner of the plate for blank calculations. Plates were then taken to Varioskan for the luciferase assay. For 24-well plates, the luciferase assay was adjusted by dispensing 500 µl of 2.16 mM luciferin in PBS (Figure 4.6). The use of 2 µl droplet in this assay was to mimic the natural size of aerosols in air. Five separated droplets were used in each well as the area size of each well allows no more than 5 droplets.

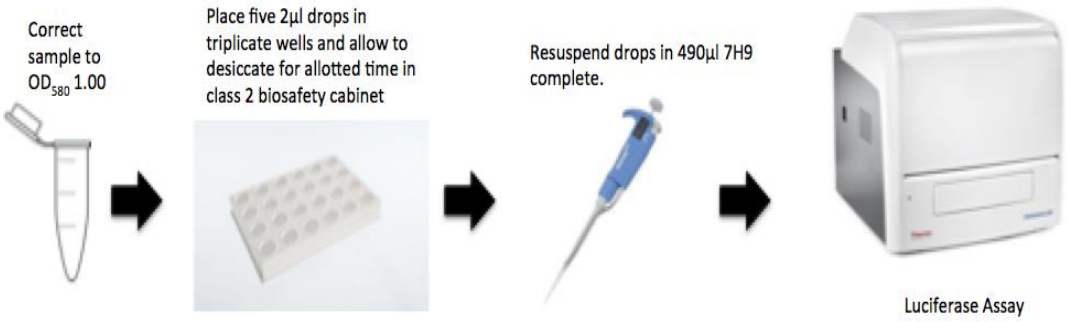


Figure 4.6: **Desiccation Assay.**
Diagram of the steps involved in the desiccation assay using 24-well white plates. (Valeria Quimper, 2019).

4.3.8 UV Dose response Assay

Bacterial inoculum were prepared following the procedure described in section 4.3.1.2. Both S and R strain samples were washed and corrected to an OD_{600} of 1.0 Then, five drops of 2 µl of both morphotypes were separately pipetted in triplicate into different wells (the wells used in this experiment were located at each corner) of a 24-well white polystyrene plate. The plates were then placed in the middle of a UV crosslinker (CL-1000, Thermo-Fisher). Then, a specific UV dose was administered using an exposure time of 10 minutes. The plate was removed from the crosslinker and the droplets were re-suspended in 490 µl of 7H9 complete medium. For blank readings, 500 µl of 7H9 complete medium was added to wells in triplicate. After that, the luciferase assay performed.

4.3.9 Measurement of CSH S and R of *Mab ffluc*

CSH assessment was performed for the *Mab* luciferase expressing strains using the MATH technique as in section 3.2.2. Firstly, CSH measurement was assessed for each strain independently. Then, the S (WT) strain was mixed with R *ffluc* strain and R (WT) strain was mixed with the S *ffluc* strain by vortexing and the MATH assay was performed on these.. The measurement of CSH for the mixed suspension was assessed using the Varioskan luminometer modality by measuring luminescence of aliquots of the aqueous phase of the MATH assay. A 24 -well white polystyrene plate was used in this experiment

and 50 µl of cell suspension before hexadecane treatment was added to 3 wells. Then, 50 µl of cell suspension post hexadecane treatment was added to a further 3 wells. 50 µl of PUM buffer was added to 3 wells as a blank.

4.4 Results

4.4.1 Survival patterns of *Mab* in aerosols

The recovery of *Mab* S and R morphotypes from nebulised equal OD mixtures, expressed as a percentage of S-CFU at baseline (T=0 immediately after the 5 min nebulisation was complete) is presented in Figure 4.8. The theoretical dilution effect due to the successive sample volumes withdrawn from the Drum and subsequent dilution with make-up air, is calculated and shown on the same graph. It should be noted that, some of the colonies appeared on the plates with unusual columnar morphology (Figure 4.7). These colonies were excluded from the cell count when the survival graph was plotted from Porton Down culture results. The rate of the decline of recovery of CFU for both morphotypes declined faster than the theoretical dilution line. The error bars of CFU determinations for the two morphotypes overlapped, indicating that the difference between them is statistically insignificant. It worth noting, that there was a sudden drop in the recovery at T15 compared to the other time points for both strains (Figure 4.8). Samples processed and cultured at Porton Down showed that the survival of the S was 60 %, while the R phenotype showed less than 10 % recovery after a period of 2 hours in aerosol. However, when the aerosol samples were re-plated at the University of Leicester, both morphotypes showed similar recovery of at least 50 % over a period of 2 hours in aerosol.

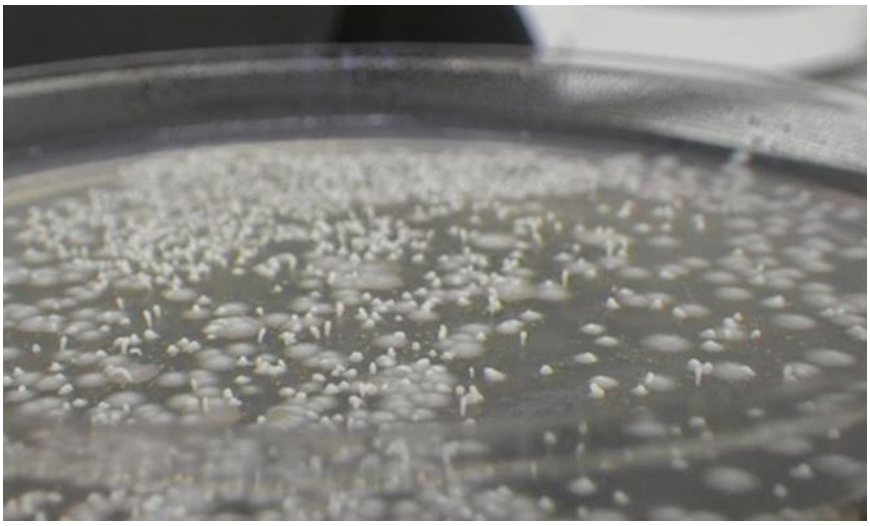


Figure 4.7**: strange morphology appeared for the colonies of *Mab* for mix culture of both phenotypes.**
These colonies were excluded from counting the cells when the survival graph was plotted from Porton Down observation

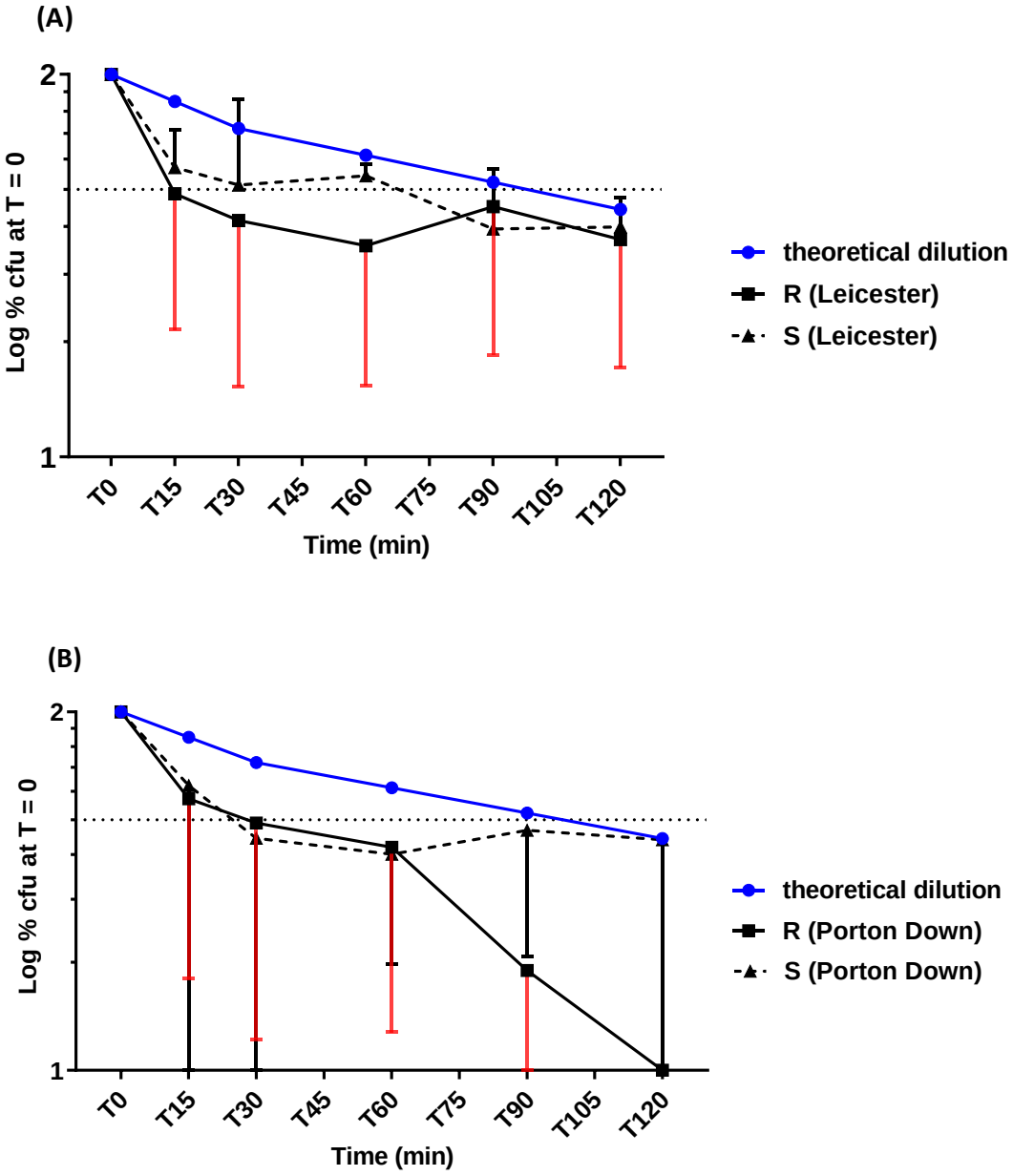


Figure 4.8: Survival of *Mab* in aerosols with the theoretical dilution.
Bacterial suspensions were prepared and nebulized for 5 min, after which the samples of aerosol were collected. S-CFU count for each sample was performed over a period of 2 hrs and theoretical dilution of the aerosol was calculated (blue line). **(A)** Leicester CFU determination **(B)** Porton Down CFU determination. Error bars represent the SD of three independent experiments. The horizontal dotted line represents 50% survival.

The survival as a percentage of the theoretical decline for both S and R morphotypes was calculated and is shown in Figure 4.9. This is defined as the (percentage recovery of the strain divided by the percentage recovery of the theoretical decline) × 100. In the Porton Down CFU determination, the S phenotype had a stability rate in aerosol of at least 60 %, while the R phenotype showed less than 10 % stability over a period of 2 hours. However, when the aerosol samples were re-plated at the University of Leicester, the S phenotype stability rate in aerosol was at least 80 %, while the R phenotype showed less than 50 % stability over a period of 2 hours. The error bars of these strains overlapped, showing that the difference between them was statistically insignificant.

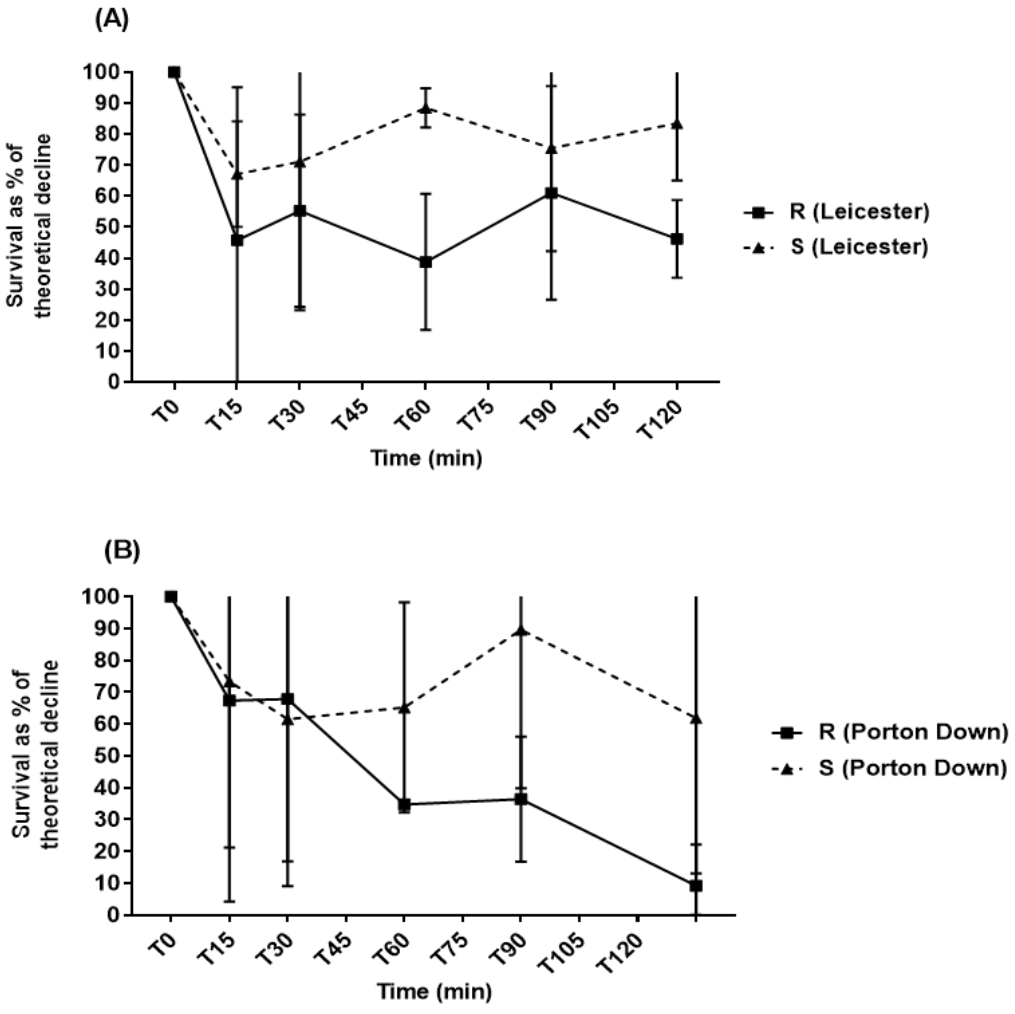


Figure 4.9: Aerosol survival stability pattern of *Mab* as a percentage of the theoretical dilution of aerosol in the Drum.

Bacterial suspensions were prepared and nebulized for 5 minutes, and aerosol samples were collected. The S-CFU count for each sample was performed over a period of 2 hours and the

theoretical dilution was calculated. **(A)** Leicester **(B)** Porton Down determination. Error bars represent the SD of three independent experiments.

4.4.2 Effect of nebulisation on *Mab* by using S-CFU counts

These experiments were done at Porton Down to investigate whether the nebulisation process using Ultrasonic NE-U780 Omron nebuliser had an effect on the integrity of aerosolised *Mab* cells. This was assessed by carrying out the S-CFU count of the cells taken from the reservoir pre- and post-nebulisation. Figure 4. 10 shows that the S-CFU count of the S morphotype was not affected by the nebulization process, and even though the S-CFU count of the R phenotype increased post-nebulisation, it was not statistically significant. Although the bacterial mixture was prepared in a 1:1 ratio of individual suspensions of similar OD, the S-CFU pre-nebulisation count was less for the R morphotype compared with the S morphotype, and this can be because the clumping of the cells. The R phenotype clumps more than S phenotype and this will be exacerbated by the lack of tween 80 in the suspension here, even though the suspensions were normalised to the same OD prior to mixing.

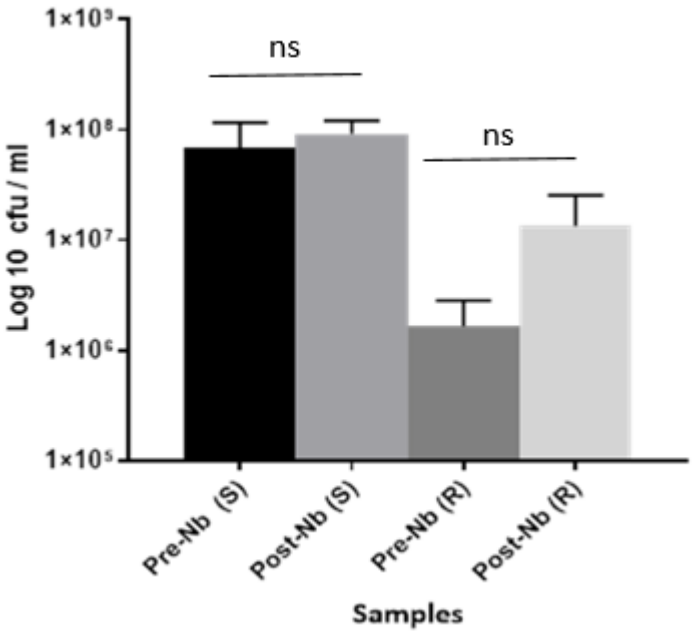


Figure 4.10: Comparison between the S-CFU counts of the suspensions inside the reservoir pre- and post-nebulisation.
These results were obtained from the nebuliser samples taken from the Porton Down experiments**.** An OD-normalised 1: 1 (v/v) mixture of S and R morphotypes was prepared and nebulised for 5 minutes. Samples pre- and post-nebulisation were collected and the S-CFU was determined using the spread plate technique. Error bars represent the SD of three independent

experiments. Statistical analysis was performed with multiple t-test using Sidak's method (ns=non-significant).

4.4.3 S-CFU counts from aerosol were not significantly affected by freezing

For further investigation into the effect of processing samples at different sites, it should be noted that samples from Porton Down were treated with 20 % (v/v) glycerol before plating to make an appropriate comparison when re-plating the samples in Leicester University and then frozen at -80°C. To see the effect of the freeze-thaw cycle, the samples were re-plated on to 7H10-OADC agar plates, and the S-CFU method was used once again, as performed at Porton Down (Figure 4.11). While consistently lower CFU values were obtained in Leicester this did not appear to affect the patterns observed.

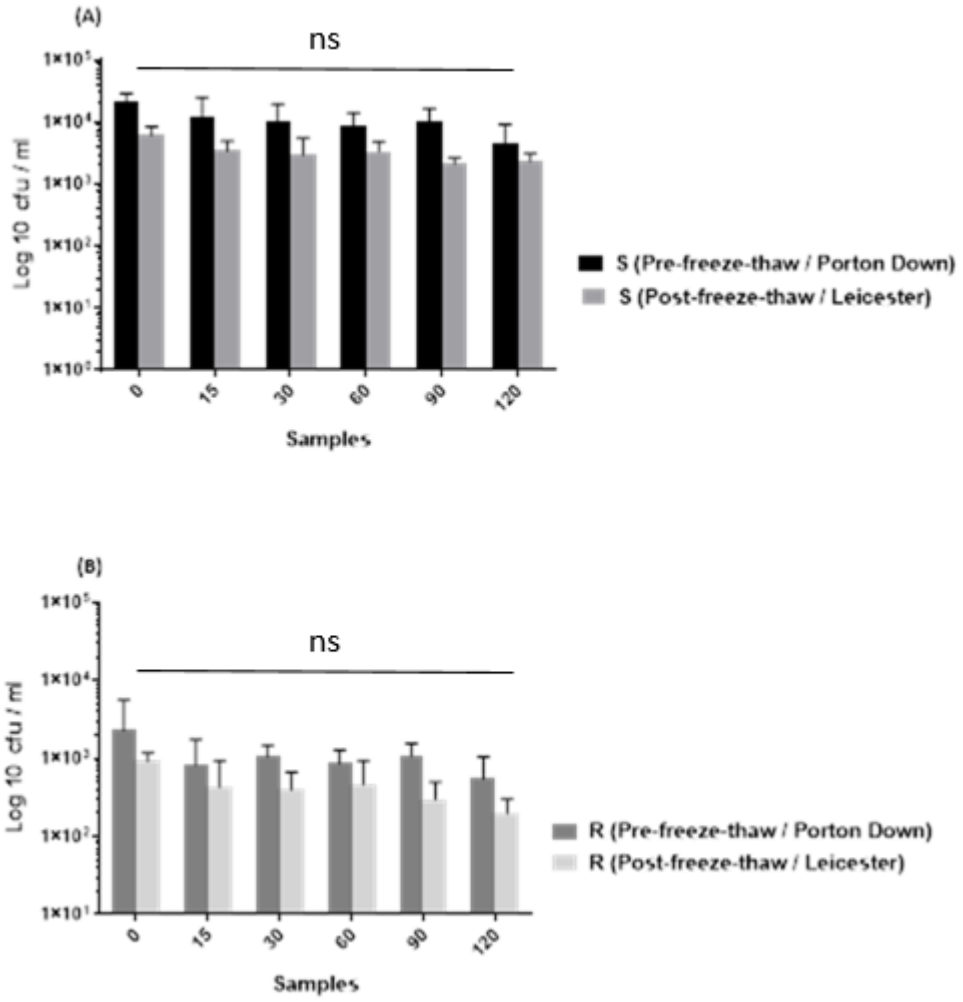


Figure 4.11: CFU counts of the aerosol samples before and after the freeze-thaw cycle.

Aerosol samples were frozen at -80°C and then transferred to Leicester. CFU assays were performed using the S-CFU technique. **(A)** S /Pre-freeze-thaw at Porton Down compared to S /Post-freeze-thaw at Leicester. **(B)** R /Pre-freeze-thaw at Porton Down compared to R /Post-freeze-thaw at Leicester. Error bars represent the SD of three independent experiments. Statistical analysis was performed with multiple t-test using Sidak's method (ns=non-significant).

4.4.4 S strain appears more hydrophobic when mixed with the R strain

The aim of this experiment was to test the hypothesis that mixing S and R might affect outcome of the aerosol experiments. Firstly, CSH was measured by using the MATH technique for each strain with measurement of the aqueous phase OD following hexadecane partitioning (Figure 4. 12). Next, the S *ffluc* was mixed by vortexing with R (WT) and S (WT) with R *ffluc* (1:1 v/v) and following hexadecane partitioning, the luminescence was then measured using the Varioskan. This allowed separate assessment of the CSH of *ffluc* strains when mixed with a wild type partner (Figure 4.12). Interestingly, *ffluc* plasmid transformation rendered the S strain more hydrophobic. However, the S *ffluc* CSH increased still further when this strain was mixed for 10 min (1:1 v/v) with the R wild type. In contrast, *ffluc* transformation of the R strain did not alter the CSH measurement but, this did appear to increase after mixing with the S (WT). However, this increase was not statically significant.

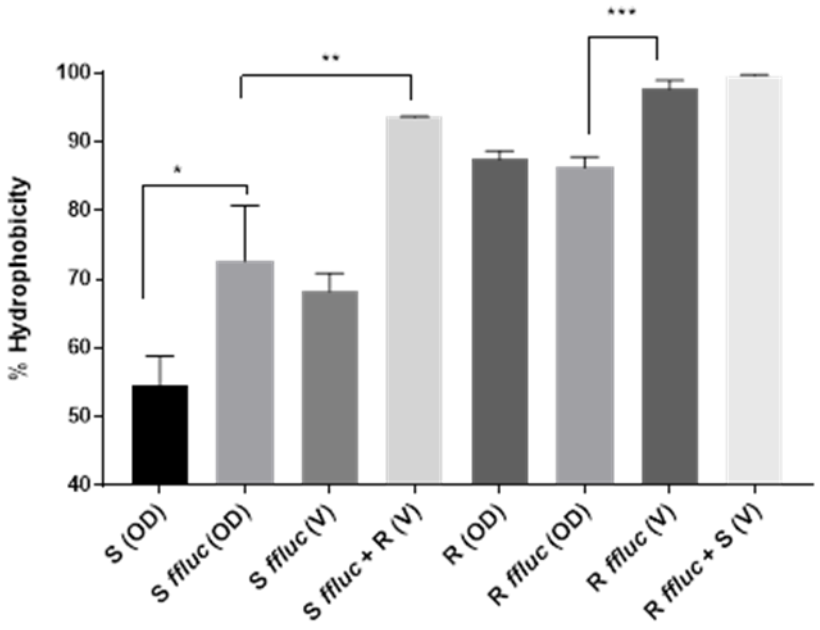


Figure 4.12: The effect of mixing S and R of *Mab* on the CSH:

Comparison between CSH of *Mab* S and R (WT) and *Mab* S and R *ffluc.* Error bars represent three biological replicates tested in triplicate and assayed by OD and luminescence measured in a Varioskan plate reader (V). The S *ffluc* transformant was mixed with R (WT) and S (WT) with R *ffluc* transformant (1:1 v/v) and the OD and luminescence measured in a Varioskan plate reader. Statistical analysis was performed with multiple t-test using Sidak's method. ($P \leq 0.05$*) ($P \leq 0.01$**) ($P \leq 0.001$***).

4.4.5 Transformation of wild type S and R to form luciferase reporter strains

Luciferase expressing *Mab* S and R strains were made by transforming the wild-type strain morphotypes with the pMV306G13 *ffluc*. Electrocompetent S and R *Mab* cells were prepared and used for electroporation with the plasmid. Transformed ells were then plated and grown in liquid medium in the presence of kanamycin selection. Cell luminescence was tested by using a luminescence assay. After the electroporation of the *Mab* R phenotype, the cells were plated and the colonies which formed were found to be a mixture of both the R and S phenotypes. The mixed S phenotypes did not have any detectable luminescence and had low RLU values. The mixed R phenotype was sub-cultured, resulting in a 10-fold luminescence drop. The mixed S colonies which were recovered had a pink hue as opposed to their usual white colour. The R phenotype was electroporated once again, and the plated transformants were all found to be the R phenotype. The luminescence of these was assessed and high RLU values were detected. Cells were sub-cultured and no decrease in luminescence was observed.

4.4.6 Comparison of the S and R phenotypes of *Mab ffluc* during desiccation

Both morphotypes of the *Mab ffluc* cells were grown following the procedure described in section 4.3.7. The survival during desiccation as a percentage of the 'before desiccation' value for both S and R phenotypes was calculated and is shown in Figure 4.13. It was calculated as follows (time point/Before Desiccation point) x 100 = % survival. Figure 4.13 shows an approximate 1-5 % difference in survival between the two morphotypes during desiccation. A paired t-test was used to analyse the results, and a p-value of 0.13 was obtained, showing that there is no statistical difference between the survival S and R strains during the desiccation survival assays. It should be noted that

preliminary desiccation experiments showed that drying time of the droplet was 45 min hence, T0 was taken after this 45 min for all experiments. The temperature was 24°C and relative humidity 24 %; the data presented in (Figure 4.13) is based on biological triplicates as well as technical triplicates.

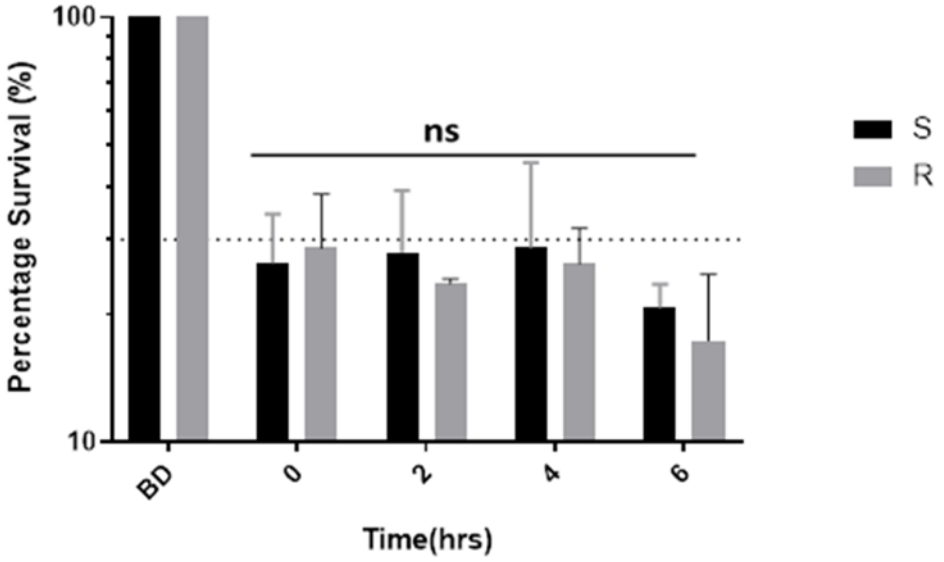


Figure 4.13: Comparison of desiccation survival of the S and R *Mab ffluc* phenotypes**.** Survival was assessed by measurements of RLU, the RLU for BD (Before Desiccation) was 1.66 x 10^5 for the S variant, and 4.40 x 10^5 for the R variant. The dotted line represents 30% survival compared with cells before desiccation. P-value of 0.13 (ns).

4.4.7 UV Dose Response Curve for *Mab ffluc* R and S Phenotype

Both the *Mab ffluc* S and R morphotypes responses were compared with each other in a UV tolerance assay. The droplets of cell suspension were placed on the plate and then the UV dose was given directly to the droplets before drying. Recovery was determined by luciferase assay. Figure 4.14 shows that S morphotype had a higher survival than R morphotype. This difference between S and R morphotypes was statically significant across all UV doses tested.

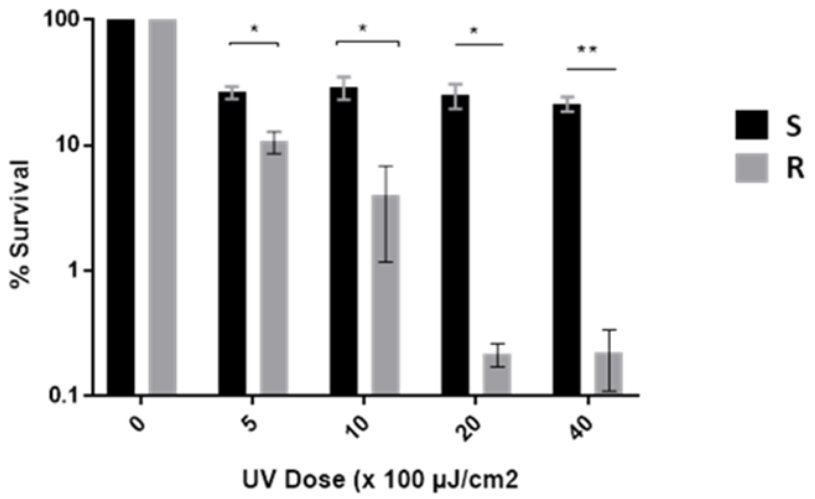


Figure 4.14: UV Radiation Survival Comparison between *Mab ffluc* S and R Phenotypes. Survival was assessed by measurements of RLU, RLU for 100% is equivalent to 2.74 x 10^5 for S variant and 1.87 x 10^6 for R variant. The results revealed significant differences between the survivals of both morphotype variants across all UV doses. ($P \leq 0.01$**), ($P < 0.05$*).

4.5 Discussion

Most of pathogens effecting the respiratory tract are transmitted by aerosols generated by multiple daily habits, such as sneezing and coughing, and even during tidal breathing sneezing (Riley *et al.* 1962; Riley *et al.* 1978). *Mtb* transmission involves three stages: aerosolisation from the host, survival while airborne, and inhalation by a new host. Three aspects of transmission have been investigated in this chapter. Firstly, the different in survival pattern of *Mab* S and R phenotypes in aerosols, secondly, the effect of desiccation on the both of strains and finally, the effect of UV exposure on the survival of *Mab*.

4.5.1 Survival patterns of *Mab* in aerosols

Previous studies into the stability of *Mtb* in aerosols have been conducted by Lever and colleagues, who demonstrated that only 10 % of the initial cells were able to form colonies (S-CFU) after 30 minutes in aerosol (Lever *et al.* 2000). This was in contrast to an earlier study done by Loudon, where *Mtb* in aerosol had a half-life time of 6 hours (Loudon *et al*. 1969). The aim of this current research was to explore the survival of *Mab*

in aerosol and the impact of *Mab* morphotype using a similar experimental model system.

This investigation was done by using the Goldberg drum system with an Ultrasonic NE-U780 Omron nebulizer at PHE Porton Down. The results gathered from the S-CFU analysis in Leicester University, showed that at least 50 % of the cells survived in aerosol (Figures 4.8 and 4.9). In previously reported survival experiments Loudon and associates (1969) reported that after 6 hrs 50 % of *Mtb* bacilli survived in aerosols (Loudon *et al.* 1969). Later, Lever and colleagues demonstrated that *Mtb* exhibited a half-life time of < 30 min within aerosols (Lever *et al.* 2000).

Gram-negative bacteria tend to better survive at low relative humidity and temperatures while Gram-positive bacteria survive better at high relative humidity, but low temperatures (Cox 1989; Grayston *et al.* 1989).The impact of both these factors is strain-dependent as some of microorganisms could survive better regardless of these factors. For example, *Staphylococcus epidermidis* which indicated similar survival patterns in aerosols at relative humidity that ranged from 20 % to 90 % (Thompson et al., 2011). For *Mtb*, it has been shown that the half-life survival in aerosol is 6 hours at 50 % relative humidity (Loudon *et al* 1969) and less than 30 minutes at 75 % relative humidity (Lever *et al*. 2000). However, no previous study has been performed with *Mab* to provide better understanding of survival mechanisms.

The reason for the greater drop in *Mab* survival at 15 min may be due to the stress of the cells in the aerosols and subsequent metabolic adaptations. A rapid initial drop in survival, followed by a slower decline is very commonly seen for lots of different bacteria. This pattern of survival was first reported by Ferry and it was proposed that the initial rapid drop was as a result of rapid desiccation and the slower decline due to oxidative stress (Ferry *et al.* 1958). It worth noting that, in CFU results obtained at Porton Down, the R phenotype showed faster declined at points 90 and 120 min; this may be because of enhanced clumping (a large variability in R CFU counts between the triplicate runs was observed). However, this observation was not confirmed with aerosol samples when they were re-plated at the University of Leicester, as both strains showed an aerosol survival half-life of at least 2 hrs post nebulisation.

4.5.2 Effect of nebulisation on *Mab* by using S-CFU counts

Most nebulisation in several aerosol survival studies of microorganisms were performed previously used Collinson jet nebulisers and it has been used (Lever *et al.* 2000); this can have different effects on cell integrity. However, this study solely used the Ultrasonic NE-U780 Omron nebuliser. For this reason, it was important to investigate the effect of this nebulizer on *Mab* cell integrity.

There was no different between S and R in recovery (Figure 4.10), indicating that cell integrity was not impacted. . The results of this experiment showed that the S-CFU count of the R phenotype increased post-nebulization compared to pre-nebulization. However, this difference was not statistically significant, leading to the rejection hypothesis that the R morphotype would be more efficiently nebulised. The aerosolisation process may have led to the breaking of the clumps in the original suspension, resulting in the production single cells. Additionally, water would have been lost from the reservoir during nebulisation through the aerosolisation process and evaporation, and this would have increased the concentration of cells in the reservoir. Similar observations were seen in a previous study with *B. subtilis* (Stone and Johnson 2002).

The possible of loss of GPLs, representing up to 85 % of the surface-exposed lipids in *Mab* S morphotype (Catherinot *et a*l. 2009) during nebulisation, might contribute to the converting S morphotype to a more R phenotype during nebulisation. This would increase the number of R phenotype cells and subsequently resulted in an over-estimation of the colonies found post nebulisation. Further investigations into the efficiency of aerosolised hydrophobic cells are required to provide solid evidence to accept or reject the above hypothesis.

4.5.3 S-CFU counts from aerosols were not significantly affected by freezing

In order to assess the effect of processing and shipping samples to Leicester and to compare with findings obtained at Porton Down, it was important to detect the effects that freezing had on the sample survival. The results were compared from direct and later plating after being treated in 20 % (v/v) glycerol and freezing.

The reason for selecting glycerol as a cryopreservation solution was because it is able to adequately store and conserve *Mtb* populations in purified sputum (Turapov *et al.* 2016). The freezing and storage of in *vitro Mtb* cultures and sputum-based samples did not influence the viability of the mycobacterial cells in several previous studies (Kim and Kubica 1972; Shu *et al.* 2012; Turapov *et al.* 2016). According to the results of this *Mab* study, the aerosol samples were not significantly affected by the freeze-thaw cycle step (Figure 4.11).

4.5.4 S strain appears more hydrophobic when mixed with the R strain

The aim of this experiment was to test the hypothesis that mixing S and R morphotypes might affect outcome of the aerosol experiments. With each strain compered separately, the CSH results showed that the S *ffluc* strain had significantly higher CSH than S (WT), while there was no difference in CSH between R (WT) and R *ffluc*. The increase in CSH of S *ffluc* compared to the S (WT) might be a result of the plasmid that used by transformation for both strains. A limitation of the experimentation is the lack of a plasmid control strain, which would show whether this is an effect of the kanamycin selection, transformation with the plasmid, or the luceriferase expression. After mixing strain S *ffluc* with R (WT) the CSH of S increased significantly. One possibility might be that some loss of GPLs from S *ffluc* occurred which were then transferred to the R (WT). This would result in the S *ffluc* cells becoming more hydrophobic, this may consequently decrease any difference in aerosolisation compared with the R strain (Figure 4.12).

4.5.5 Comparison of desiccation survival of *Mab ffluc*

It has been suggested that the *Mab* R phenotype would be more virulent, hydrophobic and aerosol transmissible when compared with the S phenotype (Jankute *et al*. 2017). Desiccation is one of the stresses factors that experienced by microorganisms in aerosol. Some microorganisms are capable of surviving almost absolute dehydration by air-drying and without loss of viability (Billi and Potts 2002). *M. tuberculosis* remains viable for around 1 week when dried as an aerosol on glass in physiological saline, but in dust, tubercle bacilli remain viable for 120 days, and the viability is extended to 2 years if stored under vegetable oil (Potts, 1994).

The aim of this experiment here, was to assess another model of transmission of *Mab* during desiccation and to measure the survival of these cells through the use an indirect viability assay using luciferase. The S morphotype had a half-life of 30.6 minutes, and a half-life of 30 minutes was calculated for the R morphotype, with no significant difference between the two variants (Figure 4.13). It should be noted that the greatest decrease in RLU was seen between the BD (Before Desiccation) and T0 time points, i.e. as the samples dried. Following desiccation, the rate of decline in RLU was much lower. This pattern has been reported previously for other bacteria (Ferry et al. 1958;(Hays *et al.* 2005).

In a previous study, the results of *M. avium* survival when desiccated on a surface showed that the mycobacterium exponentially lost viability with steady a half-life of 2.3 days (Archuleta *et al.* 2002).

The limitation of the luciferase assay used to follow response to desiccation in this study, is the lack of correlation with a direct viability assessment, such as CFU as no such determination was made. The viability assay such as (CFU) refers to the cells capable of division, with active metabolism, energy, functional cell membrane. Such cells would undergo DNA replication, would transcribe this into RNA and then translate it into proteins (Hammes *et al.* 2010). In the case of the luciferase assay, this indirect viability method that reflects the metabolic activity of the cell through its dependence on the availability of ATP (Thore *et al.* 1975; KODAKA *et al.* 1996; Venkateswaran *et al.* 2003).

4.5.6 *Mab ffluc* S phenotype has higher UV tolerance than the R Phenotype

It was previously hypothesized that the *Mab ffluc* R variant would have an increased survival rate when compared to the S variant on exposure to UV. Results showed significant differences between the survivals of both variants as UV doses increased (Figure 4.14) with the R variant being more susceptible to increased UV doses than the S variant. It is not known how this drop in RLU, relates to CFU. The difference in the survival of the S and R morphotypes in terms of desiccation and UV radiation exposure may be due to the differences in damage caused by both stressors, as either can cause indirect damage to the cells' proteins, DNA and lipids by catalysing the formation of the reactive oxygen species (ROS), which can cause oxidative stress. On the other hand, the UV-B radiation can cause direct damage to the DNA by stimulating the DNA photoproducts (Agogué *et al.* 2005). One possible explanation for this difference, is the presence of GPL in the cell wall of the S morphotype, which may absorb UV and reduce cell damage. A UV tolerance assay of *M. bovis* BCG *ffluc* showed that *M. bovis* BCG *ffluc* was highly susceptible to UV radiation (Valeria Quimper, data not shown).

4.6 Conclusions

1. There was no significant difference in survival of S and R strains in Goldberg drum aerosol experiments as determined by CFU.
2. Nebulization did not have a significant effect on cell integrity detected by CFU counts.
3. The freeze-thaw cycle did not appear to affect the patterns of survival in the aerosolisation experiments.
4. Mixing S and R strains resulted in an increased CSH in the S strain.
5. Both the *Mab ffluc* S and R phenotypes have the same level of desiccation survival.
6. The *Mab ffluc* S phenotype has higher UV tolerance than the R phenotyp

Chapter 5

Molecular mechanisms of the S to R transition of *Mab* and its associated phenotypes

5.1 Introduction

Understanding the molecular mechanisms underpinning the S to R change in *Mab* is the main theme of this chapter. Achieving this has major importance in our understanding of the organism's virulence in clinical disease, as R variants predominate in this context. The key strategies applied so far, to identify genes potentially involved in the R morphotype have been random transposon mutation and targeted symmetric recombination (Billman-Jacobe *et al.* 1999; Patterson *et al.* 2000; Sondén *et al.* 2005; Medjahed and Reyrat 2009). Although these methods have allowed identification of some key determinants of S to R transition in vitro, they have not allowed determination of how transition occurs during human infection, particularly in cystic fibrosis (Rottman *et al.* 2007; Catherinot *et al.* 2009). It has been suggested that the change from S to R happens mostly through selection during infection (Byrd and Lyons 1999; Cullen *et al.* 2000; Jönsson *et al.* 2007; Catherinot *et al.* 2009).

The methods used to address this problem are discussed below.

5.1.1 Genome and MultiLocus Sequence Typing (MLST) of *Mab*

Genotypic analysis of nontuberculous mycobacteria has proven useful not only in the research of outbreaks and pseudo-outbreaks (Wallace Jr *et al.* 1998b), but also in describing the molecular epidemiology of strains, and in evaluating clonal distribution and expansion (Arbeit *et al.* 1993; Falkinham 3rd 1996; Cangelosi *et al.* 1999; Jönsson *et al.* 2007). *M. smegmatis* was the first RGM to be sequenced and has been considered as a model mycobacterium useful in developing laboratory research. The genome of the *Mab*, type strain ATCC 19977, was published in 2009 and, with an average GC content of 64% comprises a circular chromosome of 5,067,172 base pairs (bp) (Ripoll *et al.* 2009), 1.92 Mb smaller than the genome of *M. smegmatis*. Unusually among RGM the genome is annotated with multiple virulence factors associated with *Mtb* and putative factors previously associated with non-mycobacterial genomes (Ripoll *et al.* 2009).

Whole-genome sequencing (WGS) and phylogenetic analysis are critical to understanding the population structure and genomic evolution of bacterial pathogens

as they assist high-resolution analysis of genetic variants ranging from single-nucleotide polymorphisms (SNPs) to large-scale deletions (Davidson *et al.* 2014). Intensive analysis of whole genome sequences and epidemiological data is providing new insights into the sources and dynamics of large-scale epidemics and pathogen transmission (Quan *et al.* 2018).

MLST is an established typing technique that has been used to investigate different bacterial population structures including those of *Neisseria*, *Escherichia*, *Burkholderia*, *Listeria*, *Streptococcus* and *Staphylococcus* (Enright *et al.* 2000 ; Enright *et al.* 2001; Chenal-Francisque *et al.* 2011). MLST involves sequencing of 350- to 600-bp internal fragments of several housekeeping genes (usually seven). In 2014, Macheras and his colleagues developed a multilocus sequence typing (MLST) pattern for *Mab* based on partial sequencing of *argH*, *cya*, *glpK*, *gnd*, *murC*, *pta* and *purH* (Macheras *et al.* 2014). MLST is suitable for global and local epidemiological studies because it offers data that can be easily compared between laboratories using web-based applications and without exchanging strains. Data are put in a central database and are freely available to laboratories worldwide (http://www.mlst.net/).

5.1.2 Transcriptome analysis

An organism's adaptation to its surrounding environment is a complicated process and traditional methods of analysis are insufficient. Transcriptome profiling provides key information such as transcript content, start sites, mRNA abundance and antisense RNAs (Filiatrault *et al.* 2011). The high presence of rRNA and tRNA, and relatively low amount of mRNA (about 5 % of the total cellular RNA) together with its instability, present the main difficulties in achieving thorough analysis of bacterial gene expression. Depletion of rRNA is one approach enabling more focused sequencing, although newer sequencing technologies also overcome these difficulties (He *et al.* 2010).

Apart from the quantity and size of all RNA species, reliable methods of transcriptome analysis should provide comprehensive information. RNAseq provides all aspects of transcriptome analysis. The usefulness and drawbacks of the all transcriptome analysis techniques have been discussed earlier (in section 1.9.3). To study the transcriptional

profile of *Mab* control and clinical isolates, RNAseq and real time qPCR were applied in this chapter, more explanation of this technique is provided below. Data generation (Figure 5.1 A) and data analysis (Figure 5.1 B) are the main steps in RNAseq experiments.

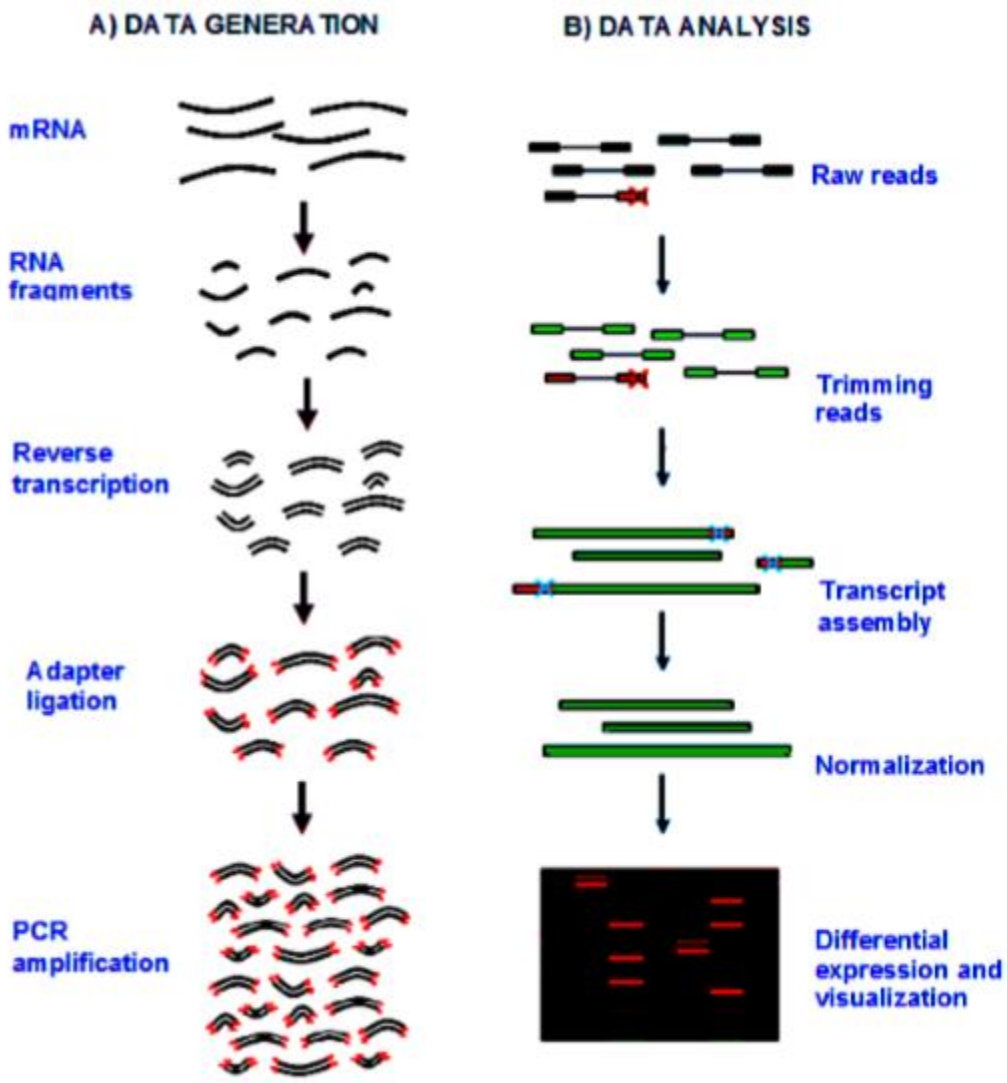


Figure 5.1: A typical RNAseq experiment.
(A) Data generation. The RNA is first extracted, after which DNA contamination is removed and short fragments of RNA are made, then reverse transcribed into cDNA. The adaptors (red) are ligated and fragment size selection is performed. Finally, the cDNAs are sequenced to generate short reads. (**B)** Data analysis. Low-quality reads and artefacts such as contaminant DNA adaptor sequences and PCR duplicates are removed after sequencing. To improve the quality of the sequence, errors (red crosses) are optionally removed. The reads are then assembled into transcripts and the errors are removed via post-assembly processes (blue crosses). The transcripts are then normalised and the expression of each transcript is measured by counting

the number of reads that align to each transcript, followed by measuring the differential expression of genes. Figure adapted from (Martin and Wang 2011).

5.1.3 Quantitative Reverse Transcription Real Time PCR (qRT-PCR)

An alternative transcript-specific targeted approach, often used either to make initial transcript analyses or to confirm features indicated by RNAseq, is qRT-PCR (Pfaffl, 2001). A number of factors may confound measurements including biological variation and non-specific variation. For example, variation in primary sample volume, RNA integrity, RNA purity, efficiency of cDNA synthesis and enzyme or primer performance (Fleige *et al.* 2006). To achieve reliable measurements a suitable mathematical model and gene or genes are needed (Pfaffl, 2000; Andersen et al. 2004).

5.1.4 Previously known genetics changes responsible for the S and R morphotype of *Mab*

Although the loss of GPL, and the implications for this for colony morphology of *Mab* were recognised several years ago (S. Howard *et al.* 2006; Catherinot *et al.* 2007; Rhoades *et al.* 2009), the work of Pawlik, et al. explored and identified the genetic mechanisms by which GPL synthesis, or transport, is actually affected in the R mutants of *Mab*. In their genome-wide approach they searched for potential genetic changes that impacted on colony morphology and reported that these were all restricted to the genes of the GPL locus (Pawlik *et al.* 2013). The spectacular transcriptional arrest observed in the R morphotype due to a CG nucleotide insertion at the 5' end of the *mps1* gene, is one striking example how a putative replication error, which occurred during infection (Rottman *et al*. 2007; Catherinot *et al.* 2009), may create loss-of-function mutants that can outgrow the wild-type strain. Identifying stable genetic changes as the apparent cause of transcriptional arrest and production of the proteins involved in GPL synthesis, argues against the R / S morphotypes being easily interchangeable. Pawlik *et al*. demonstrated that the S morphotypes of *Mab* represent the wild-type strains, whereas the R morphotypes are genetically clearly defined mutants that undergo loss of function due to genetic changes (Pawlik *et al.* 2013).

Aim

Based on previously reported data (Pawlik *et al*. 2013), we aimed to investigate genetic determinants responsible for the S to R switching of *Mab* and associated changes in gene expression.

Objectives

1. To interrogate the *Mab* strains studied here, for previously reported, and novel SNPs and other polymorphisms associated with S to R transition; both targeted PCR and WGS were used to achieve this.
2. To investigate whether the S to R transition can be further explained at the transcriptional level by use of RNAseq and qRT-PCR.

5.2 Materials and Methods

5.2.1 Materials

Proteinase K

10 mg /ml stocks were prepared and stored in small aliquots at -20 °C. Each aliquot would be used once only.

0.5% Acid alcohol

To prepare 0.5 % v/v Acid alcohol solution 0.5 ml of 1 M hydrochloric acid (HCl) was added to 100 ml of ethanol (70 % v/v).

Phenol: chloroform: isoamyl alcohol 25:24:1 v/v/v

A commercially solution available was added prior to each use.

Chloroform: isoamyl alcohol 24:1 v/v

A commercially available mixture of Chloroform: isoamyl alcohol (C/I) 24:1 v/v was used in the current study.

Lysis buffer

50 mM Tris.HCl pH 8.0, 10 mM EDTA, 100 Mm NaCl was sterilised by autoclaving.

TE buffer

10 mM Tris.HCl, pH 8.0; sterilised by autoclaving.

1.4 Guanidinium thiocyanate (GTC) solution

GTC is a chaotropic substance that stabilises nucleic acids and denatures proteins (Mason *et al.* 2003). To prepare 500 ml in a Duran bottle of 5M GTC solution for RNA extraction, the following materials were added:

Table 5.1: Compositions of 5 M GTC solution

Reagent	GTC_a
Guanidine thiocyanate (GTC)	295.4 g
Sodium N-lauryl sarcosine	2.5 g
1M Tri-sodium citrate pH 7.0	12.5 ml
Tween 80	2.5 g

For a 4 M GTC preparation of the amount guanidine thiocyanate was 236.3 g in 500 ml water. 200 ml of DW was added and the materials were dissolved by incubation at 37 °C overnight shaking at 100 RPM. To get the final volume of 500ml, DW water was added to the solution. Aliquots of 20 ml 5 M GTC were made in 50 ml polypropylene conical tubes. The aliquots were kept away from heat and light. Before to use, to each aliquot, 140 µl of β-mercaptoethanol was added to reach a final concentration of 7µl / ml.

DNase I solution

DNase I powder (provided in the Qiagen RNeasy RNA extraction kit) was dissolved in 550 µl DNase-RNase-Free water. Aliquots of 10 µl were made and stored at -20˚C. When required 10 µl of DNase I solution was mixed with 70 µl RDD Buffer which was provided in the RNeasy kit.

Rotor-Gene SYBR Green (2X) master mix

A 2X SYBR Green master mix was provided by the manufacturer (ABGene) containing SYBR Green buffer, (Tris.Cl, KCl, NH_4Cl, $MgCl_2$ and Q-Bond), SYBR I Green, Taq polymerase from *Thermus aquaticus* and dNTP mix was used for the reactions.

5.2.2 Methods

5.2.2.1 DNA Extraction

To extract *Mab* genomic DNA, the procedure of Belisle and Sonnenberg (1998) was used (Belisle and Sonnenberg 1998). Firstly, cells were harvested by centrifugation at 3000 x *g* for 15 min, followed by freezing at -20°C for 4h. The pellets were thawed and suspend well in 2ml Tris EDTA (TE) pH 8. An equal volume of chloroform: methanol 2:1 v/v was added and mixed well for 10 min using a rocker platform. Two phases (organic and aqueous) were generated by centrifugation as above and the aqueous phase was carefully removed without disrupting the pellet. The cells were then dried at 55°C using a heating block then re-suspended in 2.5 ml of TE pH 8 and incubated on rocking a platform at 37 °C overnight after treatment with 10 % (v/v) of 1 M Tris pH 9, lysozyme (200 µg/ml) and RNaseA (20 µg/ml). The next day 0.1 vol of 10 % (w/v) SDS and 0.01 vol of proteinase K (20 mg/ml) was added to the suspension mixed and incubated at 55°C for 3h after which an equal volume of phenol/chloroform/isoamyl alcohol 25:24:1 (v/v/v)was added and the preparation rocked for 30 min. After centrifugation at 3000 x *g* for 30 min, the aqueous phase was removed to a fresh tube and chloroform: isoamyl alcohol 24:1 v/v added to remove the residual phenol; this was performed twice. Genomic DNA was precipitated using 0.1 vol 3 M sodium acetate pH 5.2 and 1 vol isopropanol and freezing at - 20 °C for 1h. DNA was harvested by centrifugation at 3000 x *g* for 30 min. DNA was washed with 70 % (v/v) and 95 % (v/v) ethanol, air dried and finally DNA was suspended in 10mM Tris pH 8.5 and stored at -20 °C.

Quantification of DNA was done using a NanoDrop™ 1000 Spectrophotometer. 1 µl of undiluted DNA suspension of all samples were placed on the NanoDrop™ pedestal. The DNA suspension was measured against a blank of Tris buffer. The concentration (ng/µl) and absorbance ratio (A 260/A 280) of the sample were recorded from the output.

5.2.2.2 Primer design to detect established GPL-related SNPs

The NCBI primer design tool was used (https://www.ncbi.nlm.nih.gov/tools/primer-blast/). Sequences for genes in the GPL locus were retrieved from Gen bank (NCBI). The primers were designed in two sets for target genes. One set of primers was for amplification and another for sequencing. The specificities were confirmed against *Mab* strains using *in-silico* PCR

(http://insilico.ehu.es/PCR/). The primer characteristics such as melting temperature and hairpin loop formation were examined *in-silico* with Oligonucleotide Calculator, OligoCalc (http://biotools.nubic.northwestern.edu/OligoCalc.html). Table 5.2 shows the primers designed for *MAB_4115c* for amplification and sequencing.

Table 5.2: Primers used for current study

Name	Sequence (5'->3')	Length	Tm(°C)	GC%
P1-Forward	ATCGTCAGCGACCTGCTC	18	58	61
P3-Reverse	TGTTCAGCACCAGGAAATCC	20	58	50
P4- Forward	GGATTTCCTGGTGCTGAACA	20	58.09	50
P6-Reverse	GACCACTACTCAGAAAGCCG	20	58	55
P2- Forward	ATCTATGGGGAGATCGGCT	19	57.5	53
P5_Forward	GGTCGAACTCGTCAAGAACA	20	58	50

The primers were confirmed against *Mab* strains using *In silco*-PCR amplification and primer characteristics were examined *in-silico* using Oligonucleotide Calculator (http://biotools.nubic.northwestern.edu/OligoCalc.html). Primers designed for SNP analysis in the intergenic region between *MAB_4100c* & *MAB_4101* are shown in Table 5.3. PCR was performed and the fragments were purified using ISOLATE II PCR and Gel Kit (Bioline) and sent to GATC for Sanger sequencing (https://www.eurofinsgenomics.eu/en/custom-dna-sequencing/gatc-services/).

Table 5.3: Primers used for intergenic region between *MAB_4100c* & *MAB_4101.*

Name	Sequence (5'->3')	Length	Tm (°C)	GC%
Mab4100_F	ATCTCTCACGCAGGCTCTT	19	57.7	52.6
Mab4100_R	CCGTTCAATGGTCTGTTCTC	20	58	50

5.2.2.3 PCR

To perform PCR reaction a GoTaq DNA polymerase master mix (Promega) was used. Table 5.4 details component of the 50 µL reaction mixture.

Table 5.4: Composition of master mix for 50 µL PCR reaction.

Reagent	Sample (µl)	Control (µl)
Sterile water	19	21
2xGoTaq Green Master Mix	25	25
Primer forward (10µM)	2	2
Primer reverse (10µM)	2	2
DNA (template)	2	0

General cycling conditions were as follows, Initial denaturation at 95°C for 5 min followed by 25-30 cycles of 30 sec denaturation at 95°C. Primer annealing for 30 sec at T °C, where T is used the lowest Tm of the pair of primers was used, and the extension time at 72 °C, was based on the expected product size (1 min for 1kb). Final extension was performed at 72 °C for 10 min.

5.2.2.4 Gel electrophoresis and DNA purification

Gel electrophoresis for the visualization of DNA fragments was routinely performed in agarose gels made from 1 × TAE buffer containing 0.5 µg/ml ethidium bromide. Agarose concentration used was 1 % (w/v) for the visualization of DNA fragments. GeneRuler (Thermo Scientific) was used as DNA standard. ISOLATE II PCR and Gel Kit (Bioline) was used for the purification of PCR amplified fragments from the gel according to manufacturer's instructions.

5.2.2.5 Sequencing and bioinformatics of genomic DNA

Whole genome sequencing of both (S) and (R) variants of *M. abscessus* as control and clinical isolates was performed using an Illumina platform offered by MicrobesNG at Birmingham, UK https://microbesng.com/. The bioinformatics tools for visualisation and analysis used are outlined in Table 5.5.

Table 5.5: **Summary of bioinformatics tools used in this study**

Name	Function	Reference
BLASTn	Searches DNA databases for similar DNA sequences	(Altschul *et al.* 1997)
Clustal Omega	Multiple sequence Alignment tool	(Larkin *et al.* 2007)
Primer-BLAST	A tool for finding specific primers to the PCR template	(Altschul et al., 1997)
InSilico PCR	Online program to perform inSilico PCR	(Bikandi *et al.* 2004)
OligoCalc	Oligonucleotide analysis	(Kibbe 2007)
ApE	used for assembly of sequencing traces, virtual digests, sequence annotation, and schematic sequence representation	biologylabs.utah.edu/jorgensen/wayned/ape
snpEff	Annotate variants vcf files	(Cingolani *et al.* 2012)
SPAdes	Assembly to whole genome to give a multi-fasta assembly file	(Bankevich *et al.* 2012)
Samtools	create a “pile-up” file	(Danecek *et al.* 2011)
Bwamem	Align paired F and R reads to reference genome with program BWA-mem to give “SAM” files.	(Li 2012)
Varscan	Call SNP and INDEL variants (creates a “vcf” file) using program Varscan	(Koboldt *et al.* 2012)
Quast	Check assembly quality using Quast	(Gurevich *et al.* 2013)

5.2.2.6 RNA extraction and purification

Total RNA was extracted from all *Mab* strains. These were cultured from frozen stocks on 7H10, then a single colony from each sample was inoculated into 5 ml 7H9. On the following day, the culture was normalized to an OD_{600} (0.003) and allowed to grow to exponential phase (OD at 600 nm between (0.7 and 0.8). 4 x volumes of 5 M GTC solution (5 M guanidinium thiocyanate, 0.5% sodium N-lauroyl sarcosine, 0.1 M-mercaptoethanol, 0.5 % Tween-80) was added and the cells were harvested by centrifugation at 2500 × *g* for 15 min. The supernatant was discarded, and the pellets were suspended carefully in approximately 1ml of 5M GTC solution. All samples were then stored at -80˚C. Thawed GTC/samples were transferred to a labeled 2 ml tube and centrifuged at 14,000 x g for 5 min, and the supernatant carefully removed. 400 µl of RLT buffer and 4 µl of β -mercaptoethanol were then added, followed by 250 µl of 0.1 mm lysing matrix. Cells were lysed using a FastPrep reciprocal shaker (FastPrep-24 5G, MP Biomedicals) with two runs at setting 6.5 m/s for 45s. In between runs, the samples were left on ice for 2 min. Next, the samples were centrifuged at 10,000 x *g* for 2 min, after which the supernatants were transferred to fresh 1.5 ml microfuge tubes. This extracted RNA was processed through an RNeasy Mini Column (Qiagen) following manufacturer's instructions and including an on-column DNase digestion step. According to the amount of sample transferred, an equal volume of RNase free 100 % (v/v) ethanol was added and mixed thoroughly. 700 µl of the samples were transferred to an RNeasy Mini spin column which placed in a 2 ml collection tube. Next, the lid closed gently and centrifuged for 1 min at 10,000 x *g*. This was repeated until all the sample was loaded. Next, 700 µl of RW1 buffer was added to the RNeasy spin column and centrifuged for 1 min at 10,000 x *g*. and the flow-through was discarded. 500µl buffer RPE was then added to the RNeasy spin column then centrifuged for 15 s at 10,000 x *g* to wash the spin column membrane and the flow-through was discarded. Next, 500µl of RPE buffer was added to the RNeasy spin column and centrifuged for 2 min at 10,000 x *g* to wash the spin column membrane. RNeasy spin column placed in a new 1.5 ml collection tube and 50 µl RNase-free water was added directly to the spin column membrane and the lid closed gently. Finally, for 1 min, the samples were centrifuged at 10,000 x *g* to elute the RNA.

5.2.2.7 DNase Treatment

Turbo DNase was applied to the samples to remove any remaining DNA by using the TURBO DNA-free™ Kit (Ambion, Cat; AM 1907). The procedure was followed as per the manufacturer's instructions with the following modifications; 1 µl Turbo DNase reagent and 5 µl Turbo DNase buffer were added to the extracted RNA sample. Then, the samples mixed gently and were incubated at 37 ˚C for 30 min, after which 1 µl of Turbo DNase was added. The samples were incubated for a further 30 min at 37˚C. Next, 10 µl of inactivation reagent was added and for 3 min the samples were mixed gently, and then centrifuged at 16,000 *x g* for 5 min. To avoid carryover of inactivation reagent the aqueous phase was transferred to a new microfuge tube. For more refinement of extracted RNA, an on-column DNase method was used to treat the samples with on-column DNase I digestion. This step was accomplished by using an RNeasy Mini kit (Qiagen, Cat; 74104). The protocol was followed as per the manufacturer's instructions; firstly, the amount of the sample was increased to 100 µl with RNase-free water. Next, 250 µl 100 % (v/v) ethanol and 350 µl of buffer RLT were added. The total of the sample, 700 µl was transferred to the RNeasy spin column before being placed in a 2 ml collection tube and centrifuged at 16,000 *x g* for 15 sec. then, the flow through was discarded and 500 µl RPE buffer was added to the RNeasy spin column and centrifuged at 16,000 *x g* for 15 sec. For on-column DNase digestion, 350µl of buffer RW1 was added to the RNeasy spin column then centrifuged for 15 s at 10,000 x *g* to wash the spin column membrane and the flow-through was discarded. Next, DNase incubation mix was prepared by adding 10µl of DNase I stock solution to 70 µl buffer RDD, this was then added to the RNeasy spin column and incubated at room temperature for 15 min. The flow through was discarded and 500 µl of buffer RPE was added to the RNeasy spin column and centrifuged at 16,000 *x g* for 2 min. The RNeasy spin column was placed in a new 2 ml collection tube and centrifuged at 16,000 *x g* for 1 min. Following, the RNeasy spin column was located in a new 1.5 ml collection tube and 50µl RNase-free water was added directly to the spin column membrane and centrifuged at 16,000 *x g* for 1 min to elute the RNA. Finally, by Bioanalyzer, 1 µl of the crude RNA was used for quantification of RNA (see below), and the remaining of RNA was stored at -80 °C until needed.

5.2.2.8 RNA quantification and integrity

NanoDrop™ 1000 Spectrophotometer at a wavelength of 260 nm was used for RNA quantification. 1µl of undiluted RNA suspension for all samples was placed on the pedestal and the suspension of RNA was measured against a blank of RNase-free water. The absorbance ratio (A 260/A 280) nm and concentration (ng/µl) of samples were reported from the output. The concentration of RNA was calculated using the Equation:

$$C = \frac{A \times 40 \times D}{1000}$$ where: A_{260}=1 is equivalent to 40µg/µl

C = Concentration (µg/µl)

A = Absorbance at 260nm

D = Dilution factor

1000 = Correction (converts ml to µl)

The RNA concentration and integrity of the extracted RNA was measured using the Agilent RNA 6000 Nano-kit purchased from Agilent Technologies (Kidlington, UK). The procedure was followed as per the manufacturer's instructions. Firstly, 550 µl of Agilent RNA 6000 Nano gel matrix was loaded into a spin filter column and centrifuged at 4000 RPM for 10 min. Then, 65 µl of the matrix was filtered into 0.5ml RNase-free microfuge tube. Secondly, 1 µl of RNA 6000 Nano dye concentrate (mixed thoroughly for 10 sec) was added to the 65 µl and mixed thoroughly. The gel-dye mix was then centrifuged for 10 min at 14,000 RPM.

Each RNA Nano-chip contains 16 wells. One well was marked as black G, two wells were marked as blue G, one well as a ladder and 12 wells were allocated for the samples. The gel-dye mixture of 9 µl was loaded into the blue G wells and black G. the RNA 6000 Nano marker of 5 µl was loaded into the remaining 13 wells. Next, for each RNA sample, 1 µl of the RNA ladder was loaded into the ladder well and into each of the 12 sample wells was loaded of 1 µl of the sample. Then, the chip was horizontally located onto a vortex using a device and mixed at 3000 RPM for 60 sec. Finally, the chip was carefully placed into the Agilent Bioanlyser

2100 receptacle. In the list of programs in the Agilent 2100, expert software prokaryotic – Nano RNA was selected to run the chip.

5.2.2.9 NextSeq-500 Illumina workflow for RNAseq

The RNA samples were sent for RNAseq analysis using the Illumina platform (Vertis Biotechnologie – Germany- (www.vertis-biotech.com). The base Space software is integrated with the sequencing workflow. This software is the genomics computing environment in Illumina that is used for data analysis, storage, and collaboration. Further information on the platform can be found in the NextSeq 500 System Guide (15046563 I) at http://support.illumina.com/.

The samples of RNA were fragmented by ultrasound and an oligonucleotide adapter was ligated to the 3' end of the RNA molecules. First strand cDNA synthesis was accomplished using M-MLV reverse transcriptase. The 5' Illumina TruSeq sequencing adaptor was ligated to the 3' end of the antisense cDNA. Prior to performing a sequencing run, the libraries of cDNA with the adaptors were denatured and diluted to 3 pM, after which they were loaded onto reagent cartridges in a flow cell.

Single DNA molecules were bound to the surface of the flow cell and amplified to make clusters. The amplification ranged from 14 – 28 cycles.

5.2.2.10 Sequencing and analysis

Single end sequencing of 75 bases was performed, during which clusters were imaged via 2-channel sequencing chemistry and filter combinations that were specific to each fluorescent-labelled chain terminator. The process of imaging was repeated for each cycle of sequencing. The software performed base calling, filtering and quality scoring following image analysis. As the run progressed, the software automatically transferred base calling files to Base Space for secondary analysis. Fastq files were generated for each sample, containing details of the sequencing, including the quality score. These files were transferred to Leicester for further processing.

Trimming was performed in collaboration with Dr Richard Haigh from the Department of Respiratory Sciences, Leicester University.The universal and Illumina-specific adapters were removed, furthermore any poor quality or repeating sequences using Trimmomatic version 0.32 (Bolger *et al.* 2014). The quality score was set at Phred > 33, with a minimum acceptable fragment of 25. The line of command for trimming was written in Linux-based text. The Fastq files before and after the quality control steps were visualised by the Fastq file 0.11.2 program that using the ALICE High Performance Computing Facility at the University of Leicester.

5.2.2.11 Rockhopper processing and differential expression analysis

The trimmed reads were aligned and mapped against the *Mab* genome using Rockhopper 2.0.3 for Windows (Wellesley – MA, USA). The Fastq files generated after trimming were loaded on to Rockhopper and included three replicates. For the alignment of reads and TPM (Transcripts Per Kilobase Million) analysis, Bowtie2 was used to map reads. Dynamic programming alignment was applied when the reads did not exactly match the reference genome based on the Smith–Waterman algorithm (Smith et al., 1981), this algorithm being limited to 15 % mismatches. These were corrected via insertion and deletion scores in the dynamic programming table based on the Pared quality score (Ewing et al., 1998). The transcript abundance was measured by normalising the expression levels into RPKM (Reads per Kilobase Million).

Before calculating the differential expression (DE), the variance in expression of each gene was estimated via calculating the variance of the gene‘s expression across replicates. As the variance is affected by the expression level, a highly expressed gene would have higher variance across the replicates. Therefore, the local regression model (Anders et al., 2010) was applied to the normalised counts to achieve S estimations of the variance. Then, the DE of genes between conditions was determined by performing statistical tests for the null hypothesis. Thus, a negative binomial distribution was used as a statistical model in which a two-sided p-value was computed as the probability of observing the expression levels of a gene in two conditions. To control the false discovery rate due to multiple tests performed across the set of genes, q-values were set using the Benjamini – Hochberg procedure (Benjamini et al., 1995). A gene is designated as being DE if its expression in two different conditions is different at the threshold level of $q < 0.01$.

5.2.2.12 Differential expression determined by qTR-PCR

Primers used for this work are shown in Table 5.6.

Table 5.6: **Target genes and primers used for qTR-PCR assays**

Gene Identifier	Gene Name	Gene length bp	Location	Product	Primer	Primer Sequence (5'->3')	Tm (°C)	Reference
mab_4097c	*gap*	804	4141888	Hypothetical protein	Forward Reverse	CTGTGGGTGGCACTGGTACT GCGAGCATGGTGAAGACG	56.6 59.4	Pawlick *et al.* 2013
mab_4098c	*mps2*	7746	4142854	Probable peptide synthetises NRP	Forward Reverse	GATTCTGCGGCACTGGTC GATCTCACGGCCCACATC	55.7 56.4	Pawlick *et al.* 2013
mab_4099c	*mps1*	10365	4150596	Probable non-ribosomal peptide synthetises	Forward Reverse	CAACCACTACCCGCTTGC GCACACGCTGGAACCTCT	58.0 59.7	Pawlick *et al.* 2013
16S	16S	1503	1462398	rRNA	Forward Reverse	GGACCACACACTTCATGGTG GAGTCTGGGCCGTATCTCAG	56.9 56.0	Pawlick *et al.* 2013
mab _4115c	*mmpl4B*	2964	4164729	Putative membrane protein, MmpL	Forward Reverse	GCGCTCTCGCTGGGTG CGGGAGACCAGTAGCAGG	59.8 57.3	This study

5.2.2.13 Reverse Transcription

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 lab strain and clinical samples genome. The mtGDPs, consisting of 37 heptamers and octamers, were used to prime for first-strand cDNA synthesis. The following reagents were added to RNAse-free PCR tube on ice:

Buffer	**RT (µl)**	**NoRT (µl)**
dNTPs (10mM)	1.5	1.5
Genomic directed primers (25 pmol /µl)	1	1
RNA	x 0.5 µg*	x0.5 µg*
The total volume made with H_2O	18 µl	18 µl

*The volume of RNA suspension used in the reaction varied on 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 min and then snap cooled on ice to disrupt the secondary structure of the RNA. The following reagents were added to the reaction:

Buffers	**RT (µl)**	**NoRT (µl)**
5 x buffer	6	6
0.1 M DTT	3	3
RNase-inhibitor (RNase In)	1.5	1.5

The mixture was incubated at 25 °C for 2 min to allow the primers to anneal and then placed directly on ice.

Finally, Super Script II reverse of transcriptase of 1.5 µl (300 units) was added to each reaction, the final incubation steps as per table below:

1.5 Reverse transcription		
Buffers	**RT (µl)**	**NoRT (µl)**
Superscript II transcriptase	1.5	0
RNA free water	0	1.5
Total volume	30 µl	30 µl

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 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 reverse transcriptase was inactivated by incubation at 70 °C for 15min. Following reverse transcription, samples were diluted 1:4 (v/v) with water and stored at -20 °C.

5.2.2.14 Quantitative Real-Time Polymerase Chain Reaction using SYBR Green

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

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

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

Cycling conditions (including annealing temperatures and acquisition temperatures) were common for all the reactions. The qPCRs were run on a Rotor-gene 6000 machine (Corbett Research/Qiagen). The following cycling conditions were used for each reaction:

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

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

Respectively, 12.5 µl each reaction of SYBR Green (2 x) master mix was added to 0.1 ml PCR tube. 1 µl of 5 mM forward and 1 µl of 5 mM reverse primer was mixed. The content was brought to 24 µl by adding 9.5 µl DNase - RNase free water then 1 µl of sample cDNA was added to complete the final volume of 25 µl.

5.2.2.15 qRT-PCR result analysis

The relative of gene expression analysis was calculated by the $2^{-\Delta\Delta C}{}_T$ method and modified by (Livak and Schmittgen 2001). The values of CT were normalised against a house keeping gene in this study was (16S rRNA).

Fold difference = $2^{-\Delta\Delta C}{}_T$

ΔCT sample - ΔCT calibrator = ΔΔC

ΔCT (calibrator) = CT target gene - CT reference gene

ΔCT (sample) = CT target gene - CT reference gene

CT: Cycle number that is required for the detectable signal (fluorescence) to cross over the threshold.

Sample: Test sample.

Calibrator: Control sample.

Reference: The reference gene is selected according to the stability of the level of expression during all study conditions and in both control and test sample.

After optimising the conditions, cDNA samples of paired S and R strains were analysed using *gap, mps2, mps1* and *mmpl4b* primers and 16S as a control. The values of CT were obtained after setting the threshold of each the experiment of qPCR and were then used to calculate the relative of gene expression values by using the $2^{-\Delta\Delta CT}$ method.

Statistical analysis was performed using Graph Pad Prism software (Prism 6.0.). Rotor-Gene 6000 (Corbett Life Science – Qiagen) machine and LinReg PCR software were used to run the qPCR and analyse the data respectively. The Graph Pad prism and Rotor Gene 6000 were used for visualisation of data.

5.3 Results

5.3.1 Genome interrogation of the *Mab* control strain by PCR

Previous studies have established a number of polymorphisms (insertions, deletions or SNPs) in the GPL encoding operon that associate with the R phenotype (Pawlik *et al.* 2013). In this section, PCR and sequencing were directed to determine whether these polymorphisms were present in the *Mab* ATCC control S and R strains only.

5.3.1.1 Encoding of the mycobacterial membrane protein (*mmpl4b*)

The examination started by amplifying and sequencing of *mmpl4b* from both S and R morphotypes. Due to the size of *mmpl4* (~ 3.0 kb), this was amplified in two fragments (A and B). Primers P1 and P3 were used for amplification of fragment A while P4 and P6 were used for fragment B (Figure 5.2-A). The fragments were amplified by PCR (Figure 5.2-B) followed by purification and sequencing using P1/P2 for A and P4/P5 for B (Table 5.2).

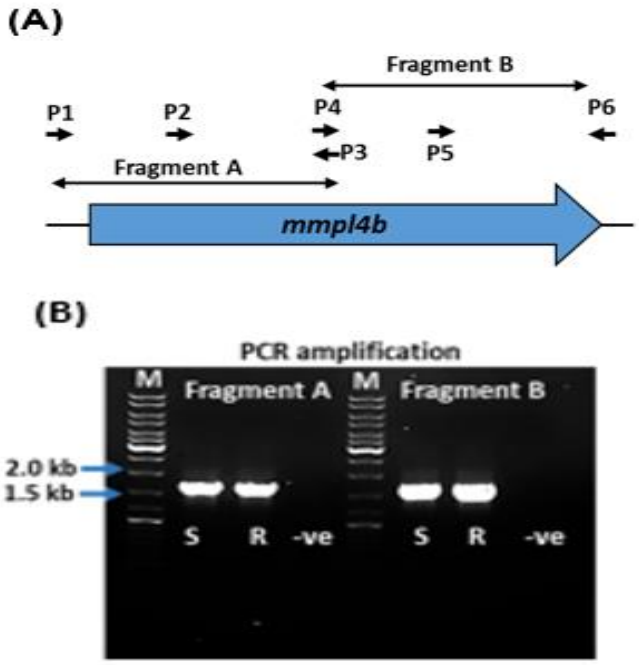


Figure 5.2 : **PCR amplification and purification of *mmpl4b* region in GPL operon**.
(A) Schematics of primers used for PCR amplification and sequencing of the amplified product. **(B)** PCR results of the amplified products. M= DNA marker, S= Smooth, R= Rough, -ve= negative control in which PCR grad water was used instead of DNA.

Sequenced files were merged in to one by removing overlapping regions among the four sections. The final sequences from both S and R were analysed by comparing them to each other, and also to the sequenced strains in NCBI. Results (Figure 5.3) showed similar sequences in both S and R that indicates that variation in the colony phenotype of these forms could not be explained by polymorphisms in this region of *mmpl4b*.

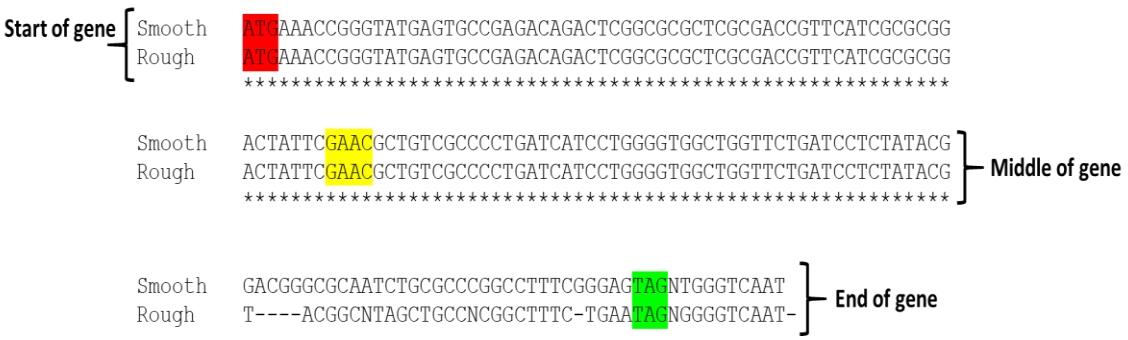


Figure 5.3: SNP detection in *mmpl4b*.
Sequence data were aligned for comparison. Red highlight represents the start codon and green indicates the stop codon. Yellow highlights show the region of adenine deletion identified by (Pawlik *et al.* 2013).

5.3.1.2 Investigation of the intergenic region between MAB_4100c & MAB_4101

Pawlik *et al*. (2013) demonstrated four regions in the GPL encoding operon affecting the S to R transition *(*Pawlick *et al*. 2013). The intergenic region between *MAB_4100c* & *MAB_4101* is shown in Figure 5.4 in which a C to T SNP at 4,161,283 was identified. Amplification and sequencing using the primers shown in Table 5.3 did not reveal any polymorphisms in this region. Further investigation was undertaken by WGS to achieve wider coverage.

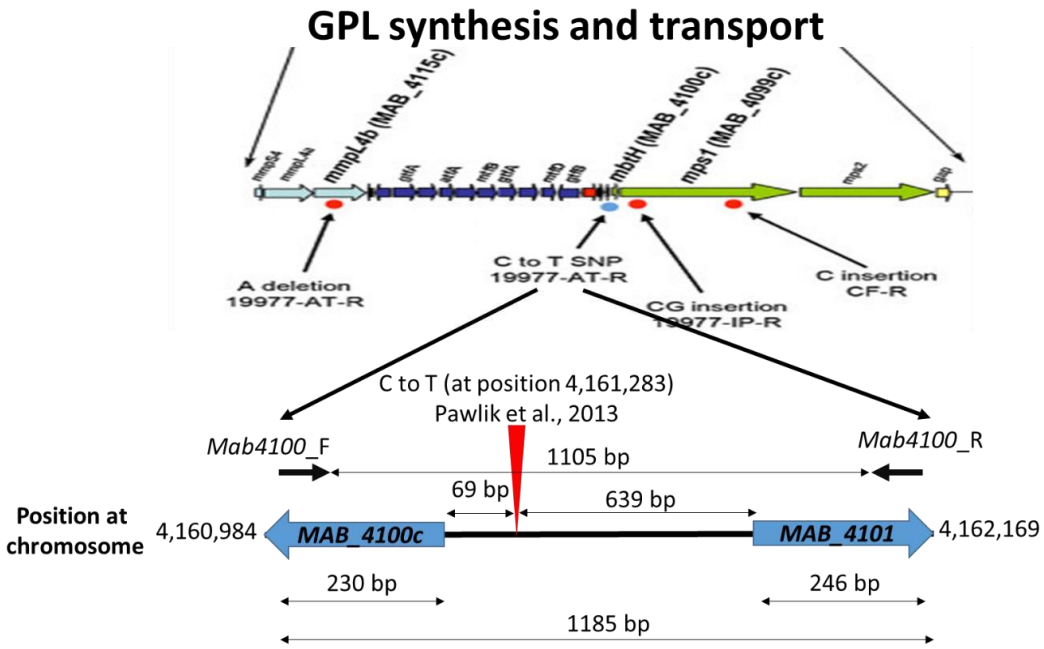


Figure 5.4: Schematics of primers used for PCR amplification and sequencing of the amplified product of intergenic region between *MAB_4100c* & *MAB_4101* in GPL operon of *Mab.*

5.3.2 Whole-genome analysis of S and R variants of *Mab* ATTCC19977 and clinical isolates

The eight selected *Mab* S and R variants, the lab control, CF patients S and R variants and S-SDS/ R-SDS were studied by whole genome sequencing. Control strains are designated S and R and patient strains by assigned numbers as previously described. High quality genomic DNA preparations were used for the downstream genomic analyses by Illumina platform-based sequencing by microbesNG in Birmingham (https://microbesng.com).

5.3.2.1 Read statistics and initial processing

Microbes NG provided read statistics and an analysis made using the Kraken program (http://ccb.jhu.edu/software/kraken) which identified the species origin of the DNA reads (Table 5.7 and 5.8). These data showed that the majority of reads in each sample (82 % to 97 %) were from *Mab*. Read statistics were assessed against the *Mab* ATCC 19977 (CIP104536T) genome sequence which includes a 5,067,172-bp chromosome and a 23,319-bp plasmid (Ripoll et al., 2009); a 23 to 100 fold coverage was achieved (Table 5.8).

Trimmomatic software was used to remove oligonucleotide linkers and poor quality sequence from the Illumina read data. Paired end reads were assembled using the SPAdes program (Bankevich et al. 2012) to give a multi-FASTA assembly file. The Quast program was used to check assembly quality (Mikheenko *et al.* 2018) and the results are shown in Table 5.9 (Bankevich *et al.* 2012). The results show that all but one of the genomes assembled well, with contig numbers between 93 and 237 and total genome sizes between 5,132,988 and 5,597,746. However, the control strain S morphotype was unusual with 899 contigs and a predicted genome size of 5,846,149 bp.

Table 5.7: The top families and genera mapped with Kraken. This data was supplied by Microbes NG.

Sample	Unclassified	Unclassified (%)	Most frequent Family	Most frequent Family (%)	Most frequent species	Most frequent species (%)
S	unclassified	5.58	Mycobacteriaceae	94.23	*Mycobacterium abscessus*	92.42
R	unclassified	0.09	Mycobacteriaceae	99.88	*Mycobacterium abscessus*	99.14
S1	unclassified	11.36	Mycobacteriaceae	88.36	*Mycobacterium abscessus*	86.18
R1	unclassified	15.41	Mycobacteriaceae	84.34	*Mycobacterium abscessus*	82.01
R2	unclassified	2.13	Mycobacteriaceae	97.81	*Mycobacterium abscessus*	97.19
S2	unclassified	2.36	Mycobacteriaceae	97.55	*Mycobacterium abscessus*	96.86
R3	unclassified	2.47	Mycobacteriaceae	97.44	*Mycobacterium abscessus*	96.76
S3	unclassified	2.22	Mycobacteriaceae	97.68	*Mycobacterium abscessus*	97.01
S4	unclassified	9.77	Mycobacteriaceae	90.01	*Mycobacterium abscessus*	87.55
R4	unclassified	10.23	Mycobacteriaceae	89.53	*Mycobacterium abscessus*	87.90
S-SDS	unclassified	10.32	Mycobacteriaceae	89.44	*Mycobacterium abscessus*	89.44
R-SDS	unclassified	12.85	Mycobacteriaceae	86.90	*Mycobacterium abscessus*	86.90

Table 5.8: The table is a summary of Illumina reads data for the *Mab* samples

Sample	Median insert size	Mean coverage	Number of reads	Number of reads w/ insert size > 300
SC	593	51.2226	743861	542158
RC	608	100.5	1189404	1042334
S1	256	23.3084	353504	148490
R1	648	36.8442	426462	376506
S2	194	94.758	1456027	452680
R2	580	24.4723	313799	194433
S3	695	32.9321	390120	303764
R3	555	30.6057	395183	237462
S4	346	60.2872	881381	447292
R4	127	36.1272	653839	171306
S-SDS	335	52.1791	720131	348031
R-SDS	643	49.5687	633776	537853

Table 5.9: Statistics for SPAdes created *Mab* genome assemblies analysed using Quast

Sample	contigs (>= 0 bp)	contigs (>=1000 bp)	Total length (>= 0 bp)	Total length (>= 1000 bp)	contigs	Largest contig	Total length	GC (%)	N50	N75	L50	L75	N's per 100 kbp
Sc	899	423	6127600	5846149	779	160639	6086770	64.01	30787	13963	52	125	0.00
Rc	237	84	5212139	5132988	140	479293	5170612	64.11	167991	90952	10	20	0.00
S1	104	73	5250922	5237113	81	460538	5242317	64.17	134169	83025	13	25	0.00
R1	107	80	5609556	5597746	88	488787	5603166	64.08	142230	60944	15	30	0.02
S2	93	74	5153461	5145295	77	377360	5147377	64.09	121037	66770	12	24	0.00
R2	101	76	5154981	5143881	81	377282	5147030	64.09	136102	83025	13	25	0.00
S3	152	107	5240553	5216692	127	374972	5230559	64.06	135905	77495	12	24	0.00
R3	175	58	5297477	5241544	92	553208	5263217	64.06	173779	64611	12	26	0.00
S3	125	84	5253564	5234812	95	308265	5241587	64.17	129438	63745	11	25	0.00
R4	206	107	5255612	5207029	141	324299	5228740	64.14	101352	55339	12	28	0.00
S-SDS	116	69	5260961	5240614	78	324319	5246291	64.16	142222	66770	12	24	0.00
R-SDS	118	93	5566297	5553712	103	372212	5560166	64.08	118622	60944	15	30	0.02

5.3.2.2 Multilocus sequence typing (MLST)

MLST was performed by submitting the assembled genome data to the Centre for Genomic Epidemiology web server (https://cge.cbs.dtu.dk/services/MLST/). The sequence types are shown in Table 5.10 (and more details are in the appendix). The findings for the control S strain are explored further in section 5.4.3.4. More details about MLST in the Appendix III Table 2.

Table 5.10: Strain types determined by MLST.

Samples	ST	Subspecies
S Control	Unknown	*Mycobacterium abscessus* ; Mixture of ST5 and ST9
R Control	ST5	*Mycobacterium abscessus*
S1,R1	ST9	*Mycobacterium abscessus* subspecies *abscessus* (Alateah et al. 2017)
S2,R2	ST5	*Mycobacterium abscessus*
S3,R3	ST5	*Mycobacterium abscessus*
S4,R4	ST9	*Mycobacterium abscessus* subspecies *abscessus* (Alateah et al. 2017)
S-SDS, R-SDS	ST9	*Mycobacterium abscessus* subspecies *abscessus* (Alateah et al. 2017)

5.3.2.3 Single-nucleotide polymorphism (SNPs) detection and annotation

In order to identify SNPs and indels in the S and R genome assemblies, I used two slightly different methodologies which are shown in the bioinformatics analysis flowchart in Figure 5.5.

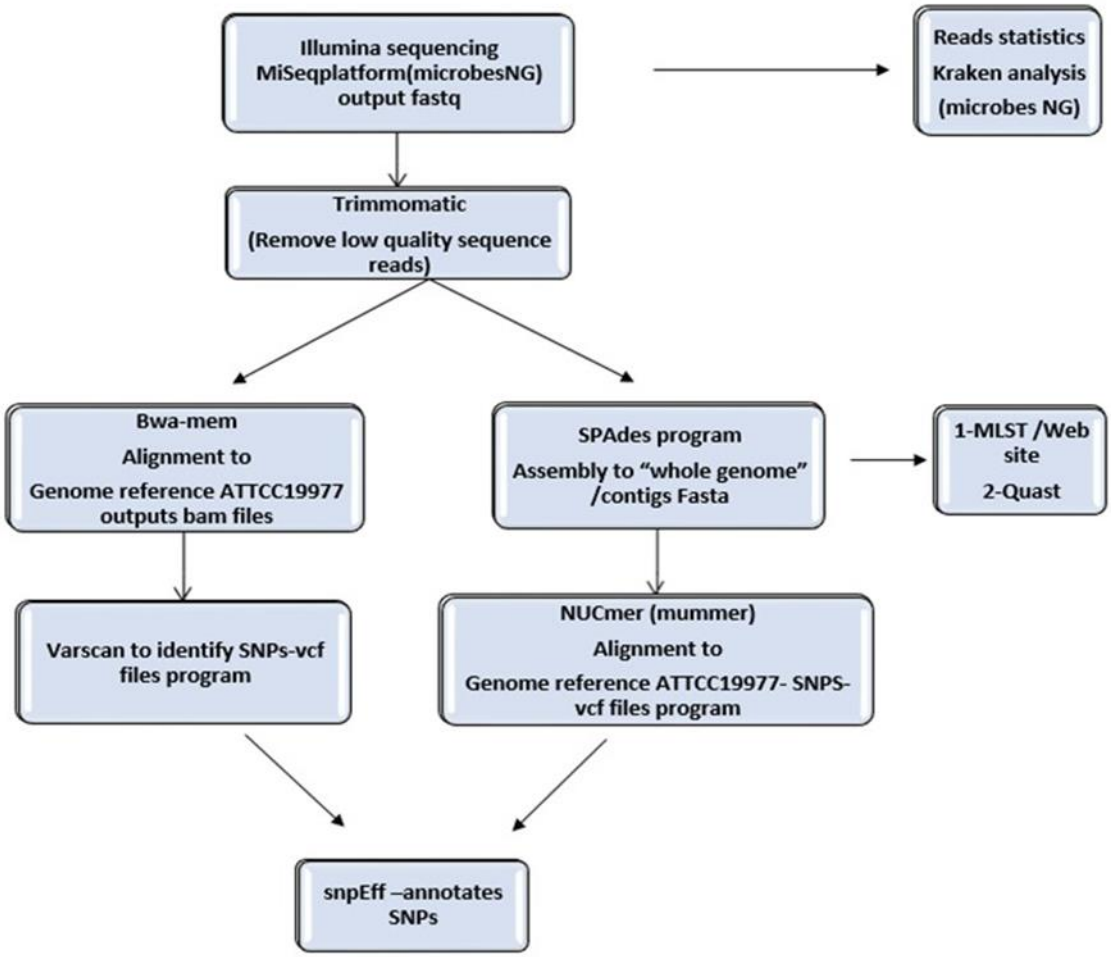


Figure 5.5: Bioinformatics analysis flow chart.

For the first method the paired reads were mapped to the reference ATCC 19977 (accession number CU458896) using BWA-mem (Li 2012) with the default parameters. The sequence alignment map output (SAM file) from BWA-mem was sorted and indexed to produce a BAM alignment file and a "pile-up" file was then created with the program suite Samtools (Danecek *et al.* 2011; Narasimhan *et al.* 2016). The pile-up file was then used to create a variant call format (VCF) file of SNPs and short indels for each sequenced isolate using the program Varscan (Koboldt *et al.* 2012) using a variant cut off of 80 %.

The second method used the NUCmer program of the Mummer package which is capable of handling 100s to 1,000s of contigs generated by assembly of shotgun sequencing and aligning them to a reference genome. Each genome assembly was NUCmer aligned to the reference ATCC 19977 and these alignment files were used to create variant call format (VCF) files using the script Mummer call-SNPs. Calling SNPs with NUCmer uses the genome assemblies (the

sets of contigs) that were made with SPAdes. These were compared to the ATCC 19977 reference genome and any differences were called as SNPs/Indels. Calling SNPs with BWA/Varscan is different, because the original sequence reads from the Illumina data – the Fastq files are used. Here, each read was mapped to the ATCC 19977 sequence and get the bam alignment file that could be examined in Artemis for visual confirmation. Because the two methods use slightly different techniques, they find different types and numbers of mutations, which is the expected consequence of the way the two programs work.

Variants in the VCF files were annotated using the program snpEff (Cingolani *et al.* 2012) with ATCC 19977 as the reference. The numbers of SNPs/indels identified by the two different methods are shown in Table 5.11. The samples identified as ST5 showed modest numbers of SNPs (8 to 408) against ATCC19977 (also ST5), while the ST9 strains showed between 19244 and 21654 differences. The only strain, which showed a large difference in the number of SNPs found by Varscan and Nucmer methods was S; this also had much higher numbers than expected when compared with the R strain that showed only 9 SNPs.

SNPs identified in this study were compared with those reported by Pawlick (Pawlick *et al*. 2013). The comparison of similarities and differences between the two studies are presented in Table 5.12.

Detailed sequence analyses identified indels in the R1 variant that were not present in S1 or genome reference, consisting of a single G insertion in MAB_4098c (*mps2*), which causes a frameshift in the 3' part of the *mps1* gene and results in the truncation of the encoded mycobacterial membrane protein MmpL4b. In addition, genome analysis identified differences outside the GPL locus recognising unique indels in the R2, variant not found in S2 or the reference genome; specifically, a single G insertion in MAB_4323 again causing a frameshift affecting a putative amino acid transporter. A further frameshift is caused by a single T insertion in MAB_4098c (*mps2*) in the R3 genome affecting a probable peptide synthetase, NRP. Finally, a single C insertion in MAB_4099c (*mps1*), which encodes a non-ribosomal peptide synthase, was found in the R4 genome. These results are summarised in Table 5.12.

Table 5.11: SNPs diversity within the *Mab* clinical isolates

Type of strain	Sequence type (ST)	No of SNPs (Varscan)	No of SNPs (Nucmer)
S	Mixture of ST5 and ST9	131	8057
R	ST5	8	9
S1	ST9	19627	21654
R1	ST9	20894	21546
S2	ST5	372	630
R2	ST5	387	629
S3	ST5	382	639
R3	ST5	408	634
S4	ST9	20621	21554
R4	ST9	19244	21596
S-SDS	ST9	20770	21584
R-SDS	ST9	20739	20736

Table 5.12: Comparison of similarities and differences between SNPs/Indels detected linked directly or indirectly to GPL regulation and production, in this study compared to Pawlick *(*Pawlick *et al. 2013*).

			This Study	Pawlick et al. 2013	
Serial Number	**Regions/Genes**	**Type of SNPs/Indel**	**Strain/s with modification**	**Strain/s with modification**	**Protein function**
1-	Intergenic (*mab_0115cmab_ 0116*)	G to T	R3	Not Reported	hypothetical protein
2-	*Intergenic (mab_0115cmab_ 0116)*	G TO A	Not detected	CF- R	hypothetical protein
3-	*mab_2372c*	C to T	Not detected	CF- R	GTP binding EngA
4-	*mab_3539*	CG deletion	R3	CF- R	Putative transcriptional regulator WhiB family
5-	*mab_4098c*	CC insertion	R	Not Reported	Probable peptide synthetase NRP
6-	*mab_4098c*	T insertion	R3	Not Reported	Probable peptide synthetase NRP
7-	*mab_4099c*	CG insertion (reverse strand)	Not detected	19977-IP-R	Probable non-ribosomal peptide synthetase
8-	*mab_4099c*	1C insertion	R4	CF- R	Probable non-ribosomal peptide synthetase
9-	*mab_4115c*	G insertion	R1	Not Reported	Putative membrane protein, MmpL
10-	*mab_4115c*	A insertion	Not detected	19977-IP-R	Putative me mbrane protein, MmpL

11-	*mab__4323*	G insertion	R2	Not Reported	Putative amino acid transporter

Surprisingly, the sequence analysis of R-SDS did not show any SNPs in the GPL encoding region. To investigate this further bam files were inspected in Artemis. A substantial deletion of ~ 7 kb from 4160000bp to 4167200bp was identified (Figure 5.6). Interestingly, this deletion knocked out genes from the region containing genes responsible for GPL production and some other genes of known or unknown proteins; details are listed in Table 5.13.

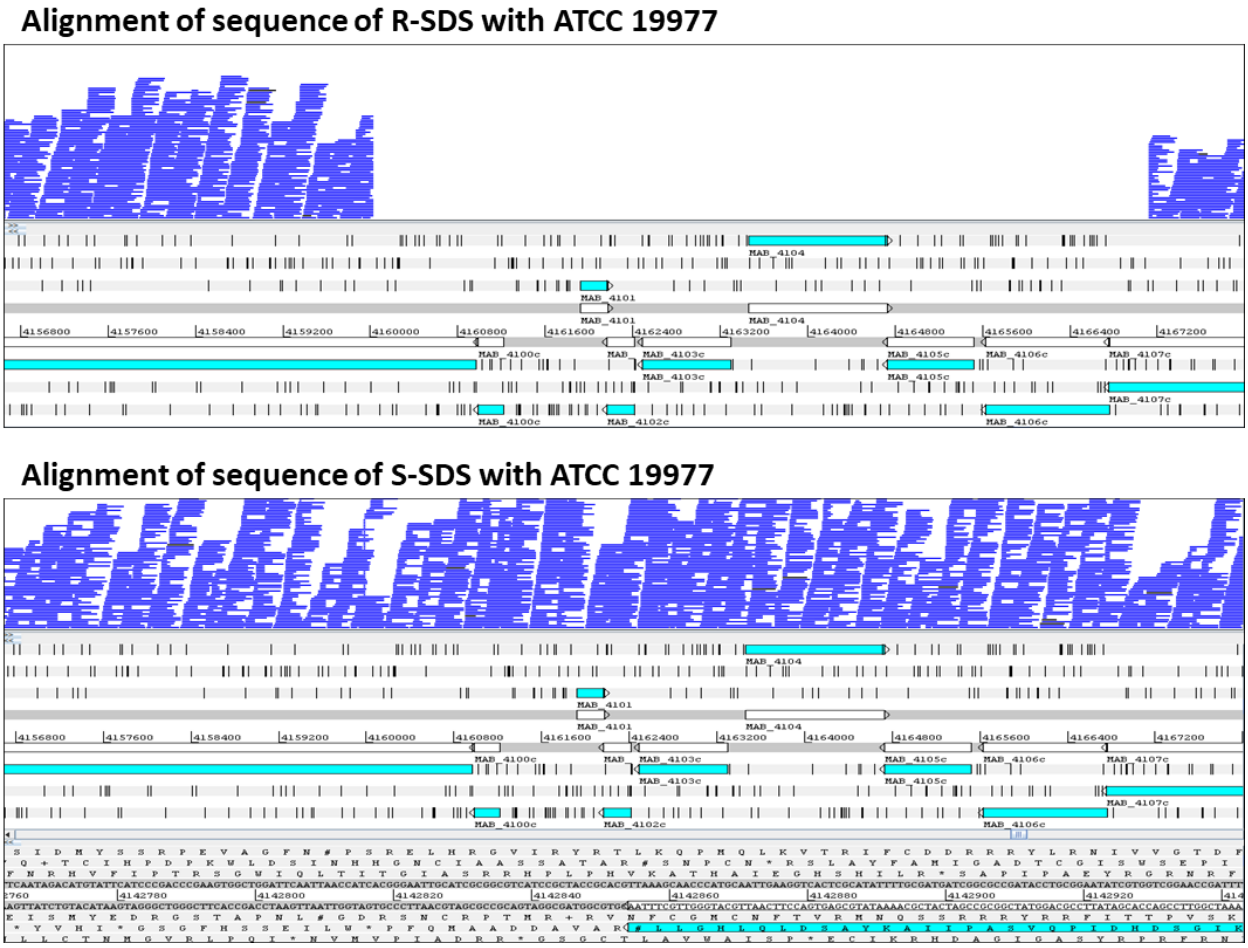


Figure 5.6: Alignment of sequence of R-SDS and S–SDS with genome reference using bam files showing the ~ 7 kb deletion in R-SDS.

Table 5.13: Deleted genes in the GPL encoding locus of the R-SDS strain.

Gene name	Location	Product	Gene length
1-MAB_4099c	4150596	Probable non-ribosomal peptide synthetase	10365 bp
2- MAB_4100c	4160984	MbtH-like protein	231 bp
3- MAB_4101	4161924	Hypothetical protein	246 bp
4- MAB_4102c	4162166	Hypothetical protein	246 bp
5- MAB_4103c	4162485	Probable methyltransferase	804 bp
6- MAB_4104	4163464	Putative glycosyltransferase GtfB	1254 bp
7- MAB_4105c	4164729	Methyltransferase MtfD	789 bp
8-MAB_4106c	4165631	Acyltransferase	1125 bp
9- MAB_4107c	4166752	Glycosyltransferase GtfA	1269 bp

5.3.2.4 Genome data for the S control strain

Scrutiny of the MLST data for the S lab control strain, which was designated unknown, revealed that the 7 housekeeping genes included alleles from both ST5 (reference genome) and from ST9 profiles, suggesting a mixture of ST5 and ST9 strains. Figure 5.7 shows a short region of the alignment of S Fastq reads to the ATCC 19977 genome reference. The BAM alignment file displayed by the Artemis program showed that equal numbers of the S reads mapped at position 1583284 have either an A (reference genome ST5) or a G (ST9) which suggests this sequencing data contains both ST5 and one ST9 and indicates the likelihood that a mixture of genomic DNA from two sources was submitted.

To determine whether the frozen stock of the S strain was mixed, PCR amplification with ST5 and ST9 primers was done on 12 well-separated colonies. The result (Figure 5.8) showed that the S control frozen culture was pure ST5, and that contamination likely occurred during the DNA preparation procedures.

The R control was confirmed as ST5 when mapped to the reference genome ATCC_19977 (Table 5.10). The number of SNPs (Varscan) after analysing was 8 (Table 5.14) while by (Nucmer) program was 11 (Table 5.15). Insertion of a C at location 4144975-caused frameshift mutation detected by both programs.

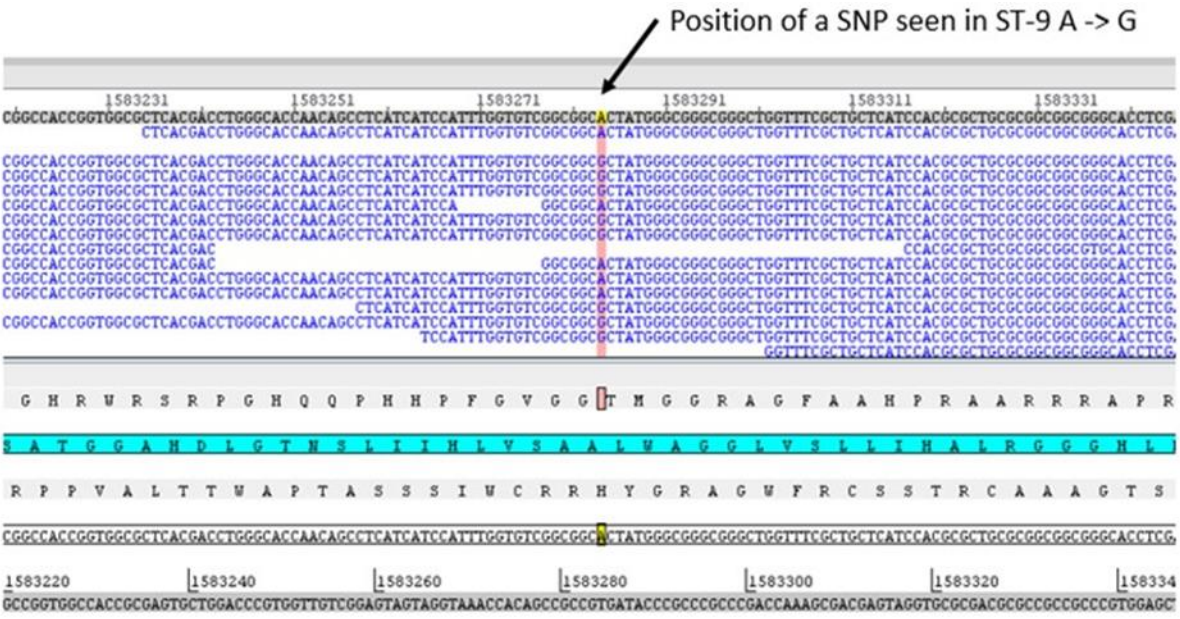


Figure 5.7: Alignment of S control Fastq reads vs ATCC_19977 - BAM alignment file viewed in Artemis.

Strain S control has individual reads which at position 1583284 have an **A** (ATCC_19977 (reference genome=ST5) or have a **G** (= ST-9).Therefore strain S control sequenced DNA appears to have contained both ST-5 and ST-9.

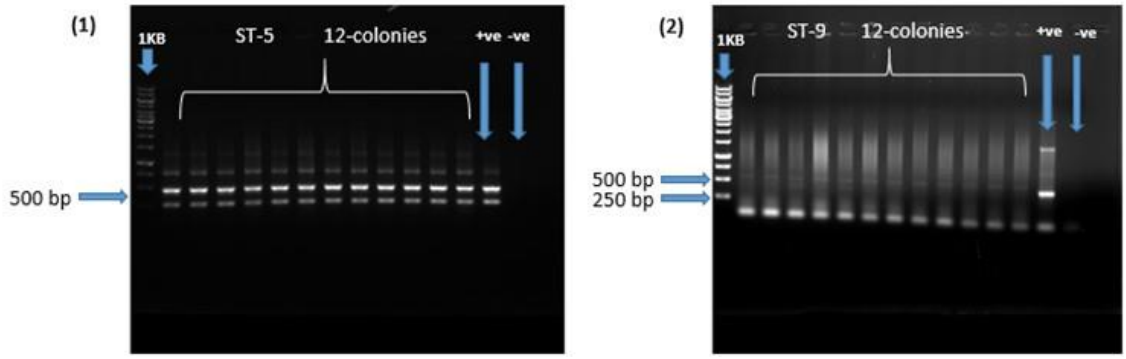


Figure 5.8: **PCR amplification and purification of ST 5 and ST9 to confirm that the stock of S control strain is pure and not contamination**.
(1)+ve control showed ST-5 as well as all other 12-colonies (expected product size 478bp). (2) ST-9 (expected product size 485bp).

Table 5.14: SNPs detected in R strain using Varscan program.

POS	REF	ALT	Effect	Importance	Gene	Locus
81392	A	AC	frameshift_variant	HIGH	MAB_0084	MAB_0084
280242	T	C	stop_lost	HIGH	MAB_0280	MAB_0280
1150461	T	A	missense_variant	MODERATE	MAB_1137c	MAB_1137c
2106418	G	A	missense_variant	MODERATE	MAB_2106c	MAB_2106c
2583761	G	A	missense_variant	MODERATE	MAB_2537c	MAB_2537c
2635541	GG	G	frameshift_variant	HIGH	MAB_2592c	MAB_2592c
4144975	C	CC	frameshift_variant	HIGH	MAB_4098c	MAB_4098c
4620270	G	C	missense_variant	MODERATE	MAB_4537c	MAB_4537c

Table 5.15: Number of SNPs detected in R strain using NUCmer program

POS	REF	ALT	Effect	Importance	Gene	Locus
81392	A	AC	frameshift_variant	HIGH	MAB_0084	MAB_0084
280242	T	C	stop_lost	HIGH	MAB_0280	MAB_0280
1150461	T	A	missense_variant	MODERATE	MAB_1137c	MAB_1137c
1786105	T	C	intergenic_region	MODIFIER	MAB_1783-MAB_1784	MAB_1783-MAB_1784
2106418	G	A	missense_variant	MODERATE	MAB_2106c	MAB_2106c
2583761	G	A	missense_variant	MODERATE	MAB_2537c	MAB_2537c
2635541	GG	G	frameshift_variant	HIGH	MAB_2592c	MAB_2592c
2896546	A	ATGCTCAAATGTGCGCAAATCGGACGATTTTGCGCACATTTGAGCATGCTCGCGCTT	intergenic region	MODIFIER	MAB_2847c-MAB_2848c	MAB_2847c-MAB_2848c
4144975	C	CC	frameshift_variant	HIGH	MAB_4098c	MAB_4098c
4356980	A	G	synonymous_variant	LOW	MAB_4284c	MAB_4284c
4620270	G	C	missense_variant	MODERATE	MAB_4537c	MAB_4537c

5.3.3 Transcript analysis comparing S and R morphotypes.

To further investigate the basis for the two phenotypes, mRNA was isolated from the *Mab* strains and analysed by qRT-PCR and RNAseq.

5.3.3.1 qRT-PCR generated expression data shows no significant downregulation of the GPL locus in R variants

Each S/R lab strain of the ATTC 19977 and the clinical strains were examined. Fold changes between S and R variants were selected and compared with fold changes observed for strains from patient 1 and patient 2 (Figure 5.9). Although differences in expression were detected, these did not attain statistical significance for the two non-ribosomal peptide synthase genes (*mps1, mps2*), the glycopeptidolipid transport gene (*gap*) and also for MAB_4115c (*mmpl4b)*.

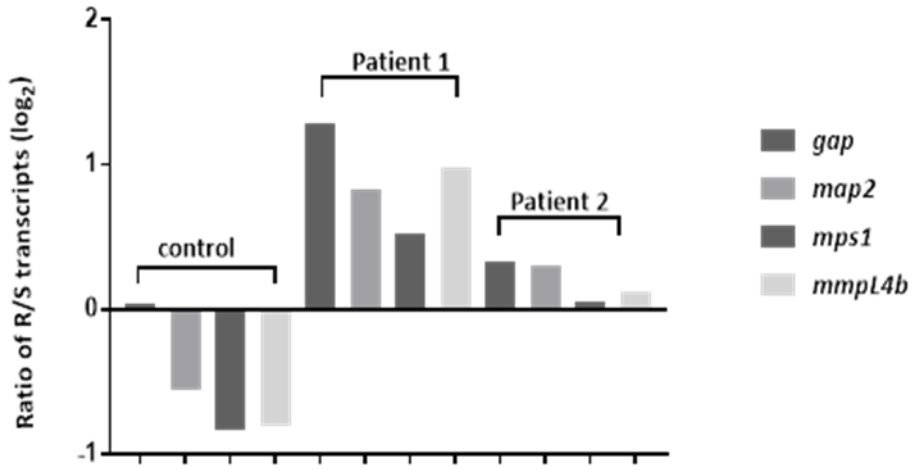


Figure 5.9: qRT-PCR results showing the ratio of gene expression R vs S.
qRT-PCR analysis was undertaken for the 4 genes of the non-ribosomal peptide synthase gene cluster (MAB_4097c, *gap*; MAB_4098c, *mps2*; MAB_4099c, *mps1*). In the GPL locus one operon and (MAB_4115c, *mmpl4b)* the encoding the transporter of GPL to the cell envelope, which for the control strain was slightly decreased according to qRT-PCR result, in R as compared with respective S variant. However, the Patient1 and Patient2 -R clinical isolates showed increased of expression compared with their respective S variants. The expression profile was performed using RNA extracted at two independent occasions and three technical replicates were used from each RNA template for each run of qRT-PCR.

5.3.3.2 RNA Sequencing results

Four RNA samples from control S and R (two independently extracted RNA from each) were chosen as being of sufficient quality and were sent for RNAseq. Libraries were made using Illumina library kits and the library pool was sequenced on Illumina NextSeq 500 system generating 74 bp read lengths.

Table 5.16: Reads information taken after RNAseq

Sample	Adaptor	PCR cycle	Total sequences
S control 1	AGTCTTCTR	13	9,344,374
S control 2	CATTCGCTR	13	7,391,468
R control 1	TCTACTCTR-	13	7,1236,42
R control 2	ATCCTGTG	13	8,769,061

The sequencing report shown in Table 5.16 was generated containing the details of sequencing results which showed the number of PCR cycles, the adaptors and total number of reads generated per sample. The total number of reads for all samples was 32,628,545 with the number of average reads per sample of 8,157,136. The reads were then calculated for base and sequence quality scores, which were detailed in Fastq files.

After trimming was completed, the good quality reads were aligned against the reference genome of *Mab* (ATCC 19977 (CIP104536T). This was done in Rockhopper version 2.0.3 and produced a mean mapping percentage of 98 % across all samples (Table 5.17).

Table 5.17: Overview of the Alignments % of Data from RNA-seq were trimmed and aligned against reference genome using Rockhopper software version 2.0.3 for windows. *Mab* S and R phenotypes showed 98 % alignments.

Sample	Total reads	Aligned reads	% Aligned reads
S control 1	9,250,854	9,088,666	98%
S control 2	7,278,717	7,160,581	98%
R control 1	7,021,146	6,881,520	98%
R control 2	7,021,146	8,566,033	98%

The transcriptome of bacteria is predominated by ribosomal RNA (rRNA), creating a technical challenge to mRNA analysis (He *et al*. 2010). The proportion of rRNA in the samples sequenced here was between 89 % and 90 %.

5.3.3.3 Differentially expressed (DE) genes

For this initial analysis of the expression data generated by Rockhopper (5.2.6.1), two rules were applied to assess potentially differentially expressed genes: 1) A change of two-fold or more between samples and 2) a q value equal to, or below 0.01. The observed differences between results obtained with qRT-PCR and RNAseq samples are summarised in Table 5.18.

5.3.3.4 Transcriptomic analysis of the *Mab* -S/R and clinical isolates by RNAseq

As depicted in Figure 5.10 RNAseq confirmed presence of mRNA transcripts for the *mps1-mps2-gap* operon in the control R variant. These findings corroborate and extend the qRT-PCR results (Table 5.18) and clearly show that the C insertion at position 4144975 in R variant seems not to cause a transcriptional arrest similar to that described by Pawlik and colleagues (Table 5.14). RNAseq analysis of the S/R pair did not link any of the other differentially regulated genes to indels or SNPs identified by the S/R genome sequencing approach. Although the *mps1-mps2- gap* operon is the key locus for GPL production, RNAseq and qRT-PCR did not show significant differences in expression profiles.

In addition, several other genes were identified as differentially expressed between R/S an analysis of RNAseq results. For example, it is observed significantly increased transcript expression levels in S strains compared with R strains, in the immediate vicinity of MAB_2254c-2259 (Figure 5.11 and Table 5.19). The genes are predicted to encode a non-ribosomal peptide or polyketide synthesis operon.

Table 5.18: **Correlation between RNAseq and qRT-PCR** results analysis

Name	Product	Expression S	Expression R	R/S RNA seq	R/S qPCR
MAB_4097c *gap*	hypothetical protein	350	325	0.93	0.99
MAB_4098c *mps2*	Probable peptide synthetase NRP	371	321	0.87	0.68
MAB_4099c *mps1*	Probable non-ribosomal peptide synthetase	555	428	0.77	0.56
MAB_4115c *mmpL4b*	Putative membrane protein, MmpL	217	215	0.99	0.57

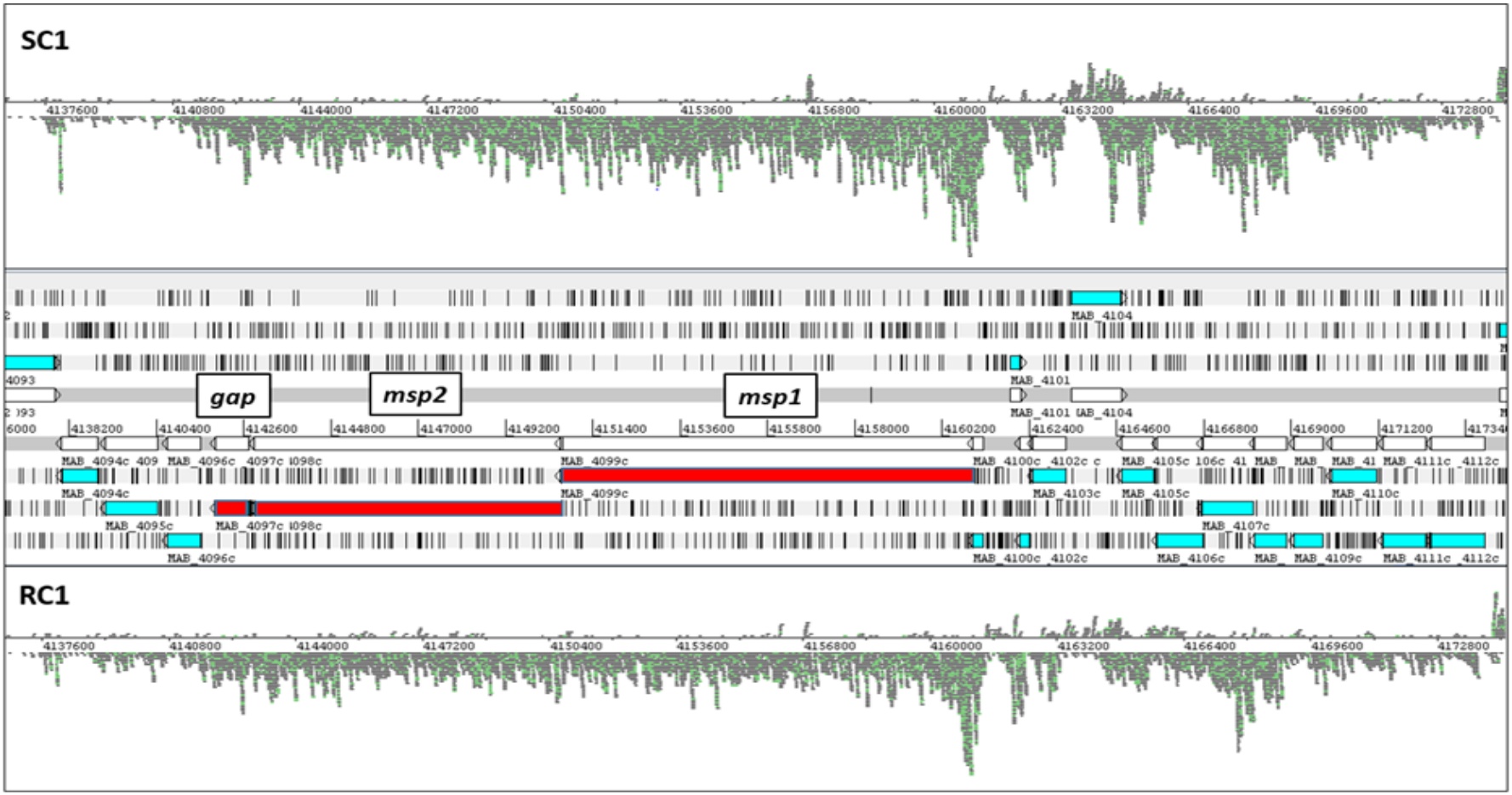


Figure 5.10: mRNA transcription of the *gap, mps2, mps1*, gene of the GPL locus in *Mab* strain R.
As shown by RNAseq read density analysis data, generated from strains S and R. The image shows the open reading frames as depicted by the Artemis (Rutherford *et al.* 2000) software.

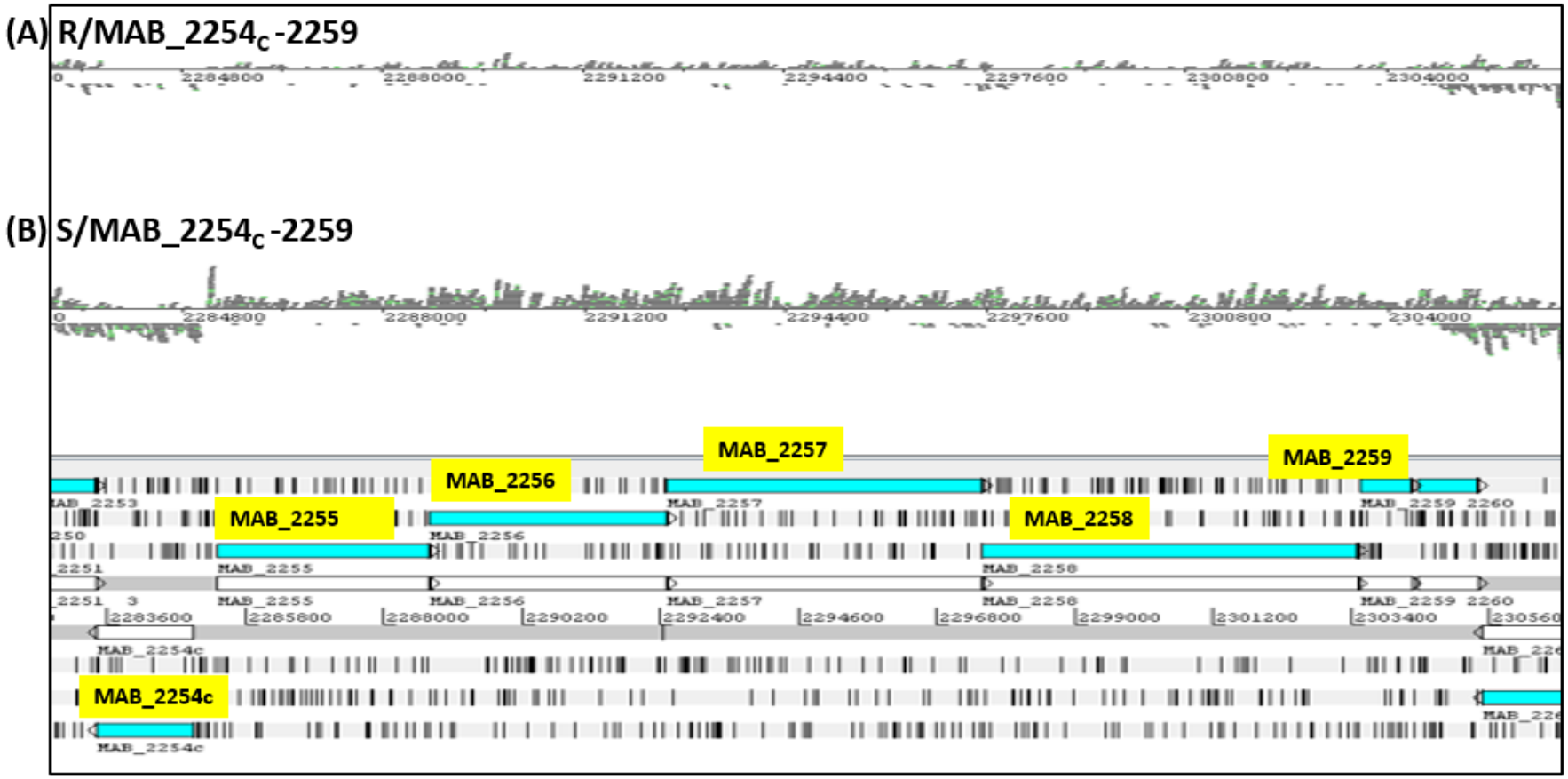


Figure 5.11: mRNA transcription of the MAB_2254c-2260 in *Mab* strain.
(A) Alignment of sequence of MAB_2254c-2259/ R to genome reference ATCC19977 **(B)** Alignment of sequence of MAB_2254c-2259 /S to genome reference ATCC19977. The image shows down regulation of MAB_2254-2259 in R. The open reading frames as depicted by the Artemis software (Rutherford *et al*. 2000).

Table 5.19: **Expression profile** of a down regulated gene cluster in R strain.

Name	Product	Expression S	Expression R	R/S	q Value S vs R
MAB_2254c	PPE family protein	77	5	0.06	2.24E-58
MAB_2255	Probable non-ribosomal peptide synthetase	66	6	0.09	1.21E-25
MAB_2256	Probable polyketide synthase	101	14	0.14	1.49E-05
MAB_2257	Probable polyketide synthase	90	12	0.13	2.09E-05
MAB_2258	Peptide synthetase and polyketide synthase	67	9	0.13	1.94E-06
MAB_2259	Putative O-methyltransferase	42	5	0.12	1.01E-16

5.4 Discussion

WGS is useful for elucidating sources and transmission dynamics of disease outbreaks, but there is relatively little information concerning pathogenic RGM. WGS provided the first convincing evidence for person-to-person transmission of *Mab*; it also solved issues of classification of *Mab* into subspecies (Adékambi *et al.* 2004).

Genetic information also provides a lot of information regarding the lifestyle of the bacteria in natural and infection conditions. WGS information was used here to investigate the genetic basis for the differences between S and R variants. As reported previously, such understanding could provide further insight into the pathogenicity of both variants of the *Mab* (Ripoll *et al.* 2007).

The GPL genetic locus has been recently described in *Mab* and *M. chelonae*, and is syntenous with the *M. smegmatis* 'GPL locus', with an extensive conservation of the predicted protein sequences (Ripoll et al., 2007). Synthesis and transport of GPLs has been the subject of multiple studies focusing on the *mbtH-mps1-mps2 gap* genomic locus (Ripoll et al., 2007). MbtH-like proteins are usually found in non-ribosomal protein synthetase clusters responsible for siderophore and antibiotic peptide synthesis (Quadri *et al*. 1998). *mps1* and *mps2* belong to the mycobacterial non-ribosomal peptide synthase (NRP) family, catalysing synthesis of the peptide moiety of the molecules (Billman-Jacobe et al., 1999). The GPL addressing protein Gap was recently described as being involved in GPL transport to the cell surface in *M. smegmatis* (Sonden et al., 2005). Moreover, MmpL4b was reported to be implicated in GPL transportation in *Mab* (Medjahed and Reyrat, 2009). Comparative transcriptomics also allowed differences in gene expression to be identified that were not related to GPLs, but linked to genes playing a role in the persistence of infection in other mycobacterial pathogens (Sherman et al., 2001; Shi and Ehrt, 2006; Gerasimova et al., 2011; Venugopal et al., 2011).

5.4.1 Genome interrogation of the *Mab* control strain by PCR

It has previously been described that insertions, deletions or SNPs in different genes of the GPL encoding operon resulted in a switch from S to R phenotype (Pawlik et al. 2013). The investigation here, started by PCR amplifying and sequencing of the *mmpl4b* gene

from strains with S and R phenotypes to detect potential deletion of adenine (A) between position 4178562 and 4178565 in *mmpl4b*; this was not detected here (Figure 5.2). The next target was the intergenic region between MAB_4100c & MAB_4101 (Figure 5.3), where C to T SNP at position 4,161,283 was targeted; again, no SNPs were found.

In view of these results, it was concluded that further PCR-sequencing based analysis would be inefficient and further analysis was pursued through WGS applied to control and clinical isolates.

5.4.2 WGS analyses

In-house written scripts were used to process the Illumina reads. Paired ends were preferred for genome assemblies. The sequences (contigs number between 93 - 237) meeting the quality threshold, were further used for investigation of genetic differences between S and R variants of *Mab.* It was immediately apparent that anomalous data was obtained from the S control strain and this is discussed further below.

Nucleotide sequence polymorphisms were investigated to understand the genetic diversities of S, R and between control and clinical isolates. The isolates were first assigned to MLST groups (section 5.3.2.2). This analysis revealed that we had two types of sequence types (ST5 and ST9) in our collection. Having two different sequence types in a small sample number, provided the opportunity to look at the genetic diversity, not only between the two variants, but also between the two sequence types of both variants. One of the major findings of this study is the number of SNPs of ST5 and ST9 strains when compared with ATCC 19977 as a reference genome. ST9 strains showed more than 20,000 SNPs compared to this reference genome. One limitation of this analysis was the lack of availability of a ST9 complete genome sequence in NCBI. The available complete genome sequence is from a ST5 strain and therefore, this might be the reason that ST9 strains end up with 100 fold more SNPs compared to ST5 isolates (section 5.3.2.3).

It was striking that the R control strain showed less than 10 SNPs against the reference genome, while the S control strain showed over 131 by Varscan and 8057 by Nucmer (Table 5.11). Nonetheless, over 92 % of the S control reads mapped to the reference genome. Further analysis (Figure 5.7) indicated that a mixture of ST5 and ST9 genomic DNA had been submitted, presumably due to a cross-contamination error in preparation, and further analysis of this dataset was not pursued. To determine whether contamination was present in the originally received S control strain, colony PCR was performed (Figure 5.8) and provided evidence that this was not the case.

5.4.2.1 Genetic diversity of S and R variants of *Mab* ATTCC 19977 and clinical isolates

Genome sequencing of clinical isolates of *Mab* revealed extensive genetic diversity across strains, with more SNPs between the reference sequence (ST5) and S and R variants (ST9). Genetic features of R mutants have been explored by several authors (S.T. Howard *et al.* 2006; Rottman *et al.* 2007), and have generally focused on GPL synthesis and transport. Genome-wide approaches have largely come to the conclusion that polymorphisms effecting the S to R transition are restricted to genes of the GPL locus, resulting in transcriptional arrest or truncation of proteins involved in GPL synthesis. The findings argue against an easy reversibility of the R/S morphotypes. Such an adaptability has previously been suggested based on observations that a single human *Mab* isolate spontaneously dissociated into an R and an S variant, where the R variant has been considered as the wild-type and the S variant as the attenuated mutant phenotype (Byrd and Lyons, 1999; Howard *et al*. 2006). In contrast, the data presented here strongly suggest that the S morphotypes of *Mab* represent the wild-type strains, whereas the R morphotypes of *Mab* are genetically clearly defined mutants that undergo loss of function due to genetic changes. These findings are consistent with results from previous mouse infection experiments using the CF strains, where an R to an S morphotype conversion was never observed (Catherinot *et al*. 2007; Rottman *et al*. 2007). These insights also allow the genetic changes in clinical isolates of *Mab* to be predicted in a genomically restricted region, which is of importance for the development

of novel molecular diagnostic tools that can rapidly identify potentially persistent *Mab* clones.

Interrogation of the GPL locus in the R strains (Table 5.12) revealed 11 polymorphisms against the reference S genome, 5 of which had not been detected by Pawlik and colleagues. Further analysis of these would be necessary to determine whether they impact on GPL production. In the R-SDS strain, a ~7kb deletion was found, starting in *mps1* (*MAB_4099c*). This would clearly impact GPL production and demonstrates yet another polymorphism effecting S to R transition, although whether this occurs in natural isolates has yet to be determined (Table 5.13).

Interestingly, none of the GPL locus polymorphisms reported by Pawlik were detected in the R control strain. Rather, a C insertion causing a frameshift in *MAB_4098c* (*mps2*) was found (Tables 5.14 and 5.15). Of the remaining detected polymorphisms, none appeared relevant to GPL production although this cannot be excluded. Possible explanations for the R control are explored further below.

5.4.3 Transcript analyses

To determine whether any of the polymorphisms detected above affected transcription of genes related to GPL expression, mRNA was isolated.

5.4.3.1 qRT-PCR analyses

After confirmation of SNPs in various genes (MAB_4097c*, gap;* MAB_4098c*, mps2;* MAB_4099c*;* MAB_4115c*, mmpl4b*) responsible directly or indirectly for GPL synthesis and transport, Pawlik and colleagues reported that the insertions and deletions they detected either downregulated, or stopped, the expression of above genes. However, unexpectedly in this study genes showed, if anything, increased expression in two patient isolates and non-significant changes compared with control strains. These results suggested that more changes than those reported by the Pawlik study can cause S to R transition.

5.4.3.2 RNA Sequencing

RNAseq was chosen in this study because of its outstanding capacity compared to other technologies for deducing and quantifying transcriptomes. This approach is capable of cataloguing all transcript species including mRNA, small RNAs and non-coding RNAs. Additionally, using RNAseq, researchers have been enabled to determine the transcriptional structure of genes and post transcriptional modification, as well as quantifying the expression levels of each transcript that changes under various conditions (Y. Wang *et al.* 2009).

The cDNA pool was sequenced on an Illumina NextSeq 500 system using 74 bp read length. This complies with the recommended range of 30-400 bp (Wang et al., 2009). Although longer reads can reduce the complexity of transcriptome assembly, shorter reads can enhance the number of high quality reads (Martin and Wang 2011). The number of PCR amplification cycles varied between samples.

The percentage of mapped reads was satisfactory. Rockhopper uses Bowtie2 as integrated alignment software which has been attested for its superior capacity compared to other alignment tools (Tjaden 2015). The results are based on mapped reads to a *Mab* reference genome and the percentage of alignment of reads for the control strain between the replicates of each strain showed satisfactory and successfully aligned reads (98 %).

The data analysis discussed here has been directed to a first level assessment specific to the S to R transition and it is recognised that many other features could be analysed.

The *mps1-mps2-gap* operon mRNA transcripts were detected in both S control and R control RNAseq data confirming and validating the qRT-PCR results (Table 5.18) and providing an evidence that the C insertion at position 4144975 did not produce a transcriptional arrest similar to that reported by Pawlik *et al* (2013). This might also suggest the possibility of production of a non-functional Mps1 protein due to alteration in amino acid sequences. Overall, no correlation was observed between RNAseq analysis with other differentially regulated genes to indels or SNPs identified by the S/R genome

sequencing approach. Conclusively, both qPCR and RNAseq analysis did not detect any significant and expected differences in expression profile of any of the previously identified genes responsible for GPL production.

Furthermore, downregulation of locus encoding MAB2254c-2259 was observed in R morphotype. The locus is composed of 6 genes (Table 5.19) and most encode putative polyketide synthases, which are multi-domain enzymes that produce polyketides. The biosynthesis of polyketides shares significant similarities with fatty acid biosynthesis (Khosla *et al.* 1999; Jenke-Kodama *et al.* 2005). The down regulation of this locus supports the overall outcome of transcriptomic analysis of variable expression of genes associated with lipid metabolism in R variants compared to S morphotypes. Another possible explanation might be provided by the detected lack of transcription in the MAB_2254*c* to MAB_2258 region (Figure 5.11 and Table 5.19). In particular, the probable non-ribosomal peptide synthetase (MAB_2255) might make a significant contribution to GPL production.

One of the major limitations of this study was that the transcriptomic analysis was only performed for RNA isolated on two independent occasions and two technical replicate cultures, and this might have affected the outcome of the expression data. For example, we only identified genes/loci which showed significant differences in up, or down-regulation between each morphotype. Such differences could be non-significant, or only slightly significant when a large sample size is used. The MAB2254c-2259 locus shows a 10-fold difference in expression. With more replicates other differentially expressed genes may have been identified.

5.4.4 Conclusions

1. Standard PCR and sequencing of amplified fragments are insufficient for SNP detection.
2. None of the GPL locus polymorphisms reported by Pawlik were detected in the R control strain by WGS. In contrast, several of these were detected in the clinical R strains and a C insertion in *mps1*, causing a frameshift, was found in the R control
3. A ~7kb deletion was detected in the R-SDS strain including genes associated to GPL metabolism.
4. RNAseq and qPCR did not show significant differences in expression profiles of genes the GPL locus.
5. Down regulation of the locus *Mab*_2254c-2259 was observed in R morphotype compared to S morphotype.

Chapter 6

General Discussion

6.1 General discussion

This study was directed to characterisation of the S and R variants of *Mab* with a major focus on CSH and its relation to propensity for bacterial aerosolisation. *Mab* is one of the major participants in infected lungs of immunocompromised patients or patients suffering with cystic fibrosis (Parkins and Floto 2015; Torres-Coy *et al.* 2016). Therefore, clinical isolates (both S and R variants) of *Mab* isolated from four cystic fibrosis patients were included in this study. Interestingly, very few differences were observed in the biological characteristics of the control and clinical isolates, indicating relatively consistent phenotypes of S and R variants.

Apart from standard differences in colony morphology, both variants showed similar strain-related patterns, with R variants all demonstrating longer lag phases (20-40 hrs) than their S counterparts (10-20hrs). Another differentiating feature was the occurrence of multiphasic growth in three of the R strains (including R control). Although overall, R strains were comparatively slow growers, several showed periods of faster growth compared to their S variants. One possibility might be the metabolic burden on the S morphotype imposed by production and transport of GPL. It was also interesting to note that 16 and 23S rRNA were expressed at significantly lower levels in the R compared to the S strain (data not shown).

As expected, GPL production was clearly observed in S, but not in R variants. Cytoplasmic lipid content LB analysis was done. However, no significant differences were observed between S and R variants. These data clarify the point that surface lipids are one of the major differences between the two variants. The effect of cell surface lipids on CSH was also studied and expected results were observed, with R strains being more hydrophobic.

TPP is a surface expressed lipid of *Mab*, specifically prominent in R strains and playing a role in pathogenesis and in cord formation (Llorens-Fons *et al.* 2017). TLC lipid profile analysis showed that both morphotypes produced a component judged to be TPP and,

similar to Burbaud (2016), analysis demonstrated that R produced relatively stronger TPP signals.

The role of extracellular lipids in mycobacterial biofilm formation has been demonstrated, however, recent genetic investigations revealed that *Mtb* biofilm formation also requires a number of additional factors that are not linked to biosynthesis or transport of cell wall lipids (Pacheco *et al.* 2013). In this study distinctive biofilm phenotypes were observed by S and R. Being more cell surface hydrophobic, R developed biofilm quicker, but this was less mature then S variants. It was assumed that the hydrophobic nature might support more adhesion and interactions between the cells and this resulted in quick biofilm formation by R variants. Perhaps the rapid phase observed in the R growth curves may be relevant in this context and could be subject further investigation.

This study also investigated the effect of antimicrobials on the mutation transition of S to R. It would appear that mutation/transition from S to R was not detected using current parameters and the spontaneous mutation frequency was below the level of detection applied here. Perhaps a combination of different stresses along with antimicrobials might enhance transition of S to R.

Understanding the physiological states of *Mab* relating to transmission has been the topic of this research. Based upon the hypothesis of Minnikin and colleagues, that CSH plays a significant role in propensity of cells for aerosolisation (Minnikin et al. 2015), the present study was directed to investigate the role of GPL/CSH in the aerosolisation *Mab*. To do so, factors affecting outcomes from the MATH technique to measure CSH were investigated. In addition, comparisons were made between the R and S control strains at whole culture and single cell levels.

The Ultrasonic NE-U780 Omron nebuliser was used to conduct the experiment for determination of the relationship between CSH and survival in Goldberg drum experiments. The principle technique to measure CSH here was MATH and it should be noted that, in contrast to other studies, a highly standardised protocol has been

established in this lab enabling comparison of outcomes across studies. Both control and clinical R morphotypes were consistently found to be more hydrophobic than S across all experiments.

It is worth noting that when mixed populations of the hydrophilic S and the hydrophobic R variants were nebulised together and CFU counts were assessed for pre and post nebulisation, no difference was observed and this did not support the first hypothesis of this study. However, the possibility that loss of GPLs, representing up to 85 % of the surface-exposed lipids in the *Mab* S strain (Catherinot *et al*. 2007), might occur during nebulisation and make the phenotype R, was considered. Further investigations into the efficiency of aerosolised hydrophobic cells are required to provide solid evidence to accept or reject the above hypothesis. One preliminary experiment was done to test the hypothesis that mixing S and R morphotypes could affect the aerosol result. In this experiment, *Mab ffluc* S and R were used to measure CSH using the hexadecane method following mixing similar to the nebulisation. The result showed an increase in S strain CSH to the same level as the R strain, providing a possible explanation of the lack of difference seen in the aerosolisation study. Time constrains prevented further investigation of this phenomenon, but it would clearly be desirable to separately establish aerosol survival patterns of S and R strains in further Goldberg drum experiments.

In addition, investigation into the survival of *Mab* under different environmental conditions could provide better insight into stability during transmission. The desiccation assay tested survival of cells against this, and other stresses including osmotic stress, nutrient starvation, cold shock, and oxidative stress. After preliminary loss of viability measured in the luciferase assay, it seems likely that the remaining viable cells were sustained by nutrients from neighbouring non-viable cells. This is a limitation of the desiccation assay in modelling aerosol survival. It was hypothesized that the R variant would have greater survival during desiccation, but no difference was found between the variants. When the survival was compared during UV tolerance assay, the results showed a significant difference between survivals of both variants with increasing doses of UV. The R variant was significantly more susceptible to the increasing

doses of UV than the S, and this might be due to greater penetration of UV into R, lacking GPL in the cell wall. Neither the desiccation, nor the UV experiments, provide support for the primary hypothesis that the R morphotype is better adapted to transmission.

WGS analysis in this study showed the variability in nucleotide sequence polymorphisms between S and R strains, that are also supported by previous study by Pawlik (Pawlik *et al.* 2013). PCR and amplified fragment sequencing appeared an inefficient approach for the determination of SNPs in the genome. The approach depends on the expectation that polymorphisms will affect a relatively small set of genomic targets. PCR based investigation might be sufficient to detect the large sequence polymorphisms such as the S to R transition detected here in the R-SDS variant, whereas WGS analysis revealed deletion of a ~7kb region including genes responsible for GPL metabolism.

Traditionally, qRT-PCR has been the method of choice for validation of results in high-throughput gene expression profiling (Wang et al., 2009c). Therefore, the effect of SNPs on gene expression was also investigated using qRT-PCR applied to four targeted genes namely, *gap*, *mps2*, *mps1* in the GPL operon and *mmpl4b*. The SNPs detected by WGS were not found to affect expression of these genes in contrast, to the Pawlik study. It is suggested that the transcript of these genes might encode non-functional proteins and therefore R variants lack ability to produce, or transport GPL.

RNAseq has proven to be a powerful tool for bacterial transcriptome assessment. However, one of main concerns in analysis of RNAseq results can be the robustness of data, due to the complexities and multi-step processing inherent to the technique. The genes were selected from samples that should have a reasonably abundant expression levels from RNAseq data. PCR targeted genes were not differentially expressed in the S and R control and two patient strains; comparison of RNAseq data and qRT-PCR results confirmed this finding in the control strain. However, RNAseq did reveal lack of transcription in another region (from MAB_2254c to MAB_2259) in the R control strain. The genes in that location were found to encode proteins associated with lipid and non-ribosomal peptide production (Table 5.19) and could be another locus contributing to transition of S to R.

Thus overall, the studies presented here, have further elucidated differences between S and R variants of *M. abscessus.* While differences in hydrophobicity have been consistently demonstrated, their linkage to enhanced transmission of R morphotypes has not been supported (Hypothesis 1), neither has the suggestion that they are more stress resistant (Hypothesis 2). The possibility that the two morphotypes might combine in transmission, however, been raised and will require further investigation.

6.2 Final Conclusions

1. Clear differences between S and R growth curves were found. The most apparent difference is the different lag phases but, there also appears to be faster and slower periods of growth as measured by optical density of the R strain.
2. R strains were consistently found to be more hydrophobic than their S counterparts.
3. Analysis of lipid bodies showed no significant difference between the control R and S strains.
4. Biofilms with different structure were produced by the control R strain compared to the S strain.
5. Analysis of R strain lipid profiles showed clear loss of in GPL with possible enhancement in TPP levels.
6. Clarithromycin and ciprofloxacin sub-inhibitory levels did not yield S variants from R, or vice versa, while a single R variant was isolated on 0.1 % (w/v) SDS.
7. In survival of the aerosol phase, there was an apparent rapid decline between T=0 and T=15 min and relatively stable levels of thereafter. On consideration of the counts done in Leicester, there is a ~50% survival over 2h compared with the theoretical maximum (as a result of dilution through sampling).
8. Both R and S *Mab ffluc* phenotypes show no significant differences in desiccation survival.
9. *Mab ffluc* R phenotype has lower UV tolerance compared with the S phenotype.
10. Standard PCR and sequencing of amplified fragments are insufficient for SNP detection.
11. None of the GPL locus polymorphisms reported by Pawlik *et al.* (2013) were detected in the R control strain by WGS. In contrast, several of these were detected in the clinical R strains and a C insertion in *mps1* causing a frameshift was found in the R control strain.
12. Transcriptional arrest of *mps1-mps2-gap* operon was not detected by RNAseq analysis and qRT-PCR.

13. Either the lack of transcription in the MAB_2254c *to* MAB_*2258* region, or the C insertion in *mps1*, provide potential explanations for the R phenotype of the control ATCC strain.

References

Adékambi, T., Berger, P., Raoult, D. and Drancourt, M. (2006) 'rpoB gene sequence-based characterization of emerging non-tuberculous mycobacteria with descriptions of Mycobacterium bolletii sp. nov., *Mycobacterium phocaicum* sp. nov. and *Mycobacterium aubagnense* sp. nov', *International Journal of Systematic and Evolutionary Microbiology*, 56(1), 133-143.

Adékambi, T., Reynaud-Gaubert, M., Greub, G., Gevaudan, M.-J., La Scola, B., Raoult, D. and Drancourt, M. (2004) 'Amoebal coculture of "*Mycobacterium massiliense*" sp. nov. from the sputum of a patient with hemoptoic pneumonia', *Journal of Clinical Microbiology*, 42(12), 5493-5501.

Agogué, H., Joux, F., Obernosterer, I. and Lebaron, P. (2005) 'Resistance of marine bacterioneuston to solar radiation', *Appl. Environ. Microbiol.*, 71(9), 5282-5289.

Aitken, M.L., Limaye, A., Pottinger, P., Whimbey, E., Goss, C.H., Tonelli, M.R., Cangelosi, G.A., Dirac, M.A., Olivier, K.N. and Brown-Elliott, B.A. (2012) 'Respiratory outbreak of *Mycobacterium abscessus* subspecies *massiliense* in a lung transplant and cystic fibrosis center', *American journal of respiratory and critical care medicine*, 185(2), 231-232.

Alateah, S., Peters, C., Dhasmana, D., Pettigrew, K., Fallon, R., Seagar, A., Sloan, D., Laurenson, I., Holden, M. and Gillespie, S. (2017) 'S92 Genomic investigation unmasks evidence of transmission across *Mycobacterium abscessus* cystic fibrosis patients', *Thorax*, 72, A56.

Alfa, M.J., Sisler, J.J. and Harding, G.K. (1995) '*Mycobacterium abscessus* infection of a Norplant contraceptive implant site', *CMAJ: Canadian Medical Association Journal*, 153(9), 1293.

Allen, J.L. (1992) 'A modified Ziehl-Neelsen stain for mycobacteria', *Medical laboratory sciences*, 49(2), 99.

Altschul, S.F., Thomas L. Madden, Alejandro A. Schäffer, Jinghui Zhang, Zheng Zhang, Miller, W. and Lipman, D.J. (1997) 'Gapped BLAST and PSI-BLAST: a new generation of protein database search programs', *Nucleic Acids Res*, 25(17), 3389-3402.

Amaral, M.D. (2005) 'Processing of CFTR: traversing the cellular maze—how much CFTR needs to go through to avoid cystic fibrosis?', *Pediatric pulmonology*, 39(6), 479-491.

Ambalam, P., Kondepudi, K.K., Nilsson, I., Wadström, T. and Ljungh, Å. (2012) 'Bile stimulates cell surface hydrophobicity, Congo red binding and biofilm formation of Lactobacillus strains', *FEMS Microbiology Letters*, 333(1), 10-19.

Andreu, N., Zelmer, A., Sampson, S.L., Ikeh, M., Bancroft, G.J., Schaible, U.E., Wiles, S. and Robertson, B.D. (2013) 'Rapid in vivo assessment of drug efficacy against *Mycobacterium tuberculosis* using an improved firefly luciferase', *Journal of Antimicrobial Chemotherapy*, 68(9), 2118-2127.

Arbeit, R.D., Slutsky, A., Barber, T.W., Maslow, J.N., Niemczyk, S., Falkinham III, J.O., O'Connor, G.T. and von Reyn, C.F. (1993) 'Genetic diversity among strains of *Mycobacterium avium* causing monoclonal and polyclonal bacteremia in patients with AIDS', *Journal of Infectious Diseases*, 167(6), 1384-1390.

Archuleta, R., Mullens, P. and Primm, T.P. (2002) 'The relationship of temperature to desiccation and starvation tolerance of the *Mycobacterium avium complex*', *Archives of microbiology*, 178(4), 311-314.

Asselineau, C. and Montrozier, H. (1976) 'Study of the biosynthesis of phleic acids, polyunsaturated fatty acids synthesised by *Mycobacterium phlei* (author's transl)', *European Journal of Biochemistry*, 63(2), 509.

Asselineau, C.P., Montrozier, H.L., PromÉ, J.C., Savagnac, A.M. and Welby, M. (1972) 'Étude d'un glycolipide polyinsaturé synthétisé par *Mycobacterium phlei*', *European Journal of Biochemistry*, 28(1), 102-109.

Bacon, J., James, B.W., Wernisch, L., Williams, A., Morley, K.A., Hatch, G.J., Mangan, J.A., Hinds, J., Stoker, N.G. and Butcher, P.D. (2004) 'The influence of reduced oxygen availability on pathogenicity and gene expression in *Mycobacterium tuberculosis*', *Tuberculosis*, 84(3-4), 205-217.

Bange, F.-C., Brown, B.A., Smaczny, C., Wallace, R.J. and Böttger, E.C. (2001) 'Lack of transmission of *Mycobacterium abscessus* among patients with cystic fibrosis attending a single clinic', *Clinical Infectious Diseases*, 32(11), 1648-1650.

Bankevich, A., Nurk, S., Antipov, D., Gurevich, A.A., Dvorkin, M., Kulikov, A.S., Lesin, V.M., Nikolenko, S.I., Pham, S. and Prjibelski, A.D. (2012) 'SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing', *Journal of Computational Biology*, 19(5), 455-477.

Baron, E. (1999) 'Bacteriology. Murray PR; Baron EJ; Pfaller MA; Tenover FC; Yolken R H. Manual of Clinical Microbiology'.

Barrow, W.W., Davis, T.L., Wright, E.L., Labrousse, V., Bachelet, M. and Rastogi, N. (1995) 'Immunomodulatory spectrum of lipids associated with *Mycobacterium avium* serovar 8', *Infection and Immunity*, 63(1), 126-133.

Barrow, W.W., De Sousa, J., Davis, T.L., Wright, E.L., Bachelet, M. and Rastogi, N. (1993) 'Immunomodulation of human peripheral blood mononuclear cell functions by defined lipid fractions of *Mycobacterium avium*', *Infection and Immunity*, 61(12), 5286-5293.

Bechara, C., Macheras, E., Heym, B. and Auffret, N. (2010) '*Mycobacterium abscessus* skin infection after tattooing: first case report and review of the literature', *Dermatology*, 221(1), 1-4.

Behr, M.A. and Falkinham Iii, J.O. (2009) 'Molecular epidemiology of nontuberculous mycobacteria', *Future microbiology*, 4(8), 1009-1020.

Belisle, J.T., Pascopella, L., Inamine, J.M., Brennan, P.J. and Jacobs, W. (1991) 'Isolation and expression of a gene cluster responsible for biosynthesis of the glycopeptidolipid antigens of *Mycobacterium avium*', *Journal of Bacteriology*, 173(21), 6991-6997.

Belisle, J.T. and Sonnenberg, M.G. (1998) 'Isolation of genomic DNA from mycobacteria' in *Mycobacteria protocols* Springer, 31-44.

BELL, A. (2013) 'Towards understanding lipid body formation in mycobacteria in Tuberculosis', *PhD, University of Leicester*.

Berleman, J.E. and Bauer, C.E. (2004) 'Characterization of cyst cell formation in the purple photosynthetic bacterium *Rhodospirillum centenum*', *Microbiology*, 150(2), 383-390.

Bermudez, L.E. and Goodman, J. (1996) '*Mycobacterium tuberculosis* invades and replicates within type II alveolar cells', *Infection and Immunity*, 64(4), 1400-1406.

Bernut, A., Dupont, C., Ogryzko, N.V., Neyret, A., Herrmann, J.-L., Floto, R.A., Renshaw, S.A. and Kremer, L. (2019) 'CFTR Protects against *Mycobacterium abscessus* Infection by Fine-Tuning Host Oxidative Defenses', *Cell reports*, 26(7), 1828-1840. e4.

Bernut, A., Herrmann, J.-L., Ordway, D. and Kremer, L. (2017) 'The diverse cellular and animal models to decipher the physiopathological traits of *Mycobacterium abscessus* infection', *Frontiers in cellular and infection microbiology*, 7, 100.

Besra, G.S., McNeil, M.R., Rivoire, B., Khoo, K.H., Morris, H.R., Dell, A. and Brennan, P.J. (1993) 'Further structural definition of a new family of glycopeptidolipids from *Mycobacterium xenopi*', *Biochemistry*, 32(1), 347-355.

Bhatnagar, S. and Schorey, J.S. (2006) 'Elevated mitogen-activated protein kinase signalling and increased macrophage activation in cells infected with a glycopeptidolipid-deficient *Mycobacterium avium*', *Cellular microbiology*, 8(1), 85-96.

Bikandi, J., San Millan, R., Rementeria, A. and Garaizar, J. (2004) 'In silico analysis of complete bacterial genomes: PCR, AFLP-PCR and endonuclease restriction', *Bioinformatics*,20(5),798-9,available: http://dx.doi.org/10.1093/bioinformatics/btg491.

Billi, D. and Potts, M. (2002) 'Life and death of dried prokaryotes', *Research in microbiology*, 153(1), 7-12.

Billman-Jacobe, H., McConville, M.J., Haites, R.E., Kovacevic, S. and Coppel, R.L. (1999) 'Identification of a peptide synthetase involved in the biosynthesis of glycopeptidolipids of *Mycobacterium smegmatis*', *Molecular microbiology*, 33(6), 1244-1253.

Binjomah, A.Z.A. (2014) Studies on the Phenotypes of *Mycobacterium tuberculosis* in Sputum, PhD thesis, University of Leicester.

Blackwood, K.S., He, C., Gunton, J., Turenne, C.Y., Wolfe, J. and Kabani, A.M. (2000) 'Evaluation of *recA* sequences for identification of Mycobacterium species', *Journal of Clinical Microbiology*, 38(8), 2846-2852.

Bloch, H., Sorkin, E. and Erlenmeyer, H. (1953) 'A Toxic Lipid Component of the Tubercle Bacillus ("Cord Factor") I. Isolation from Petroleum Ether Extracts of Young Bacterial Cultures', *American review of tuberculosis*, 67(5), 629-643.

Blossom, D., Alelis, K., Chang, D., Flores, A., Gill, J., Beall, D., Peterson, A., Jensen, B., Noble-Wang, J. and Williams, M. (2008) 'Pseudo-outbreak of *Mycobacterium abscessus* infection caused by laboratory contamination', *Infection Control & Hospital Epidemiology*, 29(01), 57-62.

Bolger, A.M., Lohse, M. and Usadel, B. (2014) 'Trimmomatic: a flexible trimmer for Illumina sequence data', *Bioinformatics*, 30(15), 2114-2120.

Brennan, P., Aspinall, G. and Shin, J. (1981) 'Structure of the specific oligosaccharides from the glycopeptidolipid antigens of serovars in the *Mycobacterium avium-Mycobacterium intracellulare-Mycobacterium scrofulaceum complex*', *Journal of Biological Chemistry*, 256(13), 6817-6822.

Brennan, P. and Goren, M. (1979) 'Structural studies on the type-specific antigens and lipids of the mycobacterium avium. *Mycobacterium intracellulare. Mycobacterium scrofulaceum serocomplex. Mycobacterium intracellulare* serotype 9', *Journal of Biological Chemistry*, 254(10), 4205-4211.

Brennan, P.J. (2003) 'Structure, function, and biogenesis of the cell wall of *Mycobacterium tuberculosis*', *Tuberculosis*, 83(1), 91-97.

Brennan, P.J. and Nikaido, H. (1995) 'The envelope of mycobacteria', *AnnualRreview of Biochemistry*, 64(1), 29-63.

Breslow, R. (1991) 'Hydrophobic effects on simple organic reactions in water', *Accounts of Chemical Research*, 24(6), 159-164.

Brimacombe, M., Hazbon, M., Motiwala, A. and Alland, D. (2007) 'Antibiotic resistance and single-nucleotide polymorphism cluster grouping type in a multinational sample of resistant *Mycobacterium tuberculosis* isolates', *Antimicrobial Agents and Chemotherapy*, 51(11), 4157-4159.

Brown-Elliott, B.A., Nash, K.A. and Wallace, R.J. (2012) 'Antimicrobial susceptibility testing, drug resistance mechanisms, and therapy of infections with nontuberculous mycobacteria', *Clinical Microbiology Reviews*, 25(3), 545-582.

Brown-Elliott, B.A. and Wallace, R.J. (2002) 'Clinical and taxonomic status of pathogenic nonpigmented or late-pigmenting rapidly growing mycobacteria', *Clinical microbiology reviews*, 15(4), 716-746.

Brown, T.A. (2016) *Gene cloning and DNA analysis: an introduction*, John Wiley & Sons.

Bryant, J.M., Grogono, D.M., Greaves, D., Foweraker, J., Roddick, I., Inns, T., Reacher, M., Haworth, C.S., Curran, M.D. and Harris, S.R. (2013) 'Whole-genome sequencing to identify transmission of *Mycobacterium abscessus* between patients with cystic fibrosis: a retrospective cohort study', *The Lancet*, 381(9877), 1551-1560.

Bryant, J.M., Grogono, D.M., Rodriguez-Rincon, D., Everall, I., Brown, K.P., Moreno, P., Verma, D., Hill, E., Drijkoningen, J. and Gilligan, P. (2016) 'Emergence and spread of a human-transmissible multidrug-resistant nontuberculous mycobacterium', *Science*, 354(6313), 751-757.

Burbaud, S., Laval, F., Lemassu, A., Daffé, M., Guilhot, C. and Chalut, C. (2016) 'Trehalose polyphleates are produced by a glycolipid biosynthetic pathway conserved across phylogenetically distant mycobacteria', *Cell chemical biology*, 23(2), 278-289.

Burdon, K.L. (1946) 'Fatty material in bacteria and fungi revealed by staining dried, fixed slide preparations', *Journal of Bacteriology*, 52(6), 665.

Byrd, T.F. and Lyons, C.R. (1999) 'Preliminary Characterization of a *Mycobacterium abscessus* Mutant in Human and Murine Models of Infection', *Infection and immunity*, 67(9), 4700-4707.

Cangelosi, G.A., Palermo, C.O., Laurent, J.-P., Hamlin, A.M. and Brabant, W.H. (1999) 'Colony morphotypes on Congo red agar segregate along species and drug susceptibility lines in the *Mycobacterium avium-intracellulare complex*', *Microbiology*, 145(6), 1317-1324.

Carson, L.A., Cusick, L.B., Bland, L. and Favero, M.S. (1988) 'Efficacy of chemical dosing methods for isolating nontuberculous mycobacteria from water supplies of dialysis centers', *Applied and Environmental Microbiology*, 54(7), 1756-1760.

Catherinot, E., Clarissou, J., Etienne, G., Ripoll, F., Emile, J.-F., Daffe, M., Perronne, C., Soudais, C., Gaillard, J.-L. and Rottman, M. (2007) 'Hypervirulence of a rough variant of the *Mycobacterium abscessus* type strain', *Infection and Immunity*, 75(2), 1055-1058.

Catherinot, E., Roux, A.-L., Macheras, E., Hubert, D., Matmar, M., Dannhoffer, L., Chinet, T., Morand, P., Poyart, C. and Heym, B. (2009) 'Acute respiratory failure involving an R variant of *Mycobacterium abscessus*', *Journal of Clinical Microbiology*, 47(1), 271-274.

Chenal-Francisque, V., Lopez, J., Cantinelli, T., Caro, V., Tran, C., Leclercq, A., Lecuit, M. and Brisse, S. (2011) 'Worldwide distribution of major clones of Listeria monocytogenes', *Emerging infectious diseases*, 17(6), 1110.

Chiaradia, L., Lefebvre, C., Parra, J., Marcoux, J., Burlet-Schiltz, O., Etienne, G., Tropis, M. and Daffé, M. (2017) 'Dissecting the mycobacterial cell envelope and defining the composition of the native mycomembrane', *Scientific Reports*, 7(1), 1-12.

Chu, Y. and Corey, D.R. (2012) 'RNA sequencing: platform selection, experimental design, and data interpretation', *Nucleic Acid Therapeutics*, 22(4), 271-274.

Cingolani, P., Platts, A., Wang, L.L., Coon, M., Nguyen, T., Wang, L., Land, S.J., Lu, X. and Ruden, D.M. (2012) 'A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff: SNPs in the genome of *Drosophila melanogaster* strain w1118; iso-2; iso-3', *Fly*, 6(2), 80-92.

Clary, G., Sasindran, S.J., Nesbitt, N., Mason, L., Cole, S., Azad, A., McCoy, K., Schlesinger, L.S. and Hall-Stoodley, L. (2018) '*Mycobacterium abscessus* smooth and rough morphotypes form antimicrobial-tolerant biofilm phenotypes but are killed by acetic acid', *Antimicrobial Agents and Chemotherapy*, 62(3), e01782-17.

Clifton, I.J., Fletcher, L.A., Beggs, C.B., Denton, M. and Peckham, D.G. (2008) 'A laminar flow model of aerosol survival of epidemic and non-epidemic strains of

Pseudomonas aeruginosa isolated from people with cystic fibrosis', *BMC Microbiology*, 8(1), 105.

Cloud, J., Neal, H., Rosenberry, R., Turenne, C., Jama, M., Hillyard, D. and Carroll, K.C. (2002) 'Identification of Mycobacterium spp. by using a commercial 16S ribosomal DNA sequencing kit and additional sequencing libraries', *Journal of Clinical Microbiology*, 40(2), 400-406.

Cortes, M.A., Nessar, R. and Singh, A.K. (2010) 'Laboratory maintenance of *Mycobacterium abscessus*', *Current Protocols in Microbiology*, 18(1), 10D. 1.1-10D. 1.12.

Costerton, J.W., Geesey, G. and Cheng, K.-J. (1978) 'How bacteria stick', *Scientific American*, 238(1), 86-95.

Costerton, J.W. and Lewandowski, Z. (1997) 'The biofilm lifestyle', *Advances in Dental Research*, 11(1), 192-195.

Costerton, J.W., Lewandowski, Z., Caldwell, D.E., Korber, D.R. and Lappin-Scott, H.M. (1995) 'Microbial biofilms', *Annual Reviews in Microbiology*, 49(1), 711-745.

Cox, C. (1989) 'Airborne bacteria and viruses', *Science Progress (1933-)*, 469-499.

Crick, F. (1970) 'Central dogma of molecular biology', *Nature*, 227(5258), 561.

Cullen, A.R., Cannon, C.L., Mark, E.J. and Colin, A.A. (2000) '*Mycobacterium abscessus* infection in cystic fibrosis: colonization or infection?', *American journal of Respiratory and Critical Care Medicine*, 161(2), 641-645.

Daffé, M., Crick, D.C. and Jackson, M. (2014) 'Genetics of capsular polysaccharides and cell envelope (glyco) lipids', *Microbiology Spectrum*, 2(4).

Daffé, M. and Draper, P. (1997) 'The envelope layers of mycobacteria with reference to their pathogenicity', *Advances in Microbial Physiology*, 39, 131-203.

Danecek, P., Auton, A., Abecasis, G., Albers, C.A., Banks, E., DePristo, M.A., Handsaker, R.E., Lunter, G., Marth, G.T. and Sherry, S.T. (2011) 'The variant call format and VCFtools', *Bioinformatics*, 27(15), 2156-2158.

Daniel, J., Deb, C., Dubey, V.S., Sirakova, T.D., Abomoelak, B., Morbidoni, H.R. and 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', *Journal of Bacteriology*, 186(15), 5017-5030.

Davidson, L.B., Nessar, R., Kempaiah, P., Perkins, D.J. and Byrd, T.F. (2011) 'Mycobacterium abscessus glycopeptidolipid prevents respiratory epithelial TLR2 signaling as measured by HβD2 gene expression and IL-8 release', *PloS One*, 6(12), e29148.

Davidson, R.M., Hasan, N.A., Reynolds, P.R., Totten, S., Garcia, B., Levin, A., Ramamoorthy, P., Heifets, L., Daley, C.L. and Strong, M. (2014) 'Genome sequencing of *Mycobacterium abscessus* isolates from patients in the United States and comparisons to globally diverse clinical strains', *Journal of Clinical Microbiology*, 52(10), 3573-3582.

De Groote, M.A., Pace, N.R., Fulton, K. and Falkinham, J.O. (2006) 'Relationships between Mycobacterium isolates from patients with pulmonary mycobacterial infection and potting soils', *Applied and Environmental Microbiology*, 72(12), 7602-7606.

Deb, C., Lee, C.-M., Dubey, V.S., Daniel, J., Abomoelak, B., Sirakova, T.D., Pawar, S., Rogers, L. and 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(6), e6077.

Dhouib, R., Ducret, A., Hubert, P., Carrière, F., Dukan, S. and Canaan, S. (2011) 'Watching intracellular lipolysis in mycobacteria using time lapse fluorescence microscopy', *Biochimica et Biophysica Acta (BBA)-Molecular and Cell Biology of Lipids*, 1811(4), 234-241.

Donlan, R.M. (2002) 'Biofilms: microbial life on surfaces', *Emerg Infect Dis*, 8(9).

Doyle, R.J. (2000) 'Contribution of the hydrophobic effect to microbial infection', *Microbes and Infection*, 2(4), 391-400.

Draper, P. (1998) 'The outer parts of the mycobacterial envelope as permeability barriers', *Front Biosci*, 3, D1253-D1261.

Drlica, K. and Zhao, X. (1997) 'DNA gyrase, topoisomerase IV, and the 4-quinolones', *Microbiol. Mol. Biol. Rev.*, 61(3), 377-392.

Edsall, J.T. (1992) 'Memories of early days in protein science, 1926-1940', *Protein Science: a publication of the Protein Society*, 1(11), 1526.

Ehrt, S., Guo, X.V., Hickey, C.M., Ryou, M., Monteleone, M., Riley, L.W. and Schnappinger, D. (2005) 'Controlling gene expression in mycobacteria with anhydrotetracycline and Tet repressor', *Nucleic Acids Research*, 33(2), e21-e21.

Embil, J., Warren, P., Yakrus, M., Stark, R., Corne, S., Forrest, D. and Hershfield, E. (1997) 'Pulmonary illness associated with exposure to *Mycobacterium-avium complex*

in hot tub water: hypersensitivity pneumonitis or infection?', *Chest*, 111(3), 813-816.

Enright, M., Day, N., Davies, C., Peacock, S. and Spratt, B. '6 2000. Multilocus Sequence Typing for Characterization of Methicillin-Resistant 7 and Methicillin-Susceptible Clones of *Staphylococcus aureus*', *J. Clin. Microbiol*, 8(38), 1008-1015.

Enright, M.C., Spratt, B.G., Kalia, A., Cross, J.H. and Bessen, D.E. (2001) 'Multilocus sequence typing of *Streptococcus pyogenes* and the relationships between emm type and clone', *Infection and Immunity*, 69(4), 2416-2427.

Esther, C.R., Esserman, D.A., Gilligan, P., Kerr, A. and Noone, P.G. (2010) 'Chronic *Mycobacterium abscessus* infection and lung function decline in cystic fibrosis', *Journal of Cystic Fibrosis*, 9(2), 117-123.

Esther Jr, C.R., Esserman, D.A., Gilligan, P., Kerr, A. and Noone, P.G. (2010) 'Chronic *Mycobacterium abscessus* infection and lung function decline in cystic fibrosis', *Journal of Cystic Fibrosis*, 9(2), 117-123.

Etienne, G., Malaga, W., Laval, F., Lemassu, A., Guilhot, C. and Daffé, M. (2009) 'Identification of the polyketide synthase involved in the biosynthesis of the surface-exposed lipooligosaccharides in mycobacteria', *Journal of Bacteriology*, 191(8), 2613-2621.

Etienne, G., Villeneuve, C., Billman-Jacobe, H., Astarie-Dequeker, C., Dupont, M.-A. and Daffe, M. (2002) 'The impact of the absence of glycopeptidolipids on the ultrastructure, cell surface and cell wall properties, and phagocytosis of *Mycobacterium smegmatis*', *Microbiology*, 148(10), 3089-3100.

Falkinham 3rd, J. (1996) 'Epidemiology of infection by nontuberculous mycobacteria', *Clinical Microbiology Reviews*, 9(2), 177.

Falkinham, J.O. (2003) 'Mycobacterial aerosols and respiratory disease', *Emerging InfectiousDdiseases*, 9(7), 763-767.

Falkinham, J.O., Norton, C.D. and LeChevallier, M.W. (2001) 'Factors influencing numbers of *Mycobacterium avium, Mycobacterium intracellulare*, and other mycobacteria in drinking water distribution systems', *Appl. Environ. Microbiol.*, 67(3), 1225-1231.

Farrell Jr, R.E. (2006) 'RNA methodologies', *Reviews in Cell Biology and Molecular Medicine*.

Farrell, R. (2010) 'RNA Methodologies: A Laboratory Guide for Isolation and Characterization. 2010'.

Feazel, L.M., Baumgartner, L.K., Peterson, K.L., Frank, D.N., Harris, J.K. and Pace, N.R. (2009) 'Opportunistic pathogens enriched in showerhead biofilms', *Proceedings of the National Academy of Sciences*, 106(38), 16393-16399.

Ferry, R.M., Brown, W.F. and Damon, E.B. (1958) 'Studies of the loss of viability of stored bacterial aerosols*. II. Death rates of several non-pathogenic organisms in relation to biological and structural characteristics', *Epidemiology & Infection*, 56(1), 125-150.

Filiatrault, M.J., Stodghill, P.V., Myers, C.R., Bronstein, P.A., Butcher, B.G., Lam, H., Grills, G., Schweitzer, P., Wang, W. and Schneider, D.J. (2011) 'Genome-wide identification of transcriptional start sites in the plant pathogen Pseudomonas syringae pv. tomato str. DC3000', *PLoS One*, 6(12), e29335.

Fisher, E.J. and Gloster, H.M. (2005) 'Infection with *Mycobacterium abscessus* after Mohs micrographic surgery in an immunocompetent patient', *Dermatologic Surgery*, 31(7), 790-790.

Fleige, S., Walf, V., Huch, S., Prgomet, C., Sehm, J. and Pfaffl, M.W. (2006) 'Comparison of relative mRNA quantification models and the impact of RNA integrity in quantitative real-time RT-PCR', *Biotechnology Letters*, 28(19), 1601-1613.

Floto, R.A., Olivier, K.N., Saiman, L., Daley, C.L., Herrmann, J.-L., Nick, J.A., Noone, P.G., Bilton, D., Corris, P. and Gibson, R.L. (2016) 'US Cystic Fibrosis Foundation and European Cystic Fibrosis Society consensus recommendations for the management of non-tuberculous mycobacteria in individuals with cystic fibrosis', *Thorax*, 71(Suppl 1), i1-i22.

Fox, L.P., Geyer, A.S., Husain, S., Della-Latta, P. and Grossman, M.E. (2004) '*Mycobacterium abscessus* cellulitis and multifocal abscesses of the breasts in a transsexual from illicit intramammary injections of silicone', *Journal of the American Academy of Dermatology*, 50(3), 450-454.

Freeman, R., Geier, H., Weigel, K.M., Do, J., Ford, T.E. and Cangelosi, G.A. (2006) 'Roles for cell wall glycopeptidolipid in surface adherence and planktonic dispersal of *Mycobacterium avium*', *Applied and Environmental Microbiology*, 72(12), 7554-7558.

Galil, K., Miller, L.A., Yakrus, M.A., Wallace Jr, R.J., Mosley, D.G., England, B., Huitt, G., McNeil, M.M. and Perkins, B.A. (1999) 'Abscesses due toM*mycobacterium abscessus* linked to injection of unapproved alternative medication', *Emerging InfectiousDdiseases*, 5(5), 681.

Garton, N.J., Christensen, H., Minnikin, D.E., Adegbola, R.A. and Barer, M.R. (2002) 'Intracellular lipophilic inclusions of mycobacteria in vitro and in sputum', *Microbiology*, 148(10), 2951-2958.

Garton, N.J., Waddell, S.J., Sherratt, A.L., Lee, S.-M., Smith, R.J., Senner, C., Hinds, J., Rajakumar, K., Adegbola, R.A. and Besra, G.S. (2008) 'Cytological and transcript analyses reveal fat and lazy persister-like bacilli in tuberculous sputum', *PLoS Medicine*, 5(4), e75.

Goldberg, L., Watkins, H., Boerke, E. and Chatigny, M. (1958) 'The Use of a Rotating Drum for the. Study of Aerosols over Extended Periods of Time', *American Journal of Hygiene*, 68(1), 85-93.

Gonzalez-Santiago, T.M. and Drage, L.A. (2015) 'Nontuberculous Mycobacteria: Skin and Soft Tissue Infections', *Dermatologic Clinics*, 33(3), 563-577.

Grayston, J., Wang, S., Kuo, C. and Campbell, L. (1989) 'Current knowledge onChlamydia pneumoniae, strain TWAR, an important cause of pneumonia and other acute respiratory diseases', *European Journal of Clinical Microbiology and Infectious Diseases*, 8(3), 191-202.

Greenspan, P. and Fowler, S.D. (1985) 'Spectrofluorometric studies of the lipid probe, nile red', *Journal of Lipid Research*, 26(7), 781-789.

Greenspan, P., Mayer, E.P. and Fowler, S.D. (1985) 'Nile red: a selective fluorescent stain for intracellular lipid droplets', *The Journal of Cell Biology*, 100(3), 965-973.

Griffith, D.E. (2003) 'Emergence of nontuberculous mycobacteria as pathogens in cystic fibrosis', *American. Journal of Respiratory and Critical Care Medicine*, 167(6), 810-812.

Griffith, D.E., Aksamit, T., Brown-Elliott, B.A., Catanzaro, A., Daley, C., Gordin, F., Holland, S.M., Horsburgh, R., Huitt, G. and Iademarco, M.F. (2007) 'An official ATS/IDSA statement: diagnosis, treatment, and prevention of nontuberculous mycobacterial diseases', *American Journal of Respiratory and Critical Care Medicine*, 175(4), 367-416.

Griffith, D.E., Girard, W.M. and Wallace Jr, R.J. (1993) 'Clinical features of pulmonary disease caused by rapidly growing mycobacteria', *Am Rev Respir Dis*, 147, 1271-1278.

Gurevich, A., Saveliev, V., Vyahhi, N. and Tesler, G. (2013) 'QUAST: quality assessment tool for genome assemblies', *Bioinformatics*, 29(8), 1072-1075.

Gutiérrez, A.V., Viljoen, A., Ghigo, E., Herrmann, J.-L. and Kremer, L. (2018) 'Glycopeptidolipids, a double-edged sword of the *Mycobacterium abscessus* complex', *Frontiers in Microbiology*, 9.

Hall-Stoodley, L., Brun, O.S., Polshyna, G. and Barker, L.P. (2006) '*Mycobacterium marinum* biofilm formation reveals cording morphology', *FEMS Microbiology Letters*, 257(1), 43-49.

Hall-Stoodley, L. and Lappin-Scott, H. (1998) 'Biofilm formation by the rapidly growing mycobacterial species *Mycobacterium fortuitum*', *FEMS Microbiology Letters*, 168(1), 77-84.

Halloum, I., Carrère-Kremer, S., Blaise, M., Viljoen, A., Bernut, A., Le Moigne, V., Vilchèze, C., Guérardel, Y., Lutfalla, G. and Herrmann, J.-L. (2016) 'Deletion of a dehydratase important for intracellular growth and cording renders rough *Mycobacterium abscessus* avirulent', *Proceedings of the National Academy of Sciences*, 113(29), E4228-E4237.

Hammes, F., Berney, M. and Egli, T. (2010) 'Cultivation-independent assessment of bacterial viability' in *High Resolution Microbial Single Cell Analytics* Springer, 123-150.

Harland, C.W., Botyanszki, Z., Rabuka, D., Bertozzi, C.R. and Parthasarathy, R. (2009) 'Synthetic trehalose glycolipids confer desiccation resistance to supported lipid monolayers', *Langmuir*, 25(9), 5193-5198.

Harland, C.W., Rabuka, D., Bertozzi, C.R. and Parthasarathy, R. (2008) 'The *Mycobacterium tuberculosi*s virulence factor trehalose dimycolate imparts desiccation resistance to model mycobacterial membranes', *Biophysical Journal*, 94(12), 4718-4724.

Harris, K.A., Underwood, A., Kenna, D.T., Brooks, A., Kavaliunaite, E., Kapatai, G., Tewolde, R., Aurora, P. and Dixon, G. (2015) 'Whole-genome sequencing and epidemiological analysis do not provide evidence for cross-transmission of *Mycobacterium abscessus* in a cohort of pediatric cystic fibrosis patients', *Clinical Infectious Diseases*, 60(7), 1007-1016.

Hays, H.C., Millner, P.A., Jones, J.K. and Rayner-Brandes, M.H. (2005) 'A novel and convenient self-drying system for bacterial preservation', *Journal of Microbiological Methods*, 63(1), 29-35.

He, S., Wurtzel, O., Singh, K., Froula, J.L., Yilmaz, S., Tringe, S.G., Wang, Z., Chen, F., Lindquist, E.A. and Sorek, R. (2010) 'Validation of two ribosomal RNA removal methods for microbial metatranscriptomics', *Nature Methods*, 7(10), 807.

Heather, J.M. and Chain, B. (2016) 'The sequence of sequencers: The history of sequencing DNA', *Genomics*, 107(1), 1-8.

Hong, X. and Hopfinger, A. (2004) 'Molecular modeling and simulation of *Mycobacterium tuberculosis* cell wall permeability', *Biomacromolecules*, 5(3), 1066-1077.

Horgen, L., Barrow, E.L., Barrow, W.W. and Rastogi, N. (2000) 'Exposure of human peripheral blood mononuclear cells to total lipids and serovar-specific glycopeptidolipids from *Mycobacterium avium serovars* 4 and 8 results in inhibition of TH1-type responses', *Microbial pathogenesis*, 29(1), 9-16.

Hori, K. and Matsumoto, S. (2010) 'Bacterial adhesion: from mechanism to control', *Biochemical Engineering Journal*, 48(3), 424-434.

Howard, S., Rhoades, E., Recht, J., Pang, X., Alsup, A., Kolter, R., Lyons, C. and Byrd, T. (2006) 'Spontaneous reversion of from a smooth to a rough morphotype is associated with reduced expression of glycopeptidolipid and reacquisition of an invasive phenotype', *Microbiology*, 152, 1581-1590.

Howard, S.T. (2013) 'Recent progress towards understanding genetic variation in the *Mycobacterium abscessus* complex', *Tuberculosis*, 93, S15-S20.

Howard, S.T., Rhoades, E., Recht, J., Pang, X., Alsup, A., Kolter, R., Lyons, C.R. and Byrd, T.F. (2006) 'Spontaneous reversion of *Mycobacterium abscessus* from a smooth to a rough morphotype is associated with reduced expression of glycopeptidolipid and reacquisition of an invasive phenotype', *Microbiology*, 152(6), 1581-1590.

Irani, V.R., Lee, S.-H., Eckstein, T.M., Inamine, J.M., Belisle, J.T. and Maslow, J.N. (2004) 'Utilization of a ts-sacB selection system for the generation of a *Mycobacterium avium serovar*-8 specific glycopeptidolipid allelic exchange mutant', *Annals of Clinical Microbiology and Antimicrobials*, 3(1), 18.

Jankute, M., Nataraj, V., Lee, O.Y.-C., Wu, H.H., Ridell, M., Garton, N.J., Barer, M.R., Minnikin, D.E., Bhatt, A. and Besra, G.S. (2017) 'The role of hydrophobicity in tuberculosis evolution and pathogenicity', *Scientific Reports*, 7(1), 1315.

Jenke-Kodama, H., Sandmann, A., Müller, R. and Dittmann, E. (2005) 'Evolutionary implications of bacterial polyketide synthases', *Molecular Biology and Evolution*, 22(10), 2027-2039.

Jeon, K., Kwon, O.J., Lee, N.Y., Kim, B.-J., Kook, Y.-H., Lee, S.-H., Park, Y.K., Kim, C.K. and Koh, W.-J. (2009) 'Antibiotic treatment of M*ycobacterium abscessus* lung disease: a retrospective analysis of 65 patients', *American journal of respiratory and Critical Care medicine*, 180(9), 896-902.

Jönsson, B., Ridell, M. and Wold, A.E. (2013) 'Phagocytosis and cytokine response to rough and smooth colony variants of *Mycobacterium abscessus* by human peripheral blood mononuclear cells', *Apmis*, 121(1), 45-55.

Jönsson, B.E., Gilljam, M., Lindblad, A., Ridell, M., Wold, A.E. and Welinder-Olsson, C. (2007) 'Molecular epidemiology of *Mycobacterium abscessus*, with focus on cystic fibrosis', *Journal of Clinical Microbiology*, 45(5), 1497-1504.

Kahana, L.M., Kay, J.M., Yakrus, M.A. and Waserman, S. (1997) '*Mycobacterium avium complex* infection in an immunocompetent young adult related to hot tub exposure', *Chest*, 111(1), 242-245.

Kauzmann, W. (1959) 'Some factors in the interpretation of protein denaturation' in *Advances in Protein Chemistry* Elsevier, 1-63.

Kham-ngam, I., Chetchotisakd, P., Ananta, P., Chaimanee, P., Reechaipichitkul, W., Lulitanond, V., Namwat, W. and Faksri, K. (2019) 'Differentiation between persistent infection/colonization and re-infection/re-colonization of *Mycobacterium abscessus* isolated from patients in Northeast Thailand', *Infection, Genetics and Evolution*, 68, 35-42.

Khermosh, O., Weintroub, S., Topilsky, M. and Baratz, M. (1979) '*Mycobacterium abscessus* (*M. chelonei*) infection of the knee joint: report of two cases following intra-articular injection of corticosteroids', *Clinical orthopaedics and related research*, 140, 162-168.

Khosla, C., Gokhale, R.S., Jacobsen, J.R. and Cane, D.E. (1999) 'Tolerance and specificity of polyketide synthases', *Annual Review of Biochemistry*, 68(1), 219-253.

Kibbe, W.A. (2007) 'OligoCalc: an online oligonucleotide properties calculator', *Nucleic Acids Res*, 35(Web Server issue), W43-6, available: http://dx.doi.org/10.1093/nar/gkm234.

Kim, T.H. and Kubica, G.P. (1972) 'Long-term preservation and storage of mycobacteria', *Appl. Environ. Microbiol.*, 24(3), 311-317.

Koboldt, D.C., Zhang, Q., Larson, D.E., Shen, D., McLellan, M.D., Lin, L., Miller, C.A., Mardis, E.R., Ding, L. and Wilson, R.K. (2012) 'VarScan 2: somatic mutation and copy number alteration discovery in cancer by exome sequencing', *Genome Research*, 22(3), 568-576.

KODAKA, H., FUKUDA, K., MIZUOCHI, S. and HORIGOME, K. (1996) 'Adenosine triphosphate content of microorganisms related with food spoilage', *Japanese Journal of Food Microbiology*, 13(1), 29-34.

Koh, W.-J., Lee, J.H., Kwon, Y.S., Lee, K.S., Suh, G.Y., Chung, M.P., Kim, H. and Kwon, O.J. (2007) 'Prevalence of gastroesophageal reflux disease in patients with nontuberculous mycobacterial lung disease', *CHEST Journal*, 131(6), 1825-1830.

Krasowska, A. and Sigler, K. (2014) 'How microorganisms use hydrophobicity and what does this mean for human needs?', *Frontiers in Cellular and Infection Microbiology*, 4.

Kreutzfeldt, K.M., McAdam, P.R., Claxton, P., Holmes, A., Seagar, A.L., Laurenson, I.F. and Fitzgerald, J.R. (2013) 'Molecular longitudinal tracking of *Mycobacterium abscessus* spp. during chronic infection of the human lung', *PloS One*, 8(5), e63237.

Krisko, A. and Radman, M. (2010) 'Protein damage and death by radiation in Escherichia coli and Deinococcus radiodurans', *Proceedings of the National Academy of Sciences*, 107(32), 14373-14377.

Krzywinska, E., Bhatnagar, S., Sweet, L., Chatterjee, D. and Schorey, J.S. (2005) 'Mycobacterium avium 104 deleted of the methyltransferase D gene by allelic replacement lacks serotype-specific glycopeptidolipids and shows attenuated virulence in mice', *Molecular Microbiology*, 56(5), 1262-1273.

Kukurba, K.R. and Montgomery, S.B. (2015) 'RNA sequencing and analysis', *Cold Spring Harbor Protocols*, 2015(11), pdb. top084970.

Kumar, S., Kirha, T.J. and Thonger, T. (2014) 'Toxicological effects of sodium dodecyl sulfate', *J Chem Pharm Res*, 6, 1488-1492.

Kuo, Y.-M., Cheng, A., Wu, P.-C., Hsieh, S.-C., Hsieh, S.-M., Hsueh, P.-R. and Yu, C.-L. (2011) 'Disseminated *Mycobacterium abscessus* infection and showerheads, Taiwan', *Emerging infectious diseases*, 17(11), 2077.

Kusunoki, S. and Ezaki, T. (1992) 'Proposal of *Mycobacterium peregrinum* sp. nov., nom. rev., and Elevation of *Mycobacterium chelonae* subsp. abscessus (Kubica et al.) to Species Status: *Mycobacterium abscessus* comb. nov', *International Journal of Systematic and Evolutionary Microbiology*, 42(2), 240-245.

Lai, C.-C., Tan, C., Lin, S., Liu, W., Liao, C., Huang, Y. and Hsueh, P. (2011) 'Clinical significance of nontuberculous mycobacteria isolates in elderly Taiwanese patients', *European Journal of Clinical Microbiology & Infectious Diseases*, 30(6), 779-783.

Lai, K.K., Brown, B.A., Westerling, J.A., Fontecchio, S.A., Zhang, Y. and Wallace, R.J. (1998) 'Long-term laboratory contamination by *Mycobacterium abscessus* resulting in two pseudo-outbreaks: recognition with use of random amplified

polymorphic DNA (RAPD) polymerase chain reaction', *Clinical Infectious Diseases*, 27(1), 169-175.

Langtry, H.D. and Brogden, R.N. (1997) 'Clarithromycin', *Drugs*, 53(6), 973-1004.

Larkin, M.A., Blackshields, G., Brown, N.P., Chenna, R., McGettigan, P.A., McWilliam, H., Valentin, F., Wallace, I.M., Wilm, A., Lopez, R., Thompson, J.D., Gibson, T.J. and Higgins, D.G. (2007) 'Clustal W and Clustal X version 2.0', *Bioinformatics*, 23(21), 2947-8, available: http://dx.doi.org/10.1093/bioinformatics/btm404.

Lebre, P.H., De Maayer, P. and Cowan, D.A. (2017) 'Xerotolerant bacteria: surviving through a dry spell', *Nature Reviews Microbiology*, 15(5), 285-296.

Lever, M., Williams, A. and Bennett, A. (2000) 'Survival of mycobacterial species in aerosols generated from artificial saliva', *Letters in Applied Microbiology*, 31(3), 238-241.

Li, H. (2012) 'Exploring single-sample SNP and INDEL calling with whole-genome de novo assembly', *Bioinformatics*, 28(14), 1838-1844.

Livak, K.J. and Schmittgen, T.D. (2001) 'Analysis of relative gene expression data using real-time quantitative PCR and the 2− ΔΔCT method', *methods*, 25(4), 402-408.

Llorens-Fons, M., Pérez-Trujillo, M., Julián, E., Brambilla, C., Alcaide, F., Byrd, T.F. and Luquin, M. (2017) 'Trehalose Polyphleates, External Cell Wall Lipids in *Mycobacterium abscessus*, Are Associated with the Formation of Clumps with Cording Morphology, Which Have Been Associated with Virulence', *Frontiers in Microbiology*, 8, 1402.

Loudon, R.G., Bumgarner, L.R., Lacy, J. and Coffman, G.K. (1969) 'Aerial Transmission of Mycobacteria 1, 2', *American Review of Respiratory Disease*, 100(2), 165-171.

Louw, G., Warren, R., Van Pittius, N.G., McEvoy, C., Van Helden, P. and Victor, T. (2009) 'A balancing act: efflux/influx in mycobacterial drug resistance', *Antimicrobial Agents and Chemotherapy*, 53(8), 3181-3189.

Macheras, E., Konjek, J., Roux, A.-L., Thiberge, J.-M., Bastian, S., Leão, S.C., Palaci, M., Sivadon-Tardy, V., Gutierrez, C. and Richter, E. (2014) 'Multilocus sequence typing scheme for the *Mycobacterium abscessus complex*', *Research in Microbiology*, 165(2), 82-90.

Macheras, E., Roux, A.-L., Ripoll, F., Sivadon-Tardy, V., Gutierrez, C., Gaillard, J.-L. and Heym, B. (2009) 'Inaccuracy of single-target sequencing for discriminating species of the *Mycobacterium abscessus* group', *Journal of Clinical Microbiology*, 47(8), 2596-2600.

Mangion, D., Arnold, D.R., Cameron, T.S. and Robertson, K.N. (2001) 'The electron transfer photochemistry of allenes with cyanoarenes. Photochemical nucleophile–olefin combination, aromatic substitution (photo-NOCAS) and related reactions', *Journal of the Chemical Society, Perkin Transactions 2*, (1), 48-60.

Mangione, E.J., Huitt, G., Lenaway, D., Beebe, J., Bailey, A., Figoski, M., Rau, M.P., Albrecht, K.D. and Yakrus, M.A. (2001) 'Nontuberculous mycobacterial disease following hot tub exposure', *Emerging infectious diseases*, 7(6), 1039.

Martin, J.A. and Wang, Z. (2011) 'Next-generation transcriptome assembly', *Nature Reviews Genetics*, 12(10), 671.

Martinez, J. and Baquero, F. (2000) 'Mutation frequencies and antibiotic resistance', *Antimicrobial Agents and hemotherapy*, 44(7), 1771-1777.

Mason, P., Neilson, G., Dempsey, C., Barnes, A. and Cruickshank, J. (2003) 'The hydration structure of guanidinium and thiocyanate ions: implications for protein stability in aqueous solution', *Proceedings of the National Academy of Sciences*, 100(8), 4557-4561.

Maurer, F.P., Bruderer, V.L., Ritter, C., Castelberg, C., Bloemberg, G.V. and Böttger, E.C. (2014) 'Lack of antimicrobial bactericidal activity in *Mycobacterium abscessus*', *Antimicrobial Agents and Chemotherapy*, 58(7), 3828-3836.

Maxson, S., Schutze, G.E. and Jacobs, R.F. (1994) '*Mycobacterium abscessus* osteomyelitis: treatment with clarithromycin', *Infectious Diseases in Clinical Practice*, 3(3), 203-206.

Medjahed, H., Gaillard, J.-L. and Reyrat, J.-M. (2010) '*Mycobacterium abscessus*: a new player in the mycobacterial field', *Trends in Microbiology*, 18(3), 117-123.

Medjahed, H. and Reyrat, J.-M. (2009) 'Construction of *Mycobacterium abscessus* defined glycopeptidolipid mutants: comparison of genetic tools', *Applied and Environmental Microbiology*, 75(5), 1331-1338.

Metchock, B., Nolte, F. and Wallace Jr, R.J. (1999) 'Mycobacterium'. p. 399-437. *In* P. R. Murray, E. J. Baron, M. A. Pfaller, F. C. Tenover, and R. H. Yolken (ed.), Manual of clinical microbiology, 7th ed. ASM Press, Washington, D.C.

Mikheenko, A., Prjibelski, A., Saveliev, V., Antipov, D. and Gurevich, A. (2018) 'Versatile genome assembly evaluation with QUAST-LG', *Bioinformatics*, 34(13), i142-i150.

Miles, A.A., Misra, S. and Irwin, J. (1938) 'The estimation of the bactericidal power of the blood', *Epidemiology & Infection*, 38(6), 732-749.

Mills, J.W., Schwiebert, E.M. and Stanton, B.A. (1994) 'The cytoskeleton and membrane transport', *Current Opinion in Nephrology and Hypertension*, 3(5), 529-534.

Minnikin, D., Minnikin, S.M., Goodfellow, M. and Stanford, J. (1982) 'The mycolic acids of *Mycobacterium chelonei*', *Microbiology*, 128(4), 817-822.

Minnikin, D.E., Lee, O.Y., Wu, H.H., Nataraj, V., Donoghue, H.D., Ridell, M., Watanabe, M., Alderwick, L., Bhatt, A. and Besra, G.S. (2015) 'Pathophysiological implications of cell envelope structure in *Mycobacterium tuberculosis* and related taxa', *Tuberculosis*.

Mnaimneh, S., Davierwala, A.P., Haynes, J., Moffat, J., Peng, W.-T., Zhang, W., Yang, X., Pootoolal, J., Chua, G. and Lopez, A. (2004) 'Exploration of essential gene functions via titratable promoter alleles', *Cell*, 118(1), 31-44.

Moore, M. and Frerichs, J.B. (1953) 'An Unusual Acid-Fast Infection of the Knee with Subcutaneous, Abscess-Like Lesions of the Gluteal Region: Report of a Case with a Study of the Organism, *Mycobacterium abscessus*, n. sp. 1', *Journal of Investigative Dermatology*, 20(2), 133-169.

Narasimhan, V., Danecek, P., Scally, A., Xue, Y., Tyler-Smith, C. and Durbin, R. (2016) 'BCFtools/RoH: a hidden Markov model approach for detecting autozygosity from next-generation sequencing data', *Bioinformatics*, 32(11), 1749-1751.

Nataraj, V., Pang, P.-c., Haslam, S.M., Veerapen, N., Minnikin, D.E., Dell, A., Besra, G.S. and Bhatt, A. (2015) 'MKAN27435 is required for the biosynthesis of higher subclasses of lipooligosaccharides in *Mycobacterium kansasii*', *PloS One*, 10(4), e0122804.

Nessar, R., Cambau, E., Reyrat, J.M., Murray, A. and Gicquel, B. (2012) '*Mycobacterium abscessus*: a new antibiotic nightmare', *Journal of Antimicrobial Chemotherapy*, 67(4), 810-818.

Nessar, R., Reyrat, J.-M., Davidson, L.B. and Byrd, T.F. (2011a) 'Deletion of the *mmpL4b* gene in the *Mycobacterium abscessus* glycopeptidolipid biosynthetic pathway results in loss of surface colonization capability, but enhanced ability to replicate in human macrophages and stimulate their innate immune response', *Microbiology*, 157(4), 1187-1195.

Nessar, R., Reyrat, J.M., Murray, A. and Gicquel, B. (2011b) 'Genetic analysis of new 16S rRNA mutations conferring aminoglycoside resistance in *Mycobacterium abscessus*', *Journal of Antimicrobial Chemotherapy*, 66(8), 1719-1724.

Nikaido, H., Kim, S.H. and Rosenberg, E.Y. (1993) 'Physical organization of lipids in the cell wall of *Mycobacterium chelonae*', *Molecular Microbiology*, 8(6), 1025-1030.

Ojha, A., Anand, M., Bhatt, A., Kremer, L., Jacobs Jr, W.R. and Hatfull, G.F. (2005) 'GroEL1: a dedicated chaperone involved in mycolic acid biosynthesis during biofilm formation in mycobacteria', *Cell*, 123(5), 861-873.

Ojha, A.K., Baughn, A.D., Sambandan, D., Hsu, T., Trivelli, X., Guerardel, Y., Alahari, A., Kremer, L., Jacobs, W.R. and Hatfull, G.F. (2008) 'Growth of *Mycobacterium tuberculosis* biofilms containing free mycolic acids and harbouring drug-tolerant bacteria', *Molecular Microbiology*, 69(1), 164-174.

Pacheco, S.A., Hsu, F.-F., Powers, K.M. and Purdy, G.E. (2013) 'MmpL11 protein transports mycolic acid-containing lipids to the mycobacterial cell wall and contributes to biofilm formation in *Mycobacterium smegmatis*', *Journal of Biological Chemistry*, 288(33), 24213-24222.

Palmer, R.J. and White, D.C. (1997) 'Developmental biology of biofilms: implications for treatment and control', *Trends in Microbiology*, 5(11), 435-440.

Parish, T. and Stoker, N.G. (1998) 'Electroporation of mycobacteria' in *Mycobacteria protocols* Springer, 129-144.

Park, I.K., Hsu, A.P., Tettelin, H., Shallom, S.J., Drake, S.K., Ding, L., Wu, U.-I., Adamo, N., Prevots, D.R. and Olivier, K.N. (2015) 'Clonal diversification and changes in lipid traits and colony morphology in *Mycobacterium abscessus* clinical isolates', *Journal of Clinical Microbiology*, 53(11), 3438-3447.

Parker, B.C., Ford, M.A., Gruft, H. and Falkinham III, J.O. (1983) 'Epidemiology of infection by nontuberculous mycobacteria: IV. Preferential aerosolization of *Mycobacterium intracellulare* from natural waters', *American Review of Respiratory Disease*, 128(4), 652-656.

Parkins, M.D. and Floto, R.A. (2015) 'Emerging bacterial pathogens and changing concepts of bacterial pathogenesis in cystic fibrosis', *Journal of Cystic Fibrosis*, 14(3), 293-304.

Patterson, J.H., McConville, M.J., Haites, R.E., Coppel, R.L. and Billman-Jacobe, H. (2000) 'Identification of a methyltransferase from *Mycobacterium smegmatis* involved in glycopeptidolipid synthesis', *Journal of Biological Chemistry*, 275(32), 24900-24906.

Pawlik, A., Garnier, G., Orgeur, M., Tong, P., Lohan, A., Le Chevalier, F., Sapriel, G., Roux, A.L., Conlon, K. and Honoré, N. (2013) 'Identification and characterization of the genetic changes responsible for the characteristic smooth-to-rough morphotype alterations of clinically persistent *Mycobacterium abscessus*', *Molecular Microbiology*, 90(3), 612-629.

Peters, D.H. and Clissold, S.P. (1992) 'Clarithromycin', *Drugs*, 44(1), 117-164.

Peyron, P., Vaubourgeix, J., Poquet, Y., Levillain, F., Botanch, C., Bardou, F., Daffé, M., Emile, J.-F., Marchou, B. and Cardona, P.-J. (2008) 'Foamy macrophages from tuberculous patients' granulomas constitute a nutrient-rich reservoir for M. tuberculosis persistence', *PLoS Pathogens*, 4(11), e1000204.

Pourshafie, M., Ayub, Q. and Barrow, W. (1993) 'Comparative effects of *Mycobacterium avium* glycopeptidolipid and lipopeptide fragment on the function and ultrastructure of mononuclear cells', *Clinical & Experimental Immunology*, 93(1), 72-79.

Qadri, F., Hossain, S.A., Ciznár, I., Haider, K., Ljungh, A., Wadstrom, T. and Sack, D.A. (1988) 'Congo red binding and salt aggregation as indicators of virulence in Shigella species', *Journal of Clinical Microbiology*, 26(7), 1343-1348.

Quan, T.P., Bawa, Z., Foster, D., Walker, T., del Ojo Elias, C., Rathod, P., Iqbal, Z., Bradley, P., Mowbray, J. and Walker, A.S. (2018) 'Evaluation of whole-genome sequencing for mycobacterial species identification and drug susceptibility testing in a clinical setting: a large-scale prospective assessment of performance against line probe assays and phenotyping', *Journal of Clinical Microbiology*, 56(2), e01480-17.

Rajagopal, S., Sudarsan, N. and Nickerson, K.W. (2002) 'Sodium dodecyl sulfate hypersensitivity of clpP and clpB mutants of Escherichia coli', *Appl. Environ. Microbiol.*, 68(8), 4117-4121.

Ramaswamy, S. and Musser, J.M. (1998) 'Molecular genetic basis of antimicrobial agent resistance in *Mycobacterium tuberculosis*: 1998 update', *Tubercle and Lung disease*, 79(1), 3-29.

Ramaswamy, S.V., Reich, R., Dou, S.-J., Jasperse, L., Pan, X., Wanger, A., Quitugua, T. and Graviss, E.A. (2003) 'Single nucleotide polymorphisms in genes associated with isoniazid resistance in *Mycobacterium tuberculosis*', *Antimicrobial Agents and Chemotherapy*, 47(4), 1241-1250.

Rastogi, N. and Barrow, W. (1994) 'Cell envelope constituents and the multifaceted nature of *Mycobacterium avium* pathogenicity and drug resistance', *Research in Microbiology*, 145(3), 243-252.

Recht, J., Martínez, A., Torello, S. and Kolter, R. (2000) 'Genetic Analysis of Sliding Motility in *Mycobacterium smegmatis*', *Journal of Bacteriology*, 182(15), 4348-4351.

Renna, M., Schaffner, C., Brown, K., Shang, S., Tamayo, M.H., Hegyi, K., Grimsey, N.J., Cusens, D., Coulter, S. and Cooper, J. (2011) 'Azithromycin blocks autophagy and

may predispose cystic fibrosis patients to mycobacterial infection', *The Journal of Clinical Investigation*, 121(9), 3554-3563.

Rhoades, E.R., Archambault, A.S., Greendyke, R., Hsu, F.-F., Streeter, C. and Byrd, T.F. (2009) '*Mycobacterium abscessus* glycopeptidolipids mask underlying cell wall phosphatidyl-myo-inositol mannosides blocking induction of human macrophage TNF-α by preventing interaction with TLR2', *The Journal of Immunology*, 183(3), 1997-2007.

Rickman, O.B., Ryu, J.H., Fidler, M.E. and Kalra, S. (2002) 'Hypersensitivity pneumonitis associated with *Mycobacterium avium complex* and hot tub use', in *Mayo Clinic Proceedings*, Elsevier, 1233-1237.

Riley, E., Murphy, G. and Riley, R. (1978) 'Airborne spread of measles in a suburban elementary school', *American Journal of Epidemiology*, 107(5), 421-432.

Riley, R., Mills, C., O'grady, F., Sultan, L., Wittstadt, F. and Shivpuri, D. (1962) 'Infectiousness of air from a tuberculosis ward: ultraviolet irradiation of infected air: comparative infectiousness of different patients', *American Review of Respiratory Disease*, 85(4), 511-525.

Ripoll, F., Deshayes, C., Pasek, S., Laval, F., Beretti, J.-L., Biet, F., Risler, J.-L., Daffé, M., Etienne, G. and Gaillard, J.-L. (2007) 'Genomics of glycopeptidolipid biosynthesis in *Mycobacterium abscessus* and *M. chelonae*', *BMC genomics*, 8(1), 1.

Ripoll, F., Pasek, S., Schenowitz, C., Dossat, C., Barbe, V., Rottman, M., Macheras, E., Heym, B., Herrmann, J.-L. and Daffé, M. (2009) 'Non mycobacterial virulence genes in the genome of the emerging pathogen *Mycobacterium abscessus*', *PloS one*, 4(6), e5660.

Rosenberg, M., Gutnick, D. and Rosenberg, E. (1980) 'Adherence of bacteria to hydrocarbons: a simple method for measuring cell-surface hydrophobicity', *FEMS Microbiol Lett*, 9(1), 29-33.

Rottman, M., Catherinot, E., Hochedez, P., Emile, J.-F., Casanova, J.-L., Gaillard, J.-L. and Soudais, C. (2007) 'Importance of T cells, gamma interferon, and tumor necrosis factor in immune control of the rapid grower *Mycobacterium abscessus* in C57BL/6 mice', *Infection and Immunity*, 75(12), 5898-5907.

Roux, A.-L., Viljoen, A., Bah, A., Simeone, R., Bernut, A., Laencina, L., Deramaudt, T., Rottman, M., Gaillard, J.-L. and Majlessi, L. (2016) 'The distinct fate of smooth and rough *Mycobacterium abscessus* variants inside macrophages', *Open Biology*, 6(11), 160185.

Rutherford, K., Parkhill, J., Crook, J., Horsnell, T., Rice, P., Rajandream, M.-A. and Barrell, B. (2000) 'Artemis: sequence visualization and annotation', *Bioinformatics*, 16(10), 944-945.

Sánchez-Chardi, A., Olivares, F., Byrd, T.F., Julián, E., Brambilla, C. and Luquin, M. (2011) 'Demonstration of cord formation by rough *Mycobacterium abscessus* variants: implications for the clinical microbiology laboratory', *Journal of Clinical Microbiology*, 49(6), 2293-2295.

Sanguinetti, M., Ardito, F., Fiscarelli, E., La Sorda, M., D'Argenio, P., Ricciotti, G. and Fadda, G. (2001) 'Fatal pulmonary infection due to multidrug-resistant *Mycobacterium abscessus* in a patient with cystic fibrosis', *Journal of Clinical Microbiology*, 39(2), 816-819.

Sapriel, G., Konjek, J., Orgeur, M., Bouri, L., Frézal, L., Roux, A.-L., Dumas, E., Brosch, R., Bouchier, C. and Brisse, S. (2016) 'Genome-wide mosaicism within *Mycobacterium abscessus*: evolutionary and epidemiological implications', *BMC Genomics*, 17(1), 118.

Saunders, H.D. (2000) 'Does predicted rebound depend on distinguishing between energy and energy services?', *Energy Policy*, 28(6-7), 497.

Saviola, B. (2010a) 'All stressed out: mycobacterial responses to stress', *Current Research, Technology, and Education Topics in Applied Microbiology and Microbial Biotechnology.* Microbiology Book Series Edition. Mendez-Vilas, A. Editor.

Saviola, B. (2010b) 'All stressed out: mycobacterial responses to stress', *Current Research, Technology, and Education Topics in Applied Microbiology and Microbial Biotechnology.* Microbiology Book Series Edition, ed. Mendez-Vilas A., editor.(Formatex Research Center, 545-549.

Schorey, J.S. and Sweet, L. (2008) 'The mycobacterial glycopeptidolipids: structure, function, and their role in pathogenesis', *Glycobiology*, 18(11), 832-841.

Selan, L., Berlutti, F., Passariello, C., Thaller, M. and Renzini, G. (1992) 'Reliability of a bioluminescence ATP assay for detection of bacteria', *Journal of Clinical Microbiology*, 30(7), 1739-1742.

Sherratt, A.L. (2008) *Lipid bodies in mycobacteria*, PhD thesis, Dept. Infection Immunity and Inflammation, University of Leicester.

Shu, Z., Weigel, K.M., Soelberg, S.D., Lakey, A., Cangelosi, G.A., Lee, K.-H., Chung, J.-H. and Gao, D. (2012) 'Cryopreservation of *Mycobacterium tuberculosis* complex cells', *Journal of Clinical Microbiology*, 50(11), 3575-3580.

Siméone, R., Constant, P., Malaga, W., Guilhot, C., Daffé, M. and Chalut, C. (2007) 'Molecular dissection of the biosynthetic relationship between phthiocerol and phthiodiolone dimycocerosates and their critical role in the virulence and permeability of *Mycobacterium tuberculosis*', *The FEBS Journal*, 274(8), 1957-1969.

Simmon, K.E., Pounder, J.I., Greene, J.N., Walsh, F., Anderson, C.M., Cohen, S. and Petti, C.A. (2007) 'Identification of an emerging pathogen, *Mycobacterium massiliense*, by rpoB sequencing of clinical isolates collected in the United States', *Journal of Clinical Microbiology*, 45(6), 1978-1980.

Siragusa, G., Dorsa, W., Cutter, C.N., Perino, L. and Koohmaraie, M. (1996) 'Use of a newly developed rapid microbial ATP bioluminescence assay to detect microbial contamination on poultry carcasses', *Journal of Bioluminescence and Chemiluminescence*, 11(6), 297-301.

Sirisattha, S., Momose, Y., Kitagawa, E. and Iwahashi, H. (2004) 'Toxicity of anionic detergents determined by *Saccharomyces cerevisiae* microarray analysis', *Water Research*, 38(1), 61-70.

Soini, H. and Viljanen, M.K. (1997) 'Diversity of the 32-kilodalton protein gene may form a basis for species determination of potentially pathogenic mycobacterial species', *Journal of Clinical Microbiology*, 35(3), 769-773.

Sondén, B., Kocíncová, D., Deshayes, C., Euphrasie, D., Rhayat, L., Laval, F., Frehel, C., Daffé, M., Etienne, G. and Reyrat, J.M. (2005) 'Gap, a mycobacterial specific integral membrane protein, is required for glycolipid transport to the cell surface', *Molecular Microbiology*, 58(2), 426-440.

Srey, S., Jahid, I.K. and Ha, S.-D. (2013) 'Biofilm formation in food industries: a food safety concern', *Food Control*, 31(2), 572-585.

Stanley, P.E. (1989) 'A review of bioluminescent ATP techniques in rapid microbiology', *Journal of Bioluminescence and Chemiluminescence*, 4(1), 375-380.

Steingrube, V., Wallace, R., Steele, L. and Pang, Y. (1991) 'Mercuric reductase activity and evidence of broad-spectrum mercury resistance among clinical isolates of rapidly growing mycobacteria', *Antimicrobial Agents and Chemotherapy*, 35(5), 819-823.

Stern, R.C. and Rosenstein, B.J. (2000) *Cystic fFbrosis: Medical Care*, Lippincott Williams & Wilkins.

Stokes, R.W., Norris-Jones, R., Brooks, D.E., Beveridge, T.J., Doxsee, D. and Thorson, L.M. (2004) 'The glycan-rich outer layer of the cell wall of *Mycobacterium*

tuberculosis acts as an antiphagocytic capsule limiting the association of the bacterium with macrophages', *Infection and Immunity*, 72(10), 5676-5686.

Stone, R. and Johnson, D. (2002) 'A note on the effect of nebulization time and pressure on the culturability of *Bacillus subtilis* and *Pseudomonas fluorescens*', *Aerosol Science & Technology*, 36(5), 536-539.

Stoodley, P., Sauer, K., Davies, D. and Costerton, J.W. (2002) 'Biofilms as complex differentiated communities', *Annual Reviews in Microbiology*, 56(1), 187-209.

Sweet, L. and Schorey, J.S. (2006) 'Glycopeptidolipids from *Mycobacterium avium* promote macrophage activation in a TLR2-and MyD88-dependent manner', *Journal of Leukocyte Biology*, 80(2), 415-423.

Tassell, S., Pourshafie, M., Wright, E., Richmond, M. and Barrow, W. (1992) 'Modified lymphocyte response to mitogens induced by the lipopeptide fragment derived from *Mycobacterium avium* serovar-specific glycopeptidolipids', *Infection and Immunity*, 60(2), 706-711.

Teenaged, C.I. (1996) 'Infection with *Mycobacterium abscessus* Associated with Intramuscular Injection of Adrenal Cortex Extract—Colorado and Wyoming, 1995–1996', *Infection*, 45(33).

Thompson, K.-A., Bennett, A. and Walker, J. (2011) 'Aerosol survival of S*taphylococcus epidermidis*', *Journal of Hospital Infection*, 78(3), 216-220.

Thomson, R., Tolson, C., Carter, R., Coulter, C., Huygens, F. and Hargreaves, M. (2013) 'Isolation of nontuberculous mycobacteria (NTM) from household water and shower aerosols in patients with pulmonary disease caused by NTM', *Journal of ClinicalMmicrobiology*, 51(9), 3006-3011.

Thore, A., Ansehn, S., Lundin, A. and Bergman, S. (1975) 'Detection of bacteria by luciferase assay of adenosine triphosphate', *Journal of Clinical Microbiology*, 1(1), 1-8.

Tjaden, B. (2015) 'De novo assembly of bacterial transcriptomes from RNA-seq data', *Genome Biology*, 16(1), 1.

Torres-Coy, J., Rodríguez-Castillo, B., Pérez-Alfonzo, R. and De Waard, J. (2016) 'Source investigation of two outbreaks of skin and soft tissue infection by *Mycobacterium abscessus subsp. abscessus* in Venezuela', *Epidemiology & Infection*, 144(5), 1117-1120.

Tu, H.-Z., Chang, S.-H., Huaug, T.-S., Huaug, W.-K., Liu, Y.-C. and Lee, S.S.-J. (2003) 'Microscopic morphology in smears prepared from MGIT broth medium for rapid presumptive identification of *Mycobacterium tuberculosis complex*,

Mycobacterium avium complex and Mycobacterium kansasii', *Annals of Clinical & Laboratory Science*, 33(2), 179-183.

Turapov, O., O'Connor, B.D., Sarybaeva, A.A., Williams, C., Patel, H., Kadyrov, A.S., Sarybaev, A.S., Woltmann, G., Barer, M.R. and Mukamolova, G.V. (2016) 'Phenotypically adapted *Mycobacterium tuberculosis* populations from sputum are tolerant to first-line drugs', *AntimicrobialAagents and Chemotherapy*, 60(4), 2476-2483.

Turenne, C.Y., Collins, D.M., Alexander, D.C. and Behr, M.A. (2008) 'M*ycobacterium avium subsp*. *paratuberculosis* and *M. avium subsp*. avium are independently evolved pathogenic clones of a much broader group of *M. avium* organisms', *Journal of Bacteriology*, 190(7), 2479-2487.

Urbano, S.B., Albarracín, V.H., Ordoñez, O.F., Farías, M.E. and Alvarez, H.M. (2013) 'Lipid storage in high-altitude Andean Lakes extremophiles and its mobilization under stress conditions in Rhodococcus sp. A5, a UV-resistant actinobacterium', *Extremophiles*, 17(2), 217-227.

VanGuilder, H.D., Vrana, K.E. and Freeman, W.M. (2008) 'Twenty-five years of quantitative PCR for gene expression analysis', *Biotechniques*, 44(5), 619-626.

Venkateswaran, K., Hattori, N., La Duc, M.T. and Kern, R. (2003) 'ATP as a biomarker of viable microorganisms in clean-room facilities', *Journal of Microbiological Methods*, 52(3), 367-377.

Viljoen, A., Blaise, M., de Chastellier, C. and Kremer, L. (2016) 'MAB _3551c encodes the primary triacylglycerol synthase involved in lipid accumulation in *Mycobacterium abscessus*', *MolecularMmicrobiology*, 102(4), 611-627.

Viljoen, A., Gutiérrez, A.V., Dupont, C., Ghigo, E. and Kremer, L. (2018) 'A simple and rapid gene disruption strategy in *Mycobacterium abscessus*: on the design and application of glycopeptidolipid mutants', *Frontiers in Cellular and Infection Microbiology*, 8, 69.

Villeneuve, C., Etienne, G., Abadie, V., Montrozier, H., Bordier, C., Laval, F., Daffe, M., Maridonneau-Parini, I. and Astarie-Dequeker, C. (2003) 'Surface-exposed Glycopeptidolipids of *Mycobacterium smegmatis* Specifically Inhibit the Phagocytosis of Mycobacteria by Human Macrophages identification of a novel family of glycopeptidolipids ', *Journal of Biological Chemistry*, 278(51), 51291-51300.

Voskuil, M.I., Bartek, I., Visconti, K. and Schoolnik, G.K. (2011) 'The response of *Mycobacterium tuberculosis* to reactive oxygen and nitrogen species', *Frontiers in Microbiology*, 2, 105.

Wagner, D. and Young, L. (2004) 'Nontuberculous mycobacterial infections: a clinical review', *Infection*, 32(5), 257-270.

Walker, D., Nwoguh, C. and Barer, M. (1994) 'A microchamber system for the rapid cytochemical demonstration of β-galactosidase and other properties in pathogenic microbes', *Letters in Applied Microbiology*, 18(2), 102-104.

Wallace Jr, R.J., Brown, B.A. and Griffith, D.E. (1998a) 'Nosocomial outbreaks/pseudo outbreaks caused by nontuberculous mycobacteria', *Annual Reviews in Microbiology*, 52(1), 453-490.

Wallace Jr, R.J., Brown, B.A. and Griffith, D.E. (1998b) 'Nosocomial outbreaks/pseudo outbreaks caused by nontuberculous mycobacteria', *Annual Review ofMmicrobiology*, 52(1), 453-490.

Wallace, R.J., Swenson, J.M., Silcox, V.A., Good, R.C., Tschen, J.A. and Stone, M.S. (1983) 'Spectrum of disease due to rapidly growing mycobacteria', *Review of Infectious Diseases*, 5(4), 657-679.

Wang, L., Feng, Z., Wang, X., Wang, X. and Zhang, X. (2009) 'DEGseq: an R package for identifying differentially expressed genes from RNA-seq data', *Bioinformatics*, 26(1), 136-138.

Wang, Y., Li, M., Stadler, S., Correll, S., Li, P., Wang, D., Hayama, R., Leonelli, L., Han, H. and Grigoryev, S.A. (2009) 'Histone hypercitrullination mediates chromatin decondensation and neutrophil extracellular trap formation', *J Cell Biol*, 184(2), 205-213.

Watanabe, M., Aoyagi, Y., Ridell, M. and Minnikin, D.E. (2001) 'Separation and characterization of individual mycolic acids in representative mycobacteria', *Microbiology*, 147(7), 1825-1837.

Wayne, L. and Kubica, G. (1986) 'The mycobacteria, p 1435–1457', *Bergey's Manual of Systematic Bacteriology*, 2.

Wendt, S.L., George, K.L., Parker, B.C., Gruft, H. and Falkinham III, J.O. (1980) 'Epidemiology of Infection by Nontuberculous Mycobacteria: III. Isolation of Potentially Pathogenic Mycobacteria from Aerosols 1, 2', *American Review of Respiratory Disease*, 122(2), 259-263.

Williams, C.M., Cheah, E.S., Malkin, J., Patel, H., Otu, J., Mlaga, K., Sutherland, J.S., Antonio, M., Perera, N. and Woltmann, G. (2014) 'Face mask sampling for the detection of Mycobacterium tuberculosis in expelled aerosols', *PloS One*, 9(8), e104921.

Wolinsky, E. (1979) 'Nontuberculous Mycobacteria and Associated Diseases 1, 2', *American Review of Respiratory Disease*, 119(1), 107-159.

Yamazaki, Y., Danelishvili, L., Wu, M., MacNab, M. and Bermudez, L.E. (2006) '*Mycobacterium avium* genes associated with the ability to form a biofilm', *Appl. Environ. Microbiol.*, 72(1), 819-825.

Young, H.-K. (1993) 'Antimicrobial resistance spread in aquatic environments', *Journal of Antimicrobial Chemotherapy*, 31(5), 627-635.

Zhang, Y., Rajagopalan, M., Brown, B.A. and Wallace, R. (1997) 'Randomly amplified polymorphic DNA PCR for comparison of *Mycobacterium abscessus* strains from nosocomial outbreaks', *Journal of Clinical Microbiology*, 35(12), 3132-3139.

Appendix

Appendix I

Additional Table and Figures

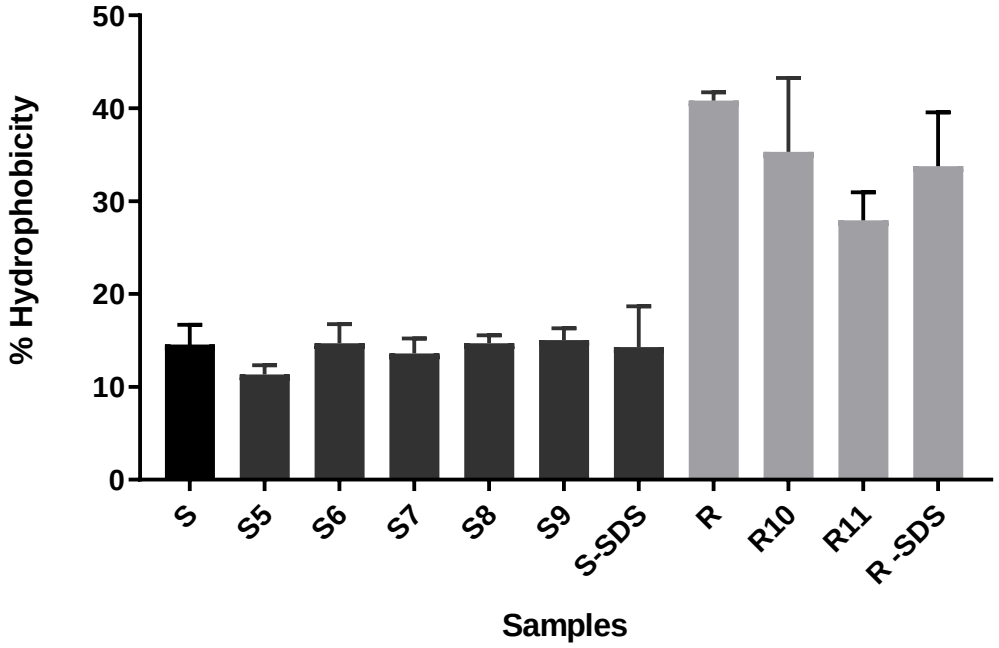


Figure 1: CSH test for *Mab* lab control strains S and R, S-SDS, R –SDS and clinical isolates (S5, S6, S7, S8, S9, R10 R11) respectively. Error bars signify the standard deviation of three independent experiments performed with triplicate technical.

Appendix II

Table 1: Total % of LB content in lab control strain and clinical isolates.

Sample	% LB
Sc	39.9
Rc	36
S1	4
R1	9
S2	64.5
R2	0.4
S3	9.7
R3	0.8
S4	10
R4	14.4

Appendix III

Table 2: MLST databases.

Reference genome of *Mab* ATCC_19977-					**ST= 5**
Locus	**%Identity**	**HSP Length**	**Allele Length**	**Gap**	**Allele**
Mab_argH	100.00	510	510	0	*Mab_argH_1*
Mab_cya	100.00	540	540	0	*Mab_cya_2*
Mab_gnd	100.00	506	506	0	*Mab_gnd_1*
Mab_murC	100.00	445	445	0	*Mab_murC_2*
Mab_pta	100.00	520	520	0	*Mab_pta_4*
Mab_purH	100.00	497	497	0	*Mab_purH_4*
Mab_rpoB	100.00	503	503	0	*Mab_rpoB_1*
***Mab* / S (lab control)**					**Unknown ST**
Locus	**%Identity**	**HSP Length**	**Allele Length**	**Gap**	**Allele**
Mab_argH	94.41	510	510	0	*Mab_argH_3*
Mab_cya	100.00	540	540	0	*Mab_cya_1*
Mab_gnd	99.80	506	506	0	*Mab_gnd_1*
Mab_murC	100.00	445	445	0	*Mab_murC_2*
Mab_pta	100.00	520	520	0	*Mab_pta_4*

Mab_purH	100.00	497	497	0	*Mab_purH_2*
Mab_rpoB	100.00	503	503	0	*Mab_rpoB_1*
***Mab* / R (lab control)**					**ST= 5**
Locus	**%Identity**	**HSP Length**	**Allele Length**	**Gap**	**Allele**
Mab_argH	100.00	510	510	0	*Mab_argH_3*
Mab_cya	100.00	540	540	0	*Mab_cya_1*
Mab_gnd	100.00	506	506	0	*Mab_gnd_4*
Mab_murC	100.00	445	445	0	*Mab_murC_3*
Mab_pta	100.00	520	520	0	*Mab_pta_1*
Mab_purH	100.00	497	497	0	*Mab_purH_2*
Mab_rpoB	100.00	503	503	0	*Mab_rpoB_1*
Patient 1	***Mab* / S**				**ST= 9**
Locus	**%Identity**	**HSP Length**	**Allele Length**	**Gap**	**Allele**
Mab_argH	100.00	510	510	0	*Mab_argH_1*
Mab_cya	100.00	540	540	0	*Mab_cya_2*
Mab_gnd	100.00	506	506	0	*Mab_gnd_1*
Mab_murC	100.00	445	445	0	*Mab_murC_2*
Mab_pta	100.00	520	520	0	*Mab_pta_4*
Mab_purH	100.00	497	497	0	*Mab_purH_4*

Mab_rpoB	100.00	503	503	0	*Mab_rpoB_1*
Patient 1	***Mab* / R**				**ST= 9**
Locus	%Identity	HSP Length	Allele Length	Gap	Allele
Mab_argH	100.00	510	510	0	*Mab_argH_1*
Mab_cya	100.00	540	540	0	*Mab_cya_2*
Mab_gnd	100.00	506	506	0	*Mab_gnd_1*
Mab_murC	100.00	445	445	0	*Mab_murC_2*
Mab_pta	100.00	520	520	0	*Mab_pta_4*
Mab_purH	100.00	497	497	0	*Mab_purH_4*
Mab_rpoB	100.00	503	503	0	*Mab_rpoB_1*
Patient 2	***Mab* / S**				**ST= 5**
Locus	%Identity	HSP Length	Allele Length	Gap	Allele
Mab_argH	100.00	510	510	0	*Mab_argH_3*
Mab_cya	100.00	540	540	0	*Mab_cya_1*
Mab_gnd	100.00	506	506	0	*Mab_gnd_4*
Mab_murC	100.00	445	445	0	*Mab_murC_3*
Mab_pta	100.00	520	520	0	*Mab_pta_1*
Mab_purH	100.00	497	497	0	*Mab_purH_2*
Mab_rpoB	100.00	503	503	0	*Mab_rpoB_1*

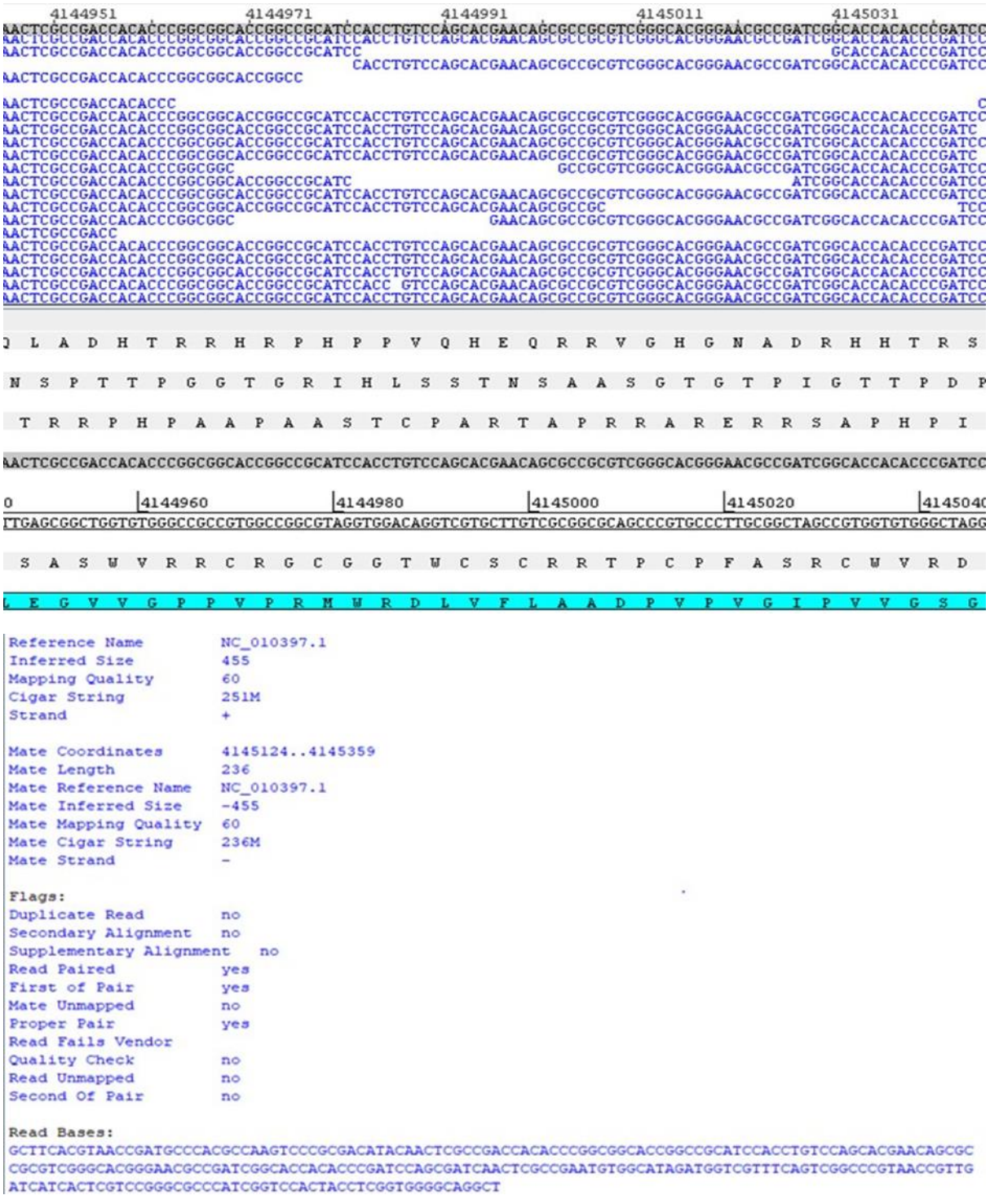
Patient 2	***Mab* /**	**R**			**ST= 5**
Locus	%Identity	HSP Length	Allele Length	Gap	Allele
Mab_argH	100.00	510	510	0	*Mab_argH_3*
Mab_cya	100.00	540	540	0	*Mab_cya12*
Mab_gnd	100.00	506	506	0	*Mab_gnd_4*
Mab_murC	100.00	445	445	0	*Mab_murC_3*
Mab_pta	100.00	520	520	0	*Mab_pta_1*
Mab_purH	100.00	497	497	0	*Mab_purH_2*
Mab_rpoB	100.00	503	503	0	*Mab_rpoB_1*
Patient 3	***Mab* / S**				**ST= 5**
Locus	%Identity	HSP Length	Allele Length	Gap	Allele
Mab_argH	100.00	510	510	0	*Mab_argH_3*
Mab_cya	100.00	540	540	0	*Mab_cya_1*
Mab_gnd	100.00	506	506	0	*Mab_gnd_4*
Mab_murC	100.00	445	445	0	*Mab_murC_3*
Mab_pta	100.00	520	520	0	*Mab_pta_1*
Mab_purH	100.00	497	497	0	*Mab_purH_2*
Mab_rpoB	100.00	503	503	0	*Mab_rpoB_1*
Patient 3	***Mab* /**	**R**			**ST= 5**

Locus	%Identity	HSP Length	Allele Length	Gap	Allele
Mab_argH	100.00	510	510	0	*Mab_argH_3*
Mab_cya	100.00	540	540	0	*Mab_cya_1*
Mab_gnd	100.00	506	506	0	*Mab_gnd_4*
Mab_murC	100.00	445	445	0	*Mab_murC_3*
Mab_pta	100.00	520	520	0	*Mab_pta_1*
Mab_purH	100.00	497	497	0	*Mab_purH_2*
Mab_rpoB	100.00	503	503	0	*Mab_rpoB_1*
Patient 4	***Mab* / S**				**ST= 9**
Locus	**%Identity**	**HSP Length**	**Allele Length**	**Gap**	**Allele**
Mab_argH	100.00	510	510	0	*Mab_argH_1*
Mab_cya	100.00	540	540	0	*Mab_cya_2*
Mab_gnd	100.00	506	506	0	*Mab_gnd_1*
Mab_murC	100.00	445	445	0	*Mab_murC_2*
Mab_pta	100.00	520	520	0	*Mab_pta_4*
Mab_purH	100.00	497	497	0	*Mab_purH_4*
Mab_rpoB	100.00	503	503	0	*Mab_rpoB_1*
Patient 4	***Mab/* R**				**ST= 9**
Locus	**%Identity**	**HSP Length**	**Allele Length**	**Gap**	**Allele**

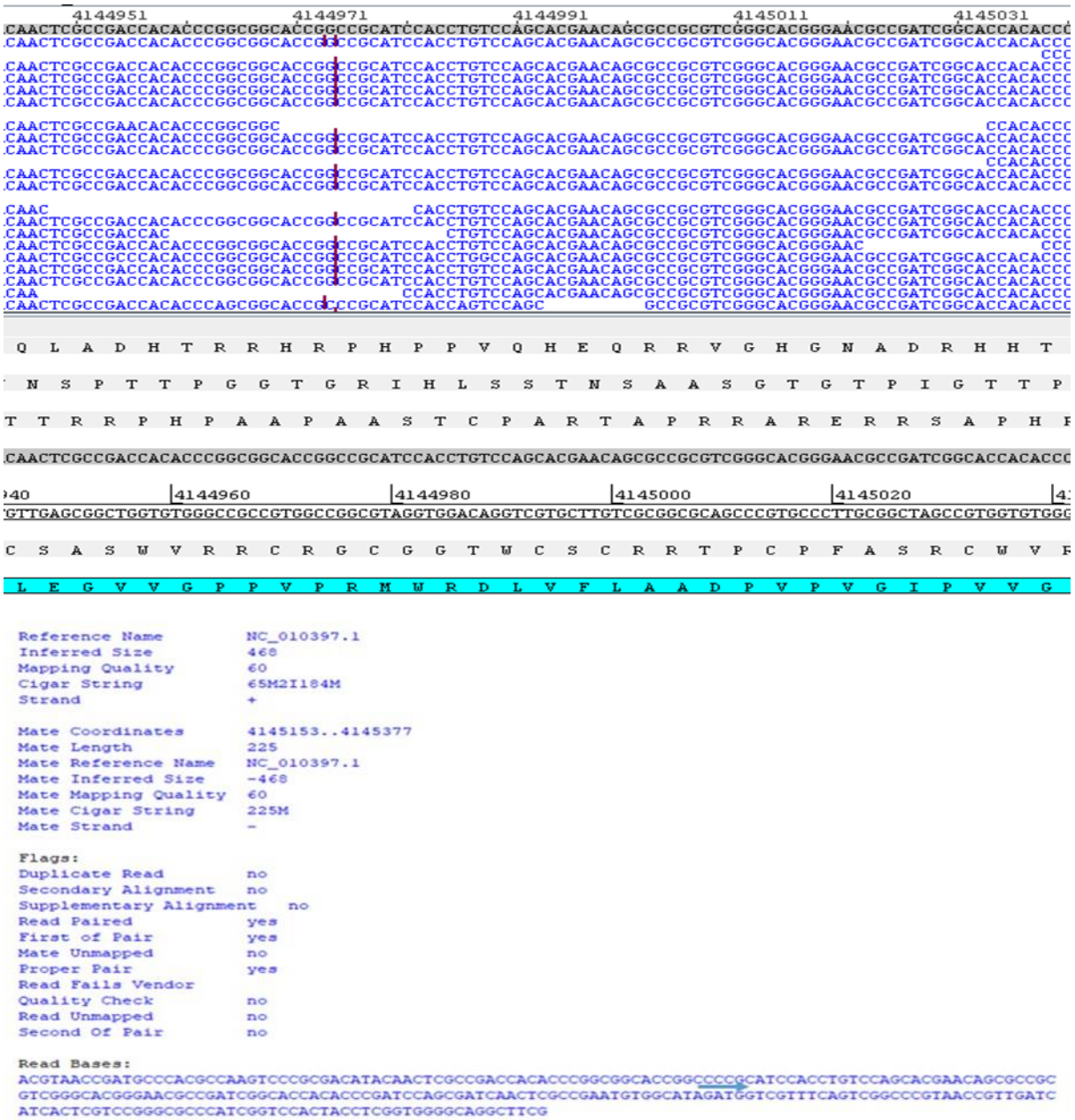
Mab_argH	100.00	510	510	0	*Mab_argH_3*
Mab_cya	100.00	540	540	0	*Mab_cya_1*
Mab_gnd	100.00	506	506	0	*Mab_gnd_4*
Mab_murC	100.00	445	445	0	*Mab_murC_3*
Mab_pta	100.00	520	520	0	*Mab_pta_1*
Mab_purH	100.00	497	497	0	*Mab_purH_2*
Mab_rpoB	100.00	503	503	0	*Mab_rpoB_1*

Appendix IV

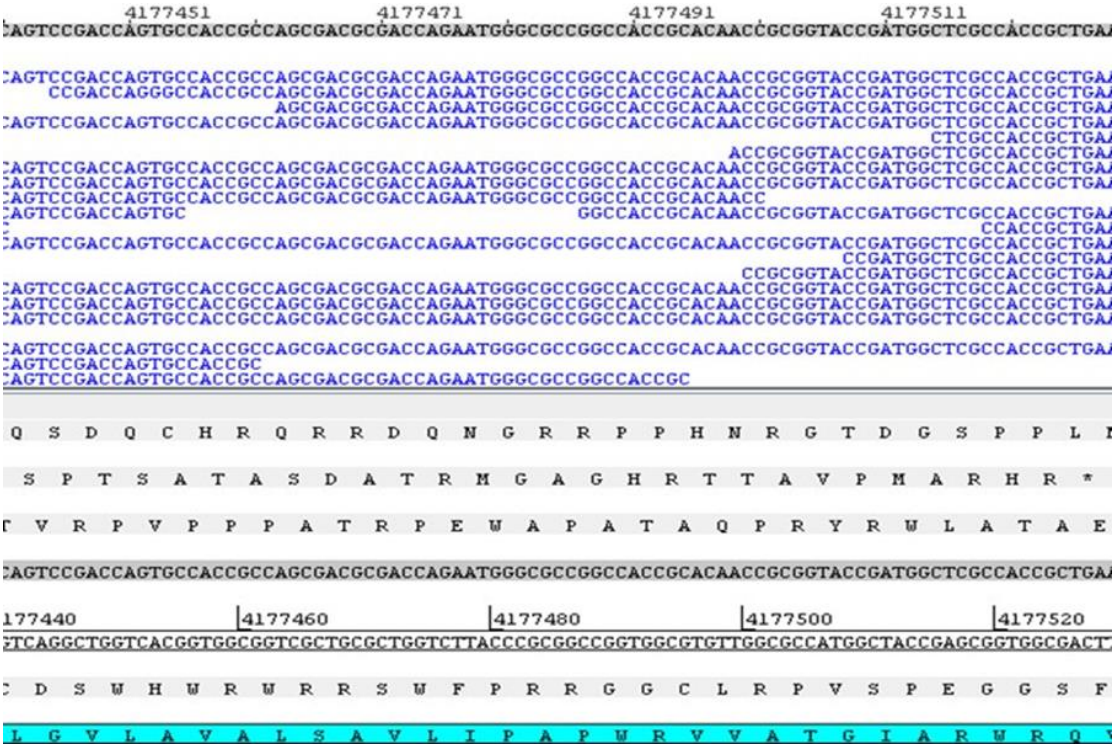
S



R


4144951
4144971
4144991
4145011
4145031
4144960
4144980
4145000
4145020
Reference Name NC_010397.1
Inferred Size 468
Mapping Quality 60
Cigar String 65M2I184M
Strand +
Mate Coordinates 4145153..4145377
Mate Length 225
Mate Reference Name NC_010397.1
Mate Inferred Size -468
Mate Mapping Quality 60
Mate Cigar String 225M
Mate Strand -
Flags:
Duplicate Read no
Secondary Alignment no
Supplementary Alignment no
Read Paired yes
First of Pair yes
Mate Unmapped no
Proper Pair yes
Read Fails Vendor
Quality Check no
Read Unmapped no
Second Of Pair no
Read Bases:
ACGTAACCGATGCCCACGCCAAGTCCCGCGACATACAACTCGCCGACCACACCCGGCGGCACCGGCCCCGCATCCACCTGTCCAGCACGAACAGCGCCGC
GTCGGGCACGGGAACGCCGATCGGCACCACACCCGATCCAGCGATCAACTCGCCGAATGTGGCATAGATGGTCGTTTCAGTCGGCCCGTAACCGTTGATC
ATCACTCGTCCGGGCGCCCATCGGTCCACTACCTCGGTGGGGCAGGCTTCG

S1



D00200:356:H5JGHBCX2:2:1212:13268:25937

File

Read Name	D00200:356:H5JGHBCX2:2:1212:13268:25937
Coordinates	4177280..4177495
Length	216
Reference Name	NC_010397.1
Inferred Size	815
Mapping Quality	60
Cigar String	216M
Strand	+
Mate Coordinates	4177844..4178094
Mate Length	251
Mate Reference Name	NC_010397.1
Mate Inferred Size	-815
Mate Mapping Quality	60
Mate Cigar String	251M
Mate Strand	-

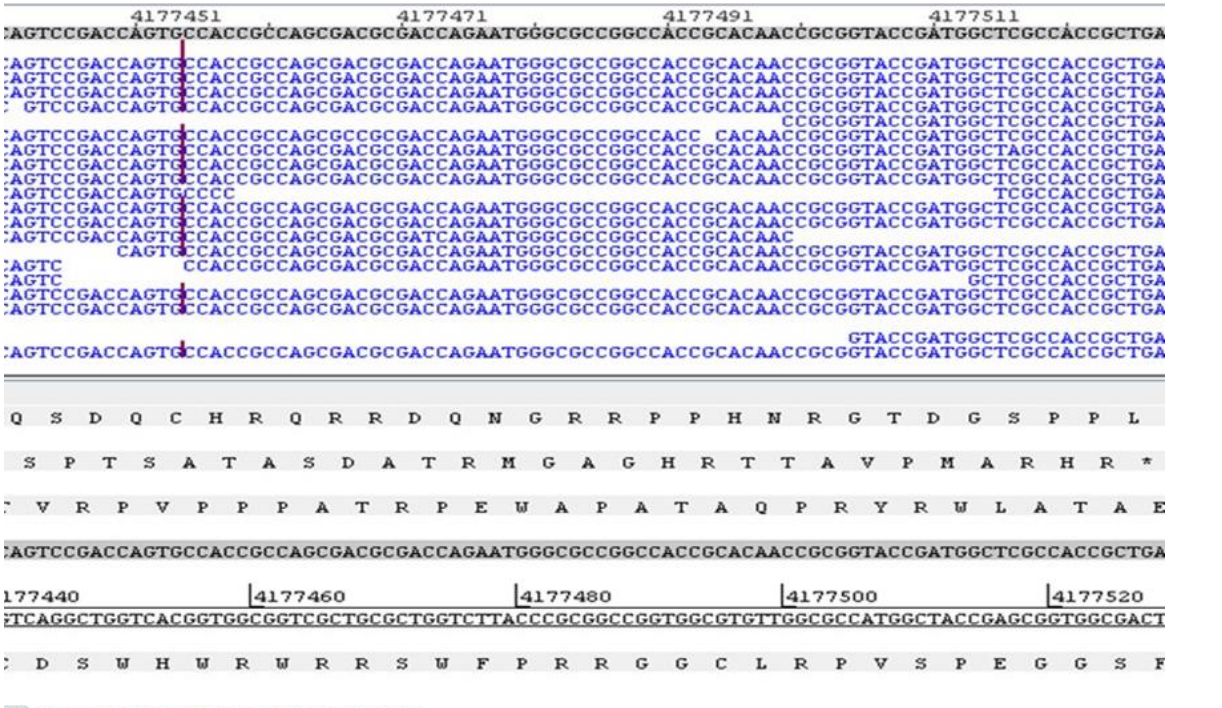
Flags:

Duplicate Read	no
Secondary Alignment	no
Supplementary Alignment	no
Read Paired	yes
First of Pair	no
Mate Unmapped	no
Proper Pair	yes
Read Fails Vendor Quality Check	no
Read Unmapped	no
Second Of Pair	yes

Read Bases:

GGGTTCCTTAGGTCCCTGTCGGTTTCGATCAACAGCACATCGGGCGTAAGACGCGACTCGGAGAAGTGCCGCTGAGCGGCAGCGAATCCCTGATTCGCCG
GGATATCGCTGGAGATATAGAGCCGATCGTTGTAGCTGGTCTTATACCCCGGGAGGGTGGACAGTCCGACCAGTGCCACCGCCAGCGACGCGACCAGAAT
GGGCGCCGGCCACCGC

R1



D00200:358:H5JKKBCX2:2:2201:10206:95479

File

Read Name	D00200:358:H5JKKBCX2:2:2201:10206:95479
Coordinates	4177303..4177551
Length	251
Reference Name	NC_010397.1
Inferred Size	-358
Mapping Quality	60
Cigar String	152M2I97M
Strand	-
Mate Coordinates	4177194..4177401
Mate Length	208
Mate Reference Name	NC_010397.1
Mate Inferred Size	358
Mate Mapping Quality	60
Mate Cigar String	208M
Mate Strand	+

Flags:

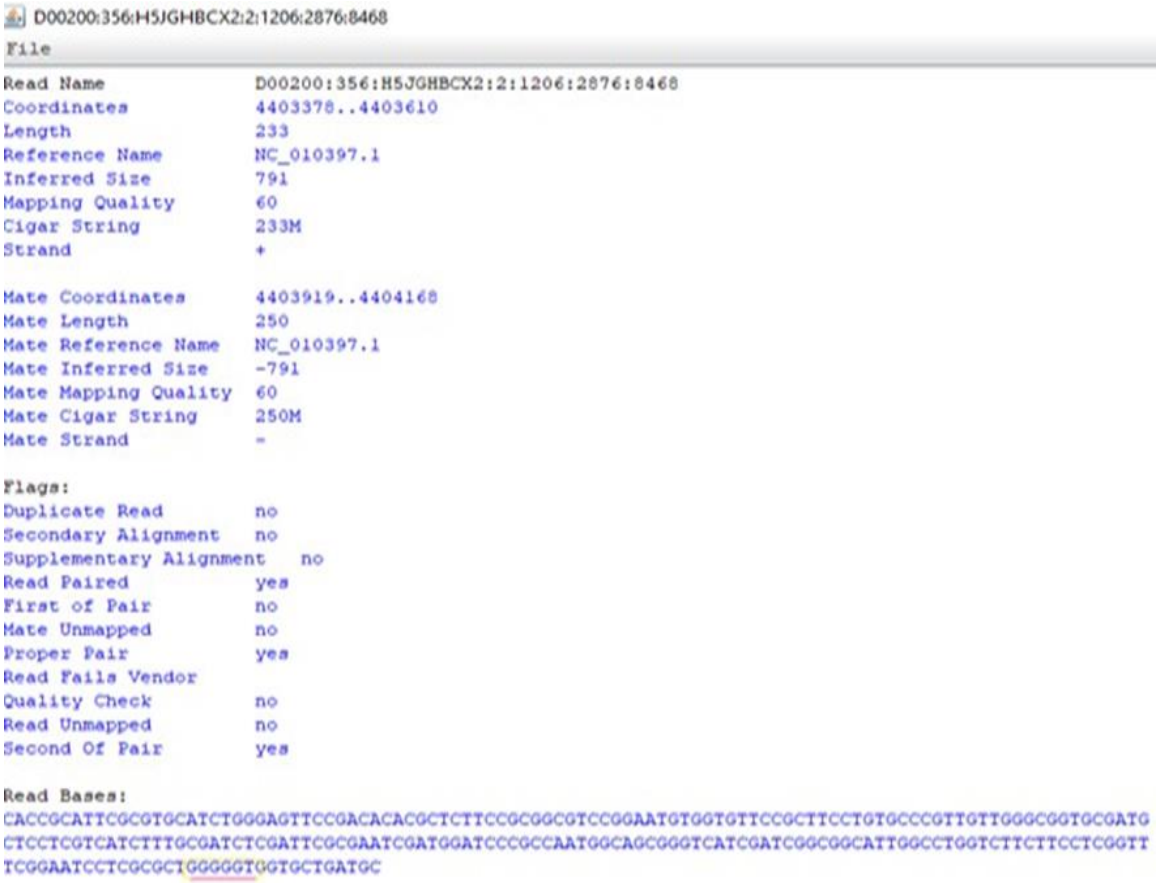
Duplicate Read	no
Secondary Alignment	no
Supplementary Alignment	no
Read Paired	yes
First of Pair	yes
Mate Unmapped	no
Proper Pair	yes
Read Fails Vendor Quality Check	no
Read Unmapped	no
Second Of Pair	no

Read Bases:

TTCGATCAACAGCACATCGGGCGTAAGACGCGACTCGGAGAAGTGCCGCTGAGCGGCAGCGAATCCCTGATTCGCCGGGATATCGCTGGAGATATAGAGC
CGATCGTTGTAGCTGGTCTTATACCCCGGGAGGGTGGACAGTCCGACCAGTGCCCCACCGCCAGCGACGCGACCAGAATGGGCGCCGGCCACCGCACAAC
CGCGGTACCGATGGCTCGCCACCGCTGAACCCTCAACTTGCGCTTGGGTTC

S2


4403581
4403601
4403621
4403641


D00200:356:H5JGHBCX2:2:1206:2876:8468
File
Read Name D00200:356:H5JGHBCX2:2:1206:2876:8468
Coordinates 4403378..4403610
Length 233
Reference Name NC_010397.1
Inferred Size 791
Mapping Quality 60
Cigar String 233M
Strand +
Mate Coordinates 4403919..4404168
Mate Length 250
Mate Reference Name NC_010397.1
Mate Inferred Size -791
Mate Mapping Quality 60
Mate Cigar String 250M
Mate Strand -
Flags:
Duplicate Read no
Secondary Alignment no
Supplementary Alignment no
Read Paired yes
First of Pair no
Mate Unmapped no
Proper Pair yes
Read Fails Vendor
Quality Check no
Read Unmapped no
Second Of Pair yes
Read Bases:

R2

D00200:356:H5JGHBCX2:1:2203:3563:74686

File

Read Name	D00200:356:H5JGHBCX2:1:2203:3563:74686
Coordinates	4403417..4403616
Length	201
Reference Name	NC_010397.1
Inferred Size	1133
Mapping Quality	60
Cigar String	177M1I23M
Strand	+
Mate Coordinates	4404318..4404549
Mate Length	232
Mate Reference Name	NC_010397.1
Mate Inferred Size	-1133
Mate Mapping Quality	60
Mate Cigar String	232M
Mate Strand	-

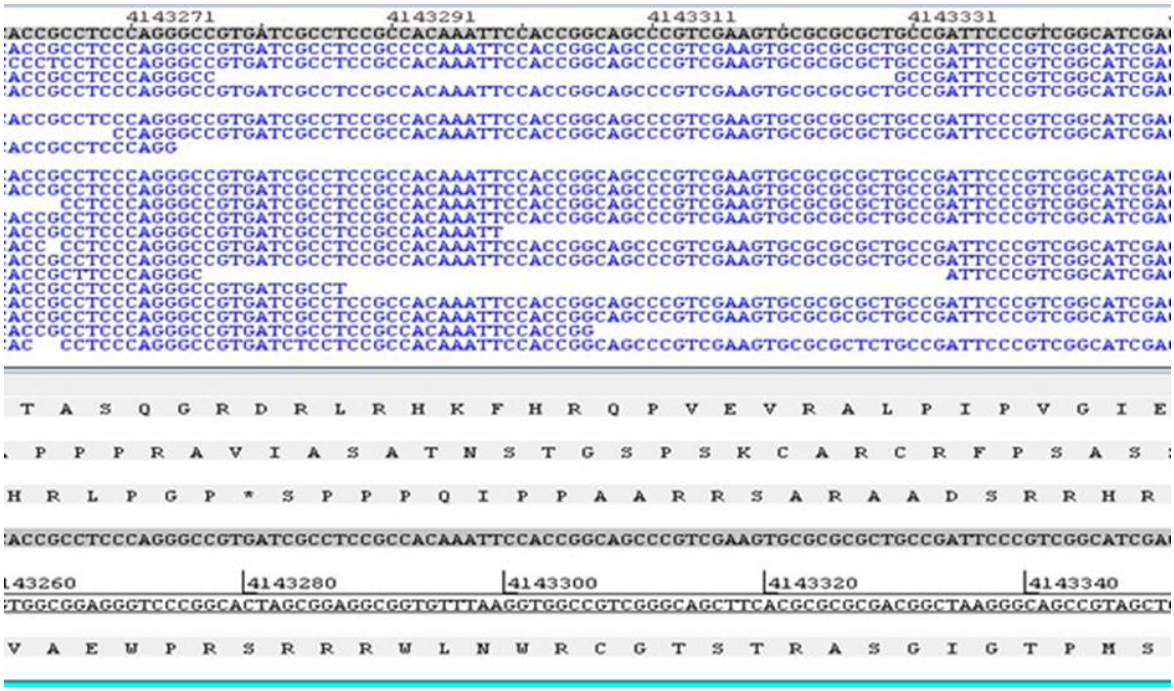
Flags:

Duplicate Read	no
Secondary Alignment	no
Supplementary Alignment	no
Read Paired	yes
First of Pair	no
Mate Unmapped	no
Proper Pair	yes
Read Fails Vendor Quality Check	no
Read Unmapped	no
Second Of Pair	yes

Read Bases:

CTTCCGCGGCGTCCGGAATGTGGTGTTCCGCTTCCTGTGCCCGTTGTTGGGCGGTGCGATGCTCCTCGTCATCTTTGCGATCTCGATTCGCGAATCCATG
GATCCCGCCAATGGCAGCGGGTCATCGATGGGCGGCATTGGCCTGGTCTTCTTCCTCGGTTTCGGAATCCTCGCGCTGGGGGGGTGGGGCTGATGCTCGTC
C

S3



Reference Name	NC_010397.1
Inferred Size	-732
Mapping Quality	60
Cigar String	188M
Strand	-
Mate Coordinates	4142722..4142968
Mate Length	247
Mate Reference Name	NC_010397.1
Mate Inferred Size	732
Mate Mapping Quality	60
Mate Cigar String	247M
Mate Strand	+

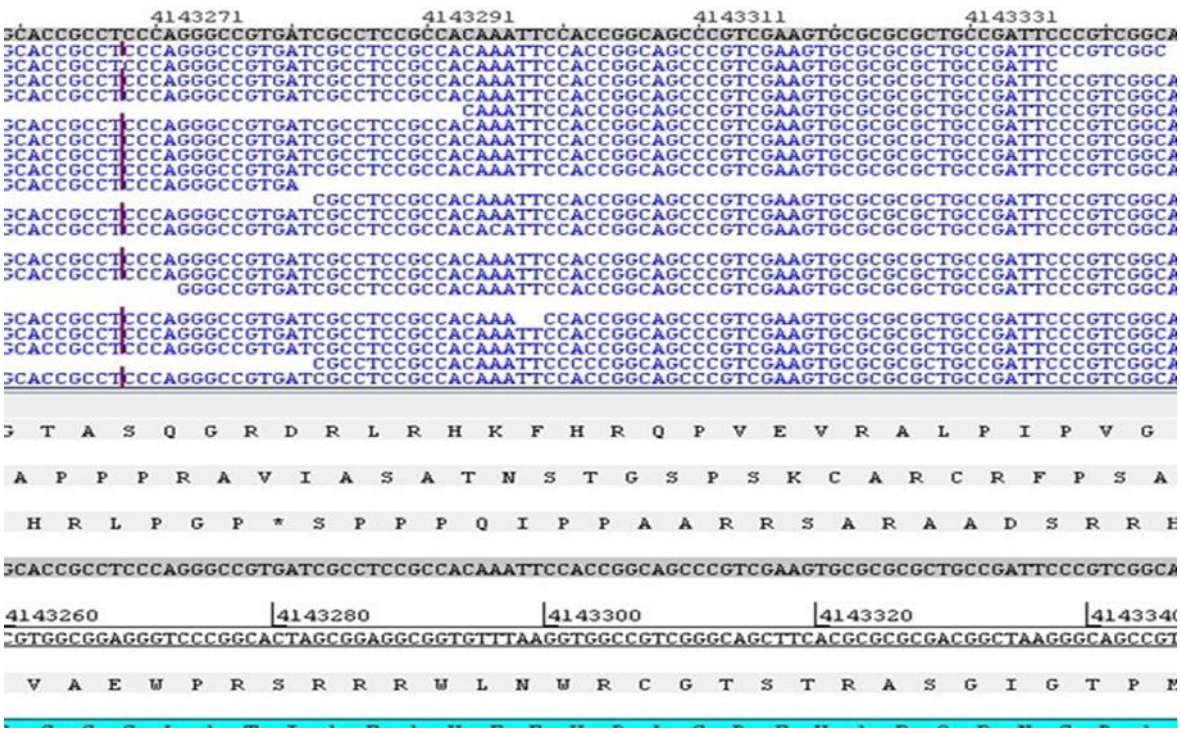
Flags:

Duplicate Read	no
Secondary Alignment	no
Supplementary Alignment	no
Read Paired	yes
First of Pair	no
Mate Unmapped	no
Proper Pair	yes
Read Fails Vendor	
Quality Check	no
Read Unmapped	no
Second Of Pair	yes

Read Bases:

CCTCCCAGGGCCGTGATCTCCTCCGCCACAAATTCCACCGGCAGCCCGTCGAAGTGCGCGCTCTGCCGATTCCCGTCGGCATCGAGCTGGTAAACGGAGT
ACGGCGCCACTCCCGTCGCGACGACGCTGAGCACCATGCGGGCCACCGTGTCCGAAGCGTTTACCTGACCGGCATAGCTCGTGTCAAC

R3



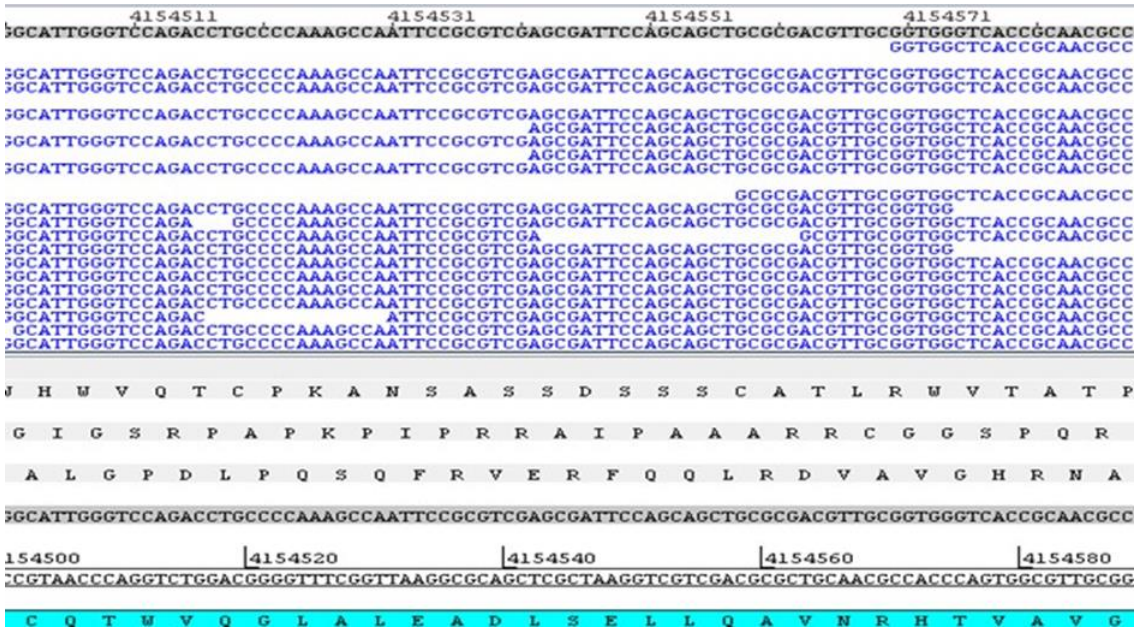
Reference Name	NC_010397.1
Inferred Size	-590
Mapping Quality	60
Cigar String	79M1I136M
Strand	-
Mate Coordinates	4142815..4143064
Mate Length	250
Mate Reference Name	NC_010397.1
Mate Inferred Size	590
Mate Mapping Quality	60
Mate Cigar String	250M
Mate Strand	+

Flags:

Duplicate Read	no
Secondary Alignment	no
Supplementary Alignment	no
Read Paired	yes
First of Pair	no
Mate Unmapped	no
Proper Pair	yes
Read Fails Vendor Quality Check	no
Read Unmapped	no
Second Of Pair	yes

Read Bases:

GCTGCCGTCGTCATGCGGATTAATCACGTGGTAGGTCTCAAAATCCGCAGATCCGTCGAATGAGCCGGCGGCACCGCCTCCCCCAGGGCCGTGATCGCCTC
CGCCACAAATTCCACCGGCAGCCCGTCGAAGTGCGCGCGCTGCCGATTCCCGTCGGCATCGAGCTGGTAAACGGAGTACGGCGCCACTCCCGTCGCGACG
ACGCTGAGCACCATGC



Reference Name	NC_010397.1
Inferred Size	691
Mapping Quality	60
Cigar String	251M
Strand	+
Mate Coordinates	4154808..4155058
Mate Length	251
Mate Reference Name	NC_010397.1
Mate Inferred Size	-691
Mate Mapping Quality	60
Mate Cigar String	251M
Mate Strand	-

Flags:

Duplicate Read	no
Secondary Alignment	no
Supplementary Alignment	no
Read Paired	yes
First of Pair	yes
Mate Unmapped	no
Proper Pair	yes
Read Fails Vendor Quality Check	no
Read Unmapped	no
Second Of Pair	no

Read Bases:

ACTTCCTCGCTGACGAGCAAGGCATGCAACTCTTCTGGTGAGCGGACAACCGCGTCCGGCACGACGAGCAATCGTCCGCCGTACAACAGCGCACCCCAGA
TTTCCCACACCGAGTAGTCGAATGCCAGCGAATGGCATTGGGTCCAGACCTGCCCCAAAGCCAATTCCGCGTCGAGCGATTCCAGCAGCTGCGCGACGTT
GCGGTGGCTCACCGCAACGCCCTTGGGAGCACCGGTAGTACCCGACGTGTA

R4

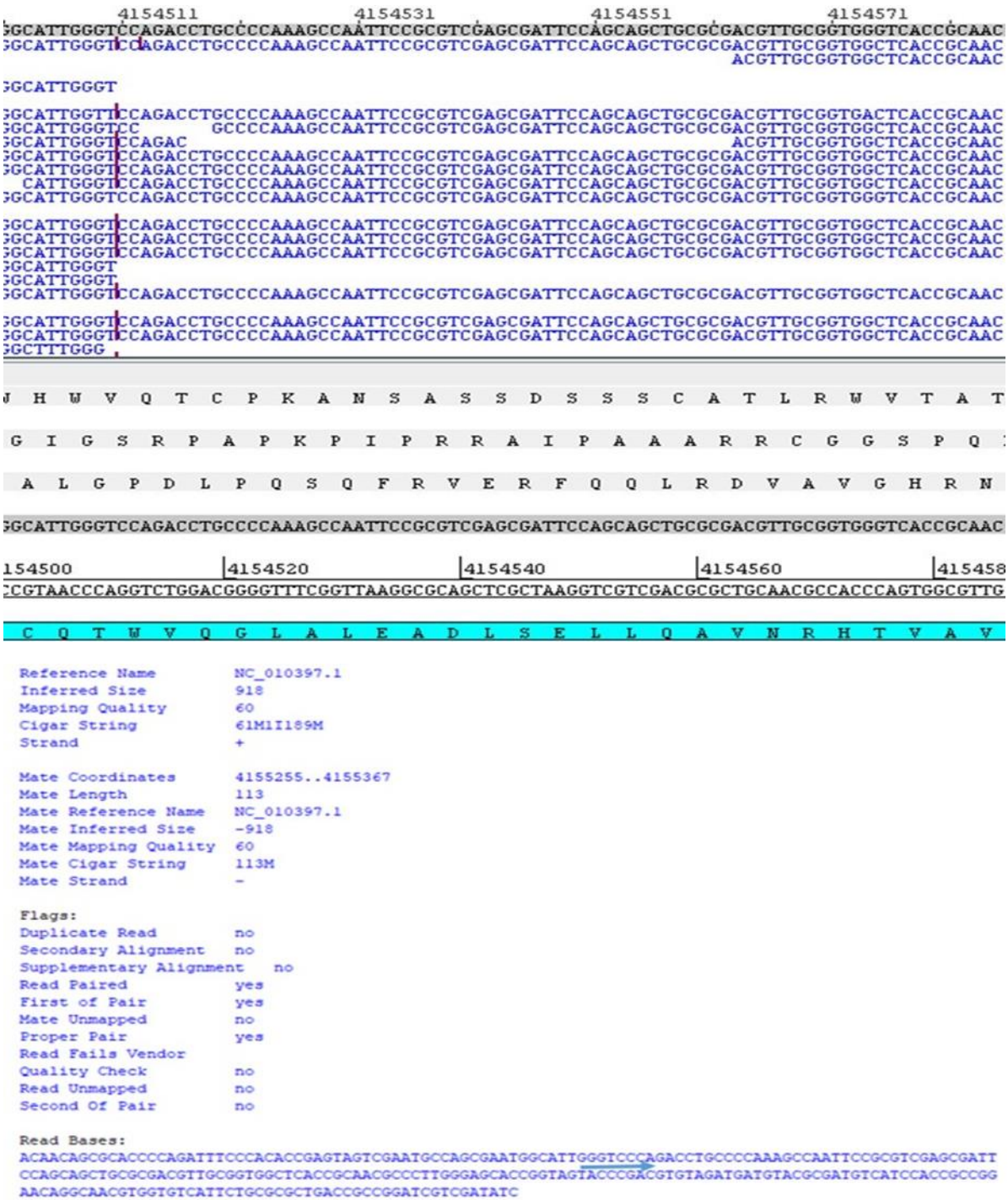


Figure2: Representation of SNPs in Artmis for S, R control and clinical isolates. The Bam files were used to visualise the genetic modification in assembled sequence using ATCC 19977 whole genome as a reference sequence. The red line represents the frame shift mutation at single nucleotide level.

Appendix V

Standard curve and melt point analysis of PCR products

After confirming the correctly size product was formed using the end point PCR *Mab* genomic DNA was used to determine the efficiency of the *gap, mps2, mps1* and *mmpl4B* primers. The best way to assess reaction efficiency is by generation of a standard curve in which five 10 fold serial dilutions were performed from 10^7 genomic copies/µl to obtain 10^6, 10^5, 10^4, 10^3,10^2 copies/µl respectively. A non-template control composed of nuclease free dH_2O was included in each run. Although the primers used were taken from another study in which their efficiency was stated as above 90% (Pawlik *et al.* 2013) it was still necessary to confirm their validity because it was planned to use the Livak method for later *gap, mps2, mps1and mmpl4B* gene expression analysis. According to Livak and Schmittgen (2001) it is recommended that all the primers have an efficiency ranging between 90 and 100 %.

In general, 100 % amplification efficiency means during each replication cycle the number of molecules of the target sequence doubles. Low reaction efficiency below 90% should be avoided because standard curves used to compare between different targets accurately. For example, when comparing two genes if one has a reaction efficiency less than the other it seems that the expression level of that gene has been reduced when actually it reduced because of the reduction in a reaction efficiency. Similarly, efficiencies above 100 % are not recommended because they are an indicator of either presence of polymerase enzyme inhibitors such as excessive amounts of DNA, or primer dimer formation. Primer dimers occur when the primers bind to each other instead of binding to the template and are amplified rather than the target gene. Figure 3, shows a standard curve of a *gap, mps2, mps1 and mmpl4B* primers pairs. It is clear from the Figure that the reaction efficiency is at least 90% and the r^2 value is greater than 0.99, which makes these primers acceptable for further use in qRT-PCR analysis according to Livak method (Livak and Schmittgen 2001).

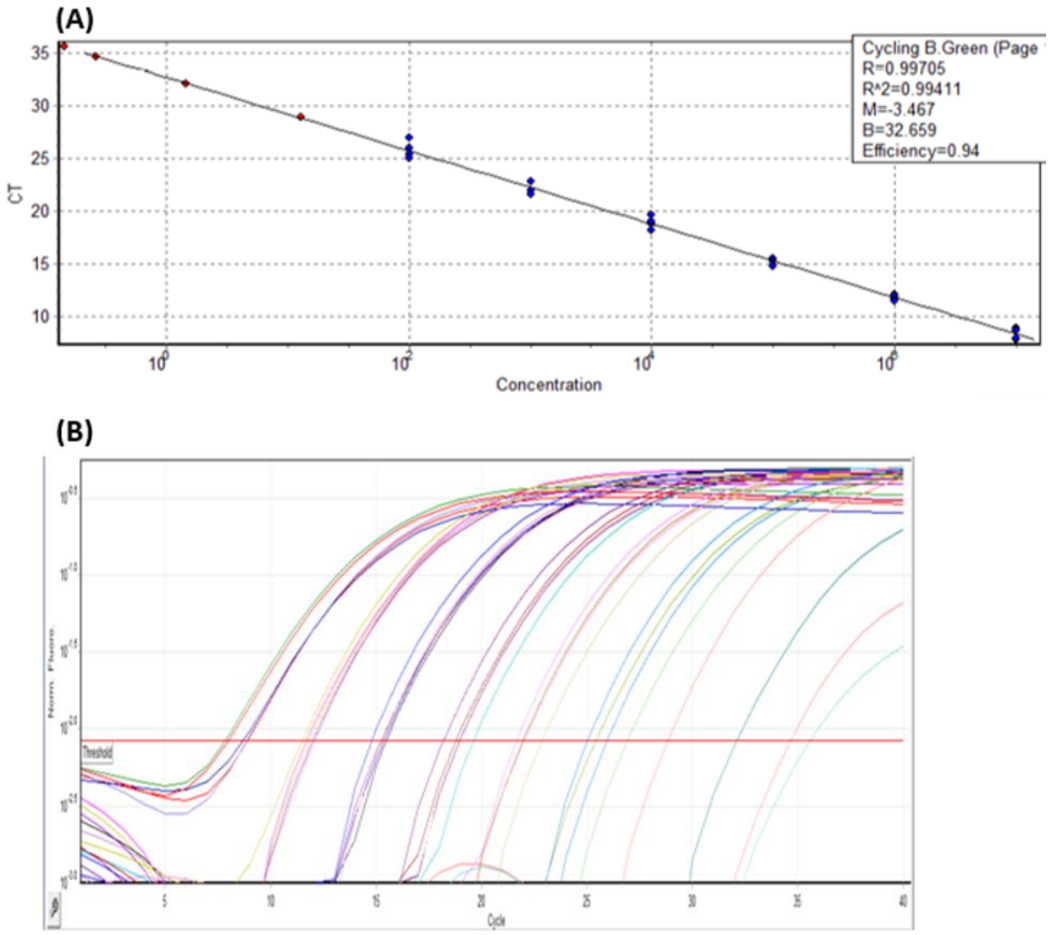


Figure 3: Standard curve and melt analysis graph *gap, mps2, mps1and mmpl4B* primers. Image **(A)** Standard curve of primers. Image. **(B)** Quantitative measurements of data for each cycle.