
UNIVERSITY OF BIRMINGHAM

# Phenotypic profiling of mycobacterial mutant libraries, through chemical-genomic screening

by

Rudi Sullivan

A thesis submitted to the University of Birmingham for the degree of

DOCTOR OF PHILOSOPHY

School of Biosciences

College of Life and Environmental Sciences

University of Birmingham

April 2024

# Abstract

The genus *Mycobacterium* encompasses numerous important human and environmental pathogens including the causative agent of tuberculosis, *Mycobacterium tuberculosis*, which is responsible for around 1.5 million deaths each year. Although curable, the treatment for *M. tuberculosis* is extensive and antibiotic resistance is an ongoing concern. Therefore, there is a demand for the development of new anti-mycobacterial drugs. Functional characterisation of mycobacterial genomes could lead to the discovery of new drug targets and help to characterise drug resistance pathways. However, mycobacterial genomes are poorly characterised with ~25% of the *M. tuberculosis* genome remaining functionally unannotated. Therefore, new methods are required for the characterisation of the outstanding genome.

Ordered mutant libraries are extremely valuable resources that not only provide a catalogue of mutants for further study but can also be combined with functional genomics to help characterise bacterial genomes. As *Mycobacterium abscessus* genetic resources are limited, we constructed an ordered *M. abscessus* mutant library. The library was constructed via random transposon mutagenesis using the phage delivery vector, ΦMycoMarT7, in which 18,432 mutants were manually ordered into 96-well plates. Although we were unable to finish the deconvolution of the library, once complete, this library will represent the largest collection of *M. abscessus* mutants in existence.

To demonstrate the utility of ordered libraries for functionally characterising bacterial genomes, we optimised a method for the chemical-genomic screening of mycobacteria using a *M. bovis* BCG mutant library. We first condensed the *M. bovis* BCG mutant library so that it was compatible with high-throughput screening. To increase experimental throughput and reduce human error, we implemented robotics to develop the chemical-genomics method. In this method, *M. bovis* BCG liquid source plates were replicated in a 384-colony format onto solid agar replica plates containing sub-inhibitory concentrations of different perturbations. Colony morphology and mutant fitness was then assessed via IRIS and ChemGAPP. To validate the method, we conducted a pilot screen using 2 library plates and 22 conditions and discovered 76 mutants had significant fitness changes in response to at least one condition, many of which had insertions in genes that had not been previously characterised. Through clustering analysis, we also identified various uncharacterised genes that may fulfil roles in important biological pathways.

Lastly, we developed a pipeline for validating the screening hits in which markerless mycobacterial gene deletion mutants were generated via ORBIT and phenotypically characterised via microscopy, growth kinetics and chemical screening. This research contributed towards the discovery of two cell-wall associated enzymes, the D,D-carboxypeptidase, LpqF, and the α-1,6 mannanase, *MSMEG_0740* (Rv0365c).

Therefore, taken together, this work outlines a method for characterising mycobacterial genomes through the generation of ordered mutant libraries, functional genomics and reverse-genetic phenotyping. Using this pipeline, we successfully

generated an ordered *M. abscessus* mutant library which will represent a useful resource for future study, we conducted a pilot functional genomic screening to phenotypically profile mutants of numerous previously uncharacterised genes and through reverse-genetics, we also discovered two new enzymes involved in the synthesis and recycling of the mycobacterial cell wall.

# Acknowledgements

I would like to acknowledge the University of Birmingham, and the BBSRC for their funding and support and I would especially like to thank both Manual Banzhaf and Patrick Moynihan for their supervision, expertise, and ongoing encouragement.

I would also like to thank my colleagues in my both the Moynihan and Banzhaf lab groups. I have thoroughly enjoyed our collaborative working and I appreciate all the support and motivation, particularly when I was stuck in a research bottleneck or things in the lab were not going to plan.

Lastly, I would also like to express my gratitude to my parents who have put up with all of my frustrations and have always been there when I needed them.

And to my partner George. Thank you for accompanying me along every step of this journey. It’s been a fantastic four years.

# Table of Contents

**Chapter 1 – Introduction** ..... 16

**Part 1 – Mycobacteria** ..... 17

1.1 An Introduction to Mycobacteria ..... 17

1.1.1 The Cytoplasmic Membrane ..... 19

1.1.2 The Cell Wall Core and The Outer Membrane ..... 20

1.1.3 Peptidoglycan ..... 20

1.1.4 Arabinogalactan ..... 24

1.1.5 Mycomembrane ..... 30

1.1.6 Capsule ..... 33

1.2 *Mycobacterium tuberculosis* ..... 34

1.3 *Mycobacterium bovis* ..... 36

1.4 Nontuberculosis Mycobacteria ..... 38

**Part 2 – Mycobacterial Genetic Engineering** ..... 41

2.1 Functional Annotation of The Mycobacterial Genome ..... 41

2.2 Introduction to Genetic Engineering in Mycobacteria ..... 41

2.3 Traditional Methods for Mycobacterial Genetic Engineering ..... 42

2.4 Specialised Transduction ..... 47

2.5 Oligonucleotide-Mediated Recombineering followed by Bxb1 Integrase Targeting (ORBIT) ..... 50

2.6 Clustered Regularly Interspaced Short Palindromic Repeats (CRISPR) ..... 50

2.6.1 CRISPR Interference and Activation ..... 54

2.7 Random Transposon Mutagenesis ..... 57

2.8 Introduction to Functional Genomics ..... 60

2.9 Transposon Site Hybridisation ..... 62

2.10 Transposon Mutagenesis with Next Generation Sequencing ..... 64

2.11 Functional Genomics for Essential Gene Mutant Libraries ..... 68

2.12 Functional Genomics Using Arrayed Libraries ..... 72

**Part 3 – Conclusion** ..... 76

**Part 4 – Aims** ..... 78

**Chapter 2 – Generating an Ordered *M. abscessus* Transposon Library** ..... 79

**Part 1 – Introduction** ..... 80

1.1 Introduction to Ordered Libraries ..... 80

1.2 Ordered Transposon Insertion Libraries ..... 81

1.3 Library Deconvolution ..... 83

1.4 Research Aims ..... 86

**Part 2 – Results** ..... 88

2.1 Selecting an Efficient Transposon Delivery System ..... 88

2.2 Optimisation of Phage Propagation ..... 89

2.3 *M. abscessus* Strain Selection ..... 90

2.4 Mycobacteriophage Selection ..... 93

2.5 Verification of Transposon Insertion ..... 93

2.6 Optimising High-Throughput Robotics for Library Construction ..... 95

2.7 Manual Library Arraying ..... 98

2.8 Cartesian Pooling for Coordinate Sequencing ..... 100

2.9 Validating the Sequencing Method ..... 101

**Part 3 – Discussion** ..... 104

3.1 Completion of the *M. abscessus* Library ..... 104

3.2 Considerations for Library Generation ..... 105

3.2.1 Library Size ..... 105

3.2.2 Phage Delivery System ..... 106

3.2.3 Strain Selection ..... 108

3.2.4 Validating Transposon Insertion Mutagenesis ..... 109

3.2.5 Colony Picking via High-Throughput Robotics ..... 110

3.2.6 Validation of Library Growth ..... 110

3.2.7 Media Selection ..... 111

3.2.8 Method for Library Deconvolution ..... 113

3.3 Limitations of Transposon Libraries ..... 114

3.4 Applications of the *M. abscessus* Mutant Library ..... 115

3.4.1 Collection of Mutants ..... 115

3.4.2 Functional Genomics ..... 116

**Part 4 – Conclusion** ..... 118

**Chapter 3 – A Functional Genomic Screening of *Mycobacterium bovis* BCG** ..... 120

**Part 1 – Introduction** ..... 121

1.1 Functional Genomics using Arrayed Mutant Libraries ..... 121

1.2 IRIS is a High Throughput Tool for Diverse Analysis of Colony Growth ..... 122

1.3 ChemGAPP ..... 124

1.4 Research Aims ..... 126

**Part 2 – Results** ..... 127

2.1 Condensing and Replicating the *M. bovis* BCG Transposon Insertion Library ..... 127

2.1.1 Selection of the Condensed *M. bovis* BCG Library ..... 128

2.1.2 Condensing the *M. bovis* BCG library ..... 130

2.2 Optimisation of Chemical-Genomic Screening in *Mycobacterium bovis* BCG ..... 136

2.2.1 Establishing the Screening Method ..... 137

2.2.2 MIC Testing of Screening Conditions ..... 140

2.4 Conducting a Pilot Test Screen ..... 140

**Part 3 – Discussion** ..... 149

3.1 Completion of the Chemical-Genomic Screening ..... 149

3.2 Considerations of the Optimised Chemical-Genomic Screening Method ..... 150

3.3 A Pilot Chemical-Genetic Screen Identifies Significant Hits in Previously Uncharacterised *M. bovis* Genes ..... 153

**Part 4 – Conclusion** ..... 159

**Chapter 4 – The Characterisation of Mycobacterial Genes Involved in Cell Wall Maintenance** ..... 161

**Part 1 – Introduction** ..... 162

1.1 The Mycobacterial Cell Envelope ..... 162

1.2 Mycobacterial β-lactamases ..... 162

1.3 AG, LM and LAM Hydrolases ..... 164

1.4 The ORBIT Method for Generating Mycobacterial Mutants ..... 166

1.5 Research Aims ..... 169

**Part 2 – Results** ..... 171

2.1 Optimising the ORBIT Method for the Generation of Mycobacterial Cell Wall Mutants ..... 171

2.1.1 Applying the ORBIT Method to *M. marinum* ..... 171

2.1.2 Validation and Optimisation of ORBIT in *M. smegmatis* ..... 173

2.1.3 Resolving ORBIT Mutants via Curing of the Payload Plasmid ..... 180

2.2 Phenotypic Characterisation of ORBIT Mutants ..... 182

2.2.1 Assaying the Growth of ORBIT Mutants ..... 182

2.2.2 Deletion of *lpqF* Reveals Enhanced Growth in Meropenem ..... 184

2.2.3 Deletion of *lpqF* Results in Altered Cell Size and Chaining ..... 193

2.2.4 Deletion of Cell Wall Glycosyl Hydrolase Family Genes Results in Reduced Cell Size ..... 197

**Part 3 – Discussion** ..... 198

3.1 Validation of the ORBIT Method ..... 199

3.2 LpqF is a D,D-carboxypeptidase ..... 203

3.3 *MSMEG_0740* is an α-mannanase ..... 207

**Part 4 – Conclusion** ..... 211

**Materials and Methods** ..... **Error! Bookmark not defined.**

1.1 Bacterial Strains, Bacteriophage and Media ..... 214

1.2 Propagation and Titration of Bacteriophage ..... 217

1.3 Resazurin Assay ..... 218

1.4 Pooled Transposon Mutagenesis Library Generation ..... 219

1.5 PCR ..... 220

1.6 Robotic Cherry-Picking ..... 222

1.7 Manual *M. abscessus* mutant library arraying ..... 222

1.8 *M. abscessus* library replication ..... 223

1.9 Cartesian Pooling ..... 223

1.10 Genomic DNA Extraction ..... 224

1.11 Coordinate Sequencing ..... 224

1.12 *M. bovis* BCG Library Condensing and Replication ..... 225

1.13 Chemical-Genetic Screening ..... 226

1.14.1 ORBIT Plasmids ..... 227

1.14.2 Oligonucleotides ..... 228

1.14.4 Markerless Gene Deletions ..... 232

1.15 Growth Assay ..... 232

1.15.1 Spot Assay ..... 232

1.15.2 Growth Curve ..... 233

1.16 Microscopy ..... 233

**General Discussion** ..... **Error! Bookmark not defined.**

6.1 Introduction ..... 235

6.2 Constructing an Ordered *M. abscessus* Transposon Insertion Library ..... 237

6.3 Chemical-Genomic Screening of *M. bovis* BCG ..... 239

6.4 Validation of Chemical-Genomic Hits and Functional Characterisation of Cell Wall Associated Enzymes ..... 241

6.5 Conclusion ..... 243

**List of References** ..... 244

**Appendix** ..... 275

## Table of Figures

Figure 1.1 The structure of the mycobacterial cell envelope. ..... 18

Figure 1.2 The biosynthesis pathway for PG. ..... 22

Figure 1.3 The biosynthesis pathway for AG. ..... 27

Figure 1.4 LM, LAM and ManLAM biosynthesis. ..... 29

Figure 1.5 The biosynthesis of mycolic acids via the FAS-I and FAS-II pathways. .. 32

Figure 1.6 Worldwide TB incidence. ..... 35

Figure 1.7 Methods for generating unmarked mutants. ..... 46

Figure 1.8 Mycobacterial recombineering. ..... 49

Figure 1.9 The CRISPR system. ..... 52

Figure 1.10 The TraSH method. ..... 63

Figure 1.11 Dub-Seq Method. ..... 71

Figure 1.12 CP-CSeq method. ..... 85

Figure 1.13 Knockout Sudoku Method. ..... 86

Figure 1.14 Zeocin™ MIC. ..... 92

Figure 1.15 Confirmation of *M. abscessus* mutants via PCR and gel electrophoresis. ..... 94

Figure 1.16 BM3-BC colony picking selection. ..... 97

Figure 1.17 Verification of *M. abscessus* mutant growth differences. ..... 99

Figure 1.18 Z master plate pooling method. ..... 101

Figure 1.19 Method for sequencing preparation. ..... 103

Figure 1.20 Selection of the condensed *M. bovis* BCG library. ..... 130

Figure 1.21 *M. bovis* BCG growth in enriched 7H9 medium. ..... 131

Figure 1.22 *M. bovis* BCG library condensing and replication method. ..... 134

Figure 1.23 Amphotericin B MIC ..... 136

Figure 1.24 Optimisation of chemical genomic screening. ..... 139

Figure 1.25 Pilot screen failed growth plates. ..... 142

Figure 1.26 Clustering analysis of mutant fitness S-scores calculated via ChemGAPP. ..... 144

Figure 1.27 Clusters of mutants with significantly altered fitness. ..... 148

Figure 1.28 ORBIT methodology for generating a targeted gene deletion. ..... 168

Figure 1.29 PCR verification of ORBIT transformants. ..... 174

Figure 1.30 PCR verification of ORBIT transformants. ..... 175

Figure 1.31 PCR verification of a Δ*MSMEG_4352* ORBIT mutant. ..... 177

Figure 1.32 PCR verification of a Δ*MSMEG_0740* ORBIT mutant. ..... 178

Figure 1.33 PCR verification Δ*MSMEG_6255 and* Δ*MSMEG_1130* ORBIT mutants. ..... 180

Figure 1.34 PCR verification of Δ*MSMEG_4352* and Δ*lpqF* ORBIT cured mutants. ..... 181

Figure 1.35 Spot assay of *M. smegmatis* ORBIT mutants ..... 183

Figure 1.36 Growth curves of *M. smegmatis* ORBIT mutants in 7H9 medium. ..... 184

Figure 1.37 A chemical genetic screening of *M. smegmatis* ORBIT mutants. ..... 186

Figure 1.38 Chemical-genetic screening of ORBIT mutants. ..... 188

Figure 1.39 Chemical-genetic screening of LpqF. ..... 190

Figure 1.40 Growth curve of Δ*lpqF* under chemical stress. ..... 191

Figure 1.41 Chemical genetic screening of *M. marinum* Δ*lpqF*. ..... 192

Figure 1.42 Microscopy data of *M. smegmatis* WT and Δ*lpqF*. ..... 195

Figure 1.43 Microscopy data of *M. marinum* WT and Δ*lpqF*. ..... 196

Figure 1.44 Analysis of ORBIT mutants via microscopy. ..... 198

## Table of Tables

Table 1.1 Targeting oligonucleotide concentrations. ..... 176

Table 1.2 The bacterial strains used in our study. ..... 214

Table 1.3 Minimal growth media. ..... 215

Table 1.4 The bacteriophages used in our study. ..... 217

Table 1.5 List of primers used for PCR. ..... 220

Table 1.6 Plasmids required for ORBIT. ..... 227

Table 1.7 ORBIT oligonucleotide sequences ..... 228

# Chapter 1

# Introduction

# Part 1 – Mycobacteria

## 1.1 An Introduction to Mycobacteria

*Mycobacterium* is a genus of bacteria closely related to *Nocardia, Rhodococcus* and *Corynebacterium,* which are all from the Phylum Actinobacteria (Goodfellow et al., 2012). This genus consists of 192 different species (Parte et al., 2020), all of which are non-motile, aerobic and non-spore forming bacilli; although it has been discovered that certain species can survive in limited oxygen environments (Lewis and Falkinham, 2015). Mycobacteria either exhibit a smooth or rough colony formation and can be classified as slowly or rapidly growing (Simner et al., 2015). However, even those considered to be rapidly growing have much longer reproductive cycles than many other bacteria. Mycobacteria can also be categorised by pathogenicity. The *Mycobacterium tuberculosis* complex encompasses species that can cause tuberculosis (TB), *Mycobacterium leprae* causes leprosy and nontuberculous mycobacteria (NTM) comprises all remaining mycobacteria (Ustinova et al., 2019).

Discovered in 1896 by Lehmann and Neumann, the name mycobacteria, meaning ‘fungus bacteria’, stems from their mould-like appearance when grown in laboratory conditions. This characteristic is a result of the thick, hydrophobic, waxy and lipid-rich mycobacterial cell envelope consisting of an inner cytoplasmic membrane, a cell wall core, an outer membrane and a capsule (figure 1.1). Due to the hydrophobic nature of the mycobacterial envelope these bacteria also tend to grow as aggregates. However,

detergents can be combined with the growth media to help separate them (Power and Hanks, 1965). Furthermore, this thick envelope prevents various agents from penetrating the cell and therefore mycobacteria cannot retain the crystal violet dye for Gram-staining and are thus considered Gram-intermediate. In addition, they are also regarded as acid-fast and are relatively resistant to antibiotics (Zhou et al., 2020).

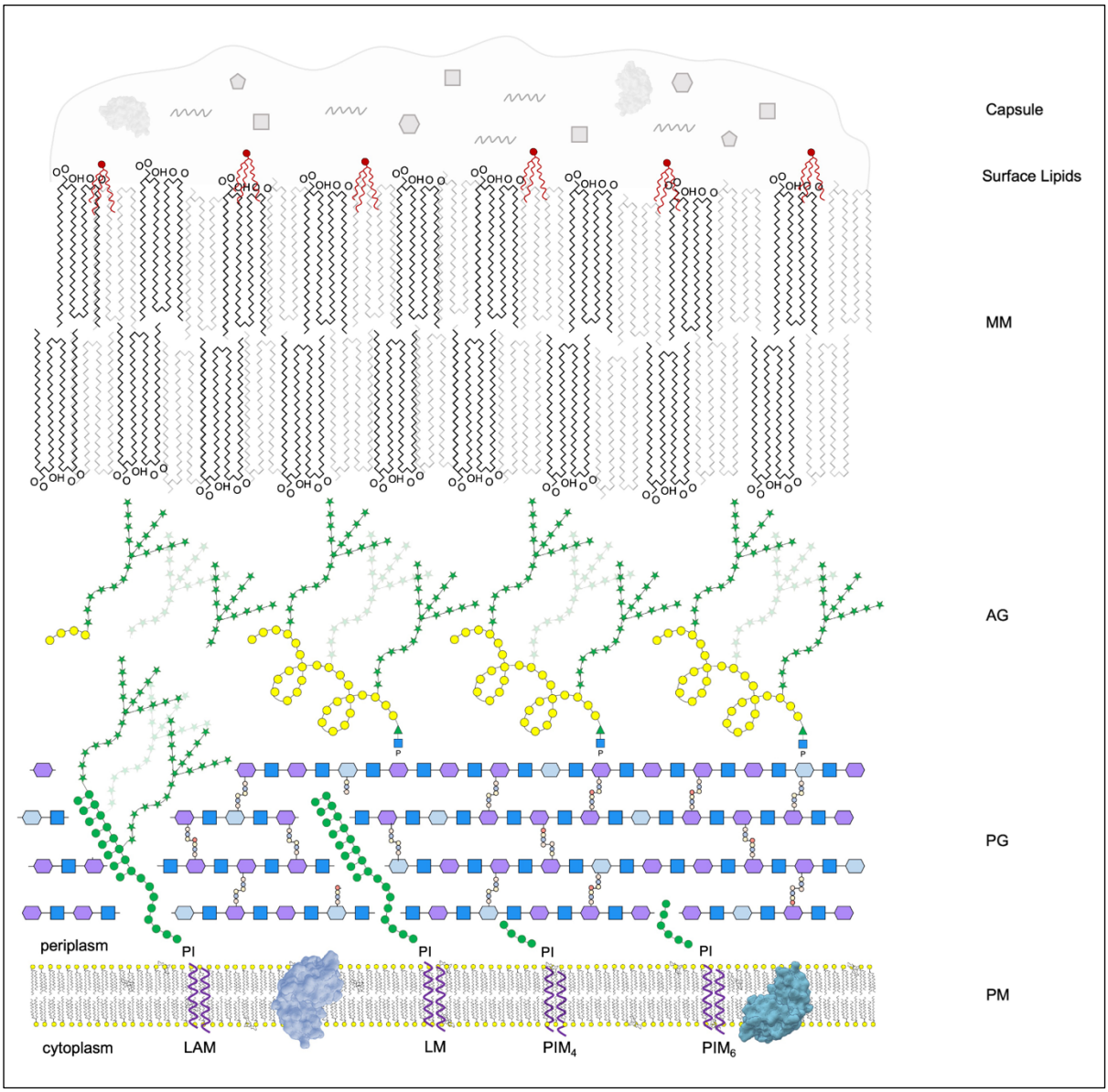



**Figure 1.1 The structure of the mycobacterial cell envelope.**

AG = arabinogalactan, LAM = lipoarabinomannan, LM = lipomannan, MM = mycomembrane, PG = peptidoglycan, PI = phosphatidylinositol, PIM = phosphatidylinositol mannosides, PM = plasma membrane (Dulberger et al., 2020).

### 1.1.1 The Cytoplasmic Membrane

The mycobacterial cytoplasmic membrane has a fundamental structure typical of other bacterial cellular membranes, consisting of a phospholipid bilayer. This membrane has various functions. It acts as a barrier to protect the bacteria from the external environment, although it has selective permeability, vital for nutrient uptake, maintaining biomolecular gradients and cellular export. The cytoplasmic membrane also plays an important role in biosynthesis, including the production of the cell wall core components. In addition, embedded within the membrane are various standard membrane lipids, proteins and lipoproteins which are essential for the respiratory chain and cell signalling (Karakousis et al., 2004). Unlike most other bacterial cytoplasmic membranes, mycobacteria express phosphatidylinositol mannosides (PIMs), lipomannans (LMs) and lipoarabinomannans (LAMs) which extend into the cell wall (Gilleron et al., 1999). PIMs, the precursor for LMs and LAMs, are extremely variable in structure (Sancho-Vaello et al., 2017). They consist of a phosphatidyl inositol (PI) base, anchored into the membrane by up to four fatty acid chains, linked with between one and six mannose residues (Ballou et al., 1963; Crellin et al., 2013). PIMs constitute around 35% and 55% of all phospholipids within the cytoplasmic membrane and the cell wall respectively with $PIM_2$ and $PIM_6$ representing the most abundant forms (Goren, 1984). These forms are generated through a series of acylation and glycosylation reactions assisted by the enzymes PimA-E and Rv2611

(Guerin et al., 2009). LMs consist of α-1,6 linked mannose chains bound to PIMs. LAMs are branches of D-arabinofuranose, capped with an additional sugar such as mannose in *M. tuberculosis* and inositol in *M. smegmatis*, attached to a LM base (Bansal-Mutalik and Nikaido, 2014). It has been discovered that these compounds provide cell wall integrity, help to strengthen the cell wall permeability barrier (Fukuda et al., 2013) and are involved in the manipulation of host immune systems (Gilleron et al., 2003; Dao et al., 2004).

### 1.1.2 The Cell Wall Core and The Outer Membrane

The mycobacterial cell wall core functions as a further barrier, protecting the cell from harmful material. It also plays a key role in cell shape and structure, cell growth, virulence and communication. Thus, given its importance, most of the cell wall core components are essential. The cell wall core is comprised of covalently bound peptidoglycan (PG), arabinogalactan (AG) and mycolic acids (MA), collectively known as the MA-AG-PG complex. These MAs, along with various surface lipids, also constitute the outer membrane, known as the mycomembrane (Dulberger et al., 2020).

### 1.1.3 Peptidoglycan

PG consists of a β-1,4-glycosidically linked *N*-acetylglucosamine (GlcNAc) and *N*-acetylmuramic acid (MurNAc) sugar backbone crosslinked with short peptide chains. In mycobacteria, PG has unusual modifications. A proportion of MurNAc can be further oxidised to form *N*-glycolylmuramic acid (*N*-MurNGlyc), suggested to improve the strength of the cell wall through increased hydrogen-bonding (Raymond et al., 2005).

In addition, in Actinobacteria, the third amino-acid in the peptide chain is variable. This amino-acid is usually *meso*-DAP in Gram-negative bacteria and mycobacteria and L-lysine in Gram-positive (Basavannacharya et al., 2010).

Although PG is an extracellular macromolecule, the initial stages of PG biosynthesis occur in the cytoplasm (figure 1.2). Firstly, the GlcNAc monomer is manufactured with the assistance of Glm operon proteins and is then capped with uridine diphosphate (UDP) (Barreteau et al., 2008). Enzymes MurA and MurB then convert a proportion of UDP-GlcNAc to UDP-MurNAc by adding a D-lactyl acid group (Benson et al., 1993; Marquardt et al., 1993). It has been discovered that mycobacteria only express a low level of MurA (Xu et al., 2014). Therefore, it has been hypothesised that MurA is either under strict regulation or its activity is assisted by other enzymes (Xu et al., 2014; Boutte et al., 2016). UDP-MurNAc can be further converted to UDP-MurNGlyc via the hydroxylase NamH (Raymond et al., 2005). Next, the lactyl acid group is phosphorylated and through the action of a series of enzymes (MurC, MurD, MurE and MurF) the phosphate group is removed and replaced with 5 amino acids (L-alanine, D-glutamic acid, *meso*-diaminopimelic acid, D-alanine and D-alanine) thus forming a UDP-MurNAc pentapeptide chain known as the park nucleotide (El Zoeiby et al., 2003). To ensure transport across the membrane, with the assistance of MraY, the park nucleotide is bound to a lipid carrier undecaprenol pyrophosphate to form lipid I, and the UDP nucleotide is displaced (Chung et al., 2013). Likewise, catalysed by MurG, the UDP nucleotide is removed from UDP-GlcNAc to form a β-1,4-glycosidic bond between GlcNAc and the MurNAc residue of lipid I to form lipid II (Mengin-Lecreulx et al., 1991). Finally, a flippase transports lipid II across the membrane and penicillin binding proteins (PBPs) help to polymerise lipid II through a process known

as glycosyltransfer before being cross-linked into the existing cell wall via transpeptidation (Sauvage et al., 2008).

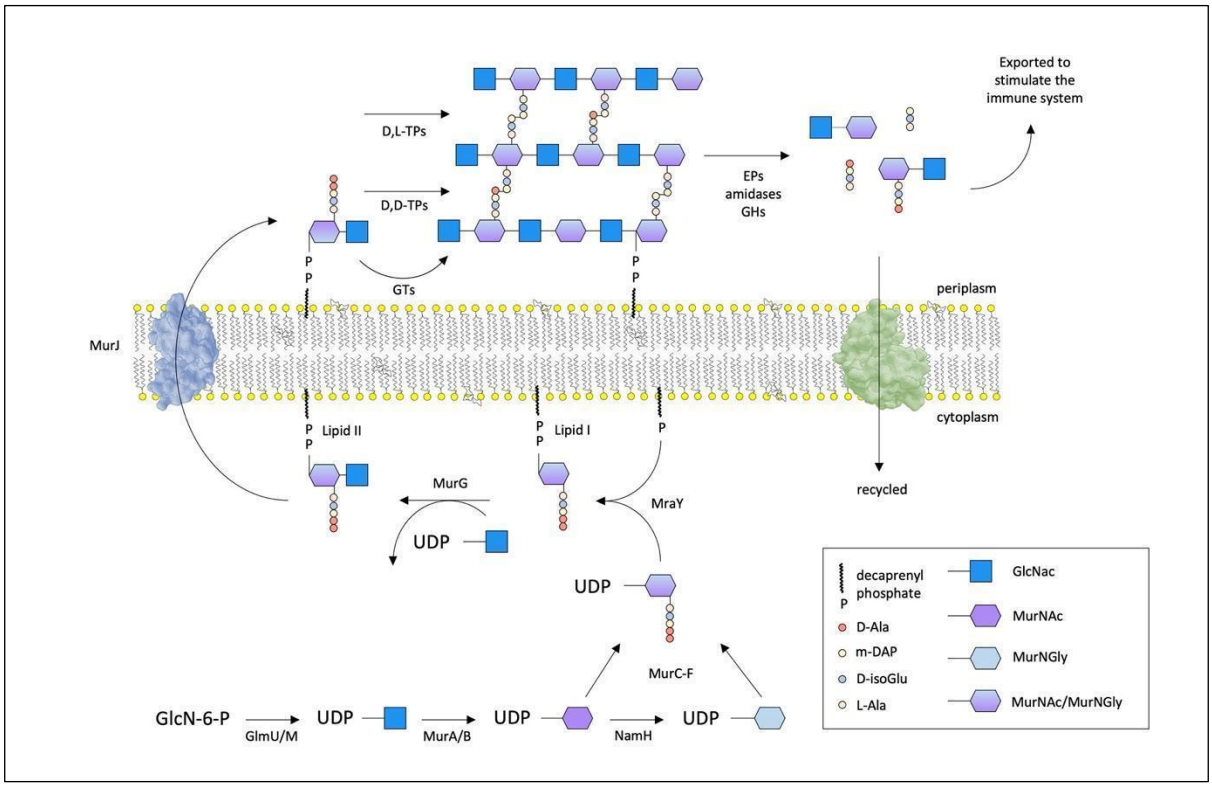



**Figure 1.2 The biosynthesis pathway for PG.**

In the cytoplasm, with the assistance of Glm operon proteins, the GlcNAc monomer is generated and capped with UDP. MurA and MurB assist in the conversion of a proportion of UDP-GlcNAc to UDP-MurNAc and, through the action of NamH, UDP-MurNAc can be further converted to UDP-MurNGlyc. Both UDP-MurNAc and UDP-MurNGlyc enter a series of enzymatic steps involving MurC-F proteins to generate the pentapeptide chains involved in PG crosslinking resulting in the formation the park nucleotide. To export the PG precursor to the periplasm, with the help of MraY, the park nucleotide is bound to a lipid carrier and lipid I is generated. Next, with the assistance of MurG, a β-1,4-glycosidic bond is formed between previously generated UDP-GlcNAc and the MurNAc residue of lipid I to form lipid II. This molecule can then be transported across the membrane via the MurJ flippase. Finally, in the periplasm, via glycosyltransfer, PBPs polymerise lipid II and crosslink the polymers into the existing cell wall via transpeptidation. GTs = glycosyltransferases, D,D-TPs = D,D-transpeptidases, D,L-TPs = D,L-transpeptidases, EPs = endopeptidases, GH = glycoside hydrolases (Abrahams and Besra, 2018).

PBPs can be monofunctional or bifunctional. Monofunctional PBPs have transpeptidase, endopeptidase and carboxypeptidase activity. D,D-carboxypeptidases cleave the terminal amino acid residues of the peptide chain, creating the substrate for D,D-transpeptidases which can then crosslink PG between the 3rd and 4th amino acid residues (Goffin and Ghuysen, 1998). Endopeptidases cleave non-terminal amino acid residues (Squeglia et al., 2019). However, in mycobacteria, the final two positions on both peptide chains are also often cleaved and a crosslink is formed between the 3rd amino acid residues instead (Lavollay et al., 2008). This reaction is catalysed by L,D-transpeptidases such as $Ldt_{Mt2}$ (Gupta et al., 2010). These enzymes are structurally distinct from PBPs and harbour a different active site conformation (Mainardi et al., 2005). The bifunctional PBPs, PonA1 and PonA2, contain a transglycosylase domain that assists in the incorporation of Lipid II into pre-existing glycan chains as well as a transpeptidase domain (Billman-Jacobe et al., 1999).

The serine/threonine protein kinase (STPK) PknB is the main regulator of PG metabolism, though its exact action is disputed (Dulberger et al., 2020). Boutte et al. (2016) initially suggested that PknB activates PG synthesis through the phosphorylation of CwlM which then binds and activates MurA. In contrast, Turapov et al., (2018) recently found that PknB detects PG crosslinking via its PASTA domain, stimulating the phosphorylation of CwlM and MurJ, a proposed Lipid II flippase. It was suggested that CwlM and MurJ then interact with FhaA and, through a mechanism that is yet to be fully established, FhaA controls PG biosynthesis. In this model, the PG biosynthesis pathway is activated by the interaction of non-phosphorylated CwlM with MurA.

Due to a lack of recycling gene homologues and the inability of mycobacteria to grow using PG as a nutrient source, the recycling of the mycobacterial cell wall components has long been under question (Abrahams and Besra, 2021). It is well known that resuscitation promoting factors (Rpfs) are able to cleave the MurNAc-GlcNAc β-1,4 glycosidic bonds and thus cleave PG (Hett et al., 2008; Van Wyk et al., 2017). However, only recently was the first evidence of PG recycling discovered: Fullam et al. (2016) found that, in *M. smegmatis*, UspABC was an essential amino-sugar transporter and suggested that de-acetylated and hydrolysed PG could be a potential substrate for the recycling of the cell wall due to the structural similarities to the confirmed UspABC binding substrate, chitobiose. Following this discovery, further evidence for PG recycling was discovered through the analysis of a homologue of the *Escherichia coli* enzyme NagA, found to catalyse the breakdown of PG and thus play a role in recycling of the cell wall. It was found that NagA also fulfilled the same role in *M. tuberculosis* (Ahangar et al., 2018). Most recently, Moynihan et al. (2019) characterised LpqI, a homologue of *nagZ*, a β-*N*-acetylglucosaminidase, discovering that this enzyme was able to break the glycosidic bond between GlcNAc-MurNAc in *M. tuberculosis* and *Mycobacterium bovis* BCG and demonstrated that these bacteria do not recycle GlcNAc which contrasts with *M. smegmatis.* Furthermore, a lysozyme homologue, Rv2525, has also been identified in *Mycobacterium tuberculosis*. However, the PG hydrolysis activity of this enzyme is yet to be established (Van Wyk et al., 2017).

### 1.1.4 Arabinogalactan

AG is a structure unique to Actinomycetes. It is a polysaccharide composed of a linear chain of galactan consisting of around 30 residues (Daffe, Brennan and McNeil, 1990), bound to three highly branched chains of arabinan, each chain also comprised of around 30 residues (Besra et al. 1995). The galactan and arabinan residues are found in the furanose ring formation, existing as Gal*f* and Ara*f* respectively. AG is covalently attached to PG via a linker unit (LU) called the rhamnose-GlcNAc linker disaccharide. The terminus of the D-arabinan chains are covalently attached to mycolic acids (MAs) fixing them to the cell wall complex (McNeil et al., 1991).

As detailed in figure 1.3, like PG, AG is initially synthesised in the cytoplasm. Firstly, through the action of the enzymes WecA and WbbL, the LU is generated (Mills et al., 2004; Jin et al., 2010). A rhamnose domain of the LU provides a binding site for Gal*f* sugar residues and, with the assistance of galactofuranolsyltransferase enzymes, GlfT1 and GlfT2, these sugar residues are polymerised to form galactan. GlfT1 is responsible for the transfer of the first two Gal*f* residues onto the LU (Alderwick et al., 2008). GlfT2 further polymerises the increasing galactan chain by adding successive Gal*f* residues with interchanging β-1,5 and β-1,6 glycosidic linkages (Rose et al. 2006). As the remainder of AG biosynthesis occurs extracellularly, it is supposed that a flippase transports this polymer across the membrane. The exact mechanism is unknown though it has been suggested that the ABC transporters, Rv3781 and Rv3783 transfer this AG precursor to the outside of the cell (Dianiskova et al., 2011). In the periplasm, D-arabinofurosyl residues from a lipid donor, decaprenylphosphoryl-D-arabinose (DPA), are added onto the galactan chain to form branched arabinose chains (Wolucka et al., 1994). DPA is generated intracellularly but it's mechanism for transport into the periplasm is unknown. It is hypothesised that a combination of six

arabinofuranosyltransferases aid in the generation of the α-1,5 linked arabinan chains. AftA catalyses the introduction of the first Ara*f* residue onto the galactan chain (Alderwick et al., 2006). EmbA and EmbB then facilitate the polymerisation of the arabanin chain and AftC-D initiate branching (Alderwick et al., 2005; Škovierova et al., 2009). By adding the final Ara*f* residues, AftB terminates the arabinan chain (Seidel et al., 2007). The two terminal residues function as an attachment site for mycolic acids (McNeil et al., 1991). To complete the AG macromolecule succinyl and galactosamine are added to the central units of the arabinan chain (Bhamidi et al., 2008). As previously mentioned, the final step involves the connecting of AG to PG, this is achieved through the assistance of a phosphotransferase, Lcp1 (Grzegorzewicz et al., 2016; Harrison et al., 2016; Baumgart et al., 2016). It was discovered that the generation of AG through this pathway is regulated by PknA and PknB (Sharma et al., 2006). These STPKs phosphorylate EmbR and, in doing so, activate the transcription of *embA-C* (Sharma et al., 2006).

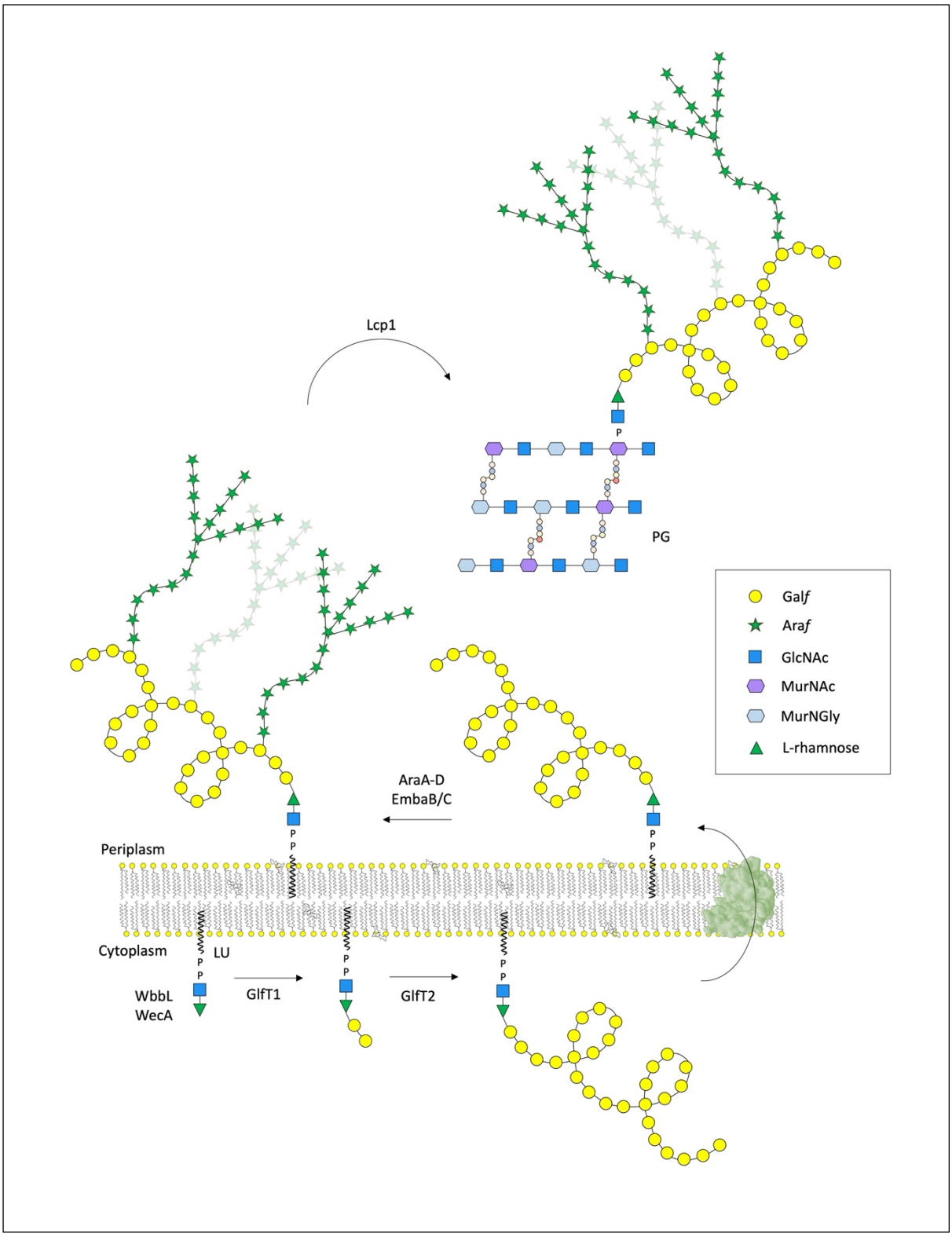



**Figure 1.3 The biosynthesis pathway for AG.**

In the cytoplasm, WecA and WbbL assist in the generation of the linker unit (LU). Extending from the LU, GlfT1 and GlfT2 catalyse the polymerisation of Gal*f* residues to form a galactan chain. By an unknown mechanism, the AG precursor is transported across the membrane. Then, in the periplasm,

a combination of various arabinofuranosyltransferases add Ara*f* residues onto the galactan chain to form AG and lastly, AG is bound to PG via Lcp1.

The biosynthesis of LAM follows a similar pathway to the generation of AG (figure 1.4). It is believed that $PIM_4$ is the PIM form that acts as the precursor for LAMs and LMs. To form LM, mannan chains are extended from the PIM through the action of mannosyltransferases MptA and MptB (Mishra et al., 2008), each chain consisting of between 21-37 α-1,6 linked mannopyranose (Man*p*) residues. MptC then assists in the branching of the mannan chains (Kaur et al., 2008). With DPA as the donor, LAM is produced through addition of between around 50-80 Ara*f* residues. The LM mannan chain is first primed through an unknown mechanism and, in contrast to AG synthesis, EmbC, instead EmbA and EmbB, then catalyses the initial α-1,5 extension of the arabinan chain (Shi et al., 2006). It is believed that AftB terminates the arabinan chain by adding Ara*f* residues via β-1,2 linkages (Jankute et al., 2017). Like AG biosynthesis, it is proposed that AftC and AftD are responsible for branching (Birch et al., 2008; Škovierova et al., 2009). LAM can be further mannosylated to form ManLAM, these α-1,2 Man*p* residue caps are formed through the action of CapA and MptC (Kaur et al., 2008). In addition, through an unknown mechanism, the caps can be further affixed with 5-deoxy-5-methyl-thio-xylofuranose residues (Turnbull et al., 2004).

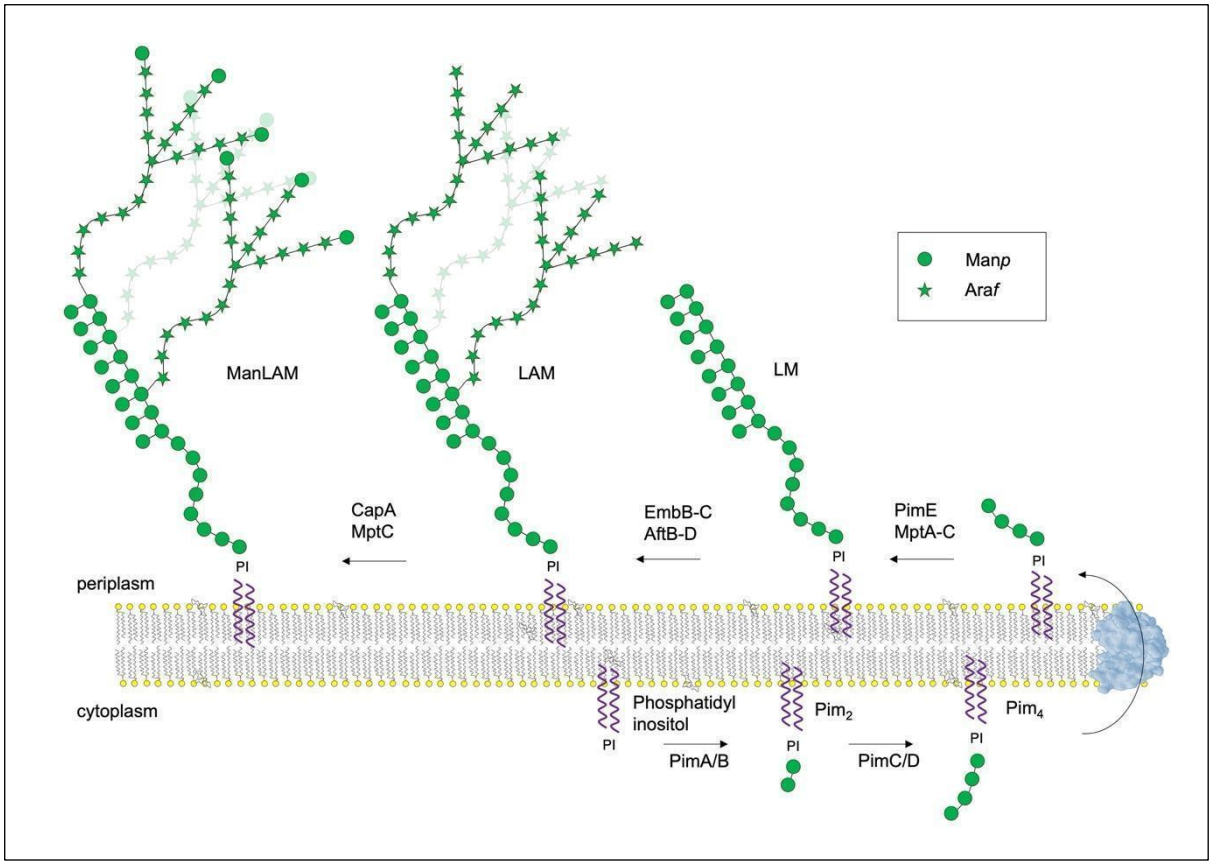



**Figure 1.4 LM, LAM and ManLAM biosynthesis.**

Assisted by PimA-D, mannan chains are added to the PI precursor to generate $Pim_4$. $Pim_4$ is transported to the periplasm where MptA and MptB add additional Man*p* resides to the growing chain and MptC generares branches resulting in the formation of LM. Next, with the assistance of EmbB and EmbC, approximately 50-80 Ara*f* residues are added to LM to form LAM. AftB-D enzymes are thought to be responsible for branching. Lastly, to produce ManLAM, MptC and CapA cap LAM with α-1,2 Man*p* residues.

Over 30 years ago, both arabinomannans and mannans were identified in the mycobacterial capsule thus demonstrating that cells are capable of cleaving AG, LM and LAM (Lemassu and Daffe, 1994). In addition, following on from this research, Rivera-Marrero et al. (2001) discovered a *M. tuberculosis* protein with α-mannosidase activity, Rv0648. However, in the past 20 years, there has been very little progress identifying the genes responsible for AG, LM and LAM cleavage and whilst there have

been some exciting recent discoveries covered in more detail in Chapter 4 Section 1.3, the mechanisms for these activities largely remain a mystery.

### 1.1.5 Mycomembrane

The mycomembrane, also known as the mycobacterial outer membrane, is characterised by an inner leaflet of long chain MAs covalently bound to AG and an outer leaflet of free lipids including MAs either existing as free mycolates or attached to trehalose monomycolates (TMMs) and trehalose dimycolates (TDMs) (Marrakchi et al., 2014). Other free lipids include phthiocerol dimycocerosates (PDIMs), acyl trehaloses, phenolic glycolipids (PGLs), lipooligosaccharides (LOSs) and glycopeptidolipids (GPLs), among others (Minnikin and Brennan, 2020). These lipids fulfil a wide range of functions such as biofilm formation, stability, immune modulation and virulence. Due to the lipid-rich nature of the mycomembrane it is extremely hydrophobic and acts as an efficient permeability barrier (Villeneuve et al., 2005).

MAs are α-alkyl-β-hydroxy fatty acids which are around 70-90 carbons in length. The general structure of MAs consists of meromycolate - a meroaldehyde chain with a length of up to 56 carbons, and an α-branch - a much shorter saturated alkyl chain with between 24 and 26 carbons (Marrakchi et al., 2014). MA biosynthesis along with the synthesis of various other mycobacterial fatty acids is conducted by fatty acid synthases (FAS) (Peterson and Bloch, 1977). FAS-I is a polypeptide able to produce fatty acids of two different chain lengths (Fernandes and Kolattukudy, 1996). To generate MAs, fatty acids are transferred to FAS-II where the chains are further extended and functional groups are added (Marrakchi et al., 2002). Polyketide

synthase (PKS) enzymes can also utilise FAS-I generated fatty acids to manufacture a range of alternative lipids. The enzymes PknA, PknB, PknD, PknF, PknH and PknL have all been implicated in the regulation of MA synthesis through the control of FAS-II pathway (Dulberger et al., 2020).

As demonstrated in figure 1.5, to commence MA biosynthesis, through a succession of enzymatic steps at the seven domains of the FAS-I protein (an acyl transferase (AT), an acyl carrier protein (ACP), a β-ketoacyl synthase (KS), a β-keto reductase (KR), a dehydratase (DH), an enoyl reductase (ER), and a malonyl/palmitoyl transferase (MPT)) fatty acid chains of 24-26 carbons and 16-18 carbons are generated from malonyl-CoA (Fernandes and Kolattukudy, 1996). These fatty acids are the precursors for the α- and meromycolate-chains; the long chain is carboxylated to form the α-chain and the shorter chain enters the FAS-II enzymatic cycle to form meromycolate (Marakchi et al., 2014). With the assistance of FabD, malonyl-CoA is first converted to malonyl-AcpM, required for the FAS-II pathway (Kremer et al., 2001). FabH, a β-ketoacyl ACP synthase, then assists in the production of β-ketoacyl-AcpM from the shorter FAS-I fatty acid chain and, through a series of keto-reduction, dehydration and enoyl- reduction, catalysed by MabA, HadAB/BC, InhA and KasA/B, meromycolate is formed (Musayev et al., 2005; Sacco et al., 2007). Meromycolate is further modified by the introduction of methoxy and keto groups and cis/trans-cyclopropanation (Batt, Minnikin and Besra, 2020). Then, assisted by FabD32, meromycolate is ligated to adenosine monophosphate (AMP) to prime the meromycolate for a condensation reaction (Abrahams and Besra, 2018). Pks13 has been identified as the condensase responsible and catalyses the production of α-alkyl-β-keto-MA from carboxyacyl-CoA (Portevin et al., 2004). Lastly, Rv2509 further

reduces α-alkyl-β-keto-MA to form mature mycolate (Bhatt et al. 2008). In the form of TMM, mycolates are transported to the mycomembrane via the flippase MmpL3 (Fay et al., 2019). Via mycolyltransferases of the Antigen 85 complex, MAs are then either attached to AG, combined with TMM to form TDM (Jackson et al., 1999; Takayama, Wang and Besra, 2005).

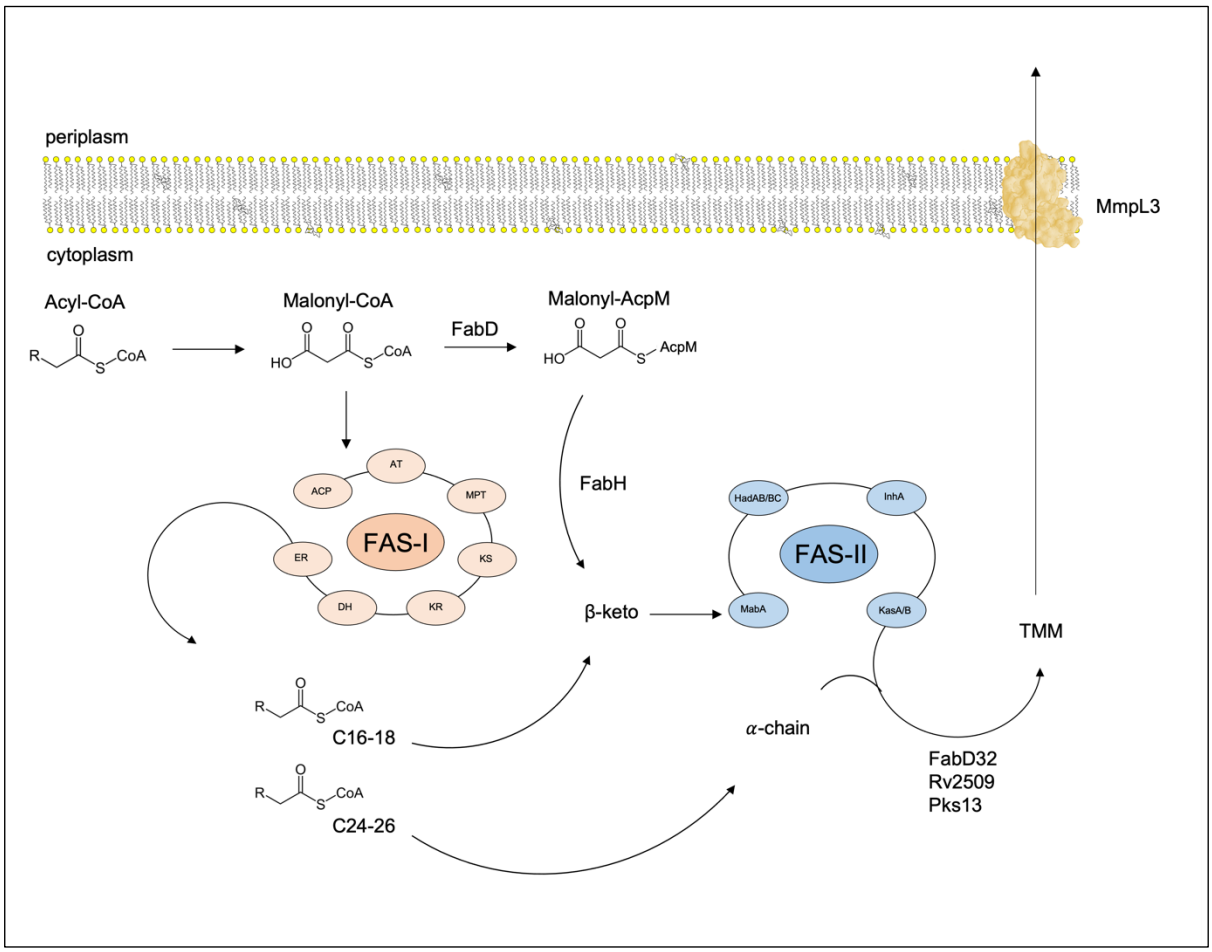



**Figure 1.5 The biosynthesis of mycolic acids via the FAS-I and FAS-II pathways.**

Malonyl-CoA undergoes a series of enzymatic steps at the FAS-I protein to produce fatty acid chains of 24-26 carbons and 16-18 carbons. Via carboxylation, the longer chain is converted into the α-chain whereas the shorter chain enters the FAS-II enzymatic cycle to form meromycolate. Assisted by FabD32 and Pks13, a condensation reaction between the α-chain and the meromycolate results in the production of α-alkyl-β-keto-MA. Lastly, Rv2509 reduces α-alkyl-β-keto-MA to form mature mycolate and a flippase, Mmpl3, transports the mature mycolates to the mycomembrane in the form of TMM.

Whilst there is evidence for the recycling of mycomembrane components including MAs, there are still gaps within our knowledge. For example, it is well documented that trehalose is recycled via the LpqY-SugA-SugB-SugC ABC transporter and although a TDM hydrolase has been discovered, there is no evidence for the recycling of MAs that are bound to AG and there remains little evidence for the existence of TMM hydrolases (Kalscheuer et al., 2010; Yang et al., 2014). Furthermore, an ABC-binding cassette transporter, Mce1, has also been implicated in the mycobacterial uptake of fatty acids. It was discovered that *M. tuberculosis* strains devoid of *mce1* lead to an accumulation of free mycolic acids and the upregulation of the fatty acid biosynthesis pathway (Cantrell et al., 2013; Queiroz et al., 2015). Recycled fatty acids can then be introduced into the β-oxidation pathway to manufacture acetyl-CoA – vital for MA synthesis. It has also been proposed that, through the action of enoyl-CoA hydratases such as EchA6, longer acyl-CoAs could be directly utilised by the FAS-II pathway without the need to for further hydrolysis via β-oxidation (Cox et al., 2016).

### 1.1.6 Capsule

The mycobacterial capsule was first described by Chapman et al. (1959) and was further confirmed via electron microscopy (EM) (Frehel et al., 1988). Nevertheless, its existence has been long under debate as only within the last 15 years was the presence of a capsule confirmed by cryoEM with previous attempts failing to image the capsular layer (Sani et al., 2010). The capsule surrounds the cell and consists of a combination of free polysaccharides and proteins as well as a small proportion of lipids (Daffe and Etienne, 1999). It is hypothesised that LM/LAM may also protrude

into the capsular layer (Ortalo-Magne et al., 1996). The capsule not only provides protection but also fulfils roles in virulence, adhesion and immune modulation (Daffe and Etienne, 1999).

## 1.2 *Mycobacterium tuberculosis*

*M. tuberculosis* is the causative agent of human TB. It is estimated that around one in every four people are infected with the bacterium although this is predominantly a latent infection with only 5-10% of those infected developing active disease (WHO, 2023). That said, the disease is still responsible for a high level of mortality and morbidity: in 2022, with the exception of COVID-19, *M. tuberculosis* was responsible for the highest number of deaths from a single infectious agent and each year causes around 1.5 million deaths (WHO, 2023).

*M. tuberculosis* is transmitted in the air via droplet nuclei and predominantly causes pulmonary disease though the infection can disseminate throughout the body (Moule and Cirillo, 2020). Latent disease is established through the containment of the infection within a mass of immune cells called a granuloma. However, the bacterium has developed many mechanisms to escape this immune response and develop into active disease (Raghuvanshi et al., 2010). The disease can affect all age groups though it is most common in adults (around 90% of cases) and especially in those that are immunocompromised, such as those diagnosed with human immunodeficiency virus (HIV) (WHO, 2023). Furthermore, a greater number of cases are reported in men and the disease is most prevalent in low/middle income countries (figure 1.6).

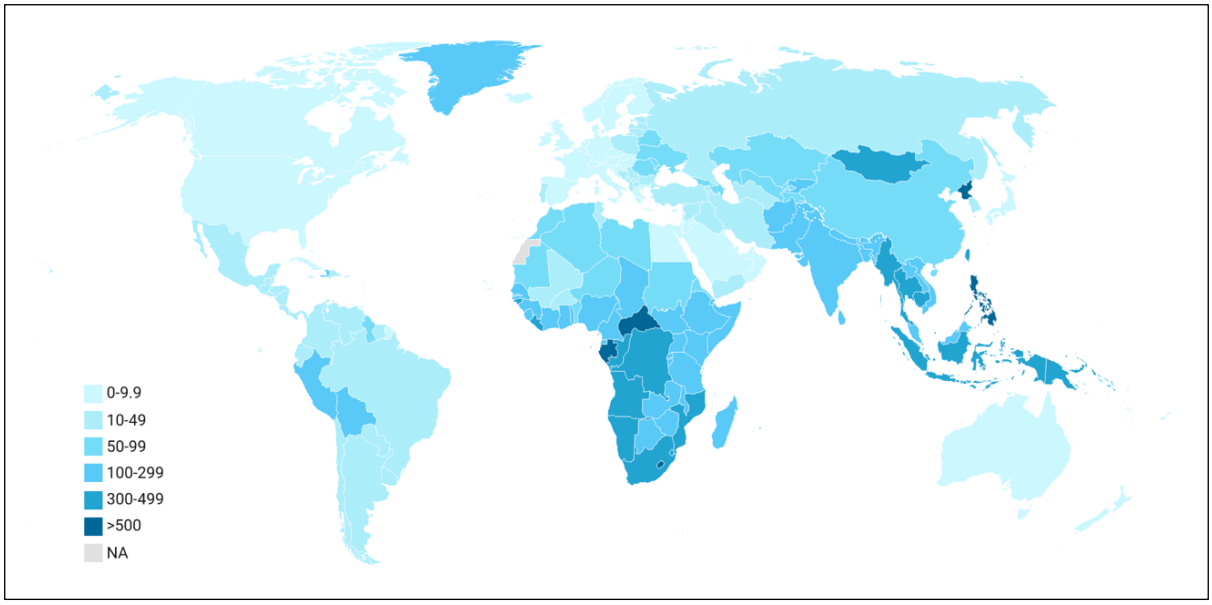



**Figure 1.6 Worldwide TB incidence**.

The worldwide annual TB incidence per 100,000 people interpreted from WHO (2023) data. The majority of cases occur in sub-Saharan Africa and South East Asia (WHO, 2023).

There are multiple TB diagnostic tests, ranging from cell culture and sputum smears to the more recently developed rapid molecular tests (Boehme et al., 2010). In addition, TB is both treatable and preventable: numerous anti-tuberculosis drugs have been discovered and a vaccine is available. However, there is a rise in multi-drug resistant (MDR) *M. tuberculosis* and the current therapeutic programmes are long and intensive with a course of four different antibiotics administered over a period of either four or six months (Wang et al., 2015). The six-month treatment uses a combination of rifampicin, isoniazid, pyrazinamide and ethambutol and is thus known as the RIPE regimen. In contrast, the 4-month treatment uses higher dosing and substitutes ethambutol with a fluoroquinolone such as moxifloxacin. For multi-drug resistant patients, bedaquiline can also be incorporated into the drug regimen (CDC, 2023). In addition, the vaccine, developed from an attenuated strain of *M. bovis*, *M. bovis* bacillus Calmette-Guérin (BCG), exhibits a limited efficiency and is only effective in children (Milstien and Gibson, 1990). For these reasons, it is essential that new TB treatments and vaccines are developed. Furthermore, in 2020, the number of TB diagnoses decreased, and the number of TB deaths increased by almost 15% (WHO, 2021). It is considered that the COVID-19 pandemic has resulted in restricted access to medical facilities and thus both the diagnosis and treatment of TB has declined (Cilloni et al., 2020). These statistics demonstrate the importance of prioritising the development of TB healthcare and maintaining the current infrastructure.

## 1.3 *Mycobacterium bovis*

*M. bovis* is the bacterium responsible for bovine TB which affects mainly cattle, bison and buffalo but can also spread to other animal hosts. *M. bovis* causes a significant

economic burden on the farming industry. In recent years, the incidence of *M. bovis* within agriculture has been on the increase and, due to infection, around 40,000 animals are slaughtered each year in the UK alone (National Statistics, 2021). Furthermore, around 30 years ago, it was estimated that *M. bovis* cost the worldwide economy around $3 billion each year (Steele, 1995). It is likely that this value has now largely increased considering rises in inflation and the expansion of the agricultural industry. Like human TB, *M. bovis* is spread via air in the form of droplets. However, bovine TB can also be spread through percutaneous transmission and ingestion. Due to the geographical separation of herds, the spread of *M. bovis* is largely attributed to wildlife carriers such as possums and badgers. Thus, preventative measures for the spread of bovine TB include the control of wildlife reservoirs, regular testing and the implementation of closed herd farming. In addition, *M. bovis* is also regarded as a zoonotic disease, though human cases are rare and usually not fatal (Hlavsa et al., 2008). Nevertheless, given the recent COVID-19 pandemic, the potential for *M. bovis* to spread to the human host is an alarming prospect.

As *M. bovis* is very closely related to *M. tuberculosis*, exhibiting a 99.95% sequence similarity (Garnier et al., 2003) it was hypothesised that an attenuated strain of *M. bovis* could serve as a TB vaccine. To achieve this, in the early 20$^{th}$ century, Albert Calmette and Camille Guérin passaged *M. bovis* ~ 230 times in ox bile to produce a resulting strain that lacked the Region of Difference-1 (RD1), a locus known to encode well characterised tuberculosis virulence factors such as the the ESX-1 protein secretion system and the ESAT-6 effector protein (Liu et al., 2009). Stemming from the original vaccine strain, coined *M. bovis* Bacille Calmette-Guérin after the original developers of the vaccine, multiple strains of *M. bovis* BCG are now in circulation.

However, due to slight changes in their genetic makeup, these strains vary in immunogenicity and virulence. Given the enhanced induction of host cytokines and T cells, the Danish 1331 strain is now considered as the preferred selection for future vaccine development. This strain has a 4.4 Mbp genome and encodes 4084 genes, as it is attenuated it is considered a Biosafety Level 2 pathogen (Borgers et al., 2019). Therefore, due to its reduced pathogenicity, its similarity to *M. tuberculosis* and its potential for vaccine production, the *M. bovis* BCG Danish strain serves as a useful model organism for *M. tuberculosis.*

## 1.4 Nontuberculosis Mycobacteria

Nontuberculosis mycobacteria (NTM) comprise all of the remaining species of mycobacteria that do not cause leprosy and are not members of the *Mycobacterium tuberculosis* complex and thus cannot cause tuberculosis in any animal or human host. There are more than 170 species of NTM most of which are free-living environmental saprophytes. However, some NTM are opportunistic pathogens and around a third have been associated with human disease (Mirsaeidi et al., 2015). Moreover, in recent years, there have been significant rises in the prevalence of NTM associated diseases and NTM-related mortalities, even in those considered to be immune competent (Ratnatunga et al., 2020). Only a small proportion of NTM are responsible for the majority (~80%) of these infections: the *Mycobacterium avium* complex, *Mycobacterium xenopi,* the *Mycobacterium fortuitum complex, Mycobacterium kansasii* and the *Mycobacterium abscessus* group (Hoefsloot et al., 2013).

*M. abscessus* is an organism of particular interest to the mycobacterial research field due to its increasing incidence in association with lung disease and nosocomial infections. *M. abscessus* infection is frequently linked with bronchiectasis and it has been reported that around 10-20% of cystic fibrosis patients will develop NTM infections, many of which can be attributed to *M. abscessus* (Johansen et al., 2020). In addition, *M. abscessus* has intrinsic resistance to many anti-TB drugs. Therefore, by studying these intrinsic mechanisms of resistance, predictions could be made about the pathways for antibiotic resistance in *M. tuberculosis*. There are 3 subspecies of *M. abscessus, M. abscessus* subsp. *abscessus*, *M. abscessus* subsp. *bolletii* and *M. abscessus* subsp. *massiliense*. These subspecies vary in virulence and drug resistance with *M. a. abscessus* infections being the most frequently reported (Lipworth et al., 2021) and *M. a. massiliense* demonstrating increased susceptibility to macrolides due to an altered version of the *erm*(41) gene (Kim et al., 2010). Furthermore, due to variation in glycopeptide constitution and thus differences in colony appearance, *M. abscessus* has 2 distinct morphotypes, smooth and rough. It has been discovered that rough morphotypes have an increased virulence due to large serpentine cord formation which stimulates an enhanced immune response and cell death (Pawlik et al., 2013). Lastly, the *M. abscessus* type strain, *M. abscessus* ATCC 19977, is comparable to *M. tuberculosis*, sharing around 85% sequence similarity. This strain harbours a larger genome (5.1 Mbp) with a total of 4920 genes but many of these genes are *M. tuberculosis* homologues (Rifat et al., 2021).

Many NTM are also significant animal pathogens, both *Mycobacterium marinum* and *Mycobacterium avium* ssp. *paratuberculosis* (MAP) cause significant economic burden on the farming industry. MAP infects predominantly cattle, sheep and goats

but has been found in other mammalian hosts (Barratt et al., 2018). Due to the reduction in the yield of milk and meat it is estimated that MAP is responsible for an annual loss of around £13 million in the UK alone (Johnes Information Centre, 2018). *M. marinum* is an aquatic organism that infects over 200 species of fresh-water fish, marine fish and amphibians including numerous species of farmed fish (Hashish et al., 2018). As zebrafish are also susceptible to *M. marinum* and as active infection shows similarities to that shown in mammalian mycobacteriosis, including TB, *M. marinum*-zebrafish infection models can be used for *in vivo* studies to determine host-pathogen interactions during mycobacterial infection (Prouty et al., 2003).

In addition, as it has comparatively quick growth and is considered non-pathogenic – requiring only Biosafety Level 1 facilities, *Mycobacterium smegmatis* is another NTM frequently used as a *M. tuberculosis* model organism.

# Part 2 – Mycobacterial Genetic Engineering

## 2.1 Functional Annotation of The Mycobacterial Genome

Due to advances in sequencing technology, there is now a plethora of mycobacterial sequence information available with around 8500 sequenced mycobacterial genomes readily accessible via NIH (National Institutes of Health, 2023). However, functional annotation of the genome is lagging with around 25% of the protein encoding genes within the *M. tuberculosis* genome still lacking functional assignment (Modlin et al., 2021). The characterisation of the remaining genome could help to elucidate pathways involved in antimicrobial resistance, lead to the discovery of new drug targets and identify orphan genes. By comparing gene homology, computational sequence analysis studies have been able to make predictions about genes but these findings are often not supported by experimental evidence. Moreover, predictions cannot be made for every gene and significant proportions of the genome remain uncharacterised. Therefore, there is great demand for new approaches to functionally annotate the remaining genome.

## 2.2 Introduction to Genetic Engineering in Mycobacteria

Genetic engineering is a powerful technique for determining gene function. Genes can be edited to tag proteins, entire gene deletions can be generated through targeted knockouts and gene expression can be either inhibited or upregulated via conditional knockdowns or overexpression. Through reverse genetics, phenotypic responses of genetically engineered strains can be analysed, linking genotype to phenotype.

However, mycobacteria are typically difficult to genetically engineer. They are slow growing, with some species taking weeks to show any evidence of colony growth, especially in particularly sick mutants (Simner et al., 2015). This results in long waiting periods before the efficiency of genetic engineering can be determined. In addition, due to the thick, hydrophobic and lipid-rich mycobacterial cell envelope, the efficacy of DNA exchange is considerably reduced as the delivery of DNA across the envelope, and thus cell transformation, is hindered (Van Kessel and Hatfull, 2007). Furthermore, mycobacteria have reduced homologous recombination (HR) efficiency (Van Kessel and Hatfull, 2008). This is due in part to the reduced reliance on HR in mycobacterial DNA repair. Instead, these bacteria demonstrate a higher dependence on alternative methods such as non-homologous end joining (NHEJ) and single-strand annealing (Gupta et al., 2011), HR efficiency is also reduced due to an increased propensity of mycobacteria to undergo illegitimate recombination events (Kalpana, Bloom and Jacobs, 1991).

## 2.3 Traditional Methods for Mycobacterial Genetic Engineering

The first techniques for introducing DNA into mycobacteria were developed over 30 years ago (Jacobs, Tuckman and Bloom, 1987). This was achieved by creating a phagemid consisting of an *Escherichia coli* cosmid packaged into a mycobacteriophage at a non-essential region. The cosmid contained an *E. coli* origin of replication and thus could replicate within *E. coli* cells. The mycobacteriophage was temperate allowing the phage to be propagated at permissive temperatures, and, at restrictive temperatures, the bacteriophage was lysogenic and thus could be introduced into the host cells and integrated into the genome. The first phage used was TM4 which was known to be lysogenic in *M. avium*. However, this phage was not lysogenic in *M. bovis* BCG and *M. smegmatis* and thus instead a L1 phage was used to increase the sensitivity of the method (Snapper et al., 1988). Similarly, Lee et al. (1991) reported the introduction of DNA via a viral insertion site, though this study used a range of integration proficient vectors to deliver the DNA into the cells. It was discovered that via its *attP* site, the L5 phage could insert at a conserved mycobacterial genomic region termed the *attB* site. Upon integration, the recombination site is split into two sites which flank the insertion, these were coined *attL* and *attR*. By constructing plasmids containing the attP site and the L5 phage integrase, an enzyme that assists in the integration of the viral DNA, *M. smegmatis* was successfully transformed to express a kanamycin resistance cassette. Though these methods were ground-breaking for introducing foreign DNA into mycobacteria, they were unable to introduce genetic alterations directly into gene targets.

Husson et al. (1990) developed the first mycobacterial genetic engineering method for targeting specific genes. By constructing a suicide vector containing a kanamycin resistance cassette with adjacent homologous arms (HA) to a target gene, through

either a single or double HR crossover event, the cassette was inserted into the target, disrupting the gene in the process, and leading to a loss of function. Later, it was hypothesised that linear DNA fragments would increase allelic exchange efficiency, and though this technique also proved successful it frequently led to illegitimate recombination in which non-homologous strands of DNA were joined together (Aldovini et al., 1993).

Whilst useful for the selection of mutants, the introduction of a resistance cassette prevents additional genetic manipulation using the same selection marker and may also cause polar effects. However, extending on the research conducted by Husson et al. (1990), Pelicic et al. (1996) developed a method for the creation of unmarked mutants. This method also uses suicide plasmids that contain an altered version of the target gene, a selection marker and a gene responsible for sensitivity to sucrose, *sacB*. Following transformation, through HR, the first crossover event occurs as the plasmid, including the antibiotic marker and the *sacB* gene, is incorporated into the genome. Successful transformants are selected through growth on the selective antibiotic. Next, by using a counter-selection of sucrose, to which the mutants are now susceptible, due to selection pressure, a second crossover event is triggered. This second event will occur either at the point of the first crossover resulting in a return to wildtype (WT), or in between homologous regions upstream and downstream of the *sacB* gene resulting in a markerless mutant (figure 1.7a). Alternative methods have also been developed using delivery vectors which are designed to contain short, repeated sequences in the direct orientation known as recombination sites. These recombination sites flank the selection marker and upon induction of a recombinase are recognised resulting in the excision of the marker through site specific

recombination (Malaga et al., 2003) (figure 1.7b). Various recombinase systems have been used to produce markerless deletions including the γδ transposon *tnpR* resolvase and res recombination sites, the *Saccharomyces cerevisiae* FLP recombinase and FLP recognition target (FRT) recombination sites and even through endogenous pathways (Malaga et al., 2003; Stephan, Stemmer and Niederweis, 2004; Cascioferro et al., 2010).

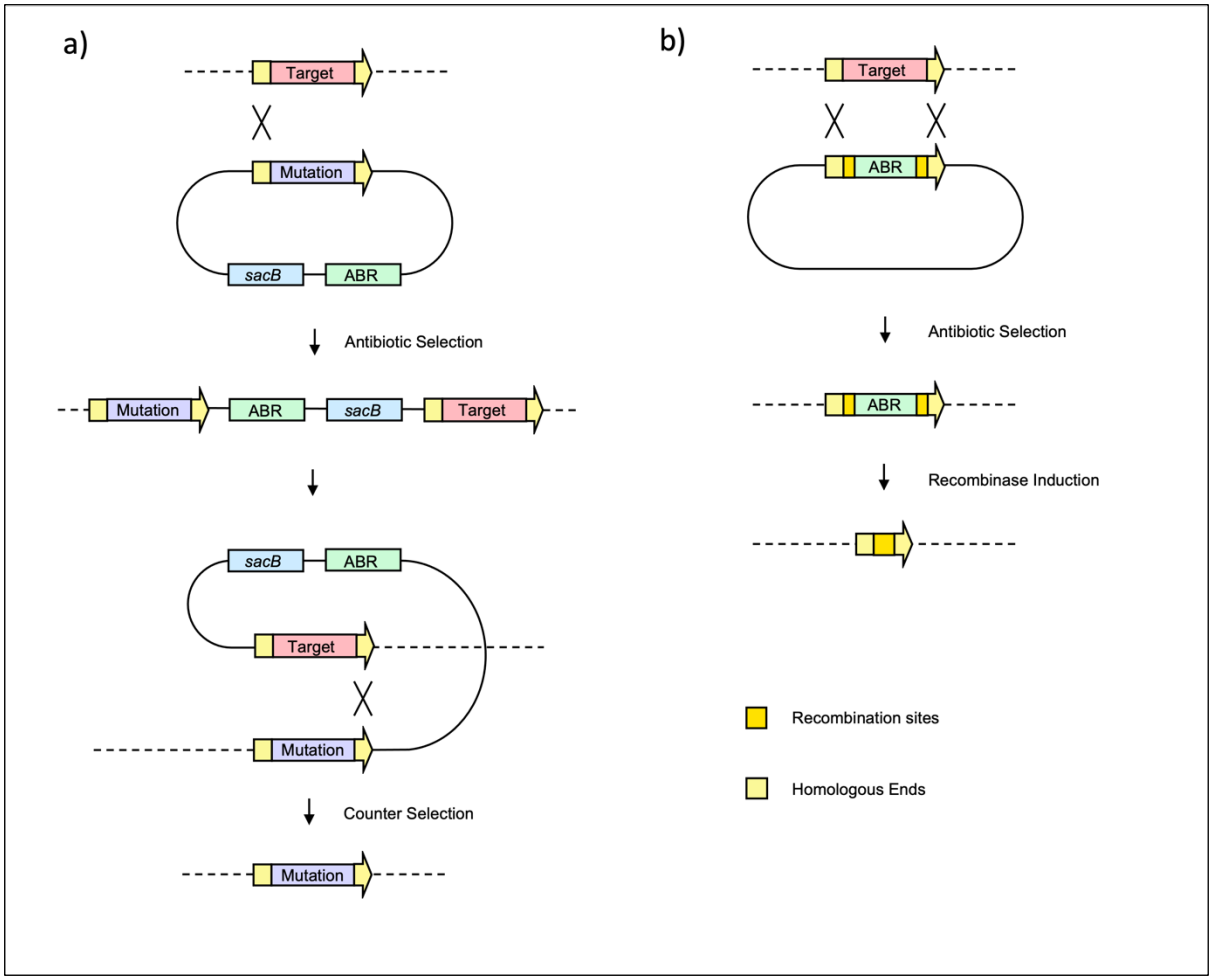



**Figure 1.7 Methods for generating unmarked mutants.**

(a) Two step allelic exchange. Through HR, a plasmid containing an altered version of the target gene, the sacB gene and an antibiotic resistance (ABR) cassette is inserted into the target site by a single crossover event occurring via the homologous ends. Cells containing the integrated plasmid are isolated via antibiotic selection. Through sucrose counter selection, a second crossover event is forced, resulting in the isolation of an unmarked mutants. (b) One step allelic exchange. Through HR, a double crossover event occurs between the target site and a plasmid containing an ABR cassette flanked by recombination sites and homologous ends. Mutants that have successfully substituted the target gene with the ABR cassette are isolated via antibiotic selection. Upon induction of a recombinase the recombination sites are restored excising the resistance cassette in the process to generate an unmarked mutant.

Though these DNA exchange techniques have been successful at generating mycobacterial mutants, these methods are often very slow mainly due to the low transformation efficiency, the reduced HR and the slow growth of these bacteria. Therefore, it is clear that faster and more robust methods are required for efficient genetic manipulation of mycobacteria.

## 2.4 Specialised Transduction

A more efficient method for generating mycobacterial mutants is through specialised transduction. This technique relies on the introduction of DNA via bacteriophage which results in nearly 100% transformation efficiency, the DNA is then incorporated into the genome via HR. The method requires the construction of a vector resulting from the ligation of a cosmid and a temperate phagemid. The cosmid contains two homologous regions to the target sequence, an antibiotic resistance cassette and a negative selection marker, recombination recognition sites, a λ cos site and OriE for replication in *E. coli*. The resulting vector is electroporated into mycobacteria. Under permissive temperatures the vector will replicate as a lytic bacteriophage or under restrictive temperatures the DNA will be incorporated into the genome (Bardarov et al., 1997).

Although specialised transduction demonstrates a much-improved transformation efficiency when compared to traditional methods, the construction of the vector is technically difficult and time-consuming, and the system still relies on the hosts HR machinery to instigate the allelic exchange.

It was discovered that phage recombineering systems such as the *E. coli* Rac prophage RecE and RecT proteins, or, the lambda red bacteriophage Exo, Beta and Gam proteins, can enhance bacterial HR. The RecE and Exo proteins act as exonucleases and bind to linear DNA substrates creating ssDNA overhangs. RecT and Beta proteins bind to the ssDNA and promote annealing of the linear DNA with complementary genomic DNA, lastly, Gam inhibits Rac prophage RecBCD nucleases from DNA degradation (Murphy, 1998). A similar system was discovered in a Che9c mycobacteriophage in which homologues of RecE and RecT, gp60 and gp61 respectively, were also found to greatly increase HR (Van Kessel and Hatfull, 2007). By placing these genes into an expression vector under the control of an inducible promoter, upon the introduction of foreign DNA, gp60 and gp61 were stimulated and thus assisted with the allelic exchange (figure 1.8). Using this method, a HR efficiency of 85% was achieved. Lastly, by combining this method with specialised transduction and thus delivering the foreign DNA via bacteriophage, the transformation efficiency was further increased. In addition, it was found that by incorporating only the gp61 gene into the genome, it was also possible to introduce ssDNA fragments with point mutations into the mycobacterial DNA though the efficiency was low (Van Kessel and Hatfull, 2008). Interestingly, it was discovered that the ssDNA was much more likely to bind to the lagging strand at the DNA replication fork.

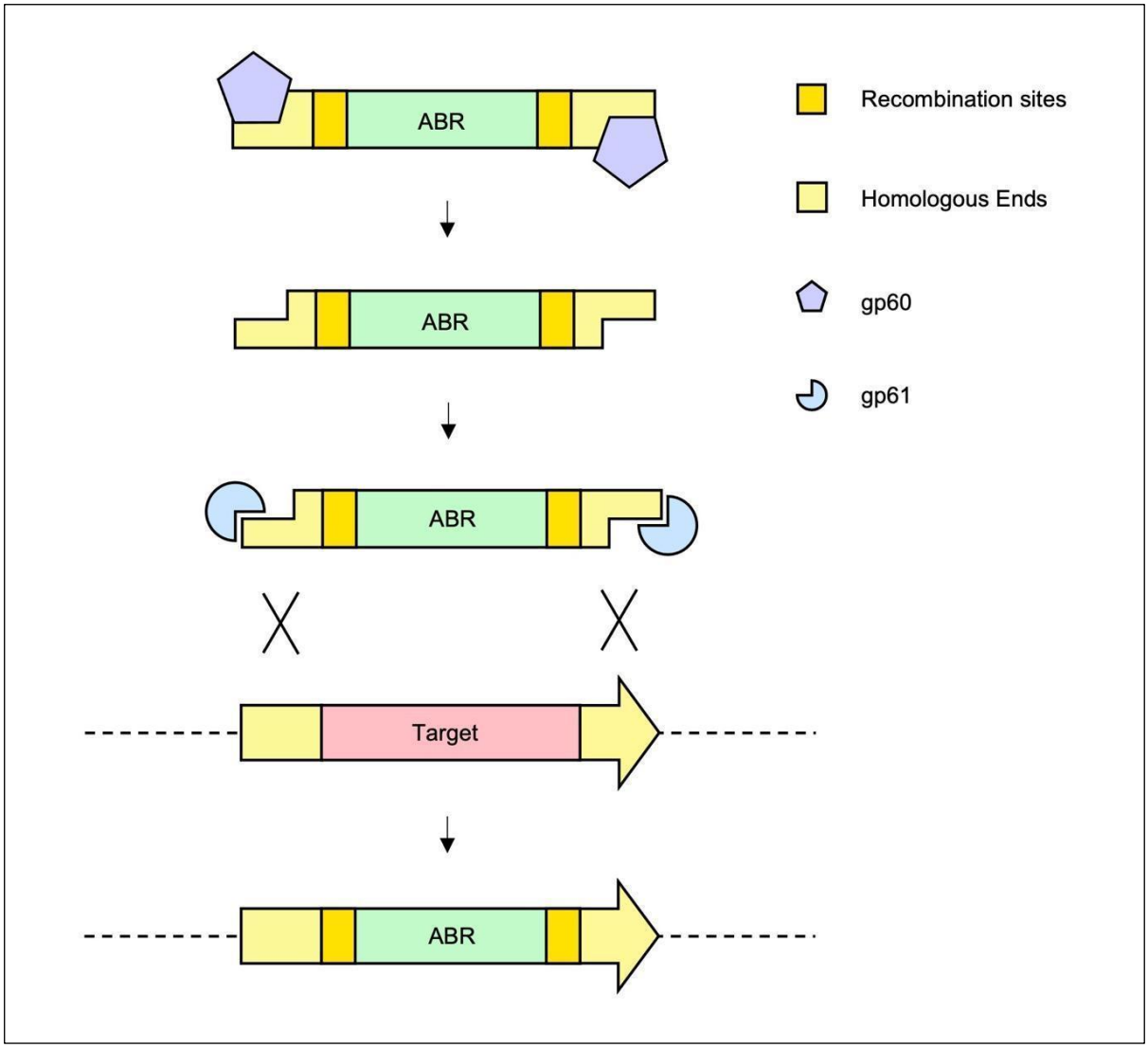



**Figure 1.8 Mycobacterial recombineering.**

Cells expressing a plasmid encoding gp60 and gp61 are transformed with a dsDNA oligonucleotide. The exonuclease gp60 binds to the oligonucleotide and degrades a strand of DNA to form a 3' overhang on each end. Gp61 binds to the ssDNA overhangs and stabilises the DNA to facilitate allelic exchange with the target site via a HR double crossover event.

## 2.5 Oligonucleotide-Mediated Recombineering followed by Bxb1 Integrase Targeting (ORBIT)

Alternatively, by coupling site-specific and homologous recombination techniques, Murphy et al. (2018) developed an improved and selectable method for genetic recombination in *M. bovis* and *M. smegmatis* called (ORBIT). Unlike previous methods, this technique generates a recombinant without the need for the construction of dsDNA plasmids or polymerase chain reaction (PCR) generated substrates and can be applied to produce chromosomal tags as well as gene deletions and substitutions. The method is described in detail in Chapter 4 Section 1.2.

It was reported that the generation of mutants via ORBIT boasted a high success rate, with the group obtaining over a hundred mutants via this method. That said, unlike phage transduction, the method necessitates an electroporation step in order to introduce the foreign DNA which could reduce the efficiency of transformation.

## 2.6 Clustered Regularly Interspaced Short Palindromic Repeats (CRISPR)

CRISPR is a defence system employed by many species of prokaryotes to prevent bacteriophage invasion (Van der Oost et al., 2009). CRISPR systems are an acquired immunity, originating from the DNA of bacteriophages that have previously infected the prokaryotic host (Van der Oost et al., 2009). They detect and destroy DNA that shares similarity and thus prevents further infection by closely related viruses

(Barrangou et al., 2007). The systems involve three key steps. Firstly, upon recognition of foreign DNA within the genome, known as proto-spacers, two Cas proteins, Cas1 and Cas2, will assist in the generation of small complementary DNA fragments that are then genomically integrated into a CRISPR array as spacers flanked by short, repeated DNA sequences (Barrangou et al., 2007). The next step requires the production of pre-crRNA, a large transcript from the CRISPR array locus. Mature crRNA is generated by dividing the pre-crRNA into short crRNAs via the endonuclease activity of a single Cas protein or a multi-protein complex (Brouns et al., 2008). Finally, Cas effector proteins and the crRNAs form a complex. To prevent self-targeting the Cas effector proteins are able to identify short sequence motifs found next to proto-spacers called proto-spacer adjacent motif (PAM) sites (Mojica et al., 2009). PAM sites are found throughout the genome and upon recognition, the CRISPR complex will bind to the DNA initiating the unwinding of the double helix. If complementary, the crRNA will attach to the now exposed DNA activating the Cas effector proteins to create a double strand break and thus degrading the target DNA (figure 1.9) (Wang et al., 2011).

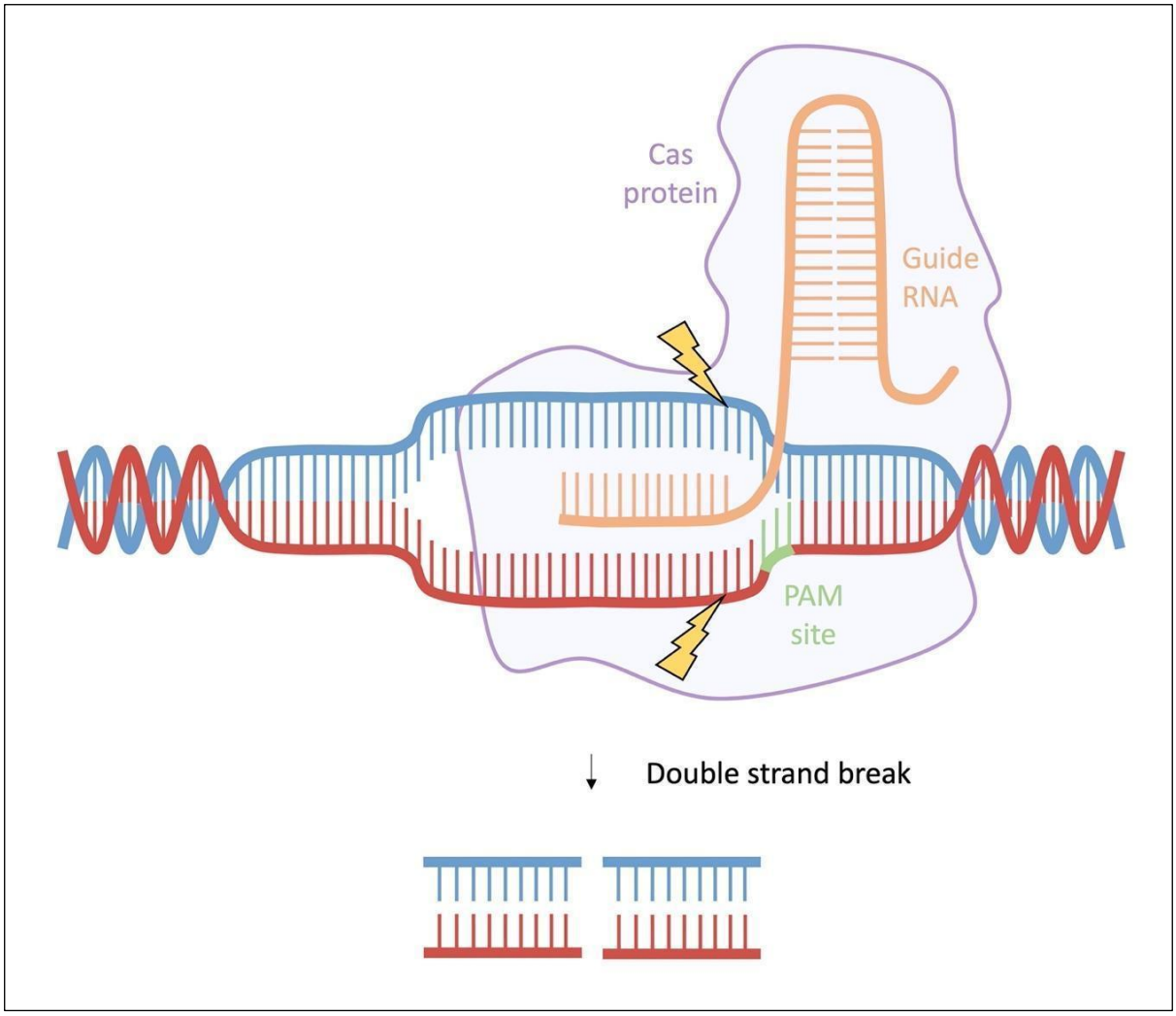



**Figure 1.9 The CRISPR system.**

The Cas protein and the guide RNA (CrRNA) form a complex. The complex binds to DNA at a PAM site triggering the uncoiling of the double helix. The guide RNA attaches to complementary DNA stimulating the Cas effector proteins to create a double strand break.

By manipulating this natural defence mechanism, CRISPR systems have been adapted as a mechanism for extremely efficient genetic engineering. Currently, the CRISPR systems can be categorised into three distinct types but the simplest is the CRISPR type II system. This system requires only a single Cas endonuclease protein, Cas9, and two crRNAs, mature crRNA and a trans-acting crRNA (tracrRNA). It was discovered that by generating a small guide RNA (sgRNA) that mimics the crRNA and tracrRNA complex, this system could be simplified to a two-component system which could degrade specific DNA and thus create targeted gene deletions (Jinek et al., 2012). However, as mycobacteria only possess the type III-A CRISPR system, the first adaptation of CRISPR for mycobacteria was initially attempted using Type II-A CRISPR system originating from *Streptococcus pyogenes*, including the Cas effector protein, SpyCas9 (Jinek et al., 2014; Rock et al., 2017). Unfortunately, SpyCas9 was found to be toxic to mycobacteria, especially when SpyCas9 expression was upregulated using a potent promoter (Sun et al., 2018). It was later discovered that the sgRNA transcriptional terminator was not functional in mycobacteria and that SpyCas9 reduced DNA repair by reducing (Bernheim et al., 2017). However, by instead utilising a *Francisella novicida* CRISPR system, a non-toxic alternative to the SpyCas9 protein, Cas12a, gene disruptions *in M. smegmatis* were successfully generated (Yan et al., 2017). To enhance HR DNA repair of the double strand breaks, gp60 and gp61 proteins were also employed resulting in a DNA deletion efficiency of 70-80%, though the efficiency was significantly reduced as the size of the deletion increased. In addition, by conducting CRISPR using oligonucleotides complementary to DNA targets, it was possible to engineer specific point mutations and chromosomal insertions of up to 1000 bp via ssDNA recombination (Yan et al., 2017).

Alternative systems have also been reported such as the use of the endogenous Class I type III-A CRISPR/Cas system found in *M. tuberculosis* and *M. bovis* (Wei et al., 2019). However, by far the most successful finding to date used a CRISPR system identified in *Streptococcus thermophilus*, that employs a St1Cas9 protein to genetically engineer pathogenic strains of mycobacteria (Meijers et al., 2020). The Cas9 protein was incorporated into the mycobacterial genome as a single copy and was placed under the control of a tetracycline promoter to reduce its toxicity. It was discovered that this method resulted in a 55% efficiency in producing loss of function knockouts in *M. marinum* and *M. tuberculosis*. Furthermore, not only can this CRISPR-Cas9 system be used to generate gene disruptions, but it can also be employed for constructing entire gene deletions by simultaneously employing two distinct sgRNAs

### 2.6.1 CRISPR Interference and Activation

Though there are numerous methods for disrupting non-essential genes in mycobacteria, it's only relatively recently that methods for interrogating essential genes have emerged (Singh et al., 2016). Hypomorphs are mutants in which target genes or proteins function at a reduced rate. Therefore, as the target is not entirely depleted, essential gene mutants can be generated and studied. Protein degradation systems can be employed to generate hypomorphs by reducing expression at the cellular level (Wei et al., 2011). These systems require a chromosomal tag bound to the gene of interest in order to label the target protein for degradation. However, this method has had limited success as not only is the efficiency of protein depletion inconsistent, but the introduction of the tag can disrupt downstream genes and interfere with the function of the target protein (Wei et al., 2011). An alternative method

for creating essential gene mutants can be achieved by replacing native gene promoters with inducible promoters thus providing control over the timing and the strength of gene expression (Freed et al., 2015). In addition, this method enables the generation of gain of function, overexpression mutants, providing a different approach for studying essential genes. Alternatively, overexpression mutants can also be generated by placing gene clones into a plasmid vector under the control of an inducible promoter (Melief et al., 2018). Nevertheless, these systems are equally inefficient as achieving expression that can rival the level of the native promoter is difficult (Parish et al., 1997).

However, by further adapting the CRISPR system, Qi et al. (2013) reported a robust method for conditionally silencing the expression of essential genes coined CRISPR interference (CRISPRi). It was discovered that by inhibiting the endonuclease activity of the Cas9 protein through the substitution of two key residues in the active site, the deactivated Cas9 protein was able to still bind to the target DNA but, instead of creating a double strand break, remained attached and interfered with transcription. Therefore, by placing deactivated *cas9* into a vector under the control of an inducible promoter the expression of the Cas9 protein and thus the silencing of the target gene could be regulated.

By using the dCas9 protein from *S. pyogenes* this technique was successfully adapted for mycobacteria (Choudary et al., 2015). Agarwal (2020) further improved the method by permitting the concurrent expression of multiple sgRNA sequences thus providing a method for analysing cumulative gene silencing. Moreover, this method was recently employed to create an entire *M. tuberculosis* knockdown library (Bosch et al., 2021).

This was achieved by identifying every PAM site in the *M. tuberculosis* genome and subsequently generating 96,700 sgRNAs complementary to the upstream target sequences. The sgRNAs were pooled and ligated into CRISPR vectors. These CRISPR vectors were introduced into *M. tuberculosis* via electroporation, induced and then grown on selective media to isolate single colonies. Lastly, the colonies were collected to form a pooled library which, through sequencing analysis, was found to successfully target 98.2% of *M. tuberculosis* ORFs and non-coding RNAs.

Alternatively, the CRISPR system can also be adapted for the generation of overexpression mutants via a technique called CRISPR activation (CRISPRa). CRISPRa was first developed in eukaryotes by fusing transcriptional activators with guide RNAs or deactivated Cas proteins (Gilbert et al., 2013). In this system, the CRISPR complex carries the transcriptional effectors to the target site to stimulate transcription. As bacteria lack efficient transcriptional activators, applying this technique to prokaryotes proved difficult. However, Dong et al. (2018) successfully generated *E. coli* overexpression mutants by constructing a synthetic promoter. Unfortunately, CRISPRa is yet to be adapted for mycobacteria but as efficient mycobacterial synthetic promoters have been recently described (Sun, Yang and Song, 2020) the method should be feasible.

Therefore, CRISPR techniques represent a simplified and efficient method for genetically engineering mycobacteria including the generation of knockdown and overexpression mutants. Furthermore, the capacity for CRISPR machinery to generate mutant libraries highlights the value of these methods for high-throughput and genome-wide study of mycobacteria.

## 2.7 Random Transposon Mutagenesis

Though it is possible to create mycobacterial mutant libraries in a gene-by-gene approach, generating mutants in bulk is a much easier strategy. Random transposon mutagenesis circumvents the need to design specific DNA constructs or oligonucleotides. Instead, entire libraries of single gene knockouts can be generated through the random integration of transposons (Borgers et al., 2020).

Transposons are mobile DNA segments able to 'jump' between different genomic loci (McClintock, 1950). Along with the necessary transposition machinery, transposons can carry additional genetic elements, such as antibiotic resistance genes, and therefore exist as an important driver of DNA exchange and thus the evolutionary process (Kleckner, Roth and Botstein, 1977). Furthermore, the insertion of transposons can also result in gene silencing. Transposons that insert into promoter regions can prevent transcription of associated genes or entire operons by disrupting the binding sites for the transcriptional machinery (Newton-Foot and van Pittius, 2013). In addition, transposons that insert directly into open reading frames may carry transcriptional terminators that will prevent gene transcription (De Lorenzo et al., 1990). Moreover, in the absence of transcriptional terminators, the transposon sequence will be transcribed along with the gene resulting in the production of truncated proteins due to the termination of translation by the presence of premature stop codons (Borgers et al., 2020). Transposition can occur via two processes: replicative or cut-and-paste transposition. Replicative transposition involves a transposon duplication event in which the transposon copy integrates into a new target

loci (Shapiro, 1979). The donor transposon is then resolved via the activity of a site-specific recombinase to restore the DNA to a single copy. Conversely, cut-and-paste transposition does not involve duplication. Instead, the transposon is cleaved from the genome and inserted into a new target site via the action of a transposase (Bender, Kuo and Kleckner, 1991).

The first mycobacterial transposon insertion libraries were generated in *M. smegmatis* using the Tn611 transposon, a Tn610 derivative originating from *M. fortuitum* which uses the replicative mechanism of transposition (Martin et al., 1990; Guilhot et al., 1994). It was previously discovered that the transposon could successfully generate transformants in *M. smegmatis* but only with a low efficiency. This was attributed to the use of non-replicating plasmids. Guilhot et al. (1994) increased the efficiency by instead incorporating the transposon into a replicating plasmid that could be conditionally controlled via temperature change. Transformants were initially obtained at a permissive temperature and then specifically selected for at a non-permissive temperature, in which the replicating plasmid was lost, ensuring that only cells with successfully integrated transposons, thus containing a selective marker, would survive. It was discovered that the transposon had no insertion bias and had a stable integration even when a selective pressure was absent.

To increase the efficiency of transposition further, transposons derived directly from *M. smegmatis* with a cut-and-paste mechanism (Tn5367, Tn5368 and Tn5370) were also developed (McAdam et al., 1995). It was discovered that these transposons had a random insertion but slightly favoured an AT rich centre (McAdam et al., 1995). It was hypothesised that packaging the transposons into temperature sensitive phage

could prove a useful method for transposon delivery efficacy. Indeed, Rybniker et al. (2003) successfully generated mutant libraries using phage delivery systems and found a significant difference in transformation efficiency depending on the type of phage used.

However, the initial attempts at random transposon mutagenesis, conducted using transposons originating from mycobacterial species, were not without limitations. For example, many of these transposons lacked functionality when employed for mutagenesis in other mycobacterial species. Furthermore, the efficacy of the transposon-transposase interactions were variable and inducing these systems introduced the possibility of activating endogenous transposons (Rubin et al., 1999).

In contrast, by utilising a transposon derived from the horn fly, *Haematobia irritans*, Rubin et al. (1999) developed a transposon from the mariner superfamily, *Himar1*, that could efficiently generate transposon mutants in mycobacteria. A significant benefit of using mariner derived transposons is the considerably reduced transposon recognition sequence, a TA dinucleotide. Furthermore, the *Himar1* system requires no host factors besides the *Himar1* transposase gene and the activity of the transposase has been further improved by generating a hyperactive version shown to further increase transposition efficacy (Lampe et al., 1999). A *Himar1* derived transposon, MycoMarT7, was generated by adding a T7 promoter and a kanamycin resistance cassette (Sassetti, Boyd and Rubin, 2001). It was discovered that in MAP, MycoMarT7 had around a 3-fold increase in transposition than Tn5367 and a greater proportion of insertions within genes (Rathnaiah et al., 2016). Unfortunately, a MycoMarT7 insertion

bias has also been observed in certain species of mycobacteria demonstrating that these systems could still be improved (Rathnaiah et al., 2016).

Whilst random transposon mutagenesis represents a rapid and effective method for creating entire libraries of mycobacterial mutants there is a clear constraint to this method: to determine the identity of individual mutants, additional genetic footprinting or sequencing steps are required. Nevertheless, robust methods have been developed for ordering and decoding pooled libraries amalgamating in the construction of an entire ordered *M. bovis* transposon library (Vandewalle et al., 2015). These techniques are covered in more detail in (Chapter 2, Section 1.3).

## 2.8 Introduction to Functional Genomics

Since mycobacteria are difficult to work with due their slow and aggregative growth as well as reduced transformation and HR efficiency, mycobacterial functional genomics methods have lagged behind other bacterial species. However, as seen by the creation of entire gene deletion libraries in *E. coli* (Baba et al., 2006) as well as transposon mutant libraries in *Staphylococcus aureus* (Fey et al., 2013), the generation of such libraries has facilitated various functional-genomic studies, leading to the characterisation of numerous previously uncharacterised genes (Typas et al., 2010; Nichols et al., 2011; Paradis-Bleau et al., 2014). Therefore, the creation of pooled or ordered mutant libraries in mycobacteria will greatly accelerate mycobacterial research by not only providing a substantial detailed resource but by also enabling the study of these organisms in bulk. In contrast to a gene-by-gene approach, by exploiting the potential of mutant libraries, functional genomics describes

the use of high-throughput omics platforms to define the roles and interactions of genes and proteins on a genome-wide scale.

Initial functional genomics capitalised on the power of pooled libraries. In these studies, entire libraries were exposed to various conditions and the relative abundance of mutants within the subsequent populations were analysed. Genes that were important for growth under specific stress conditions could be determined by detecting mutants with significantly altered abundance (Sassetti, Boyd and Rubin, 2001). Furthermore, genes could be defined as conditionally essential by identifying mutants that were completely absent from the resulting populations. Therefore, by linking genotype to phenotype, these methods were able to gain insight into gene function on a scale that could not be previously achieved (Mazurkiewicz et al., 2006).

A technique coined genetic footprinting was the first example of measuring mutant abundance in pooled libraries. In this method, individual mutants were verified via PCR using a primer specific to the gene of interest and a secondary primer specific to the transposon (Smith, Botstein and Brown, 1995). Whilst genetic footprinting has the potential to be applied on a genomic scale, this would be an extremely laborious process as only one mutant can be probed per PCR reaction (Sassetti, Boyd and Rubin, 2001).

In contrast, for concurrent analysis of mutant pools, a method called signature-tagged mutagenesis (STM) was developed (Hensel et al., 1995). In STM, ordered libraries are generated using transposons that have been tagged with short DNA barcodes. Each barcode has a variable central region, which acts as a unique identifier for each

mutant, and invariant flanking arms, which facilitate the parallel analysis of the library by permitting the co-amplification of each barcode via PCR. To compare the composition of different pools, the amplified DNA barcodes are then labelled and hybridised to a DNA chip (Hensel et al., 1995). Through this method, Camacho et al. (1999) reported the first global assessment of *M. tuberculosis* virulence factors. However, as barcoded libraries are limited in size, typically containing just 50-100 mutants, STM is not practical for conducting screens on a genome-wide, multi-condition scale (Sassetti, Boyd and Rubin, 2001).

## 2.9 Transposon Site Hybridisation

To increase screening throughput, Sassetti, Boyd and Rubin (2001) recognised that the DNA adjacent to the transposon insertion could be directly targeted for amplification thus negating the requirement for barcoded transposons and arrayed libraries. In this method, coined transposon site hybridisation (TraSH), the genomic DNA is initially digested using restriction enzymes and adapters are ligated onto the ends of the resulting DNA fragments. To enrich for transposon-adjacent DNA, PCR is performed using primers specific to the transposon and to the adapter. Lastly, the PCR products are transcribed into RNA, labelled, and then hybridised onto a microarray (Figure 1.10).

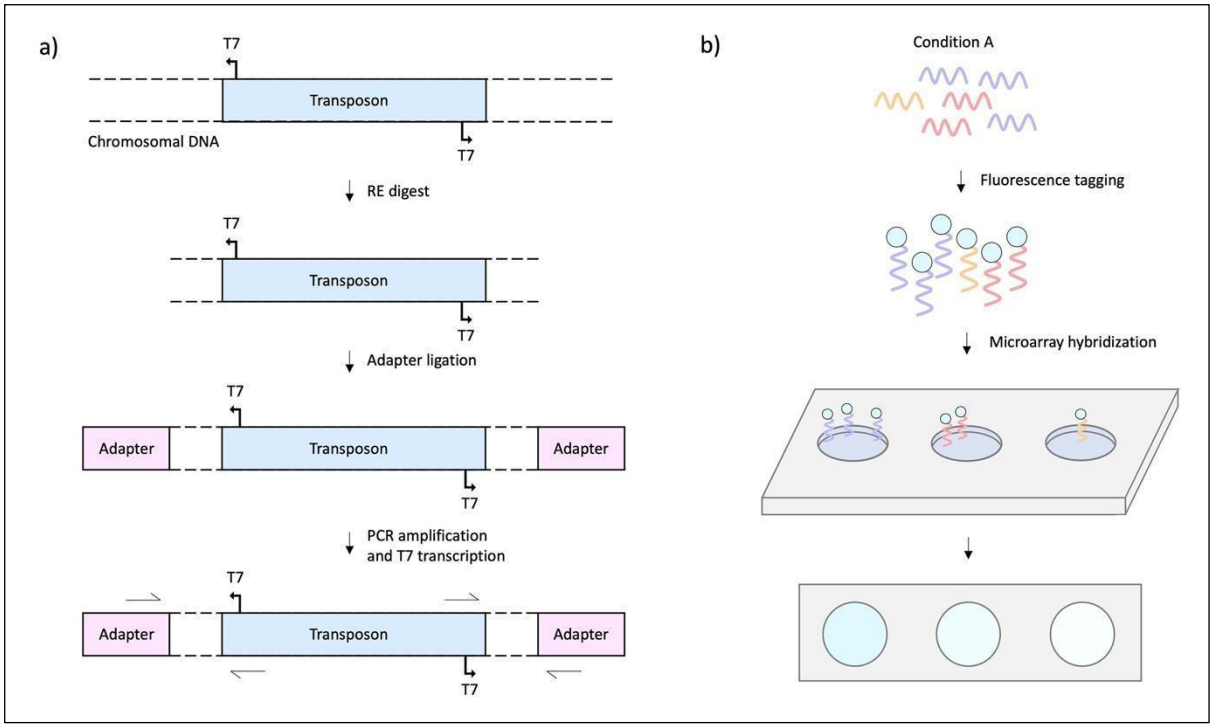



**Figure 1.10 The TraSH method.**

(a) Preparation of transposon containing DNA fragments. Genomic DNA is degraded using restriction enzymes (RE) and adapters are ligated. Transposon-adjacent DNA is amplified via PCR using transposon and adapter specific primers. *In vitro* transcription is performed using a T7 polymerase to generate RNAs. (b) Microarray. The RNAs are fluorescently labelled and hybridised to a microarray slide. The abundance of each RNA is measured via the intensity of fluorescence. (Sassetti, Boyd and Rubin, 2001).

To validate the method, the study was conducted in mycobacteria and represented the first global analysis of mycobacterial essential genes. By growing a *M. bovis* BCG transposon insertion library on standard and minimal media, enriched or depleted genes could be identified and thus genes essential to growth in a restricted environment were discovered. Furthermore, the same research group highlighted the power of this technique by identifying all of the mycobacterial essential genes *in vitro* (Sassetti, Boyd and Rubin, 2003) and defining 194 genes necessary for mycobacterial survival *in vivo* using a mouse model (Sassetti and Rubin, 2003). Many of these genes had not been previously characterised and thus represented novel candidates for drug targets. Therefore, by permitting the parallel analysis of large mutant pools, TraSH represented an important breakthrough in functional genomics. Indeed, successive studies applied this method to identify mycobacterial essential genes for slow and fast growth (Beste et al., 2009), in macrophage infection (Rengarajan et al., 2005) and for resistance to nitric oxide (Murry et al., 2009). However, the method is not without limitation. The assembly of a microarray is a laborious task and background hybridisation can obscure the presence of growth defective genes, resulting in false-negative outcomes. Furthermore, hybridisation is not an accurate method for locating the exact transposon insertion site and the method has a low resolution for quantifying differences in abundance, especially for low-copy and high-copy transcripts.

## 2.10 Transposon Mutagenesis with Next Generation Sequencing

To circumvent the requirement of a microarray for the analysis of pooled transposon libraries, in 2009, four distinct lab groups reported comparable methods, now collectively known as transposon mutagenesis with next generation sequencing (TnSeq), which utilised next generation parallel sequencing techniques to identify the exact sites of transposon insertions (Winkler et al., 2023). By detecting differences in the number of transposon insertions into a given loci, these methods enabled the quantification of gene fitness. These techniques were coined high-throughput insertion tracking by deep sequencing (HITS) (Gawronski et al., 2009), insertion sequencing (INSeq) (Goodman et al., 2009), transposon-directed insertion site sequencing (TraDIS)(Langridge et al., 2009) and transposon sequencing (Tn-Seq)(Opijnen, Bodi and Camilli, 2009). Each method was optimised in a different bacterial species but used a similar approach with the main differences occurring during the library preparation and fragmentation procedures. Firstly, the prepared transposon libraries are split into separate pools and each subjected to a different condition. Following a period of growth, like TraSH, the gDNA was fragmented, repaired, ligated to adaptors and then enriched for transposon-gDNA junctions via PCR. Lastly, the DNA is sequenced and then mapped to the genome. Genes were labelled as enriched or depleted by measuring the number of transposon insertions compared with standard conditions. Due to the improved resolution, these techniques could accurately identify slight changes (~5%) in gene fitness.

TnSeq was first applied to mycobacteria to investigate the genetic-basis for cholesterol utilisation and to update the *in vitro M. tuberculosis* essential gene list previously reported using TrASH (Sassetti, Boyd and Rubin, 2003; Griffin et al., 2011). Though the majority of the identified essential genes correlated with earlier findings and thus

further validated these methods, an additional 160 genes were discovered to be essential (Griffin et al., 2011). Extending on this research, Zhang et al. (2012) employed this approach to interrogate entire regions of the genome as opposed to individual genes and discovered various essential intergenic loci, non-coding RNAs and regulatory elements. In addition, this approach facilitated further characterisation of the essential gene list by identifying individual domains that were vital for protein function. Moreover, as previous studies had been conducted using partially saturated transposon libraries, Dejesus et al. (2017), generated a library that reached the maximum level of saturation. It was discovered that the *Himar1* transposon had an insertion bias at particular sequence motifs. Consequently, a hidden Markov model was applied which enabled the analysis of less permissive genomic loci and thus provided the most complete list of *M. tuberculosis in vitro* essentiality to date.

Whilst TnSeq is an effective tool for determining genome-wide conditional essentiality, applying this technique to large-scale screens with numerous conditions is difficult given the laborious task of sample preparation for sequencing. However, by employing a similar method to STM called random barcode transposon site sequencing (RB-TnSeq) up to 96 samples can be sequenced in parallel (Wetmore et al., 2015). In this method, random barcodes are introduced into the transposon and an initial transposon library is generated. The barcodes are mapped to the transposon insertion sites thus simplifying subsequent arrays as only the barcodes need to be sequenced instead of the transposon-gDNA junctions. Through the use of indexing primers, the samples can be sequenced in succession, therefore vastly improving the throughput (Price et al., 2018). In addition, TnSeq has also been adapted to account for population effects upon the outgrowth of a mutant library. In this method, coined Droplet TnSeq,

individual cells are packaged into oil droplets and thus can be grown in isolation (Thibault et al., 2019).

With the capacity to screen multiple conditions, various studies have used TnSeq as a method in mycobacterial drug discovery. Xu et al. (2017) screened a *M. tuberculosis* transposon library against subinhibitory concentrations of five antibiotics and discovered various cell envelope associated proteins which conferred antibiotic resistance thus highlighting the importance of the mycobacterial cell wall in antibiotic susceptibility. This study also elucidated the role of a previously uncharacterized protein, FecB, in envelope integrity. As FecB has been described as essential for *in vivo* survival, this protein was highlighted as an ideal drug target. Furthermore, Manjunatha et al. (2005) discovered the role of a protein, Fgd1, plus its associated cofactor, $F_{420}$, in pretomanid resistance and Singh et al. (2015) discovered factors involved in 5-fluorouracil resistance, helping to define its mechanism of action (MoA). In addition, Carey et al. (2018) and Akusobi et al. (2023) employed TnSeq to probe the genetic variations in clinical *M. tuberculosis* and *M. abscessus* isolates and thus identified key genes involved in antibiotic resistance as well as virulence.

Other applications of TnSeq in mycobacteria include the generation of double mutant libraries to interrogate gene-gene interactions. Through this approach, factors involved in PG synthesis (Kieser et al., 2015) and the oxidative stress response pathway were identified (Nambi et al., 2015). Furthermore, by combining TnSeq with fluorescence-activated cell sorting (FACS) changes at the cellular level could be analysed. This led to the discovery of genes involved in cellular heterogeneity and thus antibiotic susceptibility (Rego, Audette and Rubin, 2017) the discovery of mycobacterial gene

expression regulators (Smith et al., 2021) as well as the characterisation of mycobacterial D,D-transpeptidase crosslinking (Baranowski et al., 2018).

Whilst TnSeq represents an efficient tool for high-throughput functional genomics, there remain some disadvantages. Upon the outgrowth of libraries, there is a selective pressure for suppressor mutations and a bias in which strains with lower fitness are outcompeted and thus lost from the population (Gray et al., 2015). To account for the latter, adaptations to TnSeq, such as Droplet-TnSeq (Thibault et al., 2019), have been developed. However, these technologies are not widely accessible and require a further processing step which reduces throughput. In addition, although TnSeq has been combined with other quantitative methods, such as FACS, the majority of the output is sequencing data which is limited in its ability to describe changes in the behaviour of bacteria. Furthermore, although TnSeq can be employed for the discovery of essential genes, the libraries are absent of essential gene mutants and thus these cannot be extracted for further analysis.

## 2.11 Functional Genomics for Essential Gene Mutant Libraries

In contrast, to conduct high-throughput functional genomics on essential gene mutants, hypomorph libraries can be used. Perhaps, the most comprehensive mycobacterial hypomorph screen to date used a method called PROSECT (primary screening of strains to prioritise expanded chemistry and targets) (Johnson et al., 2019). In this study, barcoded hypomorphs were generated using a more traditional

approach, by tagging genes for protein degradation. In total, a library of around 2,000 hypomorphs was constructed which successfully targeted 474 *M. tuberculosis* essential genes. This library was subjected to a huge chemical collection consisting of approximately 50,000 different conditions and the hypomorph abundance was calculated via barcode sequencing. It was discovered that over 40 different compounds exhibited anti-mycobacterial action and thus demonstrated a vast increase on the hit-rate described in standard chemical screenings. This research led to the identification of anti-tuberculosis agents with a novel cellular target, *efpA*. It was discovered that resistance acquired against one *efpA* inhibitor increased sensitivity to the other, thus highlighting their potential for use in combination.

Furthermore, as CRISPRi offers a higher throughput method for generating hypomorph libraries, Li et al. (2022) recently adapted CRISPRi for mycobacterial chemical genetic screening. The method involved the sequencing of the sgRNAs to measure relative hypomorph abundance. Therefore, by subjecting the *M. tuberculosis* CRISPRi library generated by Bosch et al. (2021) to nine different chemical stresses, this study was able to discover multiple pathways involved in intrinsic drug resistance.

As an alternative to hypomorph generation, gain of function libraries can be constructed for the functional genomic screening of essential genes. Melief et al. (2018) created an overexpression library in *M. tuberculosis* by transforming cells with vectors containing genetic clones placed under the control of an inducible promoter. However, using such traditional methods for generating overexpression libraries is extremely laborious. Therefore, to increase the throughput, a technique such as CRISPRa (discussed in section 2.6.1), TraDIS-Xpress (Yasir et al., 2020) or dual

barcoded shotgun expression library sequencing (Dub-Seq) (Mutalik et al., 2018) could be employed. TraDIS-Xpress, uses the TraDIS method to extend upon the research conducted by Melief et al. (2018) to rapidly generate pooled overexpression transposon insertion libraries by incorporating inducible promoters into the transposable elements (Yasin et al., 2020). Similarly, Dub-Seq represents an efficient method for analysing the fitness of pooled overexpression libraries. In this method first described by Mutalik et al. (2018), genomic DNA is randomly sheared and each fragment is placed into an expression vector in between two short DNA sequences that function as unique barcodes. Through sequencing, the fragments are mapped to the genome and associated with the corresponding barcodes. To construct the library, cells are transformed with the vector so that each strain contains a single DNA fragment. Lastly, the library is subjected to various conditions and the barcodes are sequenced to determine the relative abundance of each strain (figure 1.11). However, though this method has proved successful in *E. coli*, it is yet to be applied to mycobacteria.

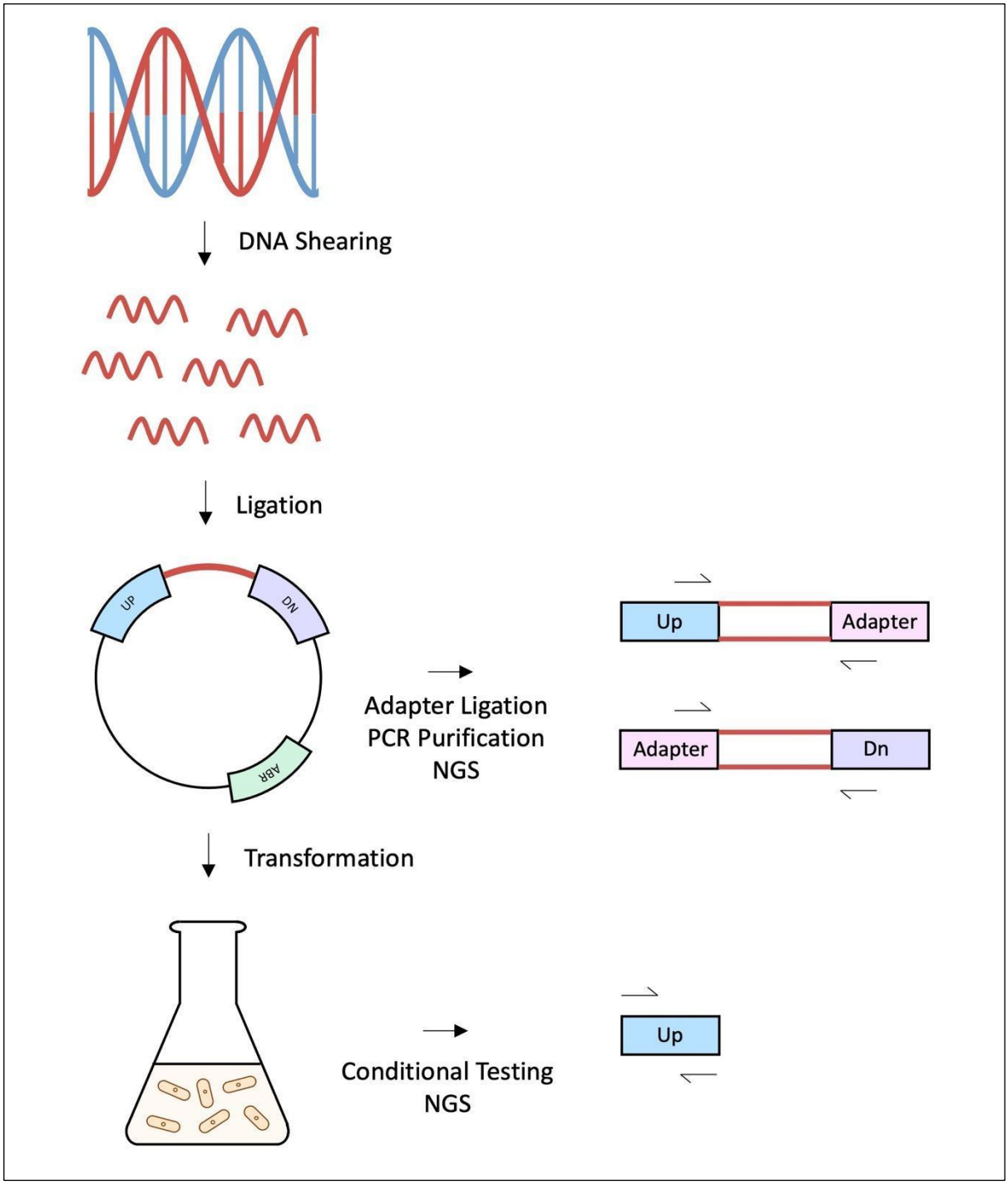



**Figure 1.11 Dub-Seq Method.**

gDNA is degraded and each fragment is ligated into an expression vector between two barcodes. Following adaptor ligation and PCR, the fragments are sequenced and mapped to the genome. To construct the library, WT cells are transformed with the expression vectors. Finally, following conditional testing the barcodes are sequenced and the abundance of each strain is calculated. NGS = Next generation sequencing (Mutalik et al., 2018).

## 2.12 Functional Genomics Using Arrayed Libraries

To diversify the data output of functional genomics there has been a recent shift away from the sequencing analysis of pooled libraries to the development of more complex assays for analysing arrayed libraries. Various studies have described intricate deconvolution techniques for generating ordered mutant libraries covered in Chapter 2, Section 1.3.

Investigating bacterial growth kinetics represents a relatively simple but effective output for analysing ordered mutant libraries. For example, Melief et al. (2018) ordered the aforementioned *M. tuberculosis* overexpression library into 96 well plates and, by measuring OD, discovered key proteins involved in D-cycloserine resistance. Similarly, McNeil et al. (2021) measured the OD and colony forming units (CFU) from an arrayed *M. smegmatis* CRISPRi library to distinguish between gene essentiality, vulnerability and lethality. Here it was discovered that whilst many genes involved in cell wall biosynthesis and core cellular processes are essential, the knockdown of genes with roles in metabolism often results in bacteriostatic phenotypes. In addition, the silencing of various metabolism-related genes previously described as non-essential were found to inhibit bacterial growth. It was hypothesised that a reduced degree of metabolic buffering could prevent growth due to a delay in metabolic remodelling. In addition, Boeck et al. (2022) monitored the growth of a collection of *M. abscessus* clinical isolates in the presence of different carbon sources, antibiotics and *in vivo*. Remarkably, when coupled with research from a conditional genome wide association study and proteome-wide structural modelling, this study identified various clinically relevant virulence factors.

For an alternative readout, ordered libraries have also been combined with metabolomic and transcriptomic study. Anglada-Girotto et al. (2022) used *M. smegmatis* and *E. coli* CRISPRi libraries to map changes in drug-induced metabolic profiles to specific genotypes. In total, a collection of 1,342 drugs were screened and around 1,000 metabolites were detected resulting in the discovery of numerous antimicrobials with unique MoAs. Furthermore, Rustad et al. (2014) generated an M. *tuberculosis* overexpression library consisting of transcription factor clones and through transcriptome analysis and CHIPseq were able to identify the regulatory networks associated with each factor. In total, nearly 10,000 regulatory consequences were discovered accounting for around 70% of the regulatory pathways in the genome.

However, it is the employment of high-throughput imaging techniques that have dominated the recent functional genomic study of arrayed bacterial libraries. Campos et al. (2018) developed a microscopy method for phenotyping the *E. coli* gene replacement library, the Keio collection (Baba et al., 2006), thus providing a valuable resource for the functional characterization of the *E. coli* genome. In this method, phase contrast and epifluorescence microscopy was conducted on agarose pads, each pad consisting of an array of 48 spotted mutants. Using a support vector machine model various morphological and cell cycle phenotypes could be analysed from the microscopy images. Furthermore, this study also determined the growth phenotypes of each strain by conducting growth curves in a microplate reader. Leading on from this work, Zahir et al. (2019) developed an increased throughput time-resolved microscopy method by imaging a growing Keio collection in glass-bottom 96-well plates. By subjecting the mutants to various antibiotics, 191 strains demonstrated

altered phenotypes highlighting the roles of various genes in antibiotic response pathways. A similar approach was also conducted for mycobacterial research in which a *M. smegmatis* CRISPRi knockdown library consisting of 263 essential gene mutants was analysed through time-lapse microscopy (de Wet et al., 2020). It was discovered that functionally related genes clustered via their corresponding mutant morphological phenotypes thus leading to the discovery of a potential restriction-modification system. Furthermore, by generating a *M. smegmatis* fluorescently labelled protein library, Judd et al. (2021) used fluorescence microscopy to determine the cellular localisation of over 1000 different proteins.

Alternatively, Typas et al. (2008) developed a high throughput imaging technique for functional genomics by analysing colony growth via a method called genetic interaction analysis technology for *E. coli* (GIANT-Coli). In this gene-gene interaction study, a library of *E. coli* double mutants were generated via a Hfr mating conjugation system and grown on solid agar plates in a 1536-colony format. To determine the fitness of each strain, the colony sizes were measured by counting the number of pixels within a defined boundary and the data was normalised to control for any plate-effects on growth. Through this relatively small dataset, the Tol-Pal system was identified as a key feature in the organisation of the cell envelope highlighting the power of this method for the functional characterisation of bacterial genomes. The same research group extended upon this research by subjecting an *E. coli* mutant library, consisting of both gene deletion and knockdown mutants, to various chemical and environmental stresses in order to measure conditional mutant fitness (Typas et al., 2010; Nichols et al., 2011). By clustering genes by their corresponding mutant phenotypic signatures genes of a similar function were identified, culminating in the

characterisation of three orphan genes, *lpoA, lpoB* and *marB,* involved in PG synthesis and antibiotic resistance respectively. Furthermore, the group also developed a tool for the improved analysis of colony growth, coined IRIS (Kritikos et al., 2017). IRIS is able to detect differences in colony area, opacity, circularity, colour, morphology and even biofilm formation and thus greatly improves the sensitivity for measuring mutant fitness. This tool has been used in a variety genome-wide functional characterization studies across a range of bacterial backgrounds helping to elucidate the roles of numerous previously uncharacterised genes (Paradis Bleau et al., 2014; Peters et al., 2016; Shiver et al., 2016). Though this method has not yet been used for the study of mycobacteria, given the success and adaptability in other bacterial species, this method could prove an extremely versatile technique for high throughput mycobacterial functional genomics.

It has been suggested that due to cross-contamination around 30% of ordered libraries are not usable (Silvis et al., 2021), though Goodall et al. (2022) developed a tool for validating ordered libraries called LI-detector and demonstrated that only 3% of the Keio library was incorrect. Furthermore, although various tools have been developed to accelerate the construction of ordered libraries and the application of more complex assays, these additional steps reduce the experimental throughput. However, this is compensated by the great diversity of data output enabling a more thorough interrogation of gene function.

# Part 3 – Conclusion

The genus *Mycobacterium* comprises nearly 200 different species of bacteria characterised by their thick and waxy cell envelope which acts as an important barrier to the external environment (Parte et al., 2020). Many species of mycobacteria are important human and animal pathogens including the causative agent of TB, *M. tuberculosis*. TB is one of the leading causes of mortality from a single infectious agent, resulting in around 1.5 million deaths each year (WHO, 2023) . Although TB is curable, the treatment regimen is long and intensive. Furthermore, rising cases of multi-drug resistance are being reported and thus there is a huge demand for the discovery of new and improved drugs. Despite being a major burden to global health and the global economy, these organisms are still poorly understood with around 25% of the genome lacking functional annotation (Modlin et al., 2021). The characterisation of the remaining genome could lead to the discovery of orphan genes, new drug targets and antibiotic resistance mechanisms. Genetic engineering is a powerful tool for deciphering gene function. However, due to the nature of the cell envelope, these bacteria are extremely difficult to genetically manipulate and traditional engineering techniques were largely inefficient (Murphy et al., 2018). Therefore, the development of tools such as ORBIT (Murphy et al., 2018) and CRISPR represents a significant advance in mycobacterial engineering providing researchers with effective and in depth methods for functionally characterising genes. Nevertheless, to interrogate gene

function on a genome-wide scale, using a gene-by-gene approach is not practical. In contrast, high-throughput functional genomics enables the functional annotation of entire genomes. Whilst TnSeq and CRISPRi are rapid tools for conducting functional genomics screens, the output is restricted to sequencing data (Griffin et al., 2011; Li et al., 2022). In contrast, by generating ordered mutant libraries, the throughput is reduced but a greater range of data outputs can be analysed to generate more comprehensive predictions on gene function (Kritikos et al., 2017).

The range of genetic tools now available to the mycobacterial research community will greatly accelerate our understanding of these bacteria. These methods can be used to identify novel drug and vaccine targets, as well as uncover important virulence and drug resistance pathways. Therefore, by helping to inform clinicians and by aiding in the discovery of new treatment options, these developments could prove an important factor in the fight against mycobacterial disease.

# Part 4 – Aims

Mycobacteria are important human, animal and environmental pathogens, contributing to considerable mortality and morbidity as well as causing huge costs to both the global health and farming industries. In addition, the current mycobacterial treatment regimens are often long and intensive and there are increasing concerns over rising drug resistance. Therefore, there is a huge demand for new approaches to combat drug resistance and to develop new therapeutic options. Unfortunately, although widely studied, these bacteria are poorly understood with much of the genome still functionally uncharacterised. By functionally annotating the remaining genome, genes involved in important biological processes and antibiotic resistance may be discovered. Furthermore, novel drug and vaccine targets could be identified. We therefore aim to develop the following pipeline for functionally annotating mycobacterial genomes to help characterise the remainder of the genome.

1. To optimise a method for constructing ordered mycobacterial mutant libraries and to generate an ordered *M. abscessus* mutant library via random transposon insertion mutagenesis.
2. To optimise a method for the functional-genomic screening of ordered mycobacterial mutant libraries and conduct a screen using a condensed *M. bovis* BCG mutant library.
3. We aim to establish a method for validating any chemical-genomic screening hits by phenotyping clean knockouts generated via ORBIT and characterise a selection of cell-wall associated proteins.

# Chapter 2

# Generating an Ordered *M. abscessus* Transposon Library

# Part 1 – Introduction

## 1.1 Introduction to Ordered Libraries

An arrayed single gene mutant library serves as an important resource for the functional analysis of microorganisms. Not only do these libraries provide a complete collection of mutants that can be selected for further study in a gene-by-gene approach, they can also be utilised for high-throughput functional genomic screens. For further information on the utility of arrayed libraries refer to Chapter 1, Section 2.12.

The first complete bacterial mutant library, coined the KEIO collection, was generated in *E. coli* by creating a single gene deletion for every known gene in the *E. coli* genome (Baba et al., 2006). The method was adapted from a technique developed in yeast by Giaever et al. (2002) in which *E. coli* cells expressing the λ Red recombinase were targeted for gene replacement via homologous recombination with a PCR product. Each PCR product contained a kanamycin resistance cassette flanked by FLP recognition (FRT) sites and homologous arms to a specific target gene. Therefore, for each gene, a unique PCR product was constructed and, through site directed recombination via the FLP recombinase, a clean knockout could be generated by excising the resistance cassette. The PCR products were designed to ensure that the entire gene was deleted in frame, in which, with the exception of overlapping genes,

only the initiation codon and the seven C-terminus codons remained. Subsequent studies have adapted this method to generate gene deletion libraries in other bacteria such as *Acinetobacter baylyi* ADP1 (De Berardinis et al., 2008), *Salmonella enterica* sv Typhimurium (Porwollik et al., 2014) and *B. subtilis* (Koo et al., 2017). The construction of these libraries has generated significant advances in the research of these bacteria, with findings from the KEIO collection being the most well documented. Since its construction, Typas et al. (2010) unearthed two key proteins involved in PG biosynthesis through a high-throughput genome-wide imaging pipeline. Furthermore, later studies employed a similar technique to identify numerous genes involved in cell envelope biogenesis and antibiotic resistance (Nichols et al., 2011; Paradis-Bleau et al., 2014). In addition, Heo et al. (2020) used a high-throughput growth assay to discover key bacterial metabolites involved in the production of algal blooms. These findings highlight the power of constructing arrayed libraries, facilitating the functional characterisation of bacterial genomes.

## 1.2 Ordered Transposon Insertion Libraries

As a unique PCR product must be generated for each gene deletion, the construction of mutant libraries through the gene replacement method is extremely laborious. As covered in Chapter 1, Section 2.6, using CRISPR machinery simplifies this process as the CRISPR method only requires the design of short sgRNAs. In addition, hypomorphs can be generated via CRISPRi thus enabling the analysis of essential genes (Bosch et al., 2021). However, by negating the need to design specific DNA constructs, random transposon mutagenesis represents a more efficient, high-

throughput method for generating entire mutant libraries. In this method, transposons are introduced into bacterial pools and will randomly integrate into the genome with the aim of obtaining a single transposon insertion per cell. Due to the insertion of premature stop codons, when a transposon integrates into an open reading frame, translation is terminated often rendering the resulting protein non-functional (Borgers et al., 2020). The level of disruption depends on the insertion site as insertions in closer proximity to the initiation codon will result in shorter polypeptide products and thus have a greater potential for disrupting protein function. Therefore, a transposon insertion into a gene does not assure successful gene knockout representing a major drawback of this method. However, this can be compensated by only selecting for mutants which harbour insertions close to the 5' region of the gene. Furthermore, as genes are not directly targeted for deletion, and as transposons exhibit an insertion bias, there is no guarantee that every non-essential gene will be disrupted. To account for this, a significantly larger library of mutants than the total number of non-essential genes can be generated. In addition, by using the Coupon Collector's Problem or a Monte Carlo simulation the optimal library size and the coverage of the genome can be predicted (Bayn et al., 2016; Borgers et al., 2020).

Pooled analysis of transposon libraries can be conducted through the sequencing and mapping of transposon insertions across the genome. However, the output is limited, relying on transposon abundance as a measure of gene fitness. Therefore, pooled analysis has been predominantly employed for the identification of essential genes (Fey et al., 2013; DeJesus et al., 2017; Goodall et al., 2018). Instead, by ordering transposon mutants into an array, more in depth analysis with a greater diversity of

output can be conducted but this does require an additional sequencing step to deconvolute the library and identify the locations of each mutant.

## 1.3 Library Deconvolution

Initial library deconvolution attempts employed a pooling method coupled with PCR screening in which PCR products were generated using primers specific to the transposon and a target gene (Lane and Rubin, 2006). To identify the location of a single mutant, two rounds of PCR were performed. The first round was conducted on pools of mutants collected from each library plate to identify the mutant origin plate. The second round was conducted on mutant pools collected from the rows and columns of the origin plate to identify the precise location. Whilst this method is beneficial for locating individual mutants within a library, as the method can only identify the location of one mutant per PCR screen, employing this method to decode an entire library would be too laborious. Instead, Fey et al. (2013) described a more efficient method by sequencing the DNA adjacent to the transposon insertion. Whilst this method was used to deconvolute a *S. aureus* library consisting of 18,767 mutants, as each mutant is sequenced independently, the method is still laborious as well as expensive.

In contrast, Vandewalle et al. (2015) employed a parallel sequencing technique combined with a library pooling method, coined Cartesian-Pooling Coordinate Sequencing (CP-CSeq), to reduce the number of sequencing steps required to decode a 96x96 *M. bovis* BCG transposon mutant library. Similar to Lane and Rubin (2006),

in this method, to locate each mutant, the library was categorised into X, Y and Z coordinates. The Z coordinate identifies the mutant origin plate whereas the X and Y coordinates identify the position of the mutant within the origin plate. Firstly, by pooling all mutants at each XY coordinate across every library plate, an initial 96-well master plate was generated. Next, by pooling all of the mutants within each library plate, a second 96-well master plate was prepared. Every column and row for both master plates were then pooled to generate 40 sequencing pools. To sequence the DNA, for each pool, the gDNA was extracted and sheared, pool-specific barcoded adapters were ligated onto the DNA fragments to enable parallel sequencing, and through PCR using primers specific to the adapter and the transposon, regions flanking the transposon insertions were amplified. Lastly, through Illumina® sequencing, the transposon insertions were mapped to the genome and the library was deconvoluted using Galaxy workflows and a BioPerl algorithm (figure 1.12). In the ideal scenario of a unique mutant, the DNA clone would only appear in 4 pools corresponding to the row and column of both master plates thus enabling precise mapping. However, due to the nature of generating arrayed random transposon insertion libraries, in which mutants are isolated from a pool, multiple copies of each mutant can exist. In these cases, a heuristic algorithm was applied to predict the mutant locations but this is not totally accurate. Nevertheless, in total, Vandewalle et al. (2015) were able to generate a deconvoluted library containing 64% of all non-essential *M. bovis* BCG genes. Furthermore, the same research group managed to later improve upon this library, achieving 83% saturation of the non-essential genome using the same method (Borgers et al., 2020).

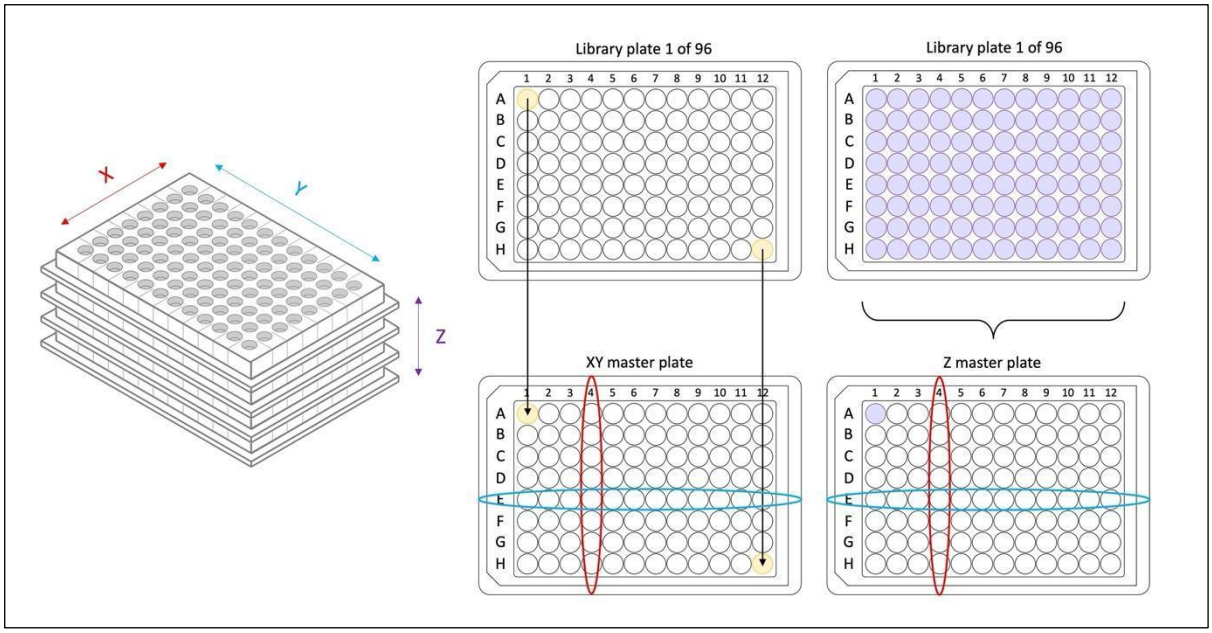



**Figure 1.12 CP-CSeq method.**

Across all library plates, mutants at each XY coordinate were pooled to generate a XY master plate (highlighted in yellow). A Z master plate was also constructed by pooling every mutant from each library plate (highlighted in purple). For both master plates, each row and column were pooled for sequencing (highlighted in blue and red) (Vandewalle et al., 2015).

As the pooling techniques described by Vandewalle et al. (2015) requires the use of robotic hardware to manage the liquid handling steps, this method is not widely accessible and can be expensive. Therefore, Baym et al. (2017) employed a similar pooling technique called Knockout Sudoku which negated the requirement of robotics. In this method, extending from a pipeline outlined by Elrich et al. (2009) and Goodman et al. (2012) the library plates were instead organised into a hypothetical grid. Pools were then generated from each corresponding row and column across the entire library as well as all of the mutants from every plate within a column or row of the grid (figure 1.13). Despite the exclusion of robotics, this method represented a rapid and

efficient method for decoding transposon libraries and culminated in the construction of a *Shewanella oneidensis* ordered library consisting of 3,667 unique mutants.

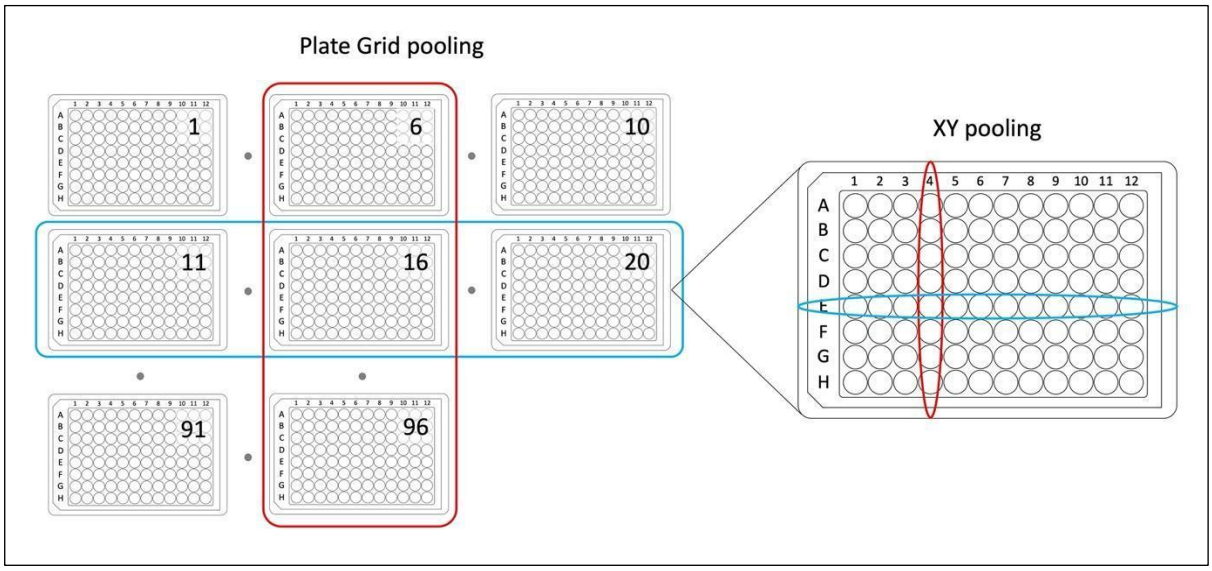



**Figure 1.13 Knockout Sudoku Method.**

The library is organised into a hypothetical grid. All of the mutants within each row and column of the grid as well as each row and column across every plate in the entire library were pooled for sequencing (highlighted in blue and red) (Baym et al., 2017).

## 1.4 Research Aims

Although described as an opportunistic pathogen, *M. abscessus* is becoming increasingly clinically relevant due to its intrinsic multi-drug resistance and rising prevalence in hospitals and lung disease patients (Johansen et al., 2020). Despite this, due to the severity of disease caused by *M. tuberculosis*, research into *M. abscessus* and other NTM is often overlooked. In addition, whilst there is an abundance of *M. abscessus* sequence data available, due to the limited mycobacterial

genetic toolbox and a lack of *M. abscessus* resources, these bacteria are still poorly understood with a large portion of the genome remaining functionally uncharacterised.

We therefore aimed to construct an ordered *M. abscessus* mutant library to enable the rapid functional annotation of the remaining genome. This will not only improve our understanding of these bacteria but could also lead to the discovery of new drug targets and antibiotic resistance mechanisms thus representing a significant asset to the *M. abscessus* research community.

# Part 2 – Results

## Contribution Statement

Sections 2.3, 2.5, 2.8 and 2.9 were conducted by Luke Kidger MSc under my supervision. Sections 2.4 and 2.7 were conducted in collaboration with Luke Kidger MSc.

## 2.1 Selecting an Efficient Transposon Delivery System

As covered in Chapter 1, Section 2.7, due to the thick and hydrophobic mycobacterial cell wall and thus reduced transformation efficiency, initial attempts at introducing transposons into mycobacteria as free transposomes or via plasmids was largely ineffective (Guilhot et al., 1994; Langridge et al., 2009). Instead, Sassetti et al. (2001) demonstrated that the use of a mycobacteriophage as a transposon delivery system could dramatically increase the efficiency of transposition. Through this method, using the ΦMycoMarT7 phagemid containing a *Himar1* derived transposon, MycoMarT7, a transposon insertion library in *M. bovis* BCG was successfully generated consisting of $10^5$ mutants. In addition, as MycoMarT7 is greater in size than the *Himar1* transposon, to increase transposition efficiency, Borgers et al. (2020) employed an alternative phage delivery system, phAE157, resulting in the construction of a *M. bovis* BCG ordered library containing 18,432 mutants.

Therefore, given the success of the phage delivery method in other mycobacterial species, we adapted this method for *M. abscessus*. To determine the optimal system for creating a library we investigated the suitability of two different *M. abscessus* subspecies, *M. a. massiliense* and *M. a. abscessus*. Furthermore, we also investigated the use of two different phage systems, ΦMycoMarT7 and phDB36. The phage phDB36 was originally designed for *M. abscessus* TnSeq experiments as it carries a *Himar1* derived transposon which was adapted to contain MmeI recognition sites (Foreman et al., 2020). Therefore, through restriction digest, fragments containing transposon adjacent DNA could be easily obtained, negating the need for PCR enrichment. As this phage has been previously employed for the generation of pooled *M. abscessus* libraries we reasoned that phDB36 could also be used for constructing an ordered library. Furthermore, in this system the kanamycin resistance gene is replaced with a Zeocin$^{TM}$ resistance cassette (Foreman et al., 2020). Therefore, as many mycobacterial genetic engineering tools involve the use of a kanamycin resistance cassette, it could be beneficial to construct a library in a different antibiotic background.

## 2.2 Optimisation of Phage Propagation

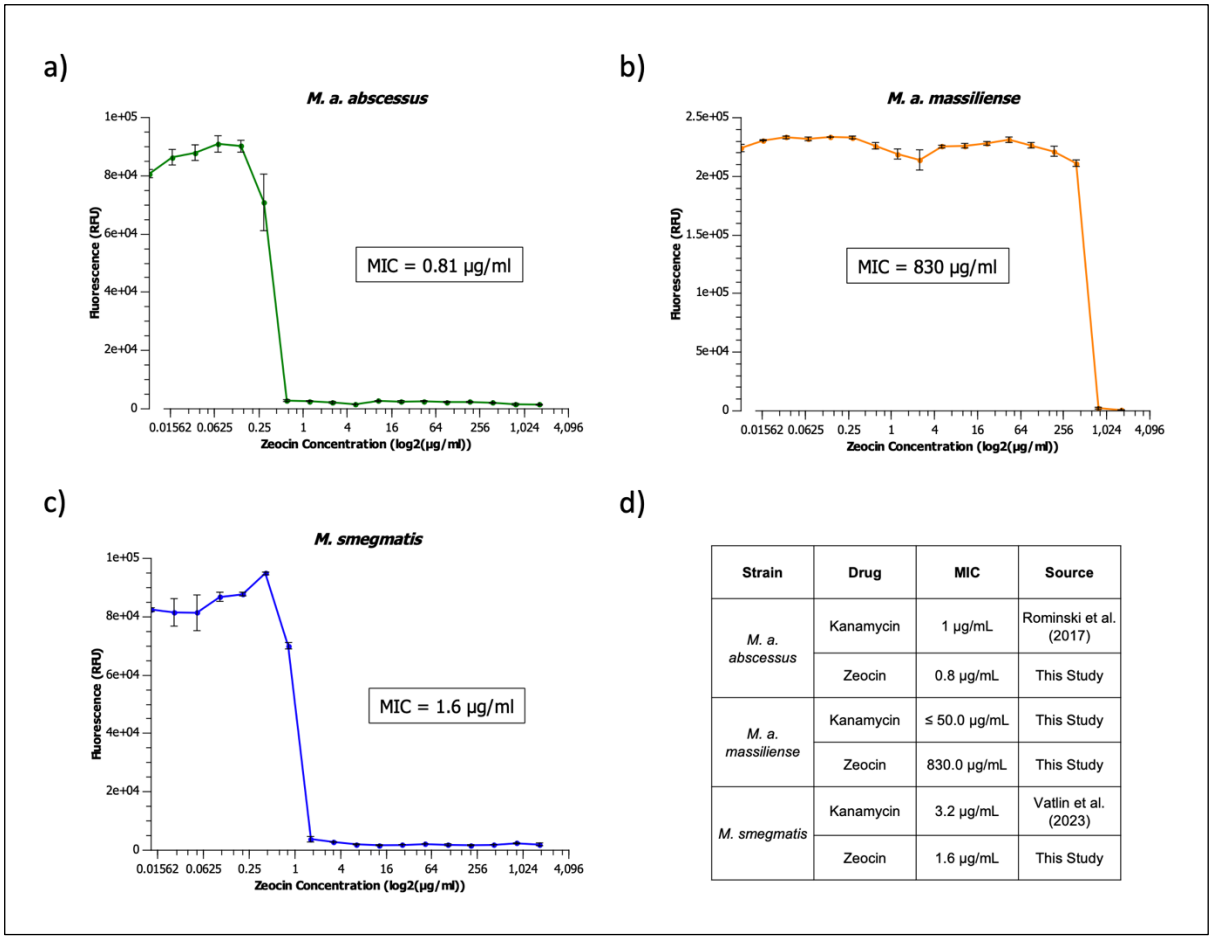
Prior to mutant generation, as a high concentration of phage is required for efficient transduction, the phage must be propagated. A titre of ≥ $5x10^{10}$ plaque forming units per mL (PFU/mL) is recommended (Siegrist and Rubin, 2009). The protocol outlined by Majumdar et al (2017) suggests that, following the incubation of phage with *M.*

*smegmatis* at the lytic-cycle inducing temperature, plates that produce a lacey plaque pattern can be gently rocked in MP buffer for a few hours at 4.0°C to produce a high titre of phage. Our initial attempts at propagating phage were conducted with *M. smegmatis* in ΦMycoMarT7 but unfortunately, we only achieved a concentration of $2.3x10^9$ PFU/mL. Therefore, further optimisation steps were conducted to increase phage titre. Through centrifugation with Pierce protein concentrators a concentration of $2.5x10^{10}$ PFU/mL was achieved but this significantly reduced the volume of phage and still failed to meet the target concentration. It has been suggested that rocking the plates for a longer period (24 h) and at varying temperatures can increase the concentration by allowing the media to shrink and expand thus releasing more phage (Bannantine et al., 2019). Through this method, outlined in Chapter 5, Section 1.2, we recovered a greater concentration of phage ($2.3x10^{10}$ PFU/mL) but again this was below the target titre. However, our highest yield was achieved by delaying the filtration of the extracted phage by 24 h. Through this adaptation, we successfully propagated both phage vectors to the required concentration. Interestingly, phDB36 had a much higher concentration ($1.8x10^{11}$ PFU/mL) than ΦMycoMarT7 ($5.7x10^{10}$ PFU/mL) suggesting that phDB36 may have greater lytic activity.

## 2.3 *M. abscessus* Strain Selection

Since the *Himar1* derived transposons carry either a Zeocin™ or kanamycin resistance cassette, to ensure that these were compatible, we next determined the minimum inhibitory concentrations (MIC) for each *M. abscessus* strain. Although previous reports have documented the MIC of kanamycin in *M. a. abscessus*

(Rominski et al., 2017), we were unable to find any data in *M. a. massiliense*. However, as we observed no WT growth on media containing 50 µg/mL of kanamycin we concluded that the MIC must be lower than or equal to this concentration and that this would be sufficient for isolating mutants. In contrast, WT *M. a. massiliense* growth was observed on media containing 50 µg/mL of Zeocin™. Therefore, as we could not find MIC data for either *M. abscessus* strain in the literature, we performed a resazurin-based fluorometric assay to confirm the MIC using *M. smegmatis* as a control. The assay, outlined in Chapter 5, Section 1.3, works on the principle that in actively growing cells, the oxidised, non-fluorescent resazurin dye will be reduced to a fluorescent resorufin dye. The level of fluorescence can then be measured to determine the proportion of live cells and MIC is determined by identifying the lowest concentration of drug that results in cell death. As observed in figure 1.14, in *M. a. abscessus* and *M. smegmatis,* Zeocin™ had an MIC of around 1 µg/mL. However, *M. a. massiliense* was extremely resistant. We therefore concluded that given the intrinsic Zeocin™ resistance, *M. a. massiliense* would not be compatible with phDB36 phage. Furthermore, although *M. a. massiliense* was susceptible to kanamycin and thus compatible with ΦMycoMarT7 we reasoned that having a library with resistance to both kanamycin and Zeocin™ could prove problematic for future study and therefore decided that *M. a. abscessus* would be the most suitable strain for library generation.




| Strain | Drug | MIC | Source |
|---|---|---|---|
| *M. a. abscessus* | Kanamycin | 1 µg/mL | Rominski et al. (2017) |
| | Zeocin | 0.8 µg/mL | This Study |
| *M. a. massiliense* | Kanamycin | ≤ 50.0 µg/mL | This Study |
| | Zeocin | 830.0 µg/mL | This Study |
| *M. smegmatis* | Kanamycin | 3.2 µg/mL | Vatlin et al. (2023) |
| | Zeocin | 1.6 µg/mL | This Study |

**Figure 1.14 Zeocin™ MIC.**

Resazurin assay to identify the Zeocin™ MIC determined via the mean + SD fluorescence across three biological replicates for (a) *M. a. abscessus*, (b) *M. a. massiliense* and (c) *M. smegmatis*. (d) A table of kanamycin and Zeocin™ MIC data for each strain based on data from the resazurin assay and literature.

## 2.4 Mycobacteriophage Selection

As *M. bovis* BCG and *M. a. abscessus* have a similar sized genome and GC-content we predicted that a *M. a. abscessus* library should generate a comparable level of saturation as a *M. bovis* BCG library of the same size. Therefore, as Borgers et al. (2020) achieved ~83% saturation of the non-essential *M. bovis* BCG genome from a library consisting of 18,432 mutants, we aimed to construct a *M. a. abscessus* library of the same size.

To generate such an extensive library, it was important that we selected a phage with a high transduction efficiency. To calculate this, as detailed in Chapter 5, Section 1.4, we incubated WT *M. a. abscessus* with either ΦMycoMarT7 or phDB36 at a variety of multiplicity of infections (MOIs) and recorded the number of successful transformants through growth on selective media. We discovered that ΦMycoMarT7 exhibited a significantly higher transduction efficiency (~2000x) indicating that, although phDB36 had a better capacity for propagation, its transposase was less active. We therefore concluded that ΦMycoMarT7 would be the most appropriate phage for generating the library.

## 2.5 Verification of Transposon Insertion

To validate the method and to confirm successful transposition of *M. abscessus* we selected three successful transformants with a variety of morphologies and conducted

PCR. Primers were designed to amplify the MycoMarT7 transposon kanamycin resistance cassette which, if present, should produce a 780bp fragment. PCR was also conducted on ΦMycoMarT7 and WT *M. a. abscessus* as positive and negative controls. As shown in figure 1.15, the PCR products were analysed by gel electrophoresis and all three transformants, along with the ΦMycoMarT7 positive control, exhibited a PCR product consistent with the expected fragment size. As no product was observed in the WT negative control, we concluded that the growth of transformants on selective media was sufficient for obtaining mutants with a successfully integrated transposon.

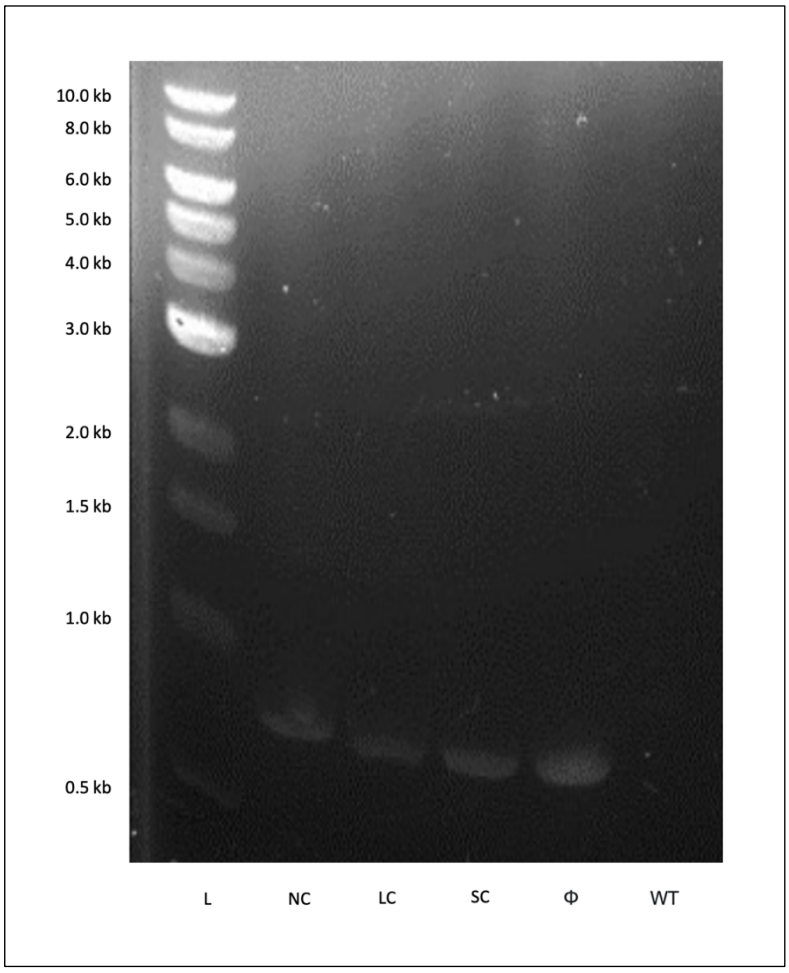



**Figure 1.15 Confirmation of *M. abscessus* mutants via PCR and gel electrophoresis.**

Gel electrophoresis of PCR products generated using primers specific to the kanamycin resistance cassette. L = 10 kb ladder, NC = normal sized ΦMycoMarT7 transduced *M. abscessus* colony, LC = large ΦMycoMarT7 transduced *M. abscessus* colony, SC = small ΦMycoMarT7 transduced *M. abscessus* colony, Φ = ΦMycoMarT7. ΦMycoMarT7 and WT were used as positive and negative controls.

## 2.6 Optimising High-Throughput Robotics for Library Construction

Picking 18,432 mutants into an ordered array is an extremely laborious process and is susceptible to human error which could result in cross-contamination. We therefore considered the use of colony picking robotics for library organisation. As the S&P Robotics BM3-BC robot boasts a 96-pin colony picking tool, we investigated the use of this machine for ordering a *M. a. abscessus* library.

As described in Chapter 5, Section 1.6, to ensure that colonies could be accurately selected, we prepared test plates containing between 100-400 well-spaced *M. a. abscessus* colonies by following a bead mixing method. By implementing the BM3-BC colony picking procedure, we were able to transfer 96 colonies into a 96-well plate containing growth supportive media with each 96-well plate taking less than 10 min to prepare (figure 1.16). Successful colony transfer was then confirmed by incubating the 96-well plates for a period of outgrowth. However, whilst this method represents an extremely fast and accurate method for picking colonies into an array, by analysing the robot colony selection criteria we identified a selection bias. Colonies that were

smaller than 1.4mm in diameter and exhibited a circular, smooth morphology were favoured over mutants exhibiting irregular growth (figure 1.16). Therefore, it was hypothesised that the robotic colony picking procedure would reduce the genomic saturation of the library by under-representing abnormal mutants.

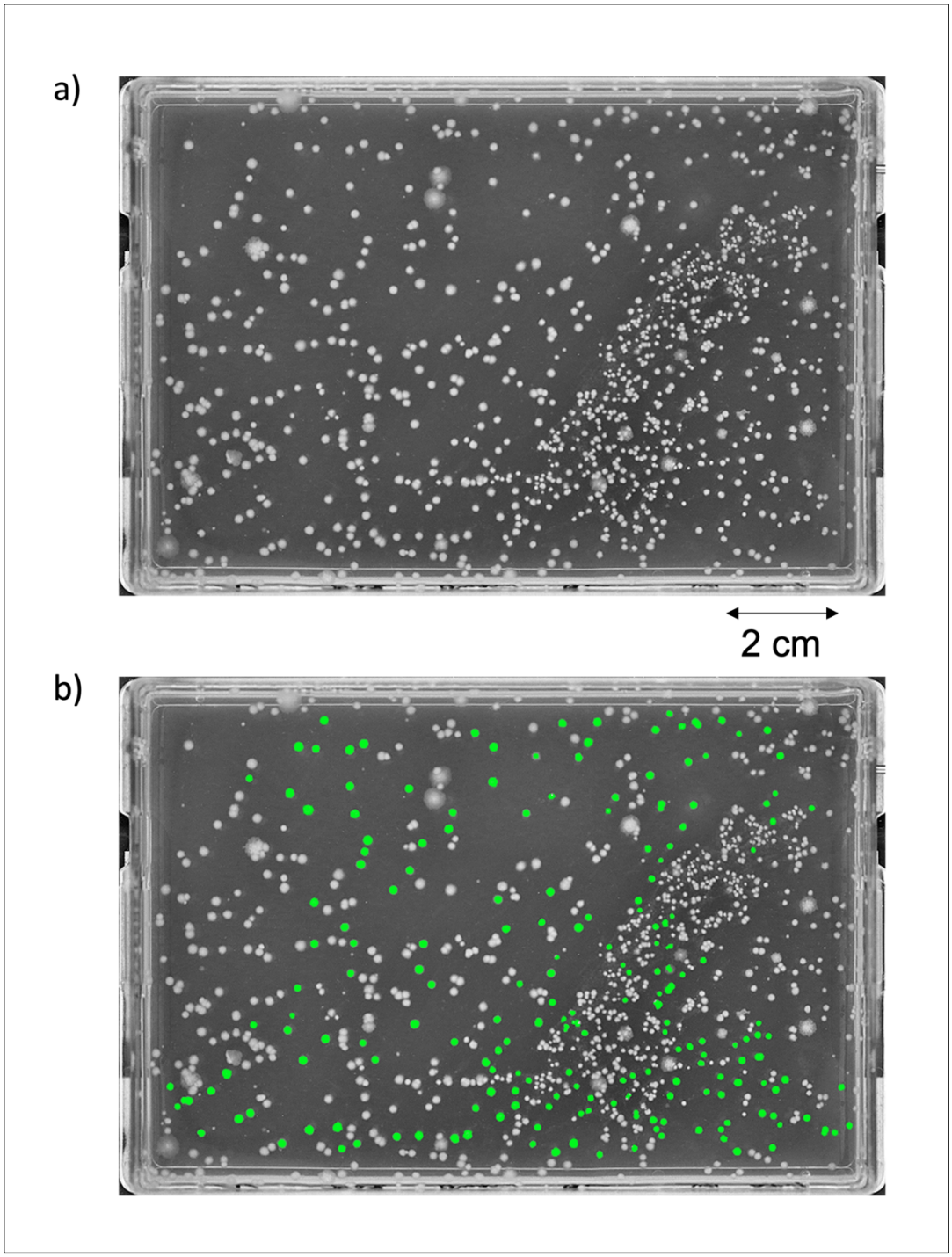



**Figure 1.16 BM3-BC colony picking selection.**

(a) Agar plate containing ΦMycoMarT7 transduced *M. abscessus* colonies prior to colony picking selection by the BM3-BC robot. (b) The BM3-BC robot selection of ΦMycoMarT7 transduced *M. abscessus* colonies (highlighted in green).

## 2.7 Manual Library Arraying

Given the selection bias of the robotic colony picking method, we instead decided that the manual picking of *M. abscessus* mutants into an ordered library was more suitable, albeit more laborious.

As detailed in Chapter 5, Section 1.7, we therefore generated a pooled *M. abscessus* mutant library by incubating ~10 x $10^{12}$ *M. a. abscessus* cells with high titre ΦMycoMarT7 at an MOI of 20 and isolated successful transductions on selective 7H9 agar plates. To maintain sterility, manual arraying of the library was conducted in BSL-2 laminar flow cabinets. The 96-well library plates were prepared by transferring mutants from transduction plates using sterilised pipette tips. With a total of 192 library plates, this method took around 3 weeks to complete resulting in the construction of a library consisting of 18,432 *M. a. abscessus* mutants. To minimise cross-contamination and selection bias, the mutants were picked with caution and at random, ensuring that only isolated colonies with a range of morphologies were selected. To facilitate growth, the library plates were incubated with breathable seals and plates were regularly assessed for contamination, though no contamination was detected. Finally, as described in Chapter 5, Section 1.8 the library was replicated using a Biomek i5 Automated Workstation with an integrated high-efficiency particulate absorbing (HEPA) filtration system to prevent contamination and stored at -80°C in 25% glycerol.

To confirm that the library could be recovered from freeze-storage, a single library plate was replicated into an intermediate plate containing enriched 7H9 media and incubated until visible growth was observed. Using the BM3-BC robot, the intermediate plate was then pinned onto enriched 7H9 agar and imaged after 40 h of incubation. As demonstrated by figure 1.17, the mutants exhibited a range of growth phenotypes. This diversity was anticipated as genetic disruptions are expected to result in a change in fitness.

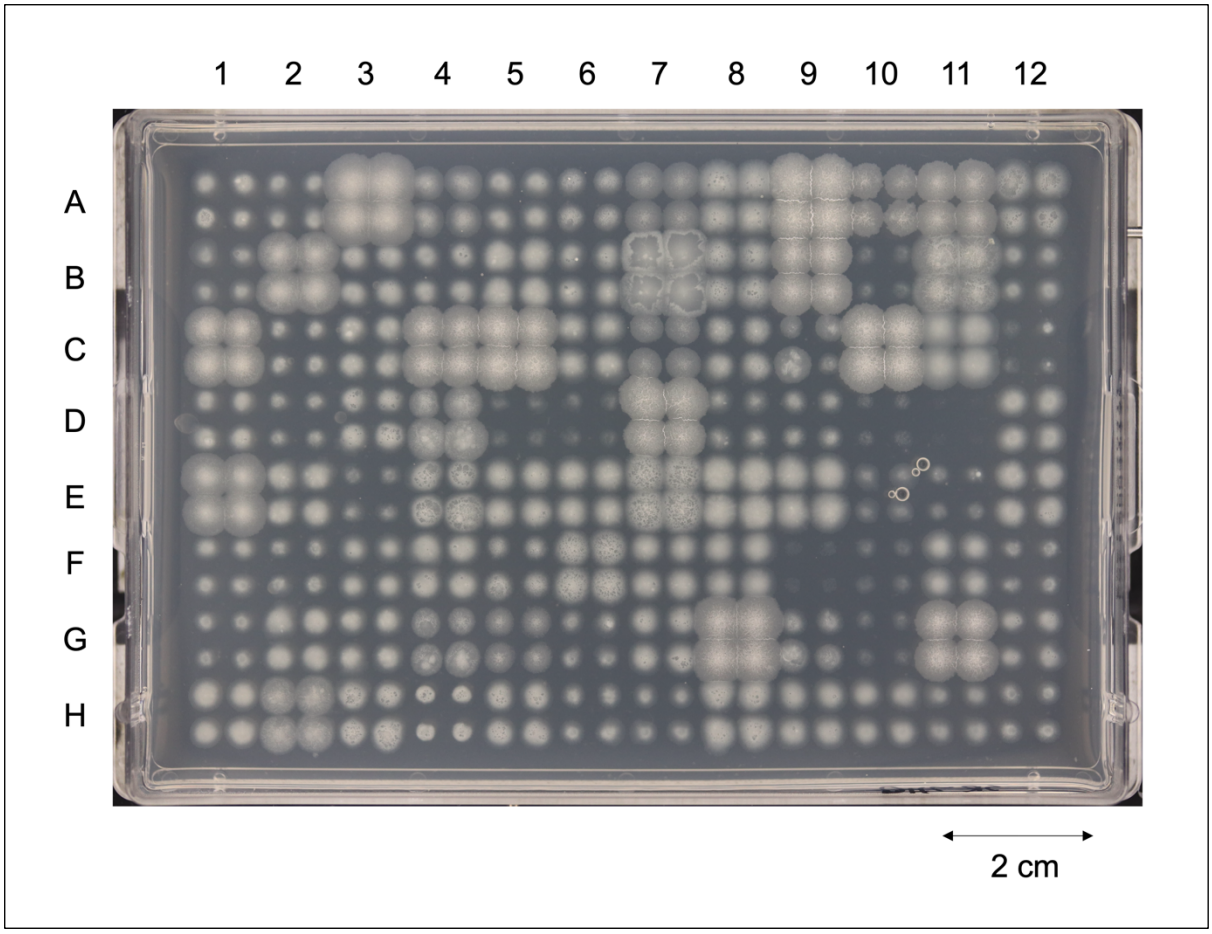



**Figure 1.17 Verification of *M. abscessus* mutant growth differences.**

A *M. abscessus* library plate replicated onto 7H9 agar in a 384-colony format with each quadrant representing four replicates of a single library strain. The plate was incubated for 3 days at 37°C before imaging using the BM3-BC robot.

## 2.8 Cartesian Pooling for Coordinate Sequencing

Given that we had access to liquid handling robotics, for increased throughput and experimental accuracy as well as reduced human error, we implemented the CP-CSeq method (Vandewalle et al., 2015) for deconvoluting our *M. a. abscessus* library. As we constructed a 192-plate library, we adapted the CP-CSeq method (Vandewalle et al., 2015; Chapter 2, Section 1.3) as outlined in Chapter 5, Section 1.9 by generating two XY and Z master plates for a total of 40 sequencing pools. Briefly, the XY master plates were prepared as previously described by pooling every mutant at each XY coordinate into the corresponding XY coordinate of the master plate. As this procedure followed the same pattern as the library replication we conducted this pooling step using the Biomek i5 Automated Workstation. In contrast, to reduce costs, we used Opentrons OT-2 liquid handling robots to generate the Z master plates. Z master plates were also prepared as previously described by creating separate pools for all of the mutants in each library plate. However, as the OT-2 robot does not have 96 channel pipettes, to achieve this we implemented an alternative pooling technique in which the contents of each library plate was first pooled into a column of an intermediate 12-column plate before being transferred into a single well of the Z master plates (figure 1.18). Lastly, the columns and rows of all four master plates were pooled to generate the final 80 sequencing pools.

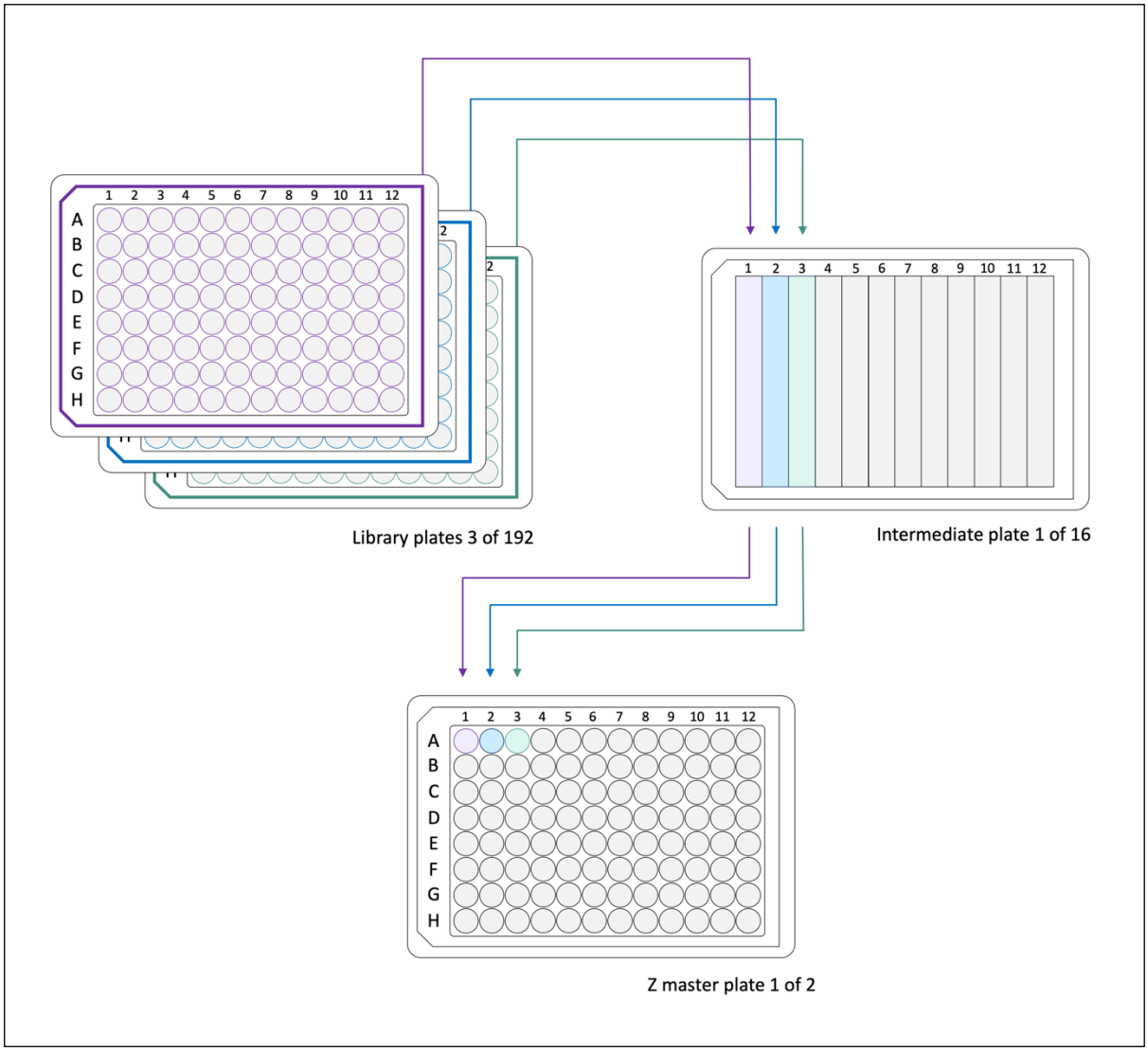



**Figure 1.18 Z master plate pooling method.**

The contents of each library plate were pooled into a single column of an intermediate plate. To generate the Z master plate, each column was then transferred to a single well of the Z master plates.

## 2.9 Validating the Sequencing Method

Before sequencing the entire library, to validate the sequencing method we first generated a small test library consisting of 384 mutants arrayed into four 96-well plates and thus generated 20 XY and four Z sequencing pools. Therefore, depending on

whether the pool originated from either a row or a column, we expected each XY sequencing pool to contain a maximum of 48 or 32 unique mutants respectively. As the Z master plate was organised into wells A1, A2, B1, B2, for each Z sequencing pool we expected a maximum of 192 unique mutants.

To prepare the pools for sequencing (Chapter 5, Section 1.11), gDNA was extracted and then fragmented via sonication. The DNA was then end repaired and dA-tailed to prime the fragments for adapter-ligation. In end repair, the DNA fragments are blunted via a T4 DNA polymerase followed by phosphorylation of the 5' region. In dA-tailing, through the assistance of a terminal transferase, the 3' region is primed with dATP. Next, through the addition of NEBNEXT Adapters for Illumina®, adapters were ligated. However, as these adapters attach as a closed loop, through the activity of two enzymes, uracil DNA glycosylase and endonuclease VIII, the loop is cleaved by removing a uracil residue and the adapters are activated. Finally, a two-step PCR method adapted from the TraDIS pipeline was conducted to enrich for transposon containing fragments and prepare the samples for Illumina® (Goodall et al., 2018). In the first step, using primers specific to the transposon and the adapter, fragments containing a transposon insertion were amplified. Due to the presence of identical inverted repeats and a T7 promoter at each pole of the MycoMarT7 transposon, transposon specific primers were designed to target a unique sequence just upstream from this region. In the second PCR step, P5 and P7 Illumina® sequences were appended to the enriched PCR fragments permitting attachment to the Illumina® flow cell for sequencing. In this case, primers were designed to target the inverted repeat to maximise the length of the gDNA sequencing reads (figure 1.19). Following

sequencing of an XY master plate column pool, 32 unique mutants were identified thus validating the method.

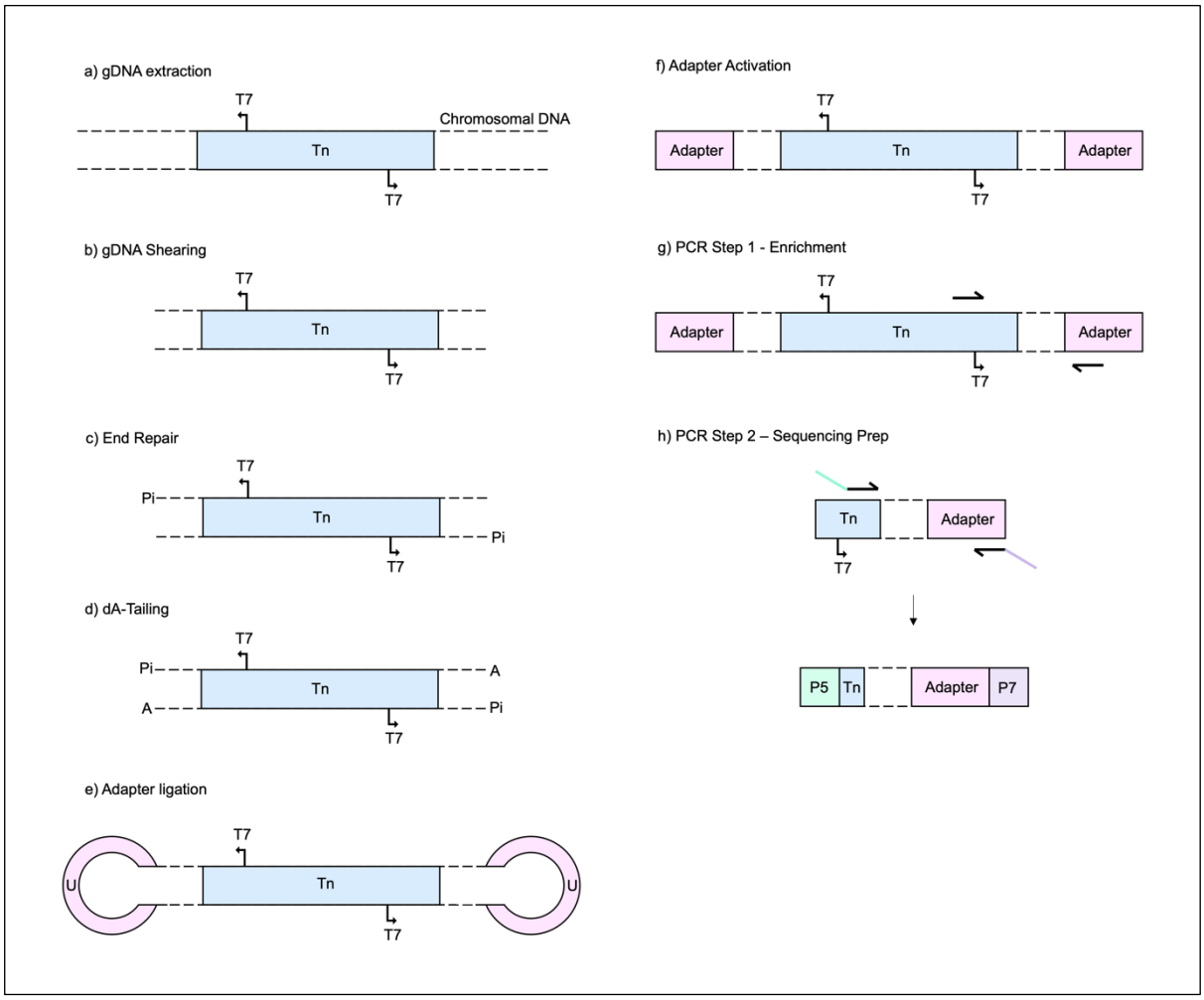



**Figure 1.19 Method for sequencing preparation.**

(a) gDNA is extracted and (b) sheared via sonication. DNA fragments are primed for adapter ligation through (c) end repair and (d) dA-tailing. (e) Adapters are ligated onto the fragments and (f) activated by removing a uracil residue. (g) Fragments containing transposon insertions are amplified and (h) prepared for Illumina® sequencing via PCR. T7 = T7 promoter, Tn = MycoMarT7 Transposon, U = Uracil.

# Part 3 – Discussion

## 3.1 Completion of the *M. abscessus* Library

Unfortunately, due to time constraints, we were unable to finish the deconvolution of the *M. abscessus* library and thus, the library currently exists as a collection of unannotated mutants. Therefore, to complete the library and thus maximise its utility, additional steps must be conducted. Firstly, the 80 library sequencing pools prepared via the CP-CSeq method must be sequenced and mapped to the genome to identify the transposon insertion sites. Then, to decode the library, the mutant locations must be identified using a deconvolution algorithm, similar to that employed by Borgers et al. (2020).

Following library completion, the library should also be validated. For example, the library saturation could be determined by identifying the number of unique insertions. A library saturation ≥ 83% would be considered extremely successful as this would match or exceed the saturation of the *M. bovis* BCG library generated by Borgers et al., (2020). In addition, to confirm that the mutant location assignment was accurate, a set of mutants from the library could be sequenced. This could be achieved via whole genome sequencing (WGS) which, although costly, could also help to identify any irregularities in the library such as multiple transposon insertion events. In contrast, as a more cost-effective alternative, a short length of transposon-adjacent DNA could be sequenced or amplified via PCR and identified via gel electrophoresis.

## 3.2 Considerations for Library Generation

Constructing a unique, comprehensive and ordered mutant library is an extensive process and thus we carefully considered the various aspects of library generation to ensure that the most efficient method was employed whilst maintaining library functionality and practicality.

### 3.2.1 Library Size

Although a gene-by-gene approach would facilitate the construction of a more complete mutant library, since the method requires the design of a unique DNA construct for each gene, it is extremely time consuming. In contrast, random transposon mutagenesis presents a rapid method for library generation that is well suited to high-throughput functional genomics and is still capable of producing mutant libraries with a high level of saturation (Borgers et al., 2020). Furthermore, gene-by-gene approaches could be used in combination with random transposon mutagenesis to complete resulting libraries and ensure that every gene is targeted for deletion. Therefore, we concluded that it would be most efficient to construct a transposon insertion library in *M. abscessus*. Through Coupon Collector's Problem, Borgers et al. (2020) previously deduced that generating a *M. bovis* BCG library larger than ~20,000 mutants would be ineffective as the library saturation begins to plateau. As *M. bovis* and *M. abscessus* have comparable genomes, we therefore determined that the generation of 18,432 mutants would be sufficient for constructing a *M. abscessus*

mutant library with ~80% saturation of the non-essential genome and that a library of this size would be both time- and cost-effective.

### 3.2.2 Phage Delivery System

The most efficient methods for constructing mycobacterial transposon insertion libraries have employed phage as transposon delivery systems. As various phage and transposon systems have been developed, by measuring the propagation and transduction efficiency, we determined the best candidate for generating an *M. abscessus* library. We considered two phage systems, phDB36, that has been previously employed to generate *M. abscessus* transposon insertion libraries, and ΦMycoMarT7, the most commonly used phage for mycobacterial TnSeq experiments.

As we obtained a phDB36 titre approximately 3x the concentration of ΦMycoMarT7, we concluded that, likely due to a stronger lytic activity, phDB36 could be propagated more efficiently (Chapter 2, Section 2.2). However, the plaque assay method that we employed for titering phage has limited accuracy. For example, as we included Tween-80 in the *M. smegmatis* growth media to prevent clumping, it is possible that the differences observed in phage titre could be due to the ineffective removal of detergent prior to incubation with phage, though, this was controlled via thorough washing of cells prior to infection. In addition, it is well documented that phage will decay over time with different species decaying at varying rates (Weld et al., 2004). For reliable results, it is therefore essential that the phage titering is conducted directly after propagation. Furthermore, it has been previously reported that the host strain can

significantly alter propagation efficiency (Lin et al., 2012). To streamline the phage propagation method, we used *M. smegmatis* as the host given that it has relatively fast growth and is categorised as a BSL-1 pathogen. However, it is conceivable that the propagation of ΦMycoMarT7 may have been more efficient in an alternative mycobacterial species. In contrast, a recent study employed a spectroscopic approach called dynamic light scattering for a more accurate measure of phage concentration (Dharmaraj et al., 2023). Not only does this method increase throughput but it can also quantify changes in phage titre and bioactivity over time. Therefore, it is expected that spectroscopic techniques will soon succeed plaque assays as the standard procedure for quantifying phage and thus should be considered for constructing future transposon insertion libraries.

Nevertheless, although phDB36 was more easily propagated, ΦMycoMarT7 demonstrated an enhanced transduction efficiency, producing around 2000x the number of successful transformants (Chapter 2, Section 2.4). Therefore, as a considerably larger quantity of phDB36 phage would be required to generate a library, we selected ΦMycoMarT7 for the *M. abscessus* library construction. That said, as we only considered two phage systems it is conceivable that other systems could be equally or even more efficacious. For example, Borgers et al. (2020) used phAE159 to deliver a large *Himar1*-derived transposon with increased efficiency in *M. bovis* BCG. This system was designed to facilitate the creation of markerless gene knockouts by incorporating FRT sites and I-SceI sites into the transposon. In the FRT system, markerless mutants were generated through FlpE/FRT-mediated recombination. The method is based on the principle that upon the activation of the

FlpE recombinase, The FRT direct repeats recombine via HR, deleting the DNA in between (Datsenko and Wanner, 2000). This method was extremely efficient but could not be used to generate a clean KO due to the creation of a scar upon the removal of the transposon (Borgers et al., 2020). In contrast, the I-SceI system has the potential to create clean deletions. In this method, activation of the meganuclease SceI results in double strand breaks at the I-SceI sites which is then repaired via NHEJ (Siegl et al., 2010). Depending on the degree of repair a variety of scars can be created, though, it was discovered that around 0.4% of unmarked mutants resulted in a clean KO (Borgers et al., 2020). As transposon insertions can cause polar effects, a tool for unmarking an entire transposon insertion library represents an extremely exciting prospect which would enable the functional analysis of mutants with improved confidence.

### 3.2.3 Strain Selection

Due to an intrinsic Zeocin™ resistance (figure 1.14), we concluded that *M. a. massiliense* was not suitable for library construction as this would result in the generation of mutants with resistance to both Zeocin™ and kanamycin. This would therefore limit future study as many mycobacterial genetic tools use kanamycin or Zeocin™ antibiotic selection markers. Furthermore, whilst *M. a. bolletii* could be a suitable strain for library generation; this subspecies was not considered as it is the least clinically relevant and is not widely researched. Therefore, given that *M. a. abscessus* was susceptible to both kanamycin and Zeocin™ and was compatible with

ΦMycoMarT7 we decided that *M. a. abscessus* would be the most appropriate strain for library construction.

### 3.2.4 Validating Transposon Insertion Mutagenesis

To verify MycoMarT7 transposition we conducted PCR followed by gel electrophoresis on three kanamycin resistant transformants. In an attempt to reflect the diversity of the population, we selected three mutants with a variety of colony morphologies. However, whilst the electrophoresis gel demonstrated that transposition was successful (figure 1.15), as this method cannot be used to identify the location of the transposon insertion, it is possible that the selected colonies originated from the same mutant. Furthermore, the sample size was too small to ascertain whether this was representative of the entire population. Indeed, it has been previously discovered that spontaneous kanamycin resistant mutants represent between 1-3% of the population (Rifat et al., 2021). We can therefore predict that in our library between 180-550 mutants do not contain a transposon insertion. Nevertheless, given that only a small portion of the library is affected, the library saturation should be relatively unchanged. In addition, as library deconvolution involves a transposon enrichment step, spontaneous mutants could be later identified and removed. Given the limitations of the PCR method, for more accurate and reliable validation of transposition, we therefore suggest a transposon sequencing method on a much larger sample size (> 100 mutants).

## 3.2.5 Colony Picking via High-Throughput Robotics

To order the *M. abscessus* library into 96-well plates we explored the use of robotics. The BM3-BC robot has a built-in colony picking tool that has been previously successful in transferring both *E. coli* and *Saccharomyces cerevisiae* colonies into well plates (Filteau et al., 2015; McCready et al., 2019). However, given the irregular growth of mycobacteria, coupled with the range of mutant morphologies, adapting this method for arraying an entire *M. abscessus* mutant library proved challenging. Although colonies were successfully transferred to 96-well plates we discovered a selection bias that favoured small, circular colonies of the smooth morphotype and thus opted for a manual colony picking method for library arraying. This bias was a result of the BM3-BC default selection criteria which is based on colony opacity, circularity and size. However, with further optimisation, this selection criteria could be altered to enable selection of larger colonies with irregular growth and thus eliminate the selection bias. We therefore suggest that this method should be considered for the generation of future libraries, offering a less labour intensive, increased throughput and more accurate alternative to manual arraying.

## 3.2.6 Validation of Library Growth

To confirm that the *M. abscessus* library could be grown from frozen stocks and to demonstrate the variation in mutant growth, we replicated a library plate onto 7H9 agar. As seen in figure 1.15, a variety of mutant phenotypes could be observed. However, we were unable to quantify these differences. Previous studies have

employed imaging analysis technologies to measure colony morphology such as IRIS covered in Chapter 1, Section 2.12 (Kritikos et al., 2017), but unfortunately, due to the plate overgrowth, these methods could not be applied to this dataset as these tools require the clear separation of colonies. This method could therefore be improved by reducing the incubation time. Nevertheless, this technique demonstrates the utility of the library, as defining mutants by their resulting phenotypes can provide insight into gene function. For example, by employing similar methods, Pawlik et al. (2013) characterised the genetic basis behind variation in *M. abscessus* morphotypes and Clary et al. (2018) discovered genes involved in biofilm formation. Furthermore, this data could be coupled with the analysis of mutant growth in liquid culture by defining growth curves for each mutant.

### 3.2.7 Media Selection

To create the most complete library, we considered which media would be the most supportive for mutants with a range of growth defects. 7H9 is the standard liquid medium for cultivating mycobacteria which is extremely nutrient rich containing a variety of vitamins and metal salts beneficial for mycobacterial growth. To generate the *M. bovis* BCG library, Borgers et al. (2020) adapted the 7H9 medium to also include additional vitamins to facilitate growth of auxotrophic mutants. We therefore formulated a similar, more cost-effective medium, coined enriched 7H9 medium, to promote the growth of mutants with metabolic defects. In contrast, for solid medium, as mycobacteria have extremely slow growth and are thus susceptible to contamination, 7H10 is usually used. 7H10 is largely comparable to 7H9 with the

addition of malachite green representing the major difference. Although malachite green has antifungal and antibacterial properties, mycobacteria are not susceptible. However, as various studies have reported the sensitivity of mycobacterial mutants to malachite green, particularly those involved in maintenance of the cell envelope (Banaei et al., 2009; Bianco et al., 2011; Kandasamy and Narayanan, 2015), we decided to exclude malachite green from our media and thus used enriched 7H9 agar instead.

During the arraying of the library we discovered that a small proportion of mutants did not grow in liquid medium. As the arraying of the library was done manually, it is plausible that certain wells were not inoculated due to human error. However, as it has been previously reported that some mycobacterial mutants cannot grow on solid medium (Borgers et al., 2020), it is conceivable that the reverse is also true. Furthermore, it is possible that spontaneous kanamycin resistant mutants were only able to survive in solid medium as agar can sequester antibiotics and thus reduce their potency. For future library generation, a flow cytometry technique could be employed to ensure that the library is maintained in liquid medium throughout its construction. For example, Shiver et al. (2021) developed a method for ordering transposon insertion libraries in anaerobic bacteria using flow sorting to transfer individual cells from a mutant pool into 96-well plates. Therefore, this method not only negated the need for library outgrowth on selective solid media, it also presented a high-throughput automated method for ordering libraries, ensuring that only a single mutant was present in each well. As manual and robotic colony picking is susceptible to error in

which more than one mutant could be picked into a single well, library construction via flow sorting represents a marked advantage.

### 3.2.8 Method for Library Deconvolution

Given that CP-CSeq (Vandewalle et al.,. 2015) had been previously employed for the generation of similarly sized mycobacterial libraries and since we had access to liquid handling robotics, to increase the experimental throughput, we opted to deconvolute the library using this method. However, as not all mutants within a library are unique, CP-CSeq cannot confirm the precise location of every mutant. Indeed, in the Borgers et al. (2020) *M. bovis* BCG library, only ~62% of the mutants were mapped to a unique position. For the remainder of the library a heuristic algorithm was applied to predict mutant locations. In contrast, the Knockout Sudoku method presents an alternative method which simplifies the pooling process by negating the need for robotics and reducing the required number of sequencing pools (Baym et al., 2016). Furthermore, it is reported that Knockout Sudoku has an increased accuracy in library deconvolution. Indeed, from a 39,918 *Shewanella oneidensis* mutant library it was discovered that 80% of the mutants could be mapped to a library location. However, 25% of these locations were multiply occupied and thus not unique (Baym et al., 2016). It would therefore be interesting to do a direct comparison between the two methods.

In addition, since we used liquid handling robots with 8-channel pipettes, an additional intermediate pooling step was required for the generation of the Z master plates. To increase the experimental throughput, we would therefore recommend that the pooling

steps are conducted using liquid handlers with 96-channel pipettes as outlined in previous studies (Baym et al., 2016; Borgers et al., 2020).

## 3.3 Limitations of Transposon Libraries

As previously covered in Chapter 2, Section 1.2, although transposon insertion mutagenesis presents a high throughput method for generating large loss-of-function libraries, unlike clean KOs generated via gene substitution, transposon insertions do not guarantee loss of function. That said, mutants with insertions closer to the initiation codon can be selected to ensure that the largest proportion of the gene is disrupted. Additionally, the insertion of transposons within an operon can result in polar effects, in which genes downstream of the insertion site are also disrupted (Van Opijnen et al., 2009). The use of transposons that do not contain transcriptional terminators should reduce polarity though this is not assured. Therefore, an additional screening could be conducted to measure the degree of polarity. For example, Van Opijnen et al. (2009) were able to assess the level of polarity in a *Streptococcus pneumoniae* transposon insertion library by comparing the fitness of operonic mutants to the fitness of downstream genes with severe growth defects. Furthermore, the MycoMarT7 transposon inserts at TA dinucleotides and, whilst these sites are found regularly across the genome, these sites may not be targeted in genes with very few TA sites or in short genes. Some genes are also relatively resistant to transposition due to the coiling of DNA and thus transposon insertion mutagenesis has an inherent bias. Lastly, as transposition results in a loss of function, essential genes cannot be targeted. For the analysis of essential genes, overexpression libraries could be generated with Dub-

Seq representing the most appropriate overexpression library construction method which is compatible with high-throughput functional genomic screening (Mutalik et al., 2018).

## 3.4 Applications of the *M. abscessus* Mutant Library

### 3.4.1 Collection of Mutants

Although *M. abscessus* is an emerging pathogen in the clinical setting, especially in association with respiratory disease, it is still poorly understood and existing genetic resources are limited (Johansen et al., 2020). Furthermore, mycobacteria are notoriously difficult to genetically engineer, and although more efficient methods have been recently developed such as CRISPR and ORBIT (Jinek et al., 2014; Murphy et al., 2018), the generation of mutants remains challenging and time-consuming. Therefore, once complete, our *M. abscessus* library should represent a significant asset to the mycobacterial research community, boasting the largest collection of *M. abscessus* mutants which can be selected for further genetic study to help functionally annotate the remaining genome. This library will therefore add to the limited but growing number of mycobacterial resources such as the Borgers et al. (2020) *M. bovis* BCG transposon insertion library and the Tuberculosis Animal Resource and Gene Evaluation Taskforce (TARGET) program which catalogues a collection of *M. tuberculosis* loss-of-function mutants.

The characterisation of the *M. abscessus* genome will not only better our understanding of these bacteria but could also lead to the discovery of new drug targets, especially if new genes involved in virulence, drug resistance and important cellular processes can be identified. Furthermore, given that *M. abscessus* has intrinsic resistance to many anti-tuberculosis drugs, and since *M. abscessus* and *M. tuberculosis* share a high degree of sequence similarity, identifying *M. abscessus* resistance pathways could also help to predict mechanisms of *M. tuberculosis* drug resistance before they occur clinically.

A common method for discovering the function of genes is through reverse genetics in which loss-of-function mutants are generated and phenotypically characterised with comparison to WT and, often, to a complemented strain. An ordered loss-of-function library is therefore extremely useful for reverse-genetics research as mutants can be individually selected from the library thus negating the need to construct loss-of-function strains, saving time, expense and resources.

### 3.4.2 Functional Genomics

Loss-of-function libraries can also be used in high-throughput functional genomic screening. In contrast to methods using pooled libraries, such as TnSeq, screening of arrayed libraries enables a greater diversity of data output, beyond sequencing data, in which entire libraries can be assayed in parallel but mutant phenotypes are measured individually.

As covered in Chapter 1, Section 2.12 ordered libraries can be used in combination with transcriptomics and metabolomics platforms to determine gene function but perhaps the most straightforward method is through bacterial growth kinetics in which the OD or CFU of each mutant is determined under a variety of conditions. Not only can this provide insight into conditional gene essentiality, vulnerability and lethality but gene function can also be predicted by constructing a phenotypic profile for each mutant in which, via clustering analysis, genes of a similar function are likely to cluster together (Sher et al., 2020). Similarly, mutant phenotypes can also be compared through imaging technologies, in which, following a period of outgrowth on a set of conditions, cell morphology or colony growth can be analysed. For example, a computational tool called IRIS, developed by (Kritikos et al., 2017), has been used in multiple functional genomic screens and can measure various aspects of colony growth through the imaging of arrayed agar plates.

# Part 4 – Conclusion

Ordered loss-of-function mutant libraries represent a valuable resource, providing a catalogue of mutants which can be individually selected for further study. In addition, when combined with high throughput omics platforms, these libraries can be used for functional genomics, helping to phenotypically profile and thus functionally characterise entire bacterial genomes. At current, random transposon mutagenesis represents the most efficient method for generating non-essential mutant libraries. In this technique, transposons are randomly integrated into the bacterial genome and due to the insertion of a premature stop codon, can result in gene disruptions (Borgers et al., 2020). To generate an ordered transposon library, mutants must be individually picked into an array and then identified via sequencing. To increase the throughput of this sequencing step, two techniques called Knockout Sodoku (Baym et al., 2016) and CP-CSeq (Vandewalle et al., 2015) were developed. In these methods, the library is divided into a series of sequencing pools and then deconvoluted using a computational algorithm.

As *M. abscessus* is an emerging clinical pathogen that is poorly understood and has limited genetic resources (Johansen et al., 2020), we generated an ordered *M. abscessus* mutant library through random transposon mutagenesis. We assessed the suitability of two phage systems for transposon delivery, phDB36 and ΦMycoMarT7, as well as two *M. abscessus* subspecies, *M. a. abscessus* and *M. a. massiliense*. Due to a significantly higher transduction efficiency, we selected ΦMycoMarT7 for the

construction of the library and as *M. a. massiliense* had an inherent resistance to Zeocin™, to avoid generating mutants with multi-drug resistance, we chose *M. a. abscessus* as the genetic background of the library. In total, the library consists of 18,432 mutants manually arrayed into 96 well plates, though, we also developed a high-throughput automated method for library arraying. To reduce human error and increase throughput we used the Cartesian Pooling method for library deconvolution and generated 80 sequencing pools. However, the library is yet to be decoded as we were unable to complete the CP-CSeq method. Nevertheless, once finished, our *M. abscessus* library will boast the largest collection of *M. abscessus* mutants to date, and, as seen through the construction of other bacterial mutant libraries (Baba et al., 2006), should greatly accelerate the functional annotation of the *M. abscessus* genome.

# Chapter 3

# A Functional Genomic Screening of *Mycobacterium bovis* BCG

# Part 1 – Introduction

## 1.1 Functional Genomics using Arrayed Mutant Libraries

While functional genomic screening via sequencing techniques such as TnSeq and Dub-Seq are comparatively quick, the output is limited as mutant fitness can only be measured by the relative abundance of transposons insertions or DNA fragments from sequencing data. In contrast, functional genomic screening of arrayed libraries enables the analysis of various aspects of mutant growth and thus enables a wider phenotypic characterisation.

Although mutant fitness of arrayed libraries can be determined via growth kinetics, like sequencing techniques, the output is still limited as only conditional alterations in growth can be monitored. Similarly, whilst mutant metabolomic or transcriptomic variations could also be analysed, only a single phenotypic measure is assessed. Alternatively, microscopy techniques have been employed to determine changes in mutant phenotypes at the cellular level. However, whilst this method has the potential to function as a highly sensitive technique for detecting minor difference between different mutant strains, it is relatively low throughput and is restricted to analysis in a 96-arrayed format. Instead, by examining colony growth, multiple phenotypic alterations can be evaluated such as colony size, shape and circularity. Furthermore, since the development of IRIS imaging analysis software (Kritikos et al., 2017), colony

growth can be analysed in much greater depth providing an even larger diversity of data output.

## 1.2 IRIS is a High Throughput Tool for Diverse Analysis of Colony Growth

IRIS is able to process colonies arrayed up to a 6,144-format and thus has the capacity for extremely high throughput. In addition, although IRIS has never been used for analysis of mycobacterial colony growth, the tool can be applied to a wide range of bacterial species and even fungi. The first step of IRIS analysis is to normalise the images of arrayed colonies, this is achieved by rotating and cropping the image, ensuring that the array is in a rectangular format. So that colonies can be analysed individually, colonies are then segmented into tiles by generating a grid. The grid is based on an average of the X and Y coordinates for each colony centre in which, colony centres are defined as the regions with the greatest level of pigmentation. Finally, colony boundaries are detected as a measure of brightness.

Not only can IRIS analyse colony size, shape and circularity, but colony opacity, colour and morphology can also be measured. As colony size, shape and circularity are measured based on pixel values within the defined colony boundary, Kritikos et al. (2017) reported that opacity can be a more sensitive measure of colony fitness as it can be used to interpret the 3-dimensional growth. This is achieved by analysing colony brightness in which raised colony regions will appear brighter and thus the

height of colonies can also be accounted for. Furthermore, colony opacity can be used to determine colony density by dividing the opacity score by the cell size score, providing an even more detailed analysis of colony growth. To determine colony morphology, IRIS analyses the colony wrinkles by identifying brightness peaks and valleys originating from the colony centre. This could be a particularly effective measure for mycobacteria as many species demonstrate wrinkled growth. In addition, some species such as *M. abscessus* switch between a rough and smooth morphotype which has been linked to variations in virulence (Pawlik et al., 2013).

IRIS can also be combined with growth assays to determine further growth phenotypes such as biofilm formation and membrane permeability. For analysis of biofilm formation, mutants can be dyed with Congo Red and Coomassie Blue in which changes in colour are detected to determine the presence of biofilms (Kritikos et al., 2017). Similarly for membrane permeability, mutants can be grown with chlorophenyl red-β-D-galactopyranoside (CPRG). CPRG can only penetrate the membrane in cells with defective membrane permeability and once in the cytoplasm, CPRG is degraded via LacZ to chlorophenyl red producing a red colour (Paradis-Bleau et al., 2014). Therefore, by detecting changes in colour, membrane permeability can also be assessed. Whilst this method cannot be adapted to mycobacteria a similar assay could be designed.

In addition, an auxiliary Kinetic Data Companion tool can be added to the IRIS repertoire which enables the evaluation of growth kinetics in which colony growth is measured over time (Kritikos et al., 2017). In comparison to static colony growth

analysis, kinetic data is more sensitive to small changes in fitness and is less susceptible to sample noise or inaccurate readouts due to variations in inoculation and asynchronous growth.

## 1.3 ChemGAPP

Whilst IRIS is a useful tool for analysing colony growth it is not able to generate mutant fitness scores and, although chemical-genomic studies have increased in popularity, it is only recently that a program for converting colony growth data has been established: ChemGAPP (Doherty et al., 2023) is an open source platform that can normalise, quality control and transform colony growth data, including IRIS data, to growth fitness scores.

As chemical-genomic screening generates large volumes of colony growth data it is subject to experimental noise. In addition, edge effects represent a major issue for analysing arrayed colony growth in which, colonies grown at the outer plate edges demonstrate enhanced growth as a result of increased space and thus availability of nutrients (Doherty et al., 2023). To account for this, ChemGAPP implements a normalisation process. In the first normalisation step, any zero values that are not reproducible across each replicate are removed as this likely signifies there was an issue during the colony arraying. Next, using a Wilcoxon rank sum test, colony growth at the outer plate edges are compared to colony growth at the plate centre. If edge effects are present, the colony growth values are scaled so that the median values for the outer edge rows and columns are equal to the plate middle mean. Finally, the

plates are globally scaled which is an important process for accurately deciphering the mutant fitness scores. This is achieved by adjusting the median of all of colony growth values within the dataset to the plate middle mean.

Other common errors that arise during chemical-genomic screening are the mislabelling of plates, the inversion of images, unequal or mis-arraying and the generation of artefacts during colony growth analysis (Doherty et al., 2023). Therefore, ChemGAPP also implements four quality control measures.

A Z-score test controls for mis-arrayed colonies or undetected growth by differentiating between no growth and colony growth artefacts. In this analysis, replicates in each condition are compared to identify missing values or outliers and a percentage of normality for each plate is then calculated. A Mann-Whitney test accounts for unequal colony arraying and mislabelled plates by determining the distribution of growth values between replicates. If the data is reproducible then the replicates should follow the same distribution. Lastly, to determine whether each condition is reproducible, a Mann-Whitney condition test and an average variance test are performed. The Mann-Whitney condition test determines the variance in the Mann-Whitney replicate distributions across all mutants in a specific condition. Whereas the average variance test measures the variance between each replicate under a specific condition and then calculates the mean variance for all replicates. Conditions with a high variance can then be removed from the dataset.

Finally, to generate the mutant fitness scores, for each mutant, using a modified T-test, ChemGAPP first calculates the mean growth value for each replicate in a specific condition and then compares this value to the mutant's median growth value across all conditions. This final value is called the S score, and an S score of < -3 or > 2 represents a significant change in fitness. Mutant growth is compared to its own average growth rather than to WT as this is a more powerful comparison since a true baseline can be established.

## 1.4 Research Aims

Since a large portion of the mycobacterial genome is yet to be functionally annotated, we aimed to optimise a chemical-genomic screening method for mycobacteria. As the analysis of colony growth results in a diverse range of data output, we investigated the use of high throughput colony arraying robots in combination with IRIS and ChemGAPP platforms to analyse mutant fitness.

In addition, given that we have access to an ordered *M. bovis* BCG mutant library, generated by Borgers et al. (2020), through chemical-genomics, we aimed to phenotypically profile *M. bovis* BCG mutants to gain insight into the function of previously uncharacterised genes and identify drug resistance pathways as well as potential drug targets.

# Part 2 – Results

## Contribution Statement

Section 2.1.1 was conducted in collaboration with Huda Ahmad and Dr Danesh Moradigaravand.

## 2.1 Condensing and Replicating the *M. bovis* BCG Transposon Insertion Library

The Borgers et al. (2020) *M. bovis* BCG Danish 1331 loss-of-function library contains 18,432 mutants. Given that chemical-genomic screening requires a significant time and financial commitment, it is not feasible to conduct a screen on a library of this size. In addition, many of the *M. bovis* BCG library strains are redundant since multiple mutants exist for the majority of successfully targeted genes, with an average of 5.63 knockout duplications per gene. The library also contains insertions into intergenic regions and thus not all mutants in the library are gene KOs and ORF integrated transposons do not guarantee gene deletion as an insertion towards the 3' region of a gene may still result in a fully functional protein. Furthermore, due to limitations with the library deconvolution method, the precise mapping of each mutant was not possible and thus many of the mutants are either not assigned a library location or are

speculatively located. Altogether, 20,731 mutants were assigned which is an additional 2,299 mutants than the total size of the library. This can be explained by the accidental picking of more than one mutant into single wells during the library construction process or the non-specific amplification of gDNA during the CP-CSeq method. We therefore concluded that before conducting the screen, the library should be condensed to contain best quality copy of each mutant ensuring that the library is of a manageable size whilst maintaining a high level of genome saturation.

### 2.1.1 Selection of the Condensed *M. bovis* BCG Library

As described in Chapter 5, Section 1.12, to select the most suitable mutants for our condensed library, we first ensured that all mutants could be found in the main library and targeted ORFs. We therefore filtered the library data to remove any mutants without a location or gene assignment. 6,973 mutants were removed leaving 13,758 mutants. Next, to ensure that we prioritised the selection of mutants which have a high likelihood of generating a loss-of-function, only mutants with a high percentage of gene deletion were selected. We therefore identified the locations of the transposon insertions and removed all mutants with insertions in the latter 20% of the target gene as there is a high chance that mutants with insertions towards the 3' end will still result in functional truncated proteins. As demonstrated by figure 1.20, this threshold was established as it maintained a large collection of mutants whilst ensuring a high confidence in the quality of the gene disruption. Without any filtering, just over 3,000 unique mutants were present. However, when the 20% cut-off threshold was applied, 2,755 mutants were selected representing 73% of the non-essential genome. Lastly,

to have the most confidence in our screening results, we selected the best quality KO duplicates, ensuring that only the mutant copies which have the greatest likelihood of a loss-of-function were chosen. To achieve this, we first selected KO duplicates with the most accurate location assignments. As the location assignments are categorised into 4 groups, unique, XY/Zok_unique,, XY/Zok_nonunique and heuristic, based on their accuracy, we therefore prioritised the selection of KO duplicates in this order, with unique being the most favourable and heuristic being the least favourable. After this process, if any KO duplicates were still present, we then selected the best copy based on the insertion site of the transposon, favouring the KO duplicate that generated the largest gene deletion.

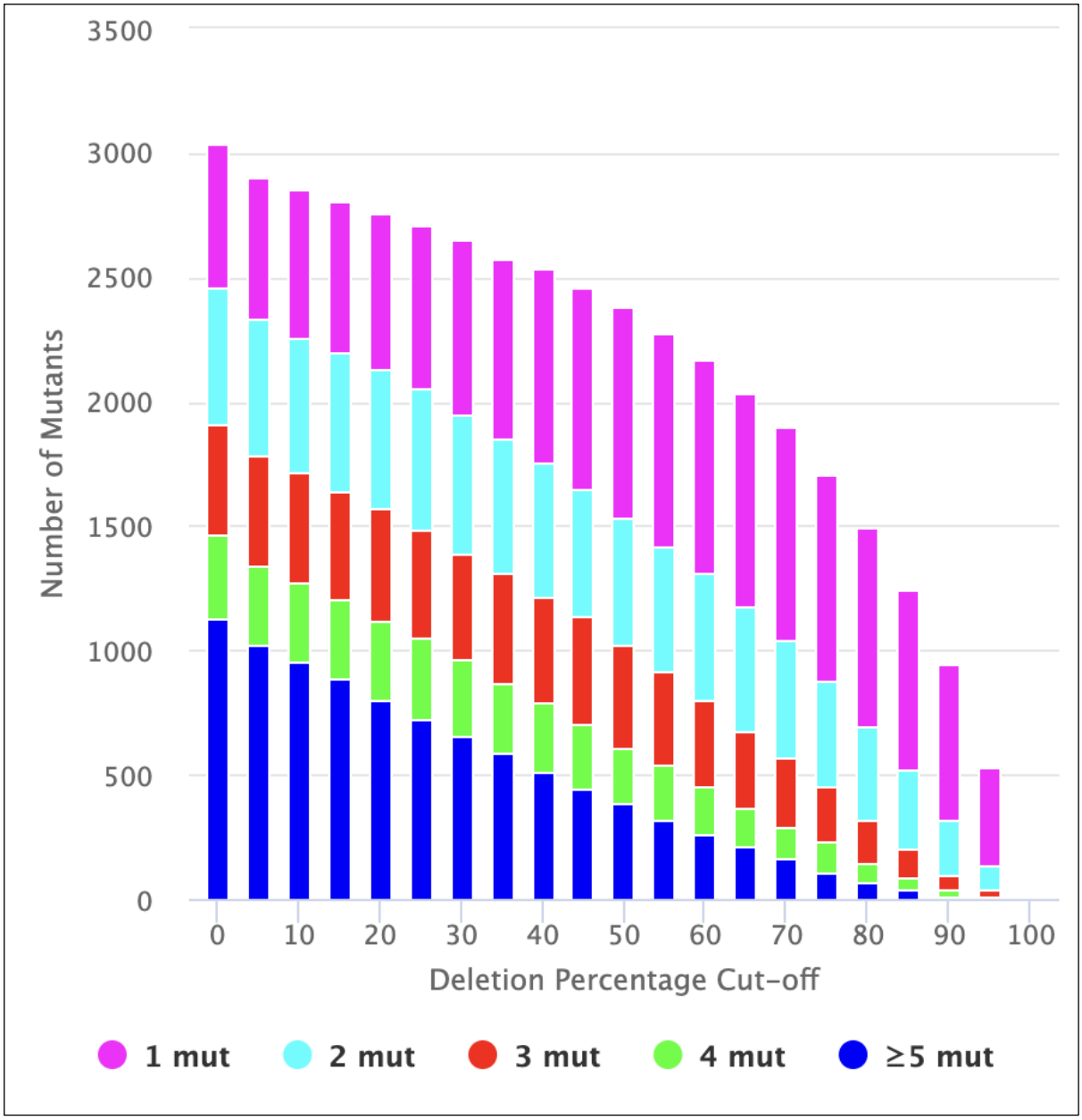



**Figure 1.20 Selection of the condensed *M. bovis* BCG library.**

The number of unique mutants remaining in the *M. bovis* BCG library as the percentage gene deletion threshold is increased. The colour code represents the number of KO duplicates per unique mutant. Figure generated by Huda Ahmad.

## 2.1.2 Condensing the *M. bovis* BCG library

Condensing a library of this size is extremely time- and cost-consuming. Therefore, it was important that the method was first optimised.

Borgers et al. (2019) reported that for the construction of the library an additionally supplemented medium was used to help support the growth of the mutants including mutants with auxotrophic or reduced growth. As outlined in Chapter 2, Section 3.2.7, we developed a similar medium, which we coined ‘enriched 7H9 medium’, differing only in glycerol concentration and the vitamin supplement. This medium was shown to successfully facilitate the growth of six mutants picked at random from the library. Furthermore, as figure 1.21. demonstrates, by conducting a growth curve with wildtype *M. bovis* BCG Danish 1331, we observed an increase in wildtype growth rate when grown in enriched 7H9 medium. We therefore concluded that the enriched 7H9 medium was sufficient for the condensing of the library.

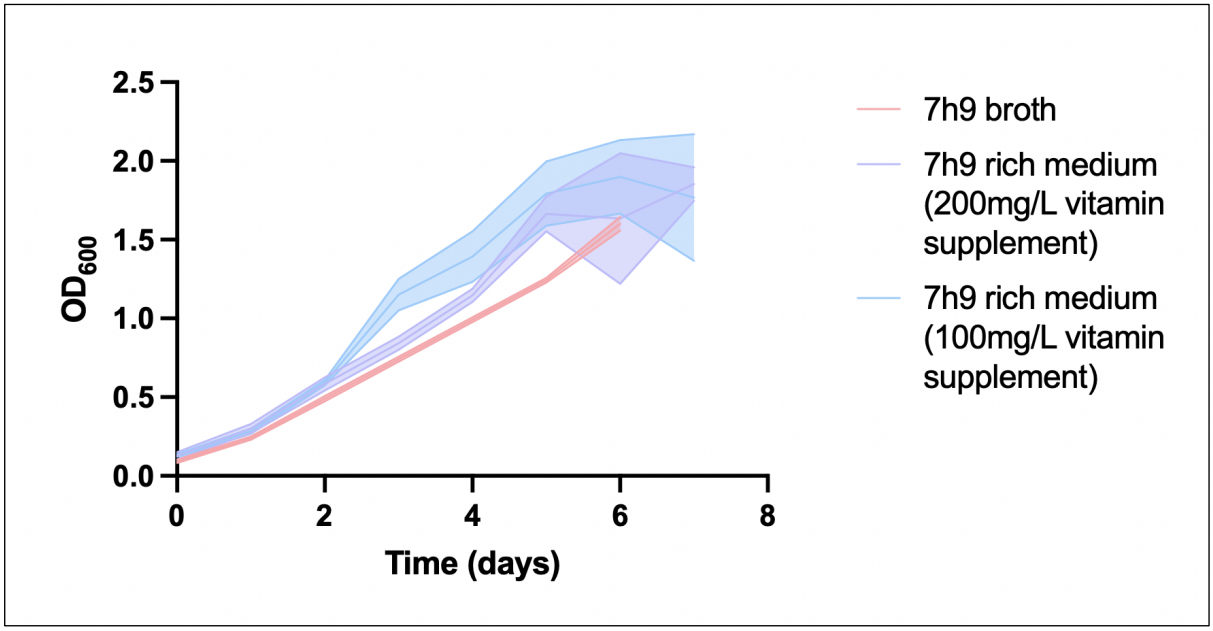



**Figure 1.21 *M. bovis* BCG growth in enriched 7H9 medium.**

Growth of WT *M. bovis* BCG Danish 1331 in 7H9 and enriched 7H9 medium supplemented with 100mg/L or 200mg/L of vitamins. For each growth curve the shaded area represents the SD between three biological replicates and the centre of measure is represented by the mean.

It was discovered that *M. bovis* BCG grown in 384-well plates had reduced and inconsistent growth. This is likely due to the waxy and slow growth of these bacteria as well as the reduced surface area resulting in decreased oxygen retention. It was therefore decided that the condensed library would be ordered into 96-well plates. By including a WT strain in each plate of the condensed library, we calculated that the resulting library would consist of just 29 96-well plates.

In order to generate multiple copies of the condensed library, we reasoned that we would require a larger volume than the capacity of wells in a standard 96-well plate. Therefore, we optimised a method for the growth of *M. bovis* BCG in deep-well plates. To validate the method, varying volumes of *M. bovis* BCG culture was added to deep well plates containing enriched 7H9 medium. To prevent contamination, deep-well plates were then sealed with either Breathe-Easy® sealing membrane or clear sealing tape. We discovered that an inoculum of > 1 µL was sufficient for inducing growth in the deep-well plates. Furthermore, we found that cultures in plates sealed with Breathe-Easy® sealing membranes had an $OD_{600nm}$ of almost double those with clear sealing tape. In addition, due to evaporation, both types of sealed plates lost around a third of their total volume. Therefore, due to the enhanced growth and comparable loss of culture volume observed with Breathe-Easy® seals, we concluded that these seals would be most appropriate for the condensing of the library however a greater volume of culture would be required to account for the loss of culture.

As it has been previously reported that the loss of culture volume via evaporation could be reduced by incubating the plates next to water reservoirs (Borgers et al., 2020) we repeated the method but employed a larger volume of enriched 7H9 medium (1.4 mL) and positioned water reservoirs adjacent to each deep-well plate. As a result, only around a quarter of the culture volume was lost via evaporation thus confirming that evaporation could be reduced via this method. Furthermore, by checking the OD at various time points, we determined that after 5 weeks of incubation, although cultures inoculated with 1 µL of WT *M. bovis* BCG had a relatively low OD, cultures with both 5 µL and 10 µL of inoculum had grown to an $OD_{600nm}$ of around 0.6, an appropriate OD for the generation of the condensed library. We therefore concluded that > 5 µL of inoculum and a minimum of 5 weeks incubation should be employed for the library condensing method.

Finally, to increase the throughput of the method we investigated the use of robotics. Protocols for condensing and replicating the library were developed for both the Opentrons OT-2 and the Beckman Coulter Biomek i5 robots. Given that the Biomek i5 was fitted with a Thermo Fisher Scientific Multipdrop Combi Reagent Dispenser and a dynamic Span-8 Multichannel pipette, it was discovered that cherry-picking from the original library into a single condensed library plate was considerably quicker than the OT-2 robot. Due to the much faster processing on the Biomek i5 robot, we therefore selected this robot for the condensing of the library. However, due to a pipetting issue, we determined that the OT-2 robot was better suited for the preparation of the replication of the library.

Following optimisation, we condensed and replicated the library and as depicted by figure 1.22, in total, we discovered that the process took 7 weeks, 1 week for the cherry-picking, 5 weeks of incubation and 1 week for the library replication.

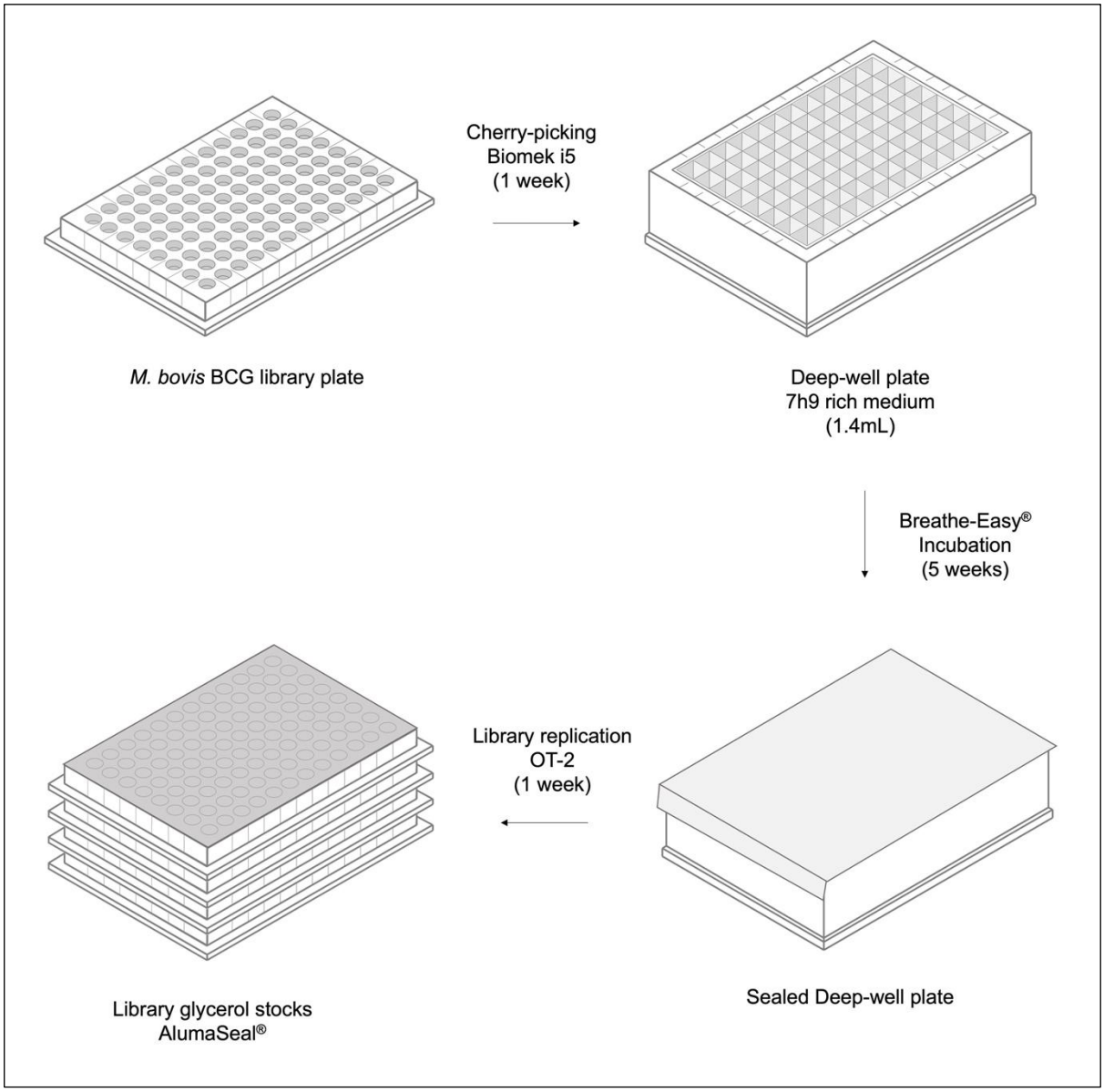



**Figure 1.22 *M. bovis* BCG library condensing and replication method.**

The highest quality KO duplicate of each unique mutant were cherry-picked from the *M. bovis* BCG library and, using the Biomek i5, organised into a deep-well plate containing 1.4mL of enriched 7h9

medium. The deep well plates were sealed with Breathe-Easy® sealing membranes and incubated for 5 weeks until an OD ~ 0.6. Lastly, the condensed library was replicated into 96-well library glycerol stocks via the OT-2 robot. The entire process took around 7 weeks to complete (Chapter 5, Section 1.12).

During the condensing of the library, we observed fungal contamination in 14 wells of the condensed library plates. Therefore, during library replication, these wells were omitted and replaced with WT. To test whether the condensed library replicates contained fungal contamination we transferred glycerol stock from three random library plates into 96-well deep-well plates containing enriched 7H9 media and monitored the growth. Unfortunately, fungal contamination was observed across ~ 15% of the plates. Therefore, to remove the fungal contamination we employed a broad spectrum anti-fungal, amphotericin B. As shown by figure 1.23, by conducting a resazurin-assay (Chapter 5, Section 1.3) with amphotericin B, we determined that WT *M. bovis* BCG growth was inherently resistant to amphotericin B with only concentrations > 1 mg/mL resulting in reduced growth. Furthermore, by growing the 3 random library plates in medium supplemented with > 25 µg/mL of amphotericin B, we discovered that fungal contamination was eradicated. We therefore concluded that the addition amphotericin B at a concentration of 25 µg/mL was suitable for the outgrowth of the library.

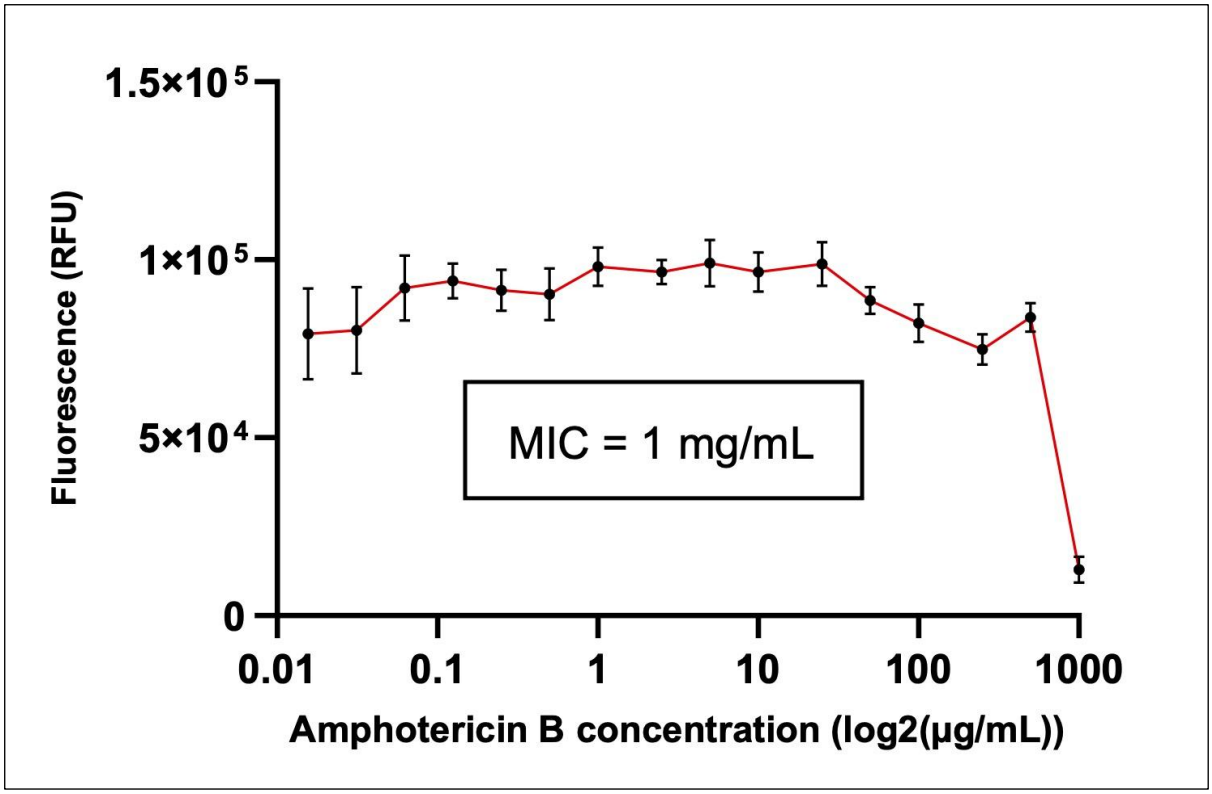



**Figure 1.23 Amphotericin B MIC**

Resazurin assay to determine the amphotericin B MIC determined via the mean + SD fluorescence for WT *M. bovis* BCG across three biological replicates.

## 2.2 Optimisation of Chemical-Genomic Screening in *Mycobacterium bovis* BCG

Although, due to their slow and aggregative growth, mycobacteria are often poorly suited for colony growth assays, given that IRIS software is extremely sensitive and is capably of analysing various aspects of colony growth, we concluded that, with optimisation, this method could be adapted for the chemical genetic screening of the *M. bovis* BCG library.

As we previously mentioned, performing a chemical genomics screen is a costly process which requires extensive resources and a large time commitment. Therefore, to maximise the efficiency of the method, before conducting the screen, we also first optimised this procedure.

### 2.2.1 Establishing the Screening Method

Given that the BM3-BC robot is designed for high-throughput chemical genomic screening through colony arraying and imaging and has the capacity to replicate 96-well source plates onto solid agar replica plates in a 96 or 384-format, we concluded that this robot would be suitable for screening the *M. bovis* BCG library.

To optimise the method we used WT *M. bovis* BCG to provide a clearer indication of growth consistency. Therefore, we arrayed WT *M. bovis* BCG into 96-well source plates and, via the BM3-BC robot, we replicated the source plates onto solid agar replica plates. Following incubation, imaging and IRIS analysis could then be conducted. Typically, as mycobacteria are slow-growing and thus require an extensive period of incubation, 7H10 is employed for growth on solid medium as it contains malachite green which has anti-microbial properties and thus prevents contamination. However, since many of the mutants in the library have severe growth defects and as the library already requires amphotericin B supplementation, we were reluctant to add any further potential stresses. Therefore, 7H9 agar was selected as the solid growth medium.

The first step of optimisation was establishing the optimal agar depth and agar concentration of the solid agar replica plates. Whilst the source plates were successfully replicated at low agar depths, following incubation, replica plates with volumes < 30 mL became dehydrated and growth was inconsistent (figure 1.24a). In addition, at agar concentrations < 2%, the replication was unsuccessful as the BM3-BC replicator head penetrated the agar (figure 1.24b). We therefore concluded that an agar volume of 40mL and a concentration of 2% was optimal for growth. Furthermore, as shown in figure 1.24c, we discovered that colony growth without a source plate mixing step resulted in irregular colony morphology and thus concluded that a mixing step was necessary for consistent growth. We also assessed whether a solid agar replica plate could be used as source plates but as shown in figure 1.24d, growth was variable. Lastly, to increase the throughput of the screen, we determined that, growth was successful and consistent in a 384-colony format (figure 1.24f) and concluded that the optimal length of incubation was 13 days. Interestingly, by analysing the growth of a subset of mutants we also discovered that the opacity and circularity scores were not as sensitive to changes in mutant fitness as colony size. This is likely due to the irregular and aggregative growth of mycobacteria as, in other backgrounds, opacity has been described as an improved readout when compared to size (Kritikos et al., 2017). We therefore selected colony size as the main measure of mutant fitness. Nevertheless, opacity or circularity could still be applied as a gauge for the reproducibility of the dataset and to identifying any outliers. An outline of the final, fully optimised chemical-genetic screening method is described in Chapter 5, Section 1.13.

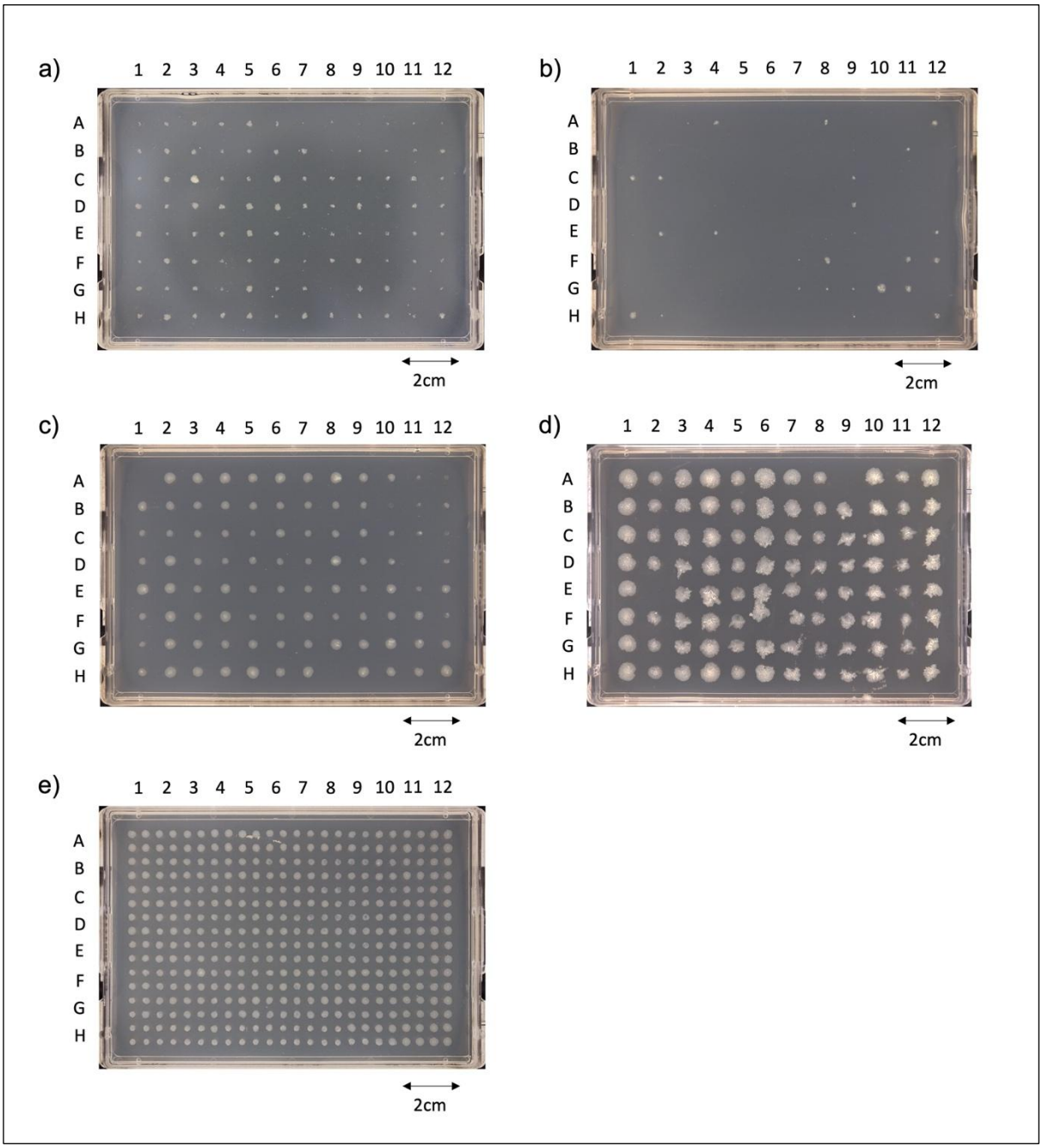



**Figure 1.24 Optimisation of chemical genomic screening.**

Solid agar plates demonstrating WT *M. bovis* BCG growth at various stages of the optimisation process. WT growth was inconsistent when (a) agar volume was below 20mL, (b) agar concentration was below 2%, (c) plates were not mixed prior to replication and (d) using a solid agar arrayed colony source plate.

When controlling for all these variables, growth was consistent and could be adapted to a 384-format for increased throughput (d). The optimal length of incubation for a 384-format plate was 13 days.

### 2.2.2 MIC Testing of Screening Conditions

To obtain the most valuable screening dataset we selected 115 conditions which target a range of cellular processes. These conditions included a variety of environmental stresses, nutrient sources, chemical stresses, and antibiotics including repurposed medical drugs as well as front-line antituberculosis agents such as rifampicin, isoniazid, ethambutol, pyrazinamide, fluoroquinolones and bedaquiline (supplementary table 1). Using MICs from the literature as a guide, by growing WT *M. bovis* BCG on 7H9 agar plates with varying concentrations for each condition, we calculated the MICs and the optimal sub-inhibitory screening concentrations (supplementary table 1). Unfortunately, although we were able to grow *M. bovis* BCG in liquid Sauton's and Roisin's minimal medium, we were unable to sustain consistent growth on solid medium and thus we could not assess growth in different carbon and nitrogen sources. In addition, we were unable to confirm the optimal concentrations for all of the conditions tested.

## 2.4 Conducting a Pilot Test Screen

To ensure that the method was successful we conducted a pilot test screen by selecting two condensed library plates at random and testing 22 diverse conditions (supplementary table 1). Colony growth values were determined using IRIS (Kritikos

et al., 2017), and the normalisation, quality control and transformation of IRIS data was conducted using ChemGAPP (Doherty et al., 2023) to generate mutant fitness S scores. Unfortunately, as shown in figure 1.25, across 6 conditions we were unable to analyse the colony growth via IRIS as not all mutants were able to grow under stress. As the first step of IRIS is to segment colonies into tiles by generating a grid and, since this grid is calculated by taking an average of the X and Y coordinates for each colony centre, even if only a small proportion of colonies do not exhibit growth, this can interfere with the grid formation and thus the plate cannot be analysed. This is clearly a drawback of the IRIS software as these plates, such as figure 1.25e-f, still have valuable data. Nevertheless, to account for this, in future screening, for the conditions that failed to support mutant growth, the condition concentration should be reduced and retested to facilitate the growth of all mutants. Interestingly, certain conditions only facilitated the growth of a single mutant indicating an enhanced resistance (figure 1.25). Δ*BCGDan_0789* (*phoP*) demonstrated resistance to amoxicillin clavulanate, dapsone and sulfamethoxazole (figure 1.25a-c) and Δ*BCGDan_0421* exhibited resistance to clofazimine (figure 1.25d).

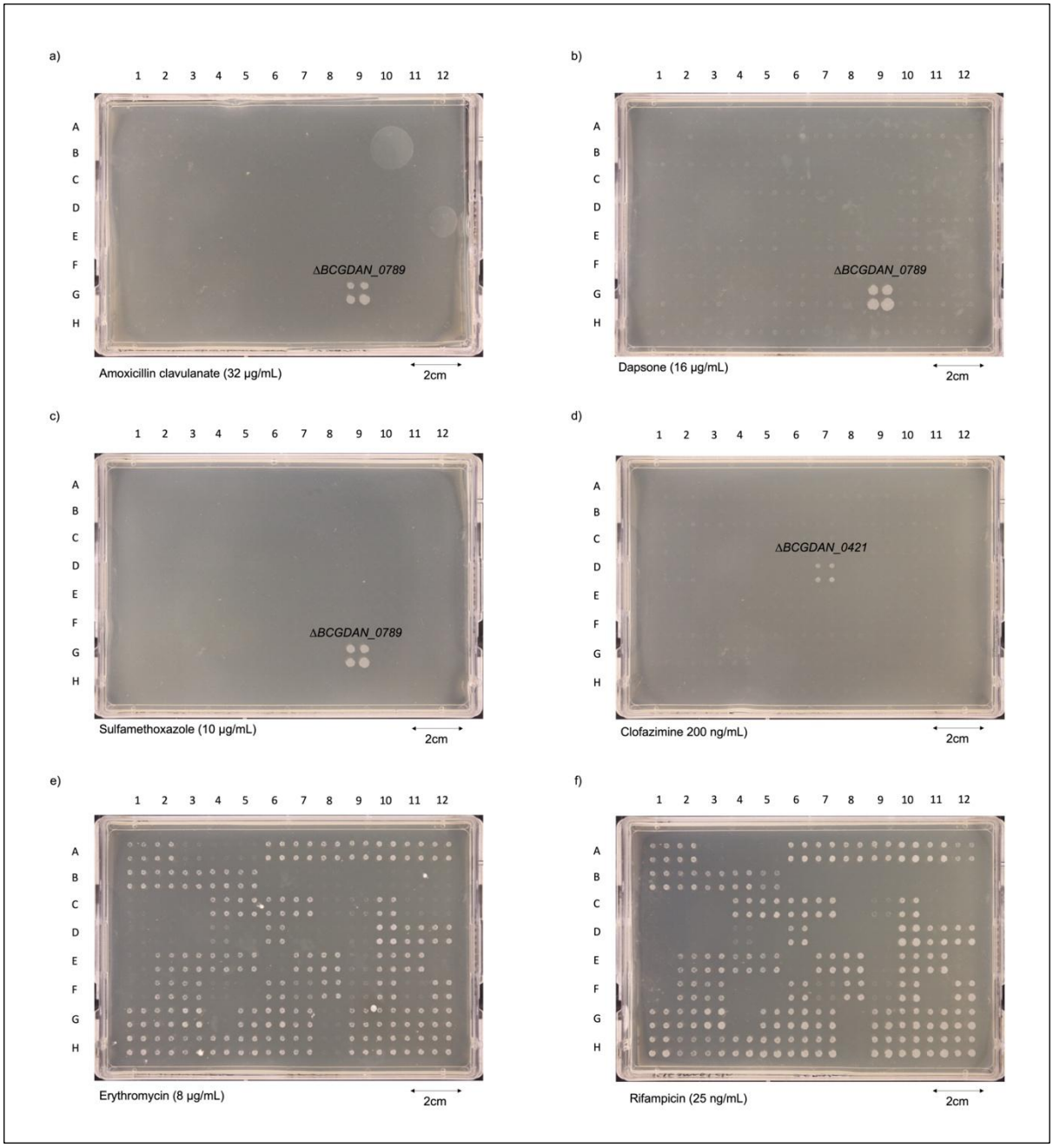



**Figure 1.25 Pilot screen failed growth plates.**

Images of condition plates that were unable to sustain consistent mutant growth and therefore could not be analysed via IRIS. (a) 32 µg/mL amoxicillin clavulanate, (b) 16 µg/mL dapsone, (c) 10 µg/mL sulfamethoxazole, (d) 200 ng/mL clofazimine, (e) 8 µg/mL erythromycin, (f) 25 ng/mL rifampicin. Δ*BCGDan_0789* and Δ*BCGDan_0421* are highlighted as these mutants were resistant to multiple drugs whilst other mutants failed to grow.

From the IRIS data from the 16 remaining conditions, through ChemGAPP analysis, fitness S-scores were generated. A normality threshold of 40%, Mann-Whitney plate threshold of 0.005846, a Mann-Whitney condition threshold of 0.006877 and an average variance threshold of 6954000.0 was applied which unfortunately removed a further 4 conditions from the dataset and thus the final dataset contained just 12 of the 22 conditions. Lastly, as shown by figure 1.26, via TreeView3, the mutants were hierarchically clustered by their fitness profiles using an uncentred Pearson Correlation.

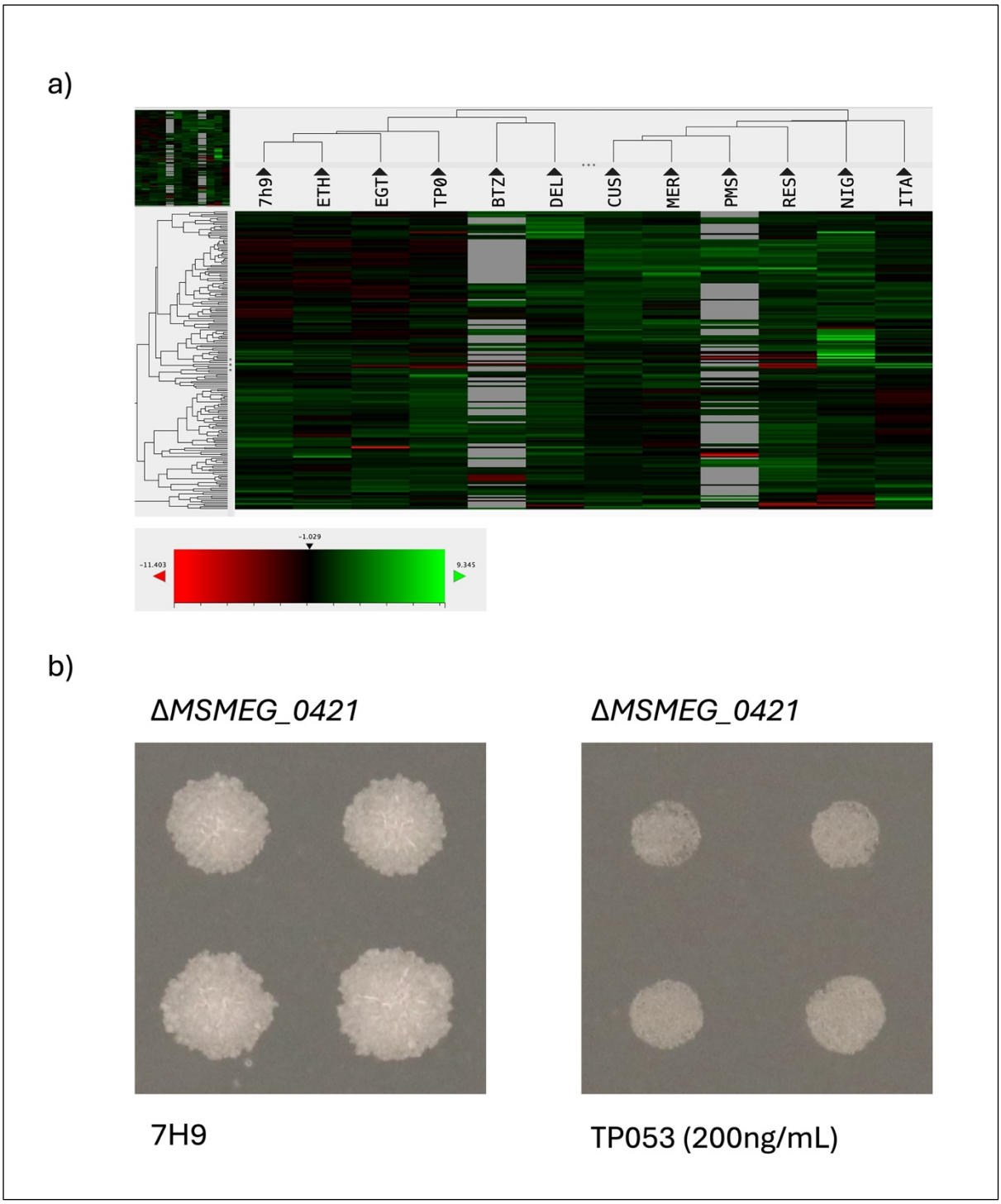



**Figure 1.26 Clustering analysis of mutant fitness S-scores calculated via ChemGAPP.**

(a) Mutants were clustered via uncentred Pearson Correlation using TreeView3. Red indicates a reduction in mutant fitness and green indicates a gain of fitness as demonstrated by the scale bar. Grey indicates missing data. The tree on the X axis represents the clustering of the mutants and the tree on the Y axis represents the clustering of the conditions based on the mutant fitness S-scores. ETH =

ethambutol, EGT = EGTA, TP0 = TP053, BTZ = BTZ043, DEL = delamanid, CUS = $CuSO_4$, MER = meropenem, RES = reserpine, NIG = nigericin, ITA = itaconic acid. (b) Δ*MSMEG_0421* growth on two selected agar plates from the chemical-genomic screening demonstrating a change in Δ*MSMEG_0421* mutant fitness in response to TP053, in which growth is significantly reduced compared to standard conditions.

Due to the small dataset, representing ~ 5% of the non-essential genome and ~ 10% of the total conditions, we did not expect to see many strong correlations between mutant phenotypes and thus did not expect the clustering analysis to identify many significant screening hits. However, as demonstrated by figure 1.26, there were many clusters of mutants that demonstrated similar phenotypic profiles. Furthermore, as shown in figure 1.27 and supplementary table 2, we identified several clusters of mutants that had significant changes in fitness in response to certain stresses. In total, 76 of the 192 strains exhibited a significant fitness change in response to at least one condition. In addition, we discovered that many of these mutants exhibited an expected phenotype thus validating the effectiveness of the screen.

For example, *BCGDan_0421 (Δfgd1)* is a delamanid activator (Manjunatha et al. (2005) and thus it was predicted that the mutant would be better at tolerating delamanid than other mutant strains. As expected, Δ*BCGDan_0421* demonstrated an increased fitness (figure 1.27a). We were also expecting Δ*fgd1* to exhibit reduced growth in the oxidative stress, PMS as *fgd1* fulfils a role in oxidative stress resistance. However, due to data loss during the data processing steps we were unable to determine the fitness of *BCGDan_0421* when grown in the presence of PMS. Interestingly, *BCGDan_1227* (*fbiC),* another delamanid activator involved in oxidative

stress resistance (Fujiwara et al., 2018) did not cluster with *fgd1.* However, whilst Δ*fbiC* also demonstrated enhanced growth in delamanid, a significantly reduced growth was observed in PMS (figure 1.27b). Therefore, it is likely that *fgd1* did not cluster with fbiC as no PMS data was available.

Similarly, multiple genes encoding ATPases and ATP transporters clustered with nigericin which is an ionophore and can disrupt ATP homeostasis (Zhu et al., 2022). Some of these ATP transporters are also associated with drug efflux such as *BCGDan_1271* (Balganesh et al., 2010). It was therefore understandable that mutant growth would be reduced in the presence of nigericin. In addition, we also identified a cluster of genes involved in lipid metabolism and cell envelope biogenesis, *BCGDAN_1061*, *BCGDAN_3516* and *BCGDAN_2128* (Wang et al., 2010; Rastogi et al., 2017; Melly et al., 2019). It was therefore comprehensible that these mutants exhibited a reduced fitness in BTZ043 which targets DprE1 involved in the synthesis of a lipid donor vital for the generation of AG (Lechartier et al., 2012).

Lastly, by identifying the genes that clustered with the mutants with expected phenotypes, we discovered many uncharacterised genes which could be worth investigating further as they may fulfil similar roles and could have an important function in drug resistance.

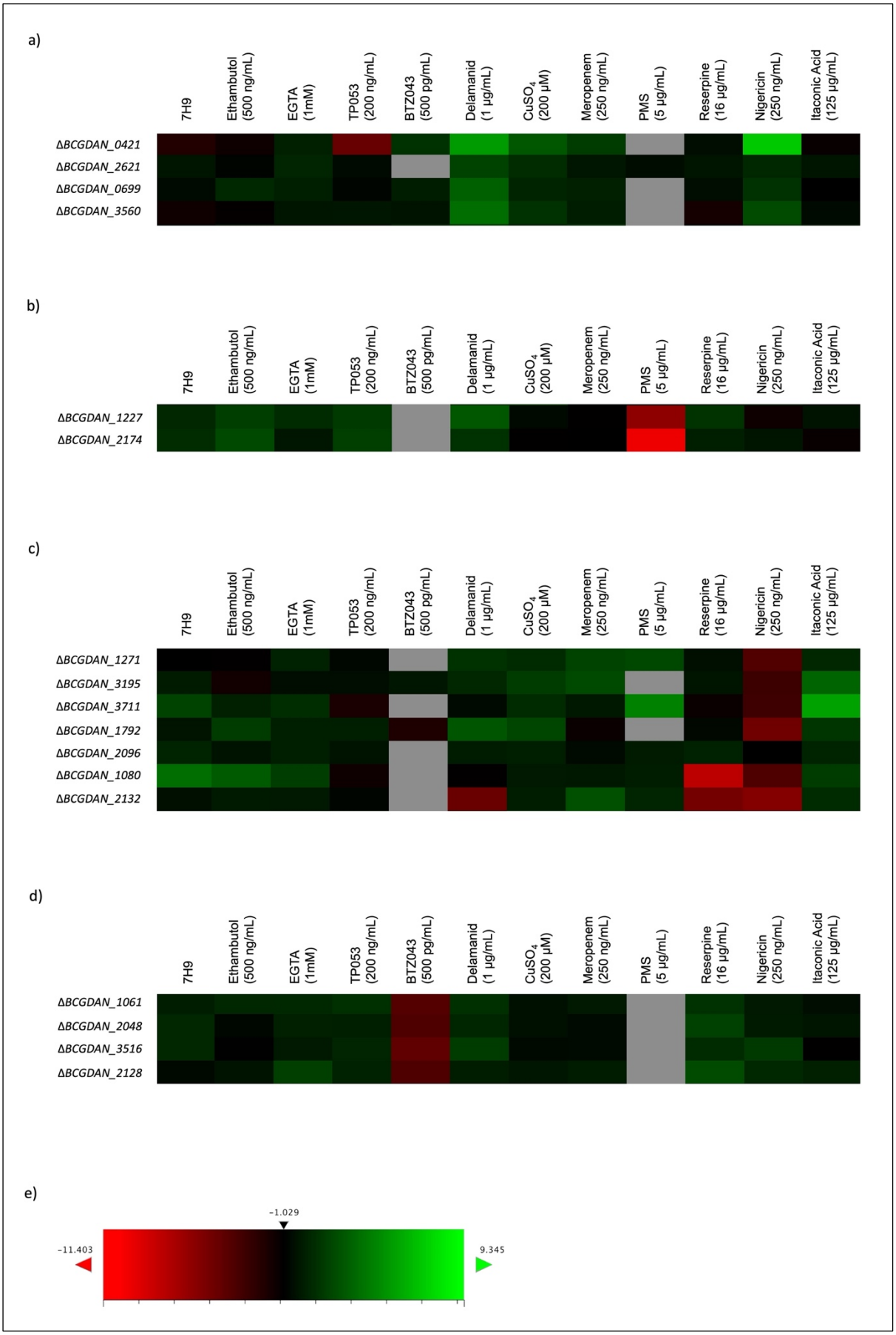

a)
7H9
Ethambutol (500 ng/mL)
EGTA (1mM)
TP053 (200 ng/mL)
BTZ043 (500 pg/mL)
Delamanid (1 µg/mL)
CuSO4 (200 µM)
Meropenem (250 ng/mL)
PMS (5 µg/mL)
Reserpine (16 µg/mL)
Nigericin (250 ng/mL)
Itaconic Acid (125 µg/mL)
ΔBCGDAN_0421
ΔBCGDAN_2621
ΔBCGDAN_0699
ΔBCGDAN_3560
b)
ΔBCGDAN_1227
ΔBCGDAN_2174
c)
ΔBCGDAN_1271
ΔBCGDAN_3195
ΔBCGDAN_3711
ΔBCGDAN_1792
ΔBCGDAN_2096
ΔBCGDAN_1080
ΔBCGDAN_2132
d)
ΔBCGDAN_1061
ΔBCGDAN_2048
ΔBCGDAN_3516
ΔBCGDAN_2128
e)
-1.029
-11.403
9.345

**Figure 1.27 Clusters of mutants with significantly altered fitness.**

Data obtained through Treeview3 clustering analysis in which mutants demonstrate altered fitness in (a) delamanid, (b) PMS, (c) nigericin and (d) BTZ043. Red indicates a reduction in mutant fitness and green indicates a gain of fitness as demonstrated by the scale bar (e). Grey indicates missing data.

# Part 3 – Discussion

## 3.1 Completion of the Chemical-Genomic Screening

Unfortunately, due to time constraints, we were only able to conduct a pilot screen and thus the full screen with the entire *M. bovis* BCG condensed library is yet to be completed. As the BM3-BC robot is capable of arraying 46 plates every 3 h and imaging 46 plates every 2 h, we estimate that the full screen, consisting of 29 source plates across 115 conditions, will take around 450 h to complete. However, as the plates must be loaded manually and the source plates must be prepared daily, accounting for working hours, around 2-3 months will be required for completion. Furthermore, an additional pre-screening period will be necessary to establish the sub-inhibitory concentrations for the remaining screening conditions.

Once the full screen is complete, to validate the resulting mutant fitness profiles, the phenotypes of significant screening hits should be compared with entire gene deletion strains. This validation step is important as, although we established a 20% gene disruption threshold for selecting the condensed library mutants, it is plausible that not all condensed library strains are null mutants (Borgers et al., 2020). Transposon insertions can also cause polar effects and inactivate neighbouring genes (Borgers et al., 2020). Therefore, some library mutants may result in multiple gene disruptions. In addition, whilst we were careful when handling the library, it is possible that the library has a degree of cross-contamination as has been reported in previous mutant

collections (Silvis et al., 2021; Goodall et al., 2023). As discussed in Chapter 1, Section 2.5 and Chapter 4, Section 1.4, ORBIT (Murphy et al., 2018) and CRISPR represent efficient methods developed for the generation of mycobacterial entire gene deletion mutants and thus could be used to construct mutant strains for validating the screen and for further elucidating the function of previously uncharacterised genes through reverse-genetic study.

## 3.2 Considerations of the Optimised Chemical-Genomic Screening Method

Although we were unable to finish the complete screen, we have outlined an optimised method for the chemical-genomic screening of ordered mycobacterial libraries. Therefore, with a few minor adjustments, mainly due to differences in growth rate, this method could be also applied to the *M. abscessus* library we generated in Chapter 2.

In the first optimisation step (Chapter 3, Section 2.1), we determined a method for condensing ordered transposon libraries to an appropriate size for chemical-genomic screening whilst maintaining a high level of genome saturation. To achieve this, we reordered the library using liquid handling robotics, selecting only a single knockout duplicate per mutant with the highest percentage of gene deletion. This ensured that all of the mutants in the condensed library had a high likelihood of loss-of-function and thus would generate the most reliable screening data. To select the condensed library dataset, we first removed any mutants that did not have a location within the library or

did not target ORFs. Next, we established a threshold for the percentage of gene disruption in which a 20% cut-off value was chosen to remove any mutants that had insertions into the latter 20% of the gene and thus a high chance of producing functional truncated proteins. Establishing this threshold was a difficult balance between selecting the highest number of mutants for the condensed library whilst ensuring that these mutants were of the highest quality. For example, we could have selected a threshold of 40% thus ensuring that the condensed library mutants had an even higher chance of generating dysfunctional proteins. However, this would have removed around 250 additional mutants from the condensed library and thus would have resulted in a non-essential genome saturation below 70%. Lastly, we selected the best quality KO duplicate for each mutant prioritising the selection of mutants that had a high confidence of library location assignment and the highest percentage of gene disruption.

Unfortunately, during the library condensing and replication we encountered a fungal contamination. Since the Biomek i5 robot has an integrated high-efficiency particulate absorbing (HEPA) filtration system and given that we supplemented the growth media with kanamycin, contamination during library condensing was controlled. Nevertheless, 14 wells demonstrated fungal growth. In addition, replication was conducted with an Opentrons OT-2 robot, which is not fitted with any contamination control measures and we discovered that ~ 15% of the replicated condensed library exhibited fungal contamination. Whilst we successfully removed the contaminants using an anti-fungal drug, we suggest that library condensing and replication should

be conducted in more sterile conditions, perhaps by employing robotics that can be contained within a biosafety cabinet.

In the second stage of optimisation (Chapter 3, Section 2.2), we established the optimal parameters for growing mycobacteria as an array of colonies. We determined that the replica plates required a 2% agar concentration to prevent the BM3-BC robot replicator head from piercing the agar and that a depth of 40 mL was necessary to prevent dehydration of the plates. We also discovered that the employment of liquid source plates resulted in the most consistent growth, as long as a mixing step was conducted prior to replication, and, to increase the throughput, these source plates could be replicated in a 384-colony format with each plate containing four replicates of each strain. Lastly, we calculated the MICs and optimal concentrations for all of the screening conditions. However, we were unable to facilitate consistent *M. bovis* growth on solid minimal media which is necessary for determining mutant fitness in response to different carbon and nitrogen sources. Given that we were able to grow *M. bovis* in liquid minimal media and since Sauton's medium has been previously employed for the growth of *M. bovis* on agar plates (Moynihan et al., 2019), we were unable to determine the reason for our failure to obtain growth. It is possible that additional nutrients are required and therefore, further optimisation of the minimal growth medium is necessary.

## 3.3 A Pilot Chemical-Genetic Screen Identifies Significant Hits in Previously Uncharacterised *M. bovis* BCG Genes

To validate the chemical-genomic screening method, we conducted a pilot screen with two condensed library source plates across 23 different conditions (Chapter 3, Section 2.4).

Unfortunately, for 6 of the conditions, the mutants had inconsistent or no growth and thus could not be analysed. Since the MIC and subinhibitory concentrations for each condition were calculated using WT *M. bovis* BCG, and, given that many mutants with the library have growth defects, it is likely that some of the screening concentrations were too high to sustain consistent growth for all mutants. We therefore propose that future MIC and subinhibitory concentrations for each condition should be determined using a condensed library source plate to ensure that alterations to mutant growth are also considered.

Nevertheless, as the WT *M. bovis* control was also unable to grow under certain conditions, it is likely that some of the previously defined optimal concentrations were inaccurate. In the pre-testing, due to human error, it is possible that the wrong concentrations were used, that the plates were mislabelled or that the media was too hot when the condition was added thus resulting in its degradation. In addition, it is plausible that some of the conditions were sequestered in the agar thus reducing the potency. It is therefore extremely important that the condition plates are prepared with

the upmost care. That said, some conditions were able to facilitate most mutant growth but could still not be analysed via IRIS. This was due to a limitation in the IRIS software as it cannot analyse plates with multiple gaps in the array since the first step of IRIS is to organise the plate into a grid by detecting the central points of each colony (Kritikos et al., 2017). We therefore propose an adaptation to IRIS in which there is an option to generate the grid manually thus allowing plates with multiple sick mutants to be analysed. Multiple conditions were also removed during the ChemGAPP data normalisation process. However, as these parameters can be adjusted it may be necessary to use a less stringent threshold for the full screen, especially given that mycobacteria demonstrate irregular colony growth.

Nonetheless, we were still able to obtain fitness data for 12 conditions and identified that 76 out of the 192 mutants had significant fitness changes in response to at least one condition.

Both Xu et al. (2017) and Li et al. (2022) have previously conducted pooled *M. tuberculosis* chemical-genomic screenings. Xu et al. (2017) performed a TnSeq experiment across 5 different conditions whereas Li et al. (2022) conducted a CRISPR KO and CRISPRi screen using 9 different antibiotics. As these studies used a variety of different conditions, only our ethambutol and meropenem data are comparable. In total, we had 12 significant hits across these conditions and, like our finding for ∆*BCGDan_0789* (*phoP*), Xu et al. (2017) also demonstrated that deletion of *M. tuberculosis phoP* resulted in an increase in fitness in meropenem. However, none of the remaining hits demonstrated significant fitness changes in homologues from the

Xu et al. (2017) and Li et al. (2022) datasets. Nevertheless, since the Xu et al. (2017) and Li et al. (2022) datasets also have many differences, including which genes are defined as essential, and given that the fitness analysis was conducted via a completely different method and in different species, we did not expect to see many similarities with such a small comparative dataset. Once the full-screen is completed however, comparison between these datasets could be a valuable tool for validating our screen.

In addition, although our clustering analysis was limited, since only 192 mutants out of the total 2,755 mutants were screened, we were able to highlight a selection of genes in significant clusters. Furthermore, as some of these genes had known roles, we were able to validate the efficiency of the screen by linking the function of these genes to expected mutant phenotypes. For example, as anticipated, the deletion of a delamanid activator, *fgd1* (Manjunatha et al. (2005), resulted in an increase in fitness in the presence of delamanid. In addition, since genes of a similar function are likely to cluster together, by identifying other mutants within these significant clusters, we highlighted various previously uncharacterised genes that could be worth investigating further.

Perhaps the most interesting discovery was the cluster of mutants susceptible to BTZ043. One of these mutants, *BCGDAN_1061* (*lpqT*), encodes a putative lipoprotein from the LpqN family (Melly et al., 2019). It has been discovered that LpqN binds to dodecyl-trehalose as well as MmpL3 and MmpL11 transporters, responsible for the export of lipid substrates to the MA layer of the cell wall (Melly et al., 2019). In addition,

deletion of *lpqN* results in altered lipid composition during biofilm maturation. It was therefore suggested that LpqN may act as a membrane fusion protein that links MmpL transporters with periplasmic proteins thus assisting in cell envelope biogenesis (Melly et al., 2019). In contrast, LpqT remains relatively uncharacterised. Whilst it has been discovered that *M. tuberculosis* LpqT also binds to MmpL11 transporters, it is yet to be determined whether LpqT can bind to lipids (Melly et al., 2019). Furthermore, although both LpqT and LpqN have been implicated in virulence by facilitating macrophage survival (Li et al., 2018; Penn et al., 2018), the function of LpqT is yet to be established. Since BTZ043 inhibits DprE1 involved in the synthesis of AG and LAM (Lechartier et al., 2012) and given that LpqT has been shown to bind to MmpL transporters, it is possible that, like lpqN, this gene may also have a role in cell envelope biogenesis and biofilm formation. Similarly, *BCGDan_2128* (*Rv2102*), which clustered with *BCGDan_1061*, has also been implicated in biofilm formation (Behura et al., 2023) though its function remains unknown and, due to a zinc-finger-like sequence signatures, it has also been predicted to fulfil a role in zinc chelation (Makarova et al., 2002). Of the remaining two BTZ043 clustered genes, the role of *BCGDan_2048* is unknown. However, *BCGDan_3516* (*Rv3371*) encodes a diacylglycerol acyltransferase linked with cell wall alterations and growth arrest during persistence (Rastogi et al., 2017). Given that the majority of these BTZ043 susceptible genes are associated with either biofilm formation or the cell wall it may be interesting to investigate whether their roles are linked to LAM and AG biosynthesis, maintenance and recycling.

We could not identify any shared function between the genes that clustered together with significant fitness changes in response to delamanid or PMS. Whilst these genes may fulfil a similar role that are yet to be established, it is also possible that, as we only screened 192 strains, these genes may not cluster together when the full screen is completed.

In contrast, two of the mutants that clustered together due to susceptibility to nigericin, had insertions in genes that fulfilled supposed roles in the cell wall. As nigericin is an ionophore that can reduce cytosol pH and reduce ATP production (Próchnicki et al., 2016), and as these genes clustered with the ATP-transporter *BCGDan_1271*, it is unclear as to why genes with roles in the cell envelope would be sensitive. That said, it has been reported that nigericin can also alter the cell membrane structure and permeability which could result in cell wall defects (Zhu et al., 2022). Furthermore, disruption to the cell envelope could also result in increased permeability thus rendering the cells more susceptible to nigericin. BCGDan_3195 (FadE24) is believed to play a role in the degradation of MAs (Wipperman et al., 2013) and BCGDan_1792 (Rv1754c) is an endo-D-arabinofuranase that catalyses the breakdown of LAM (Al-Jourani et al., 2023). Although both of these enzymes are well characterised, two other genes within the cluster do not have a defined function (*BCGDan_3711* and *BCGDan_2132*) and thus could be worth investigating further as they may fulfil similar roles.

Altogether, this pilot screen has identified various previously uncharacterised genes with significant phenotypes that could be worth researching further, many of which

may have roles in drug resistance or important biological processes such as maintenance of the mycobacterial cell wall. Therefore, this study has not only laid the foundation for future research, but, given that the pilot screen only represents around 1% of the full screen, it has also demonstrated the power of this approach for characterisation of the mycobacterial genome.

# Part 4 – Conclusion

Due to advances in sequencing technology, there is an abundance of mycobacterial sequencing data (National Institutes of Health, 2023). However, functional annotation of the genome is lagging, with around 25% remaining uncharacterised (Modlin et al., 2021). Nevertheless, chemical-genomics represents an exciting new approach for deciphering gene function on a global scale. Whilst TnSeq and CRISPRi are rapid tools for pooled functional-genomic screening, the output is limited to sequencing data (Griffin et al., 2011; Li et al., 2022). In contrast, the screening of ordered libraries facilitates a more in depth and diverse analysis of mutant fitness (Kritikos et al., 2017). Therefore, to accelerate the functional annotation of the remaining mycobacterial genome we developed a chemical-genomic screening method for ordered mycobacterial mutant libraries.

Since we had access to a *M. bovis* BCG ordered transposon insertion library, and, given that *M. bovis* is an important animal pathogen (National Statistics, 2021) which, due to genetic similarity, can be used as a model organism for *M. tuberculosis* (Kanipe and Palmer, 2020), we decided to conduct a chemical-genomic screen in *M. bovis* BCG. As transposon insertion libraries contain a considerable number of mutants, many of which are redundant (Borgers et al., 2020), we first condensed the library to include the best copy of each gene knockout duplicate by only selecting the mutants with the largest gene deletions and removing any mutants that had less than 20% gene disruption.

To optimise the screening method we employed the recently developed IRIS (Kritikos et al., 2017) and ChemGAPP (Doherty et al., 2023) chemical-genomics tools, in which various aspects of colony morphology can be analysed to determine mutant fitness. In this method, using high-throughput colony arraying and imaging robotics, *M. bovis* colonies are organised into a 384-colony format and subjected to subinhibitory concentrations across a set of conditions to assess their growth. To validate the method, we also conducted a pilot screen with 192 mutants across 22 conditions. We identified 76 mutants that had a significant fitness change in response to at least one perturbation and we discovered various mutants that had expected phenotypes thus verifying that the method could accurately identify mutant phenotypes. Furthermore, as mutants of a similar function are likely to cluster together based on their phenotypic profiles (Typas et al., 2010; Nichols et al., 2011), we were able to make assumptions about the function of previously uncharacterised genes, highlighting various avenues for further research.

Unfortunately, due to time constraints, we were unable to complete the full screen. However, since the pilot screen identified numerous screening hits and only represents around 1% of the final screening data, we have demonstrated the power of this approach for the functional characterisation of mycobacterial genomes. Since this method can assist in the annotation of previously uncharacterised genes, this approach could lead to the discovery of novel drug and vaccine targets as well as important virulence or antibiotic resistance pathways.

# Chapter 4

# The Characterisation of Mycobacterial Genes Involved in Cell Wall Maintenance

# Part 1 – Introduction

## 1.1 The Mycobacterial Cell Envelope

As mentioned in Chapter 1 Section 1.1, The mycobacterial cell envelope is unique, with only other bacteria from the Mycobacteriales order demonstrating a similar composition (Dulberger et al., 2020). The cell wall core consists of PG, AG, and MAs covalently bound to form a thick and hydrophobic barrier that protects the cells from the external environment and fulfils various roles in cell shape, growth and virulence (Abrahams and Besra, 2021). For this reason, mycobacteria are resistant to many antibiotics and thus the treatment of mycobacterial disease can be difficult. In addition, there are growing concerns over rises in drug-resistance and therefore there is a huge demand for the discovery of new drugs (Zhou et al., 2020). As the cell wall is vital for the survival of mycobacteria, the characterisation of key genes involved in the maintenance of the cell wall could lead to the discovery of new drug targets.

## 1.2 Mycobacterial β-lactamases

Mycobacterial PG is also unusual as, unlike most bacteria, in which GlcNAc is covalently bound to MurNAc, GlcNAc can also be covalently linked to *N*-MurNGlyc (Raymond et al., 2005). In addition, in Actinobacteria, the peptide chains that extend from the sugar backbone contain a variable third amino acid, though in mycobacteria

this is usually *meso*-DAP (Basavannacharya et al., 2010). With the assistance of the D,D-carboxypeptidase PBPs, in most other bacteria, the PG peptide chains are crosslinked via the 3rd and 4th amino acids. However, though these linkages still occur in mycobacteria, via the activity of LD-transpeptidases, peptide crosslinks between the 3rd amino acids are more frequently formed (Lavollay et al., 2008). These crosslinks are necessary for maintaining rod-shape particularly in mature regions of the cell wall. Without L,D-transpeptidase crosslinking, PBP-mediated cross-linking becomes more important and cells exhibit an increased sensitivity to beta-lactam drugs and environmental stresses (Baranowski et al., 2018).

As mycobacteria mainly form 3-3 PG crosslinks and as many mycobacterial species express BlaC or BlaS, which are potent β-lactamases, PBP targeting beta-lactam antibiotics are not regularly used to treat mycobacterial infections (McDonough et al., 2005). That said, though BlaS is extremely active against lysozyme, it has variable activity against different classes of β-lactams (Flores et al., 2005). It has also been discovered that BlaC only has weak activity against carbapenems and both BlaC and BlaS are inhibited by clavulanic acid (Soroka et al., 2017). Interestingly, carbapenems have also been implicated in the deactivation of L,D-transpeptidases and thus, by targeting both PBPs and L,D-transpeptidases, these beta-lactams, especially in combination with clavulanic acid, have the potential to be employed as therapeutic options for treating mycobacterial infections (Dubée et al., 2012).

Mycobacterial species, especially *M. abscessus*, also encode multiple minor β-lactamases, however, in *M. tuberculosis*, whilst multiple genes are currently annotated

as putative β-lactamases, few have been characterised (Kuman et al., 2022). For example, LpqF, which is conserved across multiple mycobacterial species, is annotated as a putative β-lactamase due to the catalytic C-terminal domain that shares homology with other β-lactamases as well as transpeptidases. By determining the activity of these minor β-lactamases, we can identify their targets and thus help to further inform clinicians on the applicability of beta-lactam antibiotics for treating mycobacterial disease.

## 1.3 AG, LM and LAM Hydrolases

As mentioned in Chapter 1 Section 1.4, AG is a polysaccharide covalently linked to both the PG and MA layers, consisting of galactan chains bound to arabinan branches (Besra et al. 1995). It is essential for the survival of mycobacteria and plays an important role in virulence and pathogenicity (Li et al., 2022). The pathways for mycobacterial AG synthesis are well characterised. However, the mechanisms for the hydrolysis and recycling of these important cell wall components are still poorly understood. For example, whilst endogenous AG degradation was discovered around 30 years ago (McNeil et al., 1994), it is only within the last five years that the enzymes responsible for the degradation of AG were identified. In 2020, Shen et al. discovered a *M. tuberculosis* exo-β-D-galactofuranase, GlfH1, that hydrolyses the β-1,5 and β-1,6 glycosidic linkages of the galactan chain. Then in 2023, by screening for AG degrading activity in gut bacteria, Al-Jourani et al. discovered two families of AG glycoside hydrolases, GH172 and GH183 which were also identified in the mycobacterial genome. The GH172 family were shown to have exo-D-arabinofuranosidase activity

whilst in contrast, the GH183 family was the first example of an enzyme with endo-D-arabinofuranase activity. Interestingly, whilst *M. tuberculosis* only expresses two GH183 family proteins, Rv3707c and Rv1754c, which hydrolyse the AG and LAM α-1,5 arabinan backbones respectively, *M. smegmatis* encodes five, *MSMEG_4352*, *MSMEG_4360*, *MSMEG_4365*, *MSMEG_2107* and *MSMEG_6255*. These enzymes are individually non-essential and thus it is likely that these enzymes are redundant as they fulfil compensatory roles. Given that these enzymes were discovered in gut bacteria which do not possess AG, it is plausible that these enzymes may also be involved in interbacterial predation. However, it is suggested that, in mycobacteria, GH172 and GH183 family proteins play an important role in the recycling and reorganising of AG and thus represent exciting new prospects as drug targets (Al-Jourani et al., 2023).

As described in Chapter 1 Section 1.1 The mycobacterial cytoplasmic membrane is also distinct from many other bacteria due to the presence of PIMs, LMs and LAMs which extend from the outer leaflet of the membrane into the periplasm and the cell wall (Gilleron et al., 1999). Whilst PIMs represent a large proportion of the phospholipids in the cytoplasmic membrane and act as precursors for LM and LAM synthesis, LMs and LAMs are large glycolipids that contain long mannan chains important for maintaining the structure of the cell wall and cell permeability (Fukuda et al., 2013; Bansal-Mutalik and Nikaido, 2014). In addition, these glycolipids also play a role in virulence, are essential for de-acidification of the phagosome and manipulate the immune response (Gilleron et al., 2003; Dao et al., 2004).

Similar to AG, the maintenance and hydrolysis of LMs and LAMs is poorly understood. Nevertheless mannans have been detected at high concentrations in the mycobacterial capsule indicating that mechanisms exist for LM and LAM cleavage and degradation (Lemassu and Daffe, 1994). In addition, three enzymes from separate GH families have been annotated as putative α-mannosidases, Rv0648, Rv0365c and Rv0584. Rv0584 and Rv0365c are yet to be fully characterised though the GH76 family protein, Rv0365c, has been implicated in macrophage survival (Van Wyk et al., 2017). In contrast, Rivera-Marrero et al. (2001) discovered that, by expressing Rv0648 in *E. coli* and using 4-methylumberlliferyl-α-D-mannopyranoside as a fluorescent substrate, the cell lysate demonstrated an 8-fold increase in α-mannosidase activity. Unfortunately, in the past 20 years there has been no advancement in the discovery of other LM and LAM mannosidases. However, the functional characterisation of Rv0365c and Rv0584 could be worth investigating further as these proteins may also fulfil key roles in LM and LAM degradation.

## 1.4 The ORBIT Method for Generating Mycobacterial Mutants

Reverse-genetics is a powerful technique for determining the roles of genes in which mutants of a target gene are generated and then phenotypically characterised to provide insight into gene function. Nevertheless, using a reverse-genetics approach to investigate mycobacteria can prove difficult as, due to the unique mycobacterial cell wall, traditional methods used for generating mutants are inefficient. However, as

mentioned in Chapter 1 Section 2.5, Murphy et al. (2018) developed a method called ORBIT for generating mycobacterial single gene deletion mutants with increased efficiency. This method was used to construct over a hundred different mycobacterial mutants in both *M. bovis* and *M. smegmatis* and it was discovered that typically, the technique would result in the isolation of 5 to 50 transformants in which 25-50% were successfully genetically engineered.

This method first requires the transformation of electrocompetent cells with a shuttle vector, pKM461, which expresses both a Che9c phage RecT annealase, a Bxb1 phage integrase under the control of an anhydrotetracycline (ATc) inducible promoter and a kanamycin resistance cassette. The transformed cells are then induced and co-electroporated with a targeting oligonucleotide and a payload plasmid, pKM464, resulting in a method requiring only a single recombination step. The targeting oligonucleotide is a chemically synthesised ssDNA containing a phage *attP* site, flanked with regions of homology to a specific DNA target. The payload plasmid contains a hygromycin resistance maker and an *attB* site. As observed in figure 1.28, through HR, on the lagging strand at the replication fork, assisted by the RecT annealase, the targeting oligonucleotide is incorporated into the genome via its homologous ends and an *attP* site is therefore deposited into the genome. Then through *attP*-*attB* site specific recombination the Bxb1 integrase initiates the incorporation of the payload plasmid. To create a gene deletion targeting oligonucleotide is designed to insert in place of the target gene. To create a chromosomal tag, a payload plasmid expressing the tag adjacent to the *attB* site is incorporated at the gene terminus. This is achieved by creating an oligonucleotide that

targets a region just before the stop codon. Successful transformants are then isolated via growth on selective media.

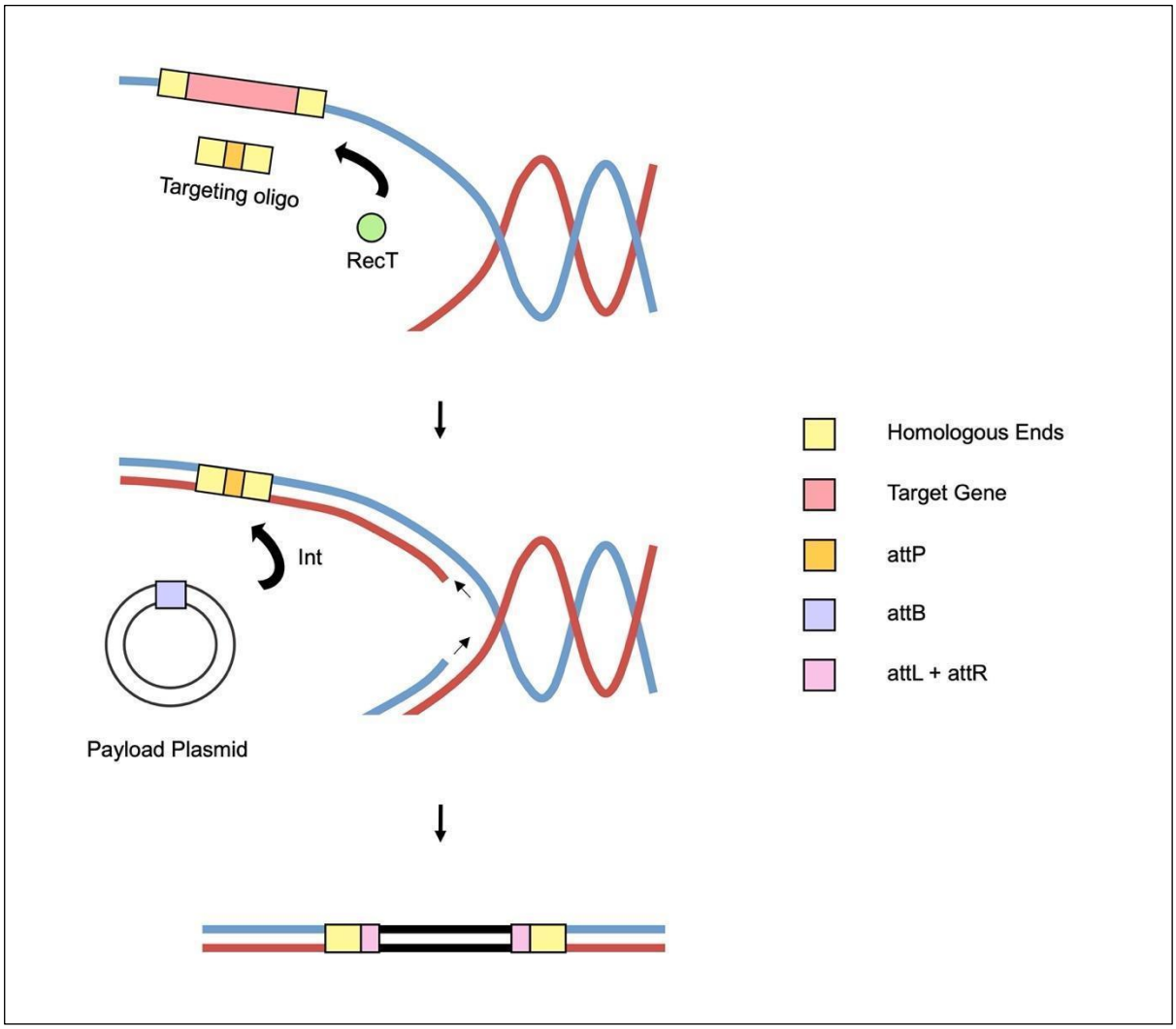



**Figure 1.28 ORBIT methodology for generating a targeted gene deletion.**

On the lagging strand at the replication fork, assisted by a Che9c RecT annealase, the targeting oligonucleotide replaces the target gene via HR, inserting an *attP* site into the genome. Via *attP-attB* site specific recombination, catalysed by the Bxb1 integrase (Int) the payload plasmid is inserted into the bacterial DNA.

To create markerless gene deletions, a secondary transformation can be performed with a Zeocin™ resistant curing plasmid, pKM512, containing Bxb1 phage integrase and gp47 (a recombination directionality factor) under the control of an ATc inducible promoter as well a copy of *sacB*. Due to plasmid incompatibility, selection via Zeocin™ removes cells containing pKM461. Upon expression of gp47, along with the Bxb1 integrase, the attP site is restored, removing the payload in the process. Lastly, through counter-selection on sucrose, the curing plasmid is removed.

## 1.5 Research Aims

Although we have developed a pipeline for condensing transposon insertion libraries to only contain mutants with a high percentage of gene deletion (Chapter 3, Section 2.1), it is possible that functional proteins could be still produced. In addition, transposon insertions can cause polar effects and thus silence downstream genes (Borgers et al., 2020). It is therefore important that any chemical-genomic screening hits are validated by verifying the phenotype of a clean gene KO. Since it has been reported that the ORBIT method is an efficient technique for generating markerless *M. smegmatis* and *M. tuberculosis* KOs, we attempted to determine whether the method could also be adapted to other mycobacterial species and thus could be employed for validating chemical-genomic screening-hits

In addition, to better understand the function of genes involved in maintenance of the mycobacterial cell wall and cell wall associated drug resistance, we attempted to generate *M. smegmatis* mutants of the putative β-lactamase, *lpqF*, homologues of the

putative mannosidases, *Rv0365c* and *Rv0648*, and members of the GH183 family for reverse genetics analysis.

# Part 2 – Results

## Contribution Statement

Section 2.2.3 was conducted in collaboration with Clare Thomas.

## 2.1 Optimising the ORBIT Method for the Generation of Mycobacterial Cell Wall Mutants

The ORBIT method is described in Murphy et al. (2018) and in Chapter 5 Section 1.14.3. Oligonucleotides sequences are listed in Chapter 5 Section 1.14.2. The method for generating markerless ORBIT mutants is described in Chapter 5 Section 1.14.4.

### 2.1.1 Applying the ORBIT Method to *M. marinum*

To validate whether ORBIT could be used to create gene deletions in other species of mycobacteria, we first chose to target the *M. marinum MMARE11_50430* and *rpsL* genes for deletion. As *MMARE11_50430* is a homologue of *Rv3707c*, we targeted this gene to investigate the roles of previously uncharacterised GH183 family proteins. We used *rspL* as a control since it has been discovered that *M. tuberculosis* strains with mutated *rpsL* have an increased resistance to streptomycin (Sreevatsan et al., 1996)

and thus successful transformants could be easily isolated via an additional streptomycin selection step.

Unfortunately, our initial attempts at ORBIT in *M. marinum* were unsuccessful as we failed to yield any *MMARE11_50430* and *rpsL* targeting transformants on the selective media.

Although it has been previously reported that the transformation of *M. marinum* is most efficient at RT (Talaat and Trucksis, 2000), as *M. smegmatis* transformation is most optimal at 4°C (Murphy et al., 2018), we also attempted the washing and electroporation steps of the ORBIT method at this temperature. Furthermore, as glycine has been reported to reduce cell growth of mycobacteria and result in the loss of PDIMs (Hermans, Boschloo and De Bont, 1990), ORBIT was also attempted without the addition of glycine. However, transformation remained unsuccessful as we were still unable to obtain any transformants on the selective media.

As ORBIT has been successfully applied to other mycobacterial species we concluded that it was unlikely that the ORBIT method could not be adapted to *M. marinum*. We therefore hypothesised that the genes targeted for deletion may be essential or have extremely defective growth. Indeed, we discovered that *M. marinum rpsL* and its *M. tuberculosis* homologue have been found to be essential via genome-wide TnSeq studies explaining why no transformants could be generated by targeting this gene (DeJesus et al., 2017; Lefrancois et al., 2023). On the other hand, *MMARE11_50430* is non-essential. However, it is still possible that strains lacking *MMARE11_50430*

could have reduced growth thus making it more difficult to obtain a knockout.

Nevertheless, we concluded that the most likely reason for our failure to obtain a *MMARE11_50430* knockout was due to experimental error or the requirement for further optimisation of the ORBIT methodology in *M. marinum*. Therefore, to ensure that the experimental procedure was robust we decided to validate and optimise the ORBIT method in *M. smegmatis* before conducting further experiments in *M. marinum*.

## 2.1.2 Validation and Optimisation of ORBIT in *M. smegmatis*

In order to assess the roles and genetic redundancy of GH183 family genes we targeted *MSMEG_4352*, *MSMEG_4360*, *MSMEG_4365*, *MSMEG_2107* and *MSMEG_6255* for deletion with the aim of generating a strain with multiple GH183 family deletions.

In contrast to *M. marinum,* initial attempts at ORBIT in *M. smegmatis* resulted in the successful isolation of transformants on selective media. However, the number of transformants was far lower than previously reported (Murphy et al., 201). Unfortunately, following PCR using primers designed to amplify the target gene, it was confirmed that the gene deletion was unsuccessful as the transformants did not harbour the payload plasmid insertion at the target site and resembled the wildtype phenotype (figure 1.29).

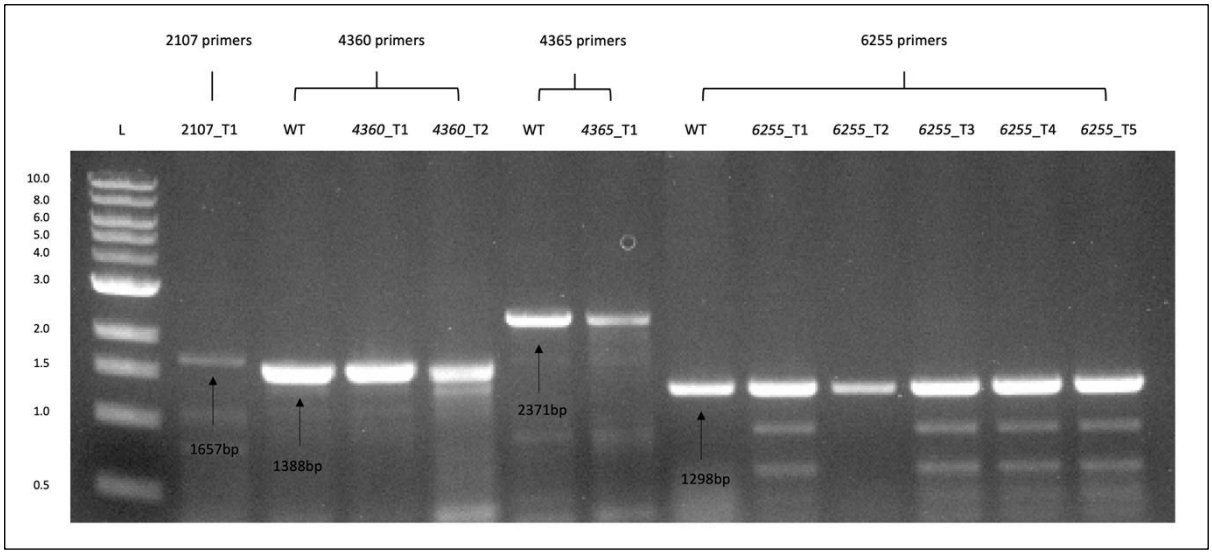



**Figure 1.29 PCR verification of ORBIT transformants.**

PCR was conducted with primers designed to amplify the respective target gene. L = ladder, T = transformant, WT = wildtype. *MSMEG_2107* targeting primers produce a 1,657 bp product for WT and a 3,319 bp product for mutants. *MSMEG_4360* targeting primers produce a 1,388 bp product for WT and a 3,372 bp product for mutants. *MSMEG_4365* targeting primers produce a 2,371 bp product for WT and a 3,454 bp product for mutants. Lastly, *MSMEG_6255* targeting primers produce a 1,298 bp product for WT and a 3,351 bp product for mutants.

Interestingly, through PCR using pKM464 targeting primers, it was discovered that transformants harboured the payload plasmid (figure 1.30). We therefore concluded that the successful isolation of transformants was the result of illegitimate recombination in which the payload plasmid had been inserted at an off-target site. We hypothesised that the failure to knockout the target gene may be due to the inefficient induction of the pKM461 plasmid or the employment of incorrect concentrations of the targeting oligonucleotide.

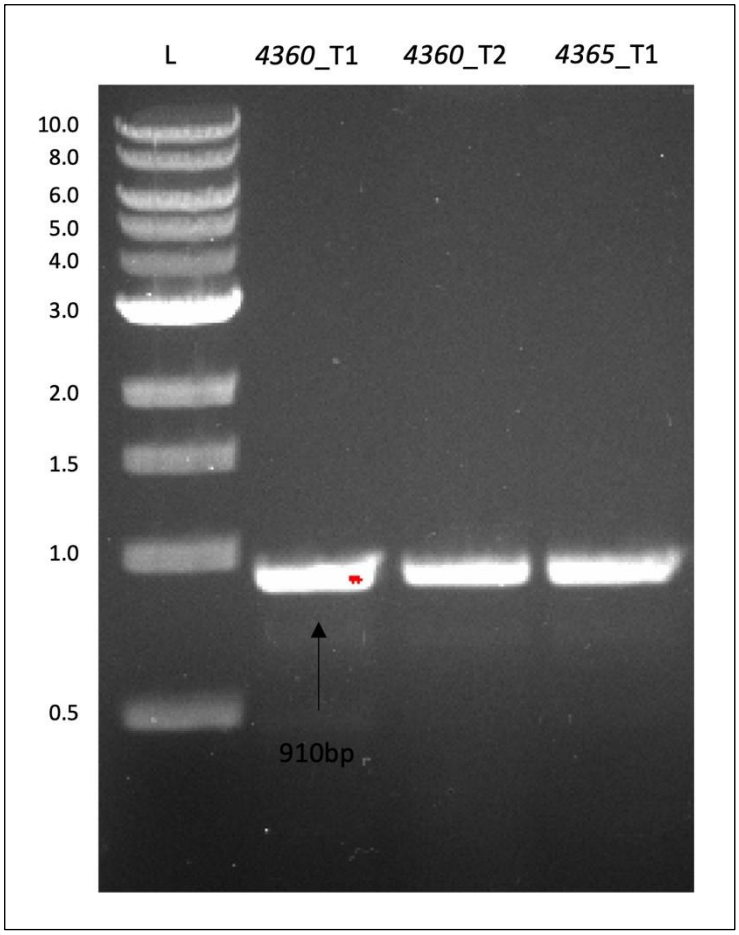



**Figure 1.30 PCR verification of ORBIT transformants.**

PCR was conducted with primers designed to amplify pKM464. L = ladder, T = transformant. Primers targeting pKM464 do not produce a product for WT and produce a 910 bp product for mutant.

To test these hypotheses, we verified the concentration of the oligonucleotides and repeated ORBIT using freshly prepared ATc at varying concentrations (250 ng/mL, 500 ng/mL, 1000 ng/mL and 2000 ng/mL). As shown in table 1.1, using a NanoDrop spectrophotometer we determined that the oligonucleotides were of the expected concentrations. We therefore discounted the oligonucleotide concentration as the cause for an inability to obtain pKM464 insertions at the target site. Furthermore, despite previous reports suggesting that in *M. smegmatis* a concentration of ATc between 50-200 ng/mL is optimal, with higher concentrations causing growth defects

(Ehrt et al., 2005), we discovered that the ATc concentration of 500ng/mL yielded the highest number of transformants.

**Table 1.1 Targeting oligonucleotide concentrations.**

Determined via a NanoDrop spectrophotometer.

| Targeting oligonucleotide | Concentration |
|---|---|
| *MSMEG_2107* | 1361.3 ng/μL |
| *MSMEG_4352* | 1316.1 ng/μL |
| *MSMEG_4360* | 1358.9 ng/μL |
| *MSMEG_4365* | 1445.2 ng/μL |
| *MSMEG_6255* | 1435.0 ng/μL |

Fortunately, upon verification of the transformants via PCR, we were able to successfully isolate a Δ*MSMEG_4352* mutant (figure 1.31a). This mutant was obtained when conducting ORBIT under the standard conditions outlined by Murphy et al. (2018) and thus we hypothesised that it was likely the ORBIT method was most efficient using the standard practice. The successful transformant was a mixed population of WT and Δ*MSMEG_4352* cells as represented by the presence of two bands on the agarose gel (figure 1.31a). To select for pure mutant the successful transformant was re-streaked on hygromycin selective media before repeating PCR and gel electrophoresis. As figure 1.31b demonstrates in lanes 1, 4, 5 and 6, pure Δ*MSMEG_4352* was obtained.

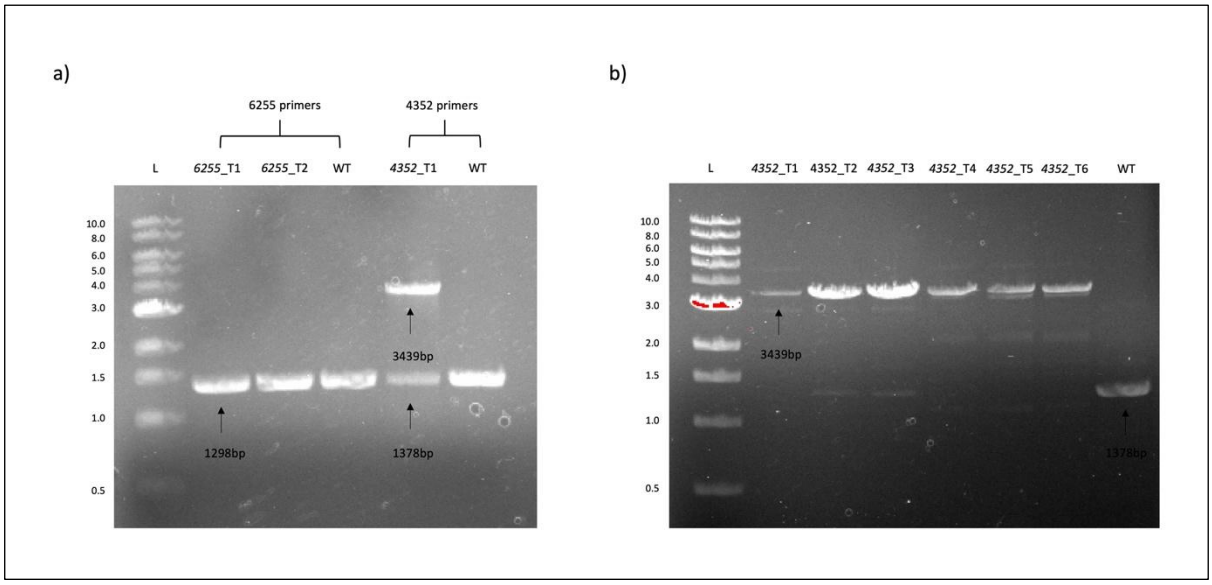



**Figure 1.31 PCR verification of a ∆*MSMEG_4352* ORBIT mutant.**

PCR was conducted with primers designed to amplify the respective target gene. L = ladder, T = transformant, WT = wildtype. *MSMEG_4352* targeting primers produce a 1,378 bp product for WT and a 3,439 bp product for mutants. *MSMEG_6255* targeting primers produce a 1,298 bp product for WT and a 3,351 bp product for mutants. a) agarose gel of PCR products from transformants grown on hygromycin selective media following ORBIT. b) agarose gel of PCR products from colonies obtained by restreaking 4352_T1.

Following the success of generating a ∆*MSMEG_4352* mutant in *M. smegmatis* we designed two further oligonucleotides to target two putative mannosidases, *MSMEG_0740* (*Rv0365c*) and *MSMEG_1341* (*Rv0648*). As figure 1.32 demonstrates, we were able to successfully generate a *MSMEG_0740* mutant via ORBIT.

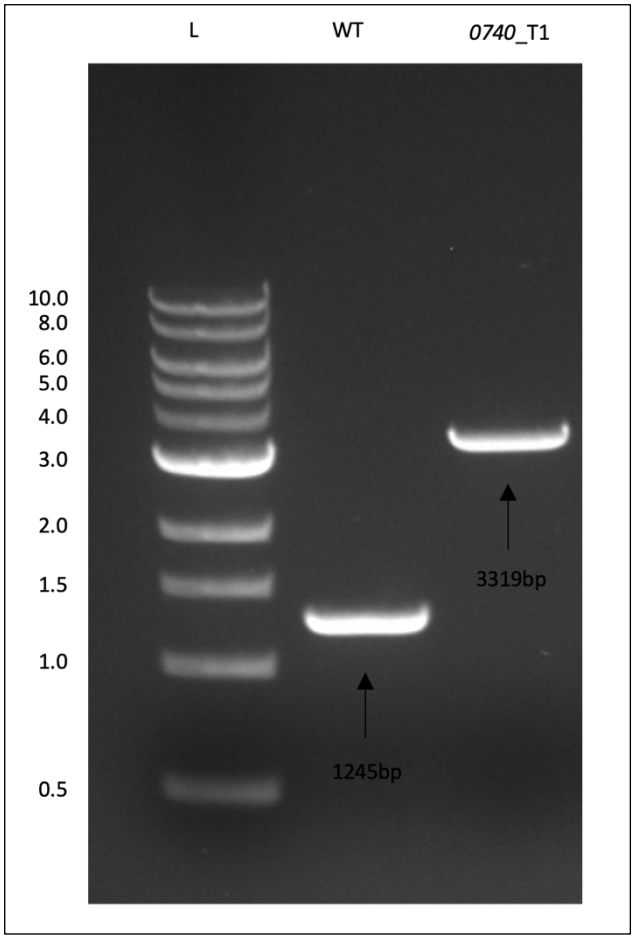



**Figure 1.32 PCR verification of a Δ*MSMEG_0740* ORBIT mutant.**

PCR was conducted with primers designed to amplify the respective target gene. L = ladder, T = transformant, WT = wildtype. *MSMEG_0740* targeting primers produce a 1,245 bp product for WT and a 3,319 bp product for mutants.

However, further attempts at generating ORBIT mutants were unsuccessful. We therefore hypothesised that the targeting oligonucleotides may have been designed incorrectly. Indeed, we discovered that all oligonucleotides had been designed to target the leading strand instead of the lagging strand and a stop codon had been incorporated into the end of the *attP* site. Furthermore, the *attP* site had been reversed in all oligonucleotides apart from the *MSMEG_2107* targeting oligonuclotide and, with the exception of *MSMEG_2107* and *MSMEG_4365* targeting oligonucleotides, the

insertion had caused a frameshift. Since it has been reported that ssDNA HR in mycobacteria is increased by up to $10^4$-fold by targeting the lagging strand (Van Kessel and Hatfull, 2008) it seems likely that the cause for our inability to generate knockouts via ORBIT was due to the targeting of the leading strand. In addition, creating frameshifts upon insertion of the payload plasmid could cause a missense and thus potential disrupt downstream genes, some of which may be essential. We therefore redesigned two oligonucleotides to target *MSMEG_6255* and *MSMEG_1130. MSMEG_1130* was targeted for deletion as a knockout strain was required to support a publication (Helena-Bueno et al., 2024).

Using the newly designed targeting oligonucleotides we successfully obtained *MSMEG_6255* and *MSMEG_1130* ORBIT mutants on our first attempt (figure 1.33). However, as we only isolated around 1-10 colonies per plate and discovered that only ~20% of these colonies were successful transformants, we were unable to achieve the same level of efficiency as Murphy et al. (2018). Nevertheless, these results confirm that the most likely reason for our failure to obtain ORBIT mutants previously was due to the incorrect construction of the targeting oligonucleotides.

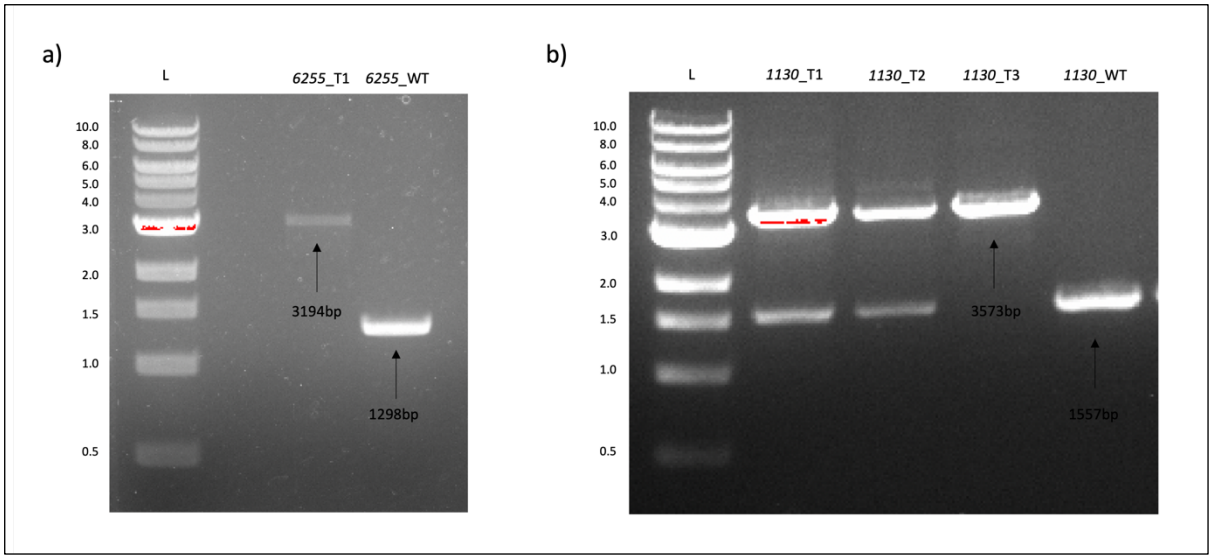



**Figure 1.33 PCR verification Δ*MSMEG_6255 and* Δ*MSMEG_1130* ORBIT mutants.**

(a) Δ*MSMEG_6255*, (b) Δ*MSMEG_1130*. PCR was conducted with primers designed to amplify the respective target gene. L = ladder, T = transformant, WT = wildtype. *MSMEG_6255* targeting primers produce a 1,298 bp product for WT and a 3,194 bp product for mutants. *MSMEG_1130* targeting primers produce a 1,557 bp product for WT and a 3,573 bp product for mutants.

## 2.1.3 Resolving ORBIT Mutants via Curing of the Payload Plasmid

Finally, we attempted to cure the ORBIT plasmids from the successfully generated mutants as well as a previously generated *M. smegmatis* mutant, Δ*lpqF*$_{smeg}$. Δ*lpqF*$_{smeg}$ was selected to investigate the role of LpqF as a potential β-lactamase. Being able to remove the ORBIT plasmids is extremely useful as we can generate clean knockouts and also re-employ the antibiotic markers for the generation of strains with multiple gene deletions. Unfortunately, the creation of markerless gene deletions in Δ*MSMEG_0740* and Δ*MSMEG_6255* was unsuccessful. Nevertheless, for Δ*MSMEG_4352* and Δ*lpqF*$_{smeg}$, we were able to successfully isolate colonies on sucrose selective media that were sensitive to both hygromycin and kanamycin.

Following PCR and gel electrophoresis we confirmed the successful curing of the Δ*MSMEG_4352* and Δ*lpqF*$_{smeg}$ mutants (figure 1.34). To easily distinguish between the marked and unmarked Δ*lpqF*$_{smeg}$ strains we renamed the unmarked Δ*lpqF*$_{smeg}$ strain Δ*lpqFu*$_{smeg}$. Interestingly, the Δ*lpqFu*$_{smeg}$ strain did not produce any PCR product (figure 1.34b) but this is likely due to the need for a longer extension time as, due to the insertion of the payload plasmid, the mutant expresses a much larger PCR product than WT.

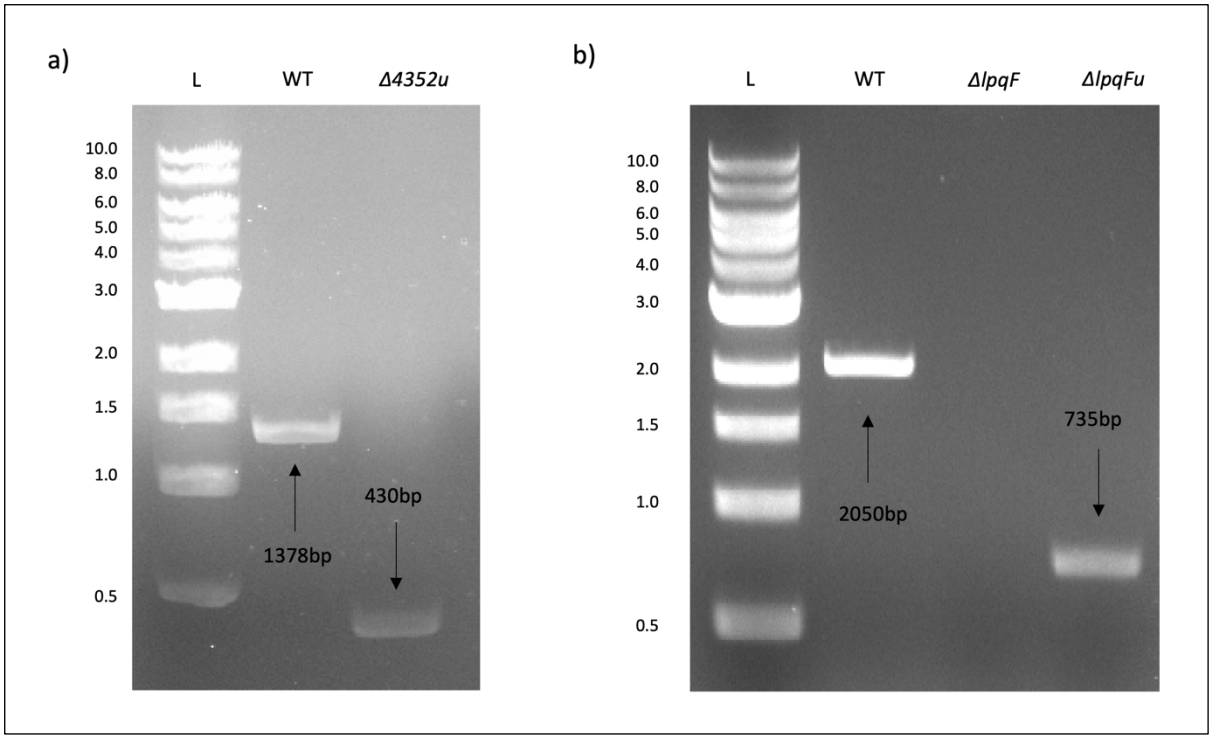



**Figure 1.34 PCR verification of Δ*MSMEG_4352* and Δ*lpqF* ORBIT cured mutants.**

(a) Δ*MSMEG_4352*, (b) Δ*lpqF*$_{smeg}$. PCR was conducted with primers designed to amplify the respective target gene. L = ladder, C = cured mutant, WT = wildtype. *MSMEG_4352* targeting primers produce a 1,378 bp product for WT and a 430 bp product for cured mutants. *lpqF*$_{smeg}$ targeting primers produce a 2,050 bp product for WT and a 735 bp product for cured mutants.

## 2.2 Phenotypic Characterisation of ORBIT Mutants

Since the ORBIT method required an extensive period of optimisation, mainly due to our own errors during the design of the targeting oligonucleotides, we decided to investigate the phenotypes of the mutants that we had successfully generated rather than attempting to generate further mutants with multiple gene deletions or adapting the ORBIT method for *M. marinum*.

### 2.2.1 Assaying the Growth of ORBIT Mutants

To determine whether the ORBIT mutants exhibited altered fitness compared to WT *M. smegmatis*, we conducted a spot assay and growth curves (Chapter 5, Section 1.15). As a control, the fitness of a strain with a known growth defect, Δ*MSMEG_4722* (Bhatt et al., 2008), was also examined. As demonstrated by figure 1.35, the ORBIT strains all displayed similar growth to WT on solid 7H9 agar. Only the Δ*MSMEG_4722* control showed reduced growth. Likewise, mutant growth in 7H9 media also mimicked WT growth apart from Δ*MSMEG_4722* (figure 1.36). Given that the *M. bovis* BCG homologue of Δ*MSMEG_0740* has been previously reported to have a prolonged lag phase and a 20% reduction in exponential phase growth rate we expected to see a similar phenotype. However, it is possible that other genes, such as *MSMEG_1341*, encoding a GH5 family α-mannosidase, fulfil a more compensatory role in *M. smegmatis* and thus *MSMEG_0740* is less important for *M. smegmatis* growth.

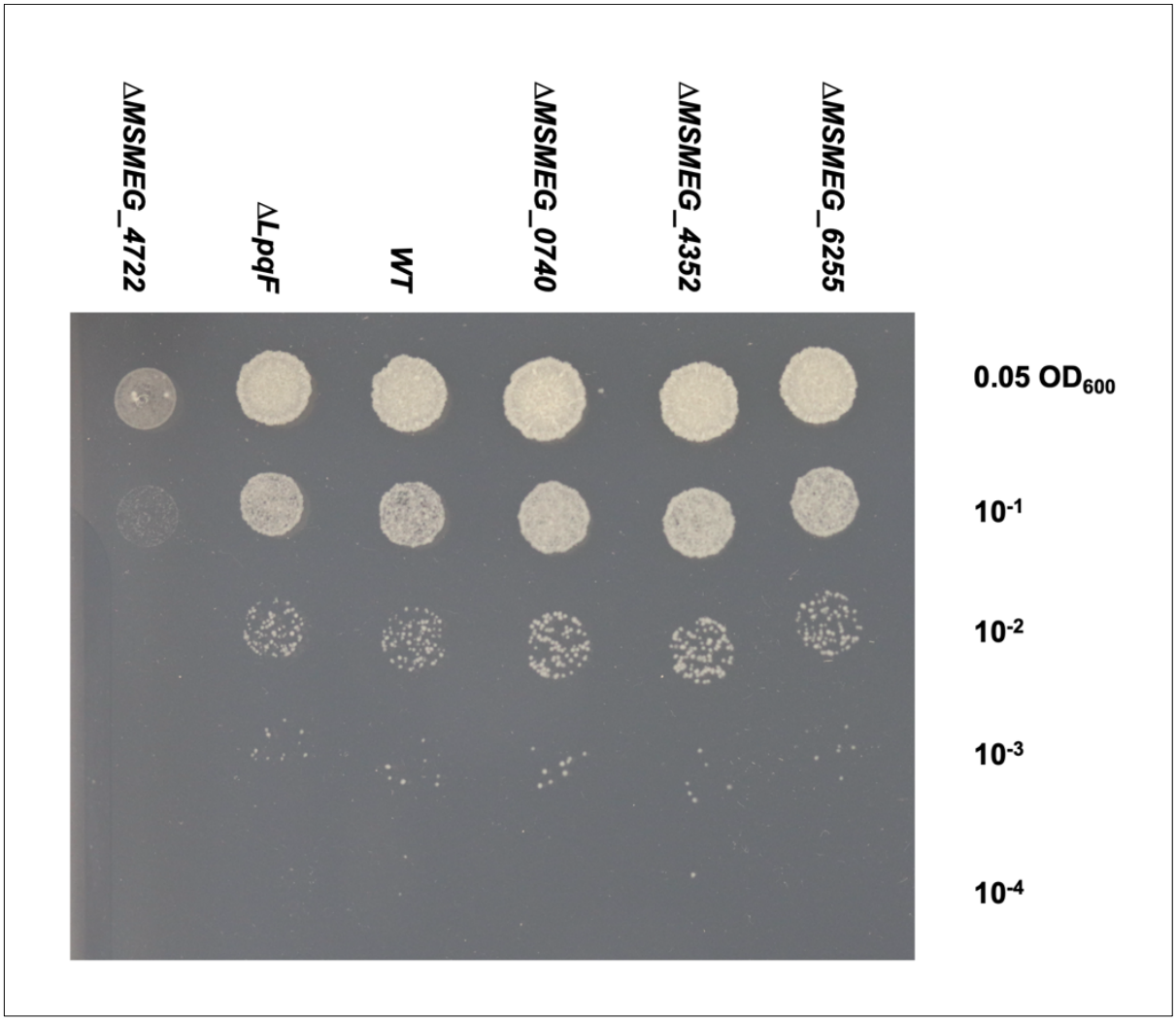



**Figure 1.35 Spot assay of *M. smegmatis* ORBIT mutants**

The growth of *M. smegmatis* ORBIT mutants determined through a spot assay on 7H9 agar. Starting $OD_{600nm}$ = 0.05. The mutants were grown for 3 days at 37°C.

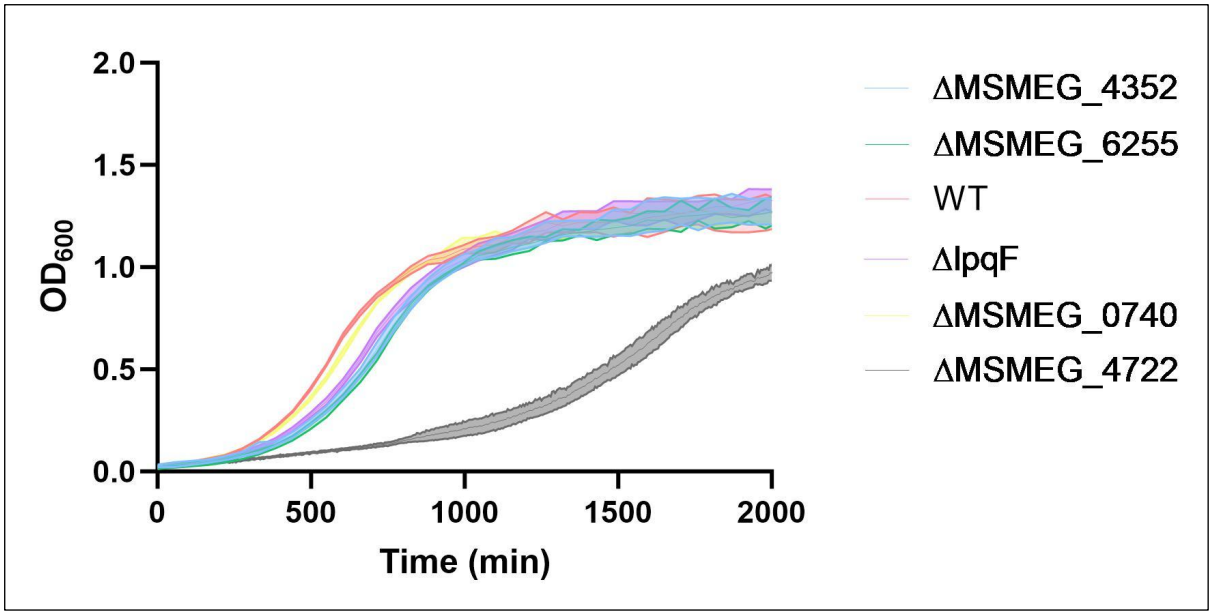



**Figure 1.36 Growth curves of *M. smegmatis* ORBIT mutants in 7H9 medium.**

The growth of *M. smegmatis* ORBIT mutants in 7H9 medium compared with WT and a defective growth positive control, ΔMSMEG_4722. For each growth curve the shaded area represents the SD between three biological replicates and the centre of measure is represented by the mean.

## 2.2.2 Deletion of *lpqF* Reveals Enhanced Growth in Meropenem

Since, under standard conditions, we were unable to identify any phenotypic differences between the ORBIT strains, we conducted a small chemical screen to observe changes in fitness in response to various stresses. This was conducted using the same method described in Chapter 3, Section 2.2 (Chapter 5, Section 1.13) in which each strain was grown on 7H9 agar plates containing sub-inhibitory concentrations of different drugs and environmental conditions before quantifying colony growth using IRIS software and generating fitness S-scores using ChemGAPP. Figure 1.37 demonstrates the growth of each strain compared to WT in which a S-score > 1 represents an increase in fitness and a score < 1 signifies a reduction in

fitness. Interestingly, in contrast to the growth curves and spot assays, under standard conditions, we detected a slight reduction in fitness of Δ*lpqF*$_{smeg}$ and Δ*MSMEG_4352* compared to WT. We therefore hypothesised that this method may have a better sensitivity to changes in mutant fitness. We also discovered that, in meropenem, Δ*MSMEG_0740* had a significant fitness benefit but unfortunately this result could not be replicated in subsequent screens. However, as demonstrated more clearly in figure 1.38, the Δ*lpqF*$_{smeg}$ fitness defect was consistently supressed when grown in the presence of meropenem or amoxicillin-clavulanic acid and partially supressed in the presence of imipenem.


**Figure 1.37 A chemical genetic screening of *M. smegmatis* ORBIT mutants.**

Colony size was calculated via IRIS and, via CHEMGAPP, mutant fitness S-scores were generated as a comparison to WT growth. An S-score < 1 (red) represents a decrease in fitness and an S-score > 1 (blue) indicates a fitness advantage.

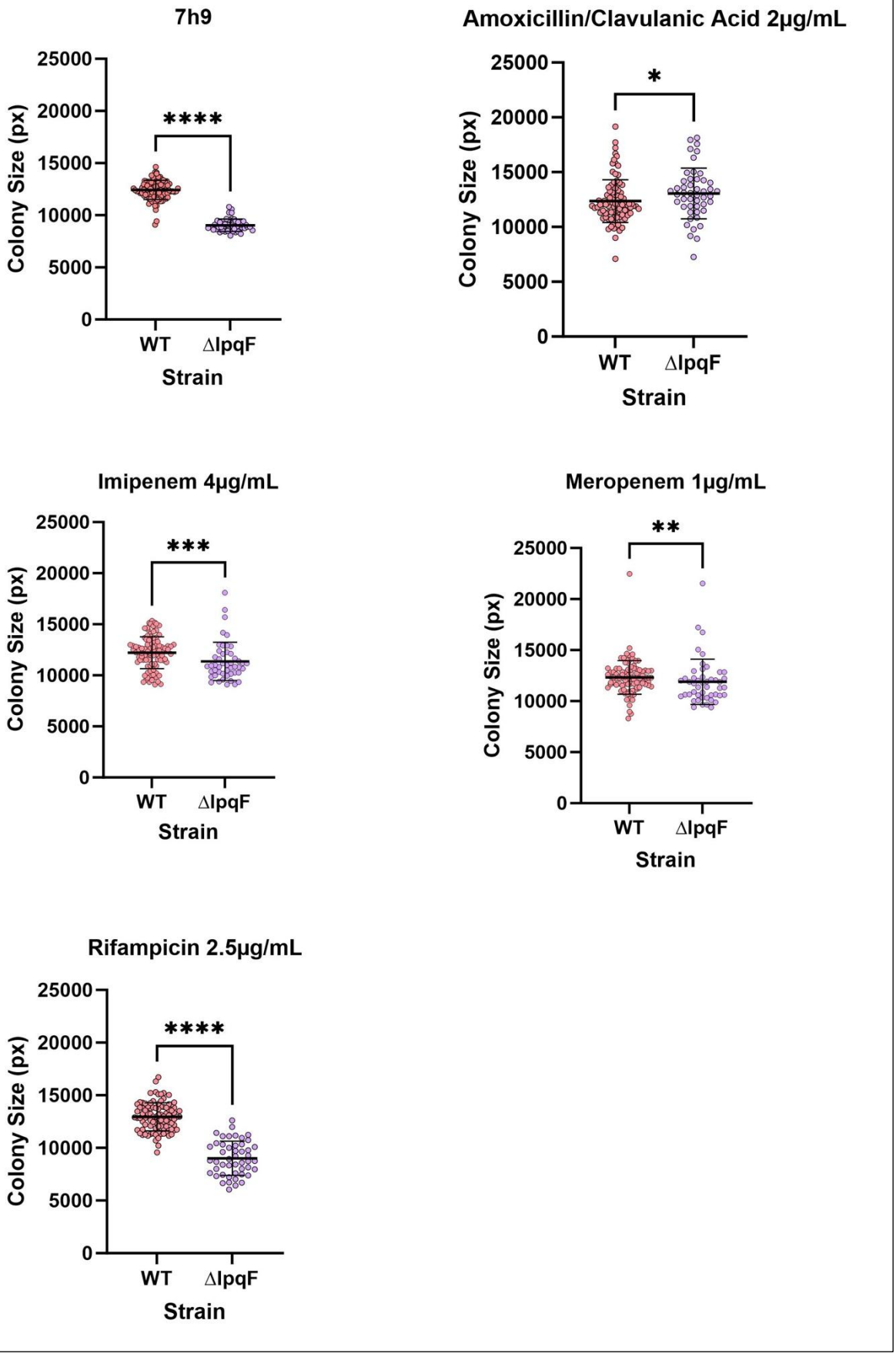

7h9
Colony Size (px)
25000
20000
15000
10000
5000
0
****
WT
ΔlpqF
Strain
Amoxicillin/Clavulanic Acid 2μg/mL
*
Imipenem 4μg/mL
***
Meropenem 1μg/mL
**
Rifampicin 2.5μg/mL
****

**Figure 1.38 Chemical-genetic screening of ORBIT mutants.**

Colony size scatter plot of *M. smegmatis* WT compared to Δ*lpqF* under different conditions. Each data point represents a technical replicate across three biological replicates. N = 96 for WT and N = 48 for Δ*lpqF*. The bars represent mean ± SD. Significance was evaluated via Mann Whitney tests for 7H9 (****P < 0.0001), amoxicillin and clavulanic acid (*P = 0.0144), imipenem (***P = 0.0003) and meropenem (**P = 0.0049) and via a paired T test for rifampicin (****P ≤ 0.0001).

To verify these findings we repeated the chemical screen using additional conditions and monitored the growth of the unmarked *lpqF* deletion mutant, Δ*lpqFu*$_{smeg}$, and a complemented *lpqF* deletion strain (Δ*lpqFc*$_{smeg}$). Δ*lpqFu*$_{smeg}$ was employed as a control to ensure that the insertion cassette was not responsible for the observed changes in mutant fitness. Δ*lpqFc*$_{smeg}$ was also used as a control to confirm that the mutant phenotype could be restored to WT upon expression of LpqF thus validating that the altered mutant fitness was not due to polar or off-target effects. The complemented strain was constructed via the transformation of Δ*lpqFu*$_{smeg}$ with a pTIC plasmid carrying a copy of the *lpqF*$_{smeg}$ gene and placed under the control of an ATc inducible promoter. To ensure that the plasmid vector does not result in any phenotypic changes, Δ*lpqFe*$_{smeg}$ was also generated. This was achieved by electroporating Δ*lpqFu*$_{smeg}$ with an empty pTIC plasmid. To induce the complement plasmid, ATc was added to plates at a concentration of 500 ng/mL. As depicted in figure 1.39, Δ*lpqFu*$_{smeg}$, Δ*lpqFc*$_{smeg}$ and Δ*lpqFe*$_{smeg}$ all exhibited reduced growth when compared with Δ*lpqF*$_{smeg}$. This result could indicate that the ORBIT payload plasmid results in a growth benefit. However, as we concluded that this would be an unlikely outcome, we checked the ordering of the source plate and discovered that the strains were arrayed with an accidental bias with Δ*lpqF*$_{smeg}$ grown at the plate edges. As

described in Chapter 3 Section 1.3, due to edge effects, colonies are much larger at the plate edges and whilst ChemGAPP can control for this, for maximum efficiency, it is vital that each strain is distributed equally. Furthermore, although Δ*lpqF*$_{smeg}$ exhibited improved fitness in the presence of meropenem, amoxicillin and imipenem, Δ*lpqFu*$_{smeg}$ only demonstrated an improved fitness in the former. It is possible that the removal of the ORBIT payload plasmid resulted in an alteration in phenotype due to polar effects. However, it is also plausible that due to the plate effects and thus reduced colony size, changes in fitness could not be as easily distinguished. In addition, despite the plate effects, it is evident that Δ*lpqfc*$_{smeg}$ was unsuccessful at complementing the WT phenotype. It is therefore likely that Δ*lpqFc*$_{smeg}$ was not efficiently induced and thus optimisation of the ATc concentration may result in a successfully complemented strain. However, it is possible that the complement strain is not functional and requires redesigning.

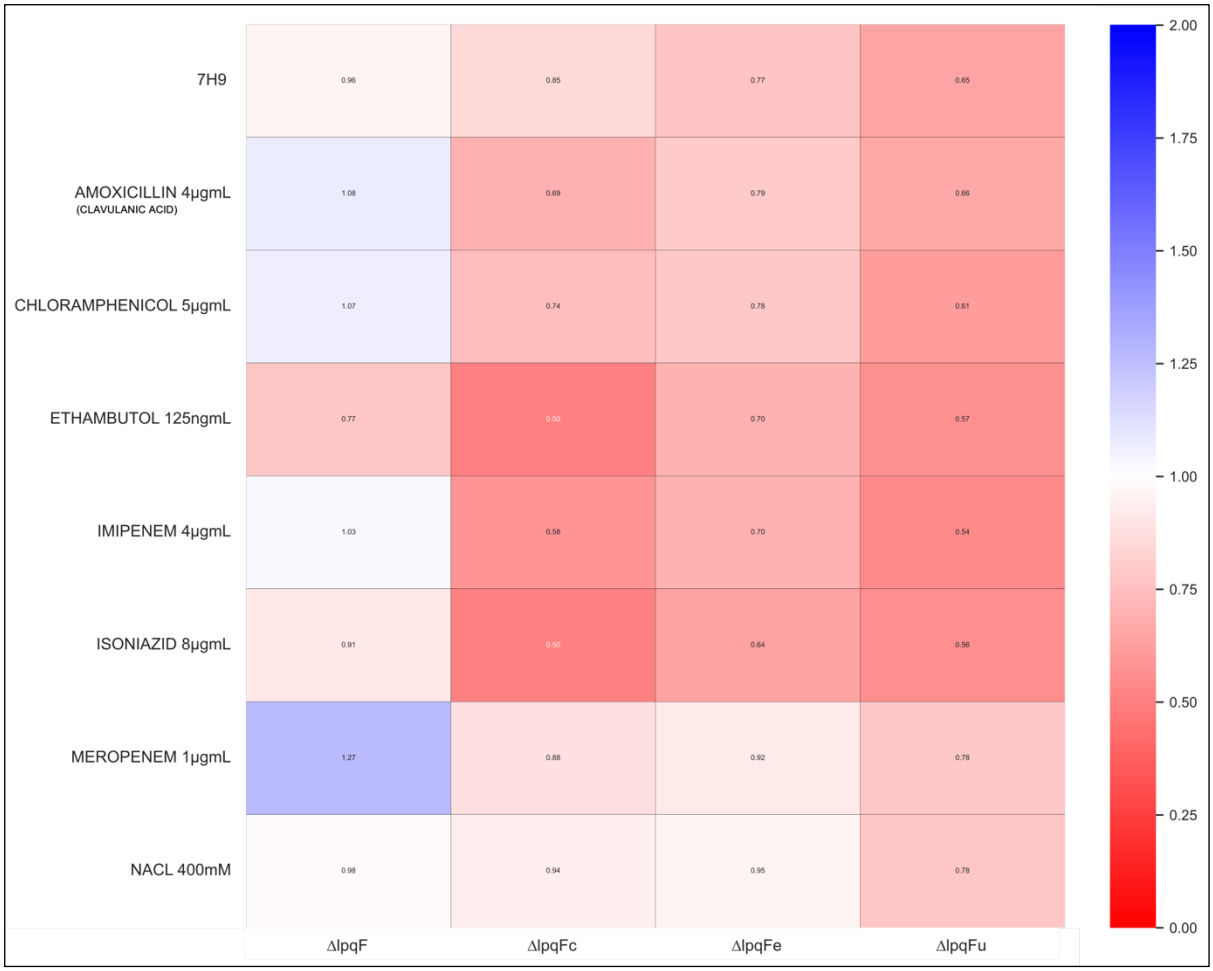



**Figure 1.39 Chemical-genetic screening of LpqF.**

*M. smegmatis* Δ*lpqF.* Δ*lpqFu* (unmarked Δ*lpqF*) and Δ*lpqFc* (complemented Δ*lpqF*) and ΔlpqFe (unmarked Δ*lpqF* expressing an empty pTIC vector) mutants were chemically screened. Colony size was calculated via IRIS and, via CHEMGAPP, mutant fitness S-scores were generated as a comparison to WT growth. An S-score < 1 (red) represents a decrease in fitness and an S-score > 1 (blue) indicates a fitness advantage.

Given the differences observed between the Δ*lpqF*$_{smeg}$ and Δ*lpqFu*$_{smeg}$ phenotypes, to further validate the results, we conducted Δ*lpqF*$_{smeg}$ and WT growth curves in subinhibitory concentrations of amoxicillin with clavulanic acid, imipenem, and meropenem. As shown by figure 1.40, when grown with imipenem or amoxicillin with clavulanic acid, the deletion of *lpqF*$_{smeg}$ resulted in a similar growth to WT. It is possible

that due to the transition to liquid medium, a different concentration of drug may be required to observe a change in growth. However, a significant difference was observed in meropenem in which WT failed to grow but Δ*lpqF*$_{smeg}$ sustained growth, albeit defectively.

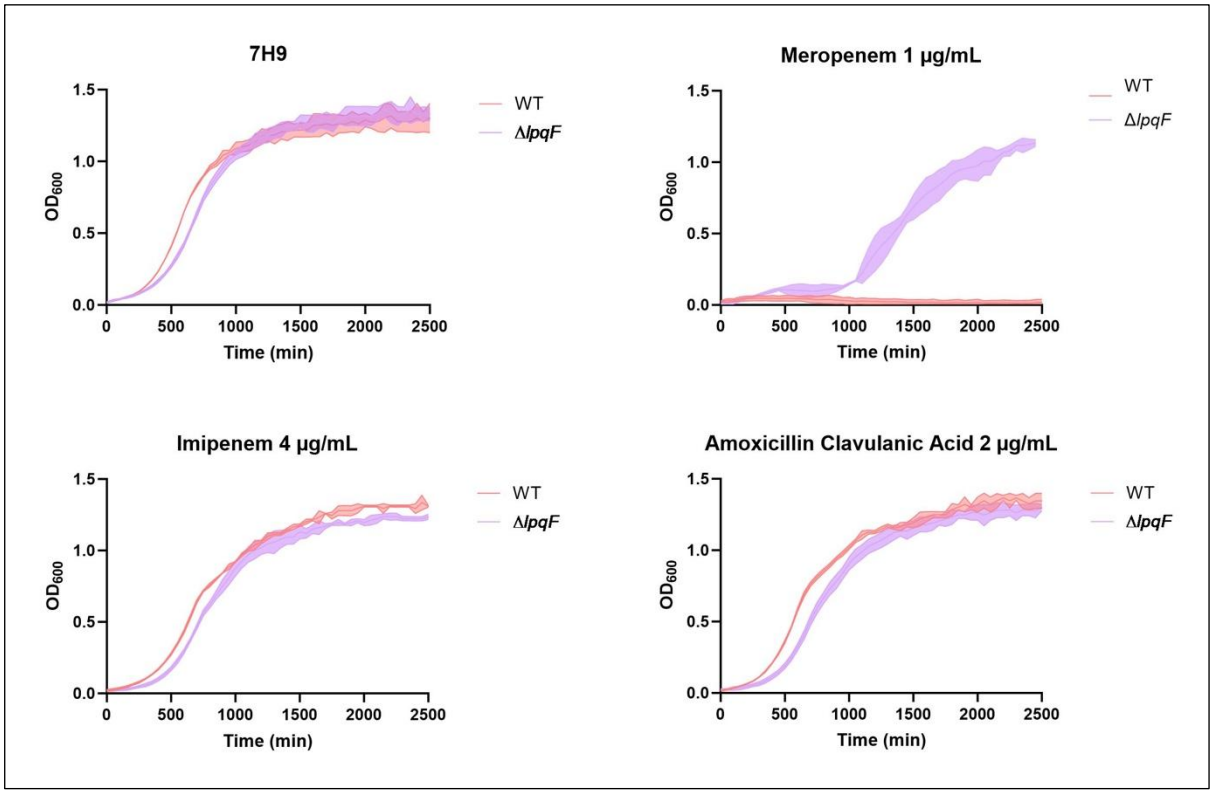



**Figure 1.40 Growth curve of Δ*lpqF* under chemical stress.**

Growth curves of WT *M. smegmatis* compared to Δ*lpqF* in 7H9 medium supplemented with different antibiotics. For each growth curve the shaded area represents the SD between three biological replicates and the centre of measure is represented by the mean.

To determine whether a similar phenotype could be observed in other mycobacterial species we also conducted a chemical screen in *M. marinum*. However, as demonstrated in figure 1.41, although Δ*lpqF*$_{mar}$ demonstrated reduced growth compared to WT, no significant change in fitness was observed across any of the

conditions. This could indicate that $LpqF_{mar}$ has a different function in *M. marinum* but the lack of phenotype could also be explained by genetic redundancy or the reduced overall growth restricting distinction between phenotypes.

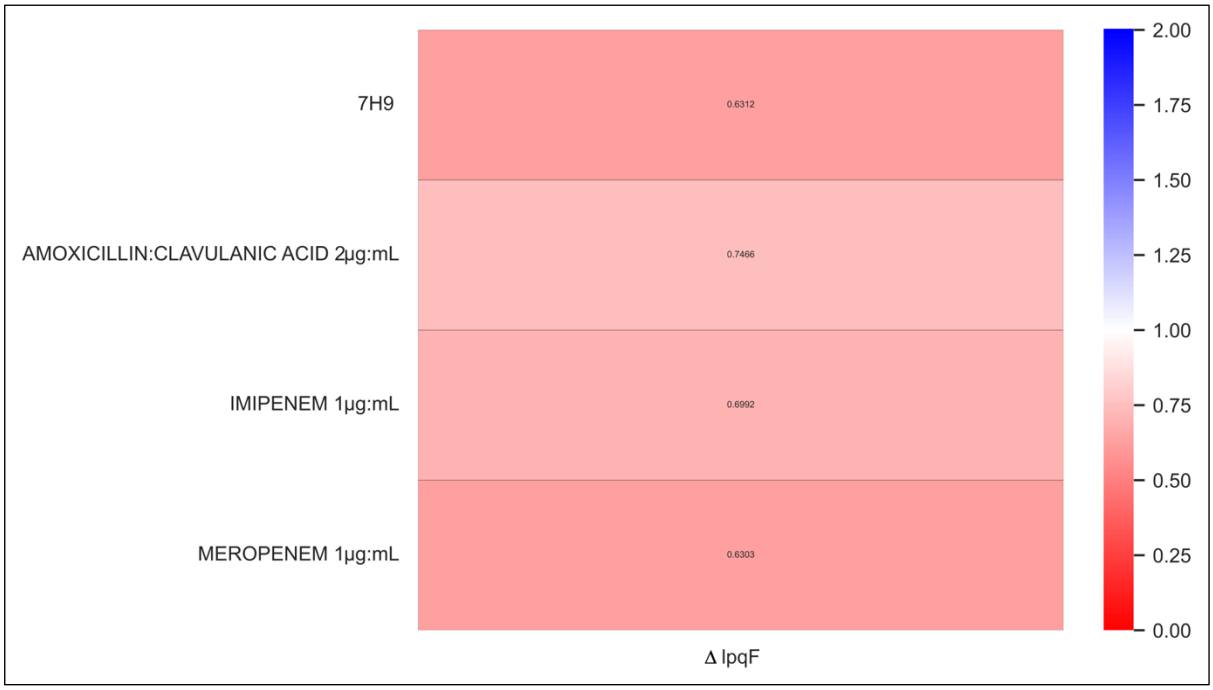



**Figure 1.41 Chemical genetic screening of *M. marinum* Δ*lpqF*.**

Colony size was calculated via IRIS and, via CHEMGAPP, mutant fitness S-scores were generated as a comparison to WT growth. An S-score < 1 (red) represents a decrease in fitness and an S-score > 1 (blue) indicates a fitness advantage.

Together, these results suggest that, although there is some uncertainty over the response of Δ$lpqF_{smeg}$ to imipenem and amoxicillin with clavulanic acid, the deletion of $lpqF_{smeg}$ is important for meropenem resistance. As $LpqF_{smeg}$ is annotated as a hypothetical β-lactamase, and, since meropenem is a carbapenem and thus resistant to many β-lactamases, this result was unexpected as the deletion of a β-lactamase should have little to no effect on growth in meropenem. It is possible that meropenem

is still sensitive to *lpqF*$_{smeg}$ but in this instance we would expect to see an increased susceptibility to meropenem in a *lpqF*$_{smeg}$ deletion strain. We therefore concluded that *lpqF* may fulfil another role. As meropenem targets L,D-transpeptidases and PBPs, in particular D,D-carboxypeptidases (Kumar et al., 2012), we hypothesised that *lpqF* may encode a PBP, an L,D-transpeptidase, or play another important role in maintenance of the PG cell wall.

## 2.2.3 Deletion of *M. marinum* and *M. smegmatis lpqF* Results in Altered Cell Size and Chaining

As described in Chapter 5, Section 1.16, to determine whether the deletion of *lpqF* resulted in a change in the cell wall, we labelled *M. smegmatis* and *M. marinum* WT and Δ*lpqF* PG with the fluorescent dye HADA and visualised mid-log phase cells via microscopy (figures 1.42 and 1.43). As figure 1.42c demonstrates, WT and Δ*lpqF*$_{smeg}$ had very similar cell lengths. Likewise, although significantly different, WT and Δ*lpqF*$_{smeg}$ also exhibited similar cell widths (figure 1.42d) and cell chaining (figure 1.42e) though Δ*lpqF*$_{smeg}$ was slightly thinner and more chained.

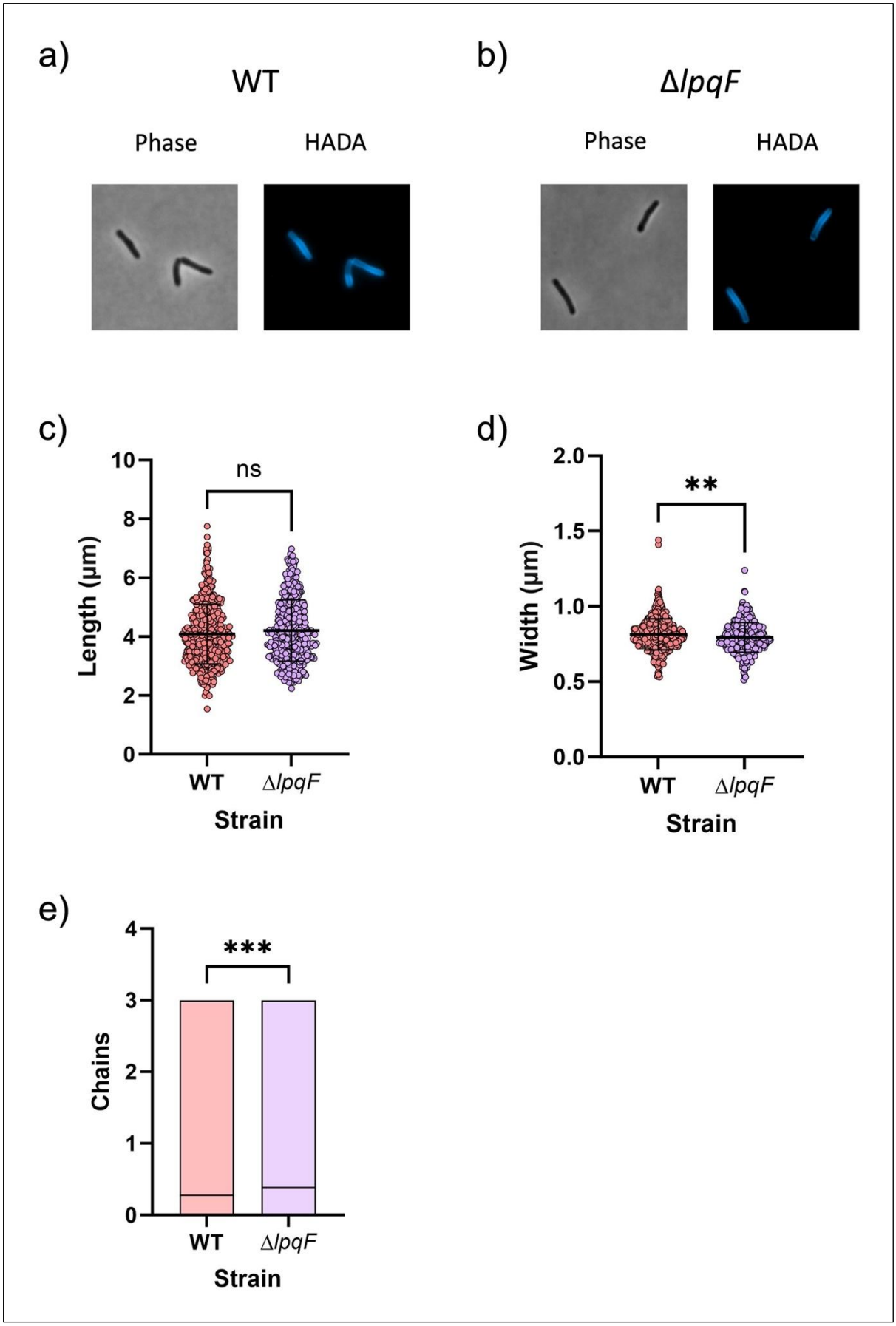

a)
WT
Phase
HADA
b)
ΔlpqF
Phase
HADA
c)
ns
Length (µm)
10
8
6
4
2
0
WT
ΔlpqF
Strain
d)
**
Width (µm)
2.0
1.5
1.0
0.5
0.0
WT
ΔlpqF
Strain
e)
***
Chains
4
3
2
1
0
WT
ΔlpqF
Strain

**Figure 1.42 Microscopy data of *M. smegmatis* WT and Δ*lpqF*.**

(a) WT cells and (b) Δ$lpqF_{smeg}$ cells were imaged via phase contrast and via the detection of HADA fluorescence. (c) the cell length (mean ± SD), (d) width (mean ± SD) and (e) the number of chains (mean) were calculated via MicrobeJ. Significance was evaluated via a Mann Whitney test for (c) P = 0.0833, (d) **P = 0.0016 and (e) ***P = 0.0006.

In contrast, Δ$lpqF_{mar}$ exhibited significantly longer, thicker and more chained cells (figure 1.43a-c). In combination, apart from cell chaining, these results demonstrate that *M. marinum* and *M. smegmatis* have conflicting phenotypes for Δ*lpqF*. It is therefore possible that these homologues fulfil a different function in each species however it is also plausible that there may be a level of redundancy. It appears that Δ$lpqF_{smeg}$ is more important for meropenem resistance whilst Δ$lpqF_{mar}$ plays a more pivotal role in cell wall maintenance.

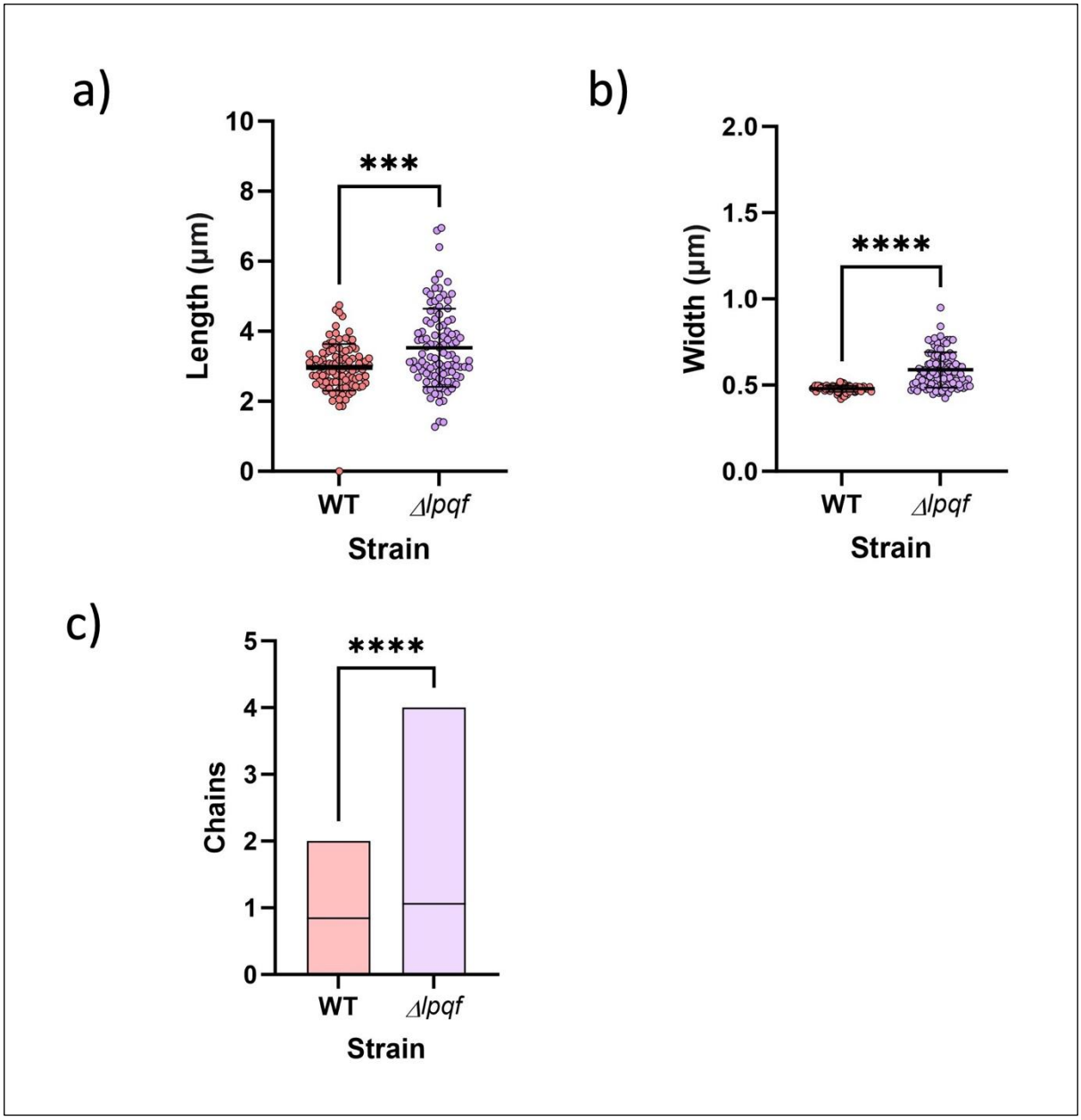



**Figure 1.43 Microscopy data of *M. marinum* WT and ∆*lpqF*.**

WT cells and ∆*lpqF*$_{mar}$ cells were imaged via phase contrast and via the detection of HADA fluorescence. (a) the cell length (mean ± SD), (b) width (mean ± SD) and (c) the number of chains (mean) were calculated via MicrobeJ. Significance was evaluated via a Mann Whitney test for (a) ***P = 0.0003, (b) ****P < 0.0001 and (c) ****P < 0.0001.

## 2.2.4 Deletion of Cell Wall Glycosyl Hydrolase Family Genes Results in Reduced Cell Size

Given that the GH183 family proteins and MSMEG_0740 are thought to play a role in the hydrolysis of cell wall components we reasoned that deletion of these genes should result in a change in cell morphology. Therefore, we also conducted phase contrast microscopy on Δ*MSMEG_0740*, Δ*MSMEG_4352* and Δ*MSMEG_6255* (Chapter 5, Section 1.16). As it was too difficult to discern the number of chains through phase contrast we only measured the cell length and width. As demonstrated in figure 1.44, all of the mutants exhibited significantly reduced cell length and width, though there was only a slight change in the size of Δ*MSMEG_0740* cells (figure 1.44d-e).

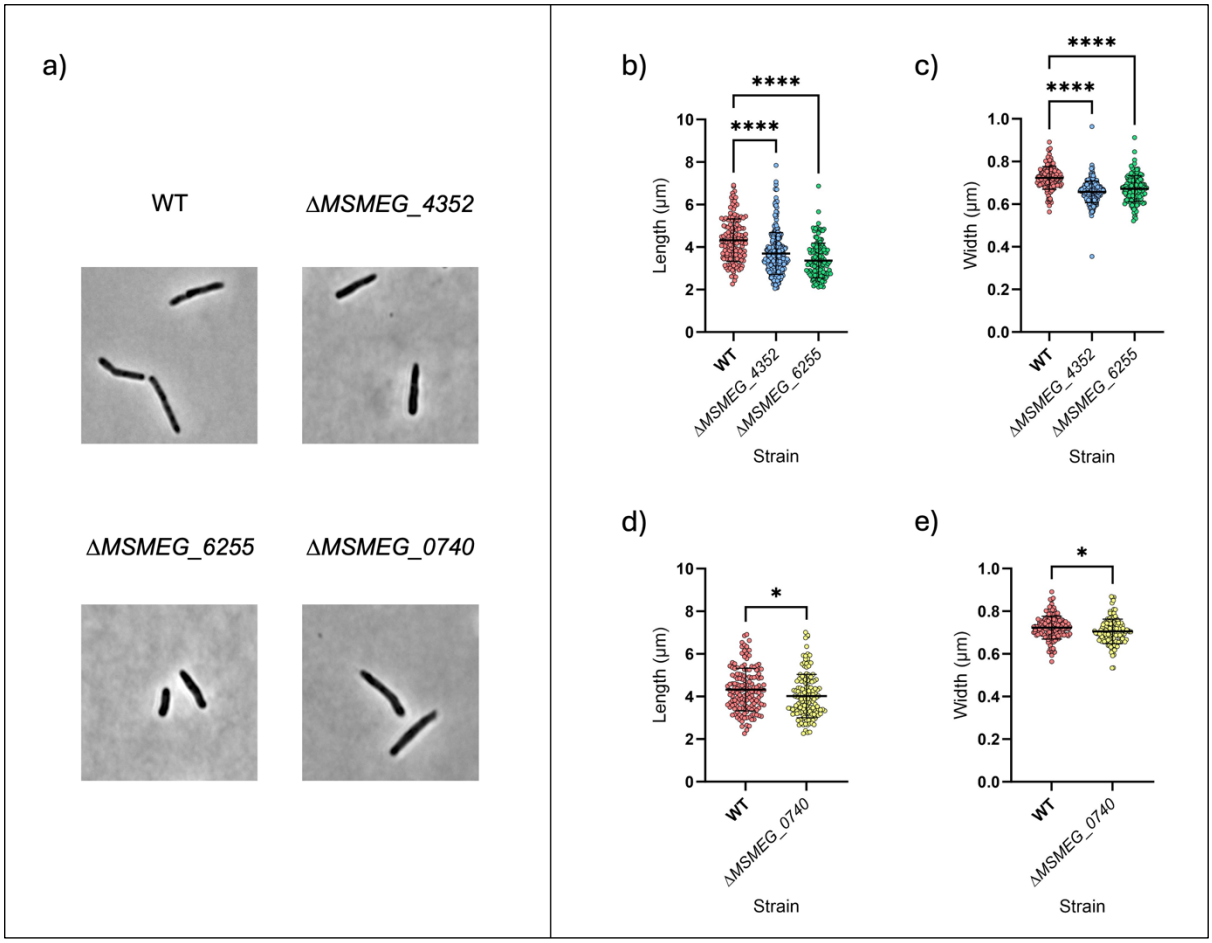



**Figure 1.44 Analysis of ORBIT mutants via microscopy.**

(a) WT, Δ*MSMEG_0740*, Δ*MSMEG_4352* and Δ*MSMEG_6255* cells imaged via phase contrast. (b) the cell length (mean ± SD) of GH183 family mutants compared with WT, (c) cell width (mean ± SD) of GH183 family mutants compared with WT, (c) the cell length (mean ± SD) of Δ*MSMEG_0740* compared with WT and, (d) the cell width (mean ± SD) of Δ*MSMEG_0740* compared with WT, calculated via MicrobeJ. Significance was evaluated via a Kruskal-Wallis test and Dunn's multiple comparison test for cell length (*P = 0.0365, ****P = <0.0001) and cell width (*P = 0.0335, ****P < 0.0001).

# Part 3 – Discussion

## 3.1 Validation of the ORBIT Method

As the ORBIT method has only been previous applied to *M. bovis* and *M. smegmatis* (Murphy et al., 2018), to discover whether ORBIT could be adapted for other mycobacterial species and whether the method could be used to efficiently generate clean KOs, necessary for validating chemical-genomic screening hits, we attempted to generate mutants in *M. marinum*. However, we were unable to obtain any transformants on selective media. In the ORBIT method, the transformation steps are conducted at RT for *M. bovis* and at 4°C for *M. smegmatis*. We therefore investigated whether performing ORBIT at RT or 4°C increased efficacy. Despite previous studies having established that *M. marinum* transformation was most effective at RT (Talaat and Trucksis, 2000), we were unable to obtain successful transformants at either temperature. Furthermore, since glycine has been reported to improve mycobacterial transformation efficiency as it reduces cell wall permeability (Murphy et al., 2018) but can also cause defective growth due to the loss of PDIMs (Hermans, Boschloo and De Bont, 1990), we also investigated whether the presence or absence of glycine improved the success of the ORBIT method. However, we were still incapable of generating transformants.

We later discovered that one of the genes targeted for deletion, *rpsL*, was essential (DeJesus et al., 2017), thus explaining why we were unable to generate a mutant. However, as *MMAR_50430* is non-essential, this did not explain why the ORBIT method was unsuccessful. We originally concluded that the most likely reason for

failure was due to experimental error but, as we demonstrated later, we were able to generate *M. smegmatis* ORBIT mutants using the standard ORBIT protocol.

Though unlikely, given the success of ORBIT in *M. smegmatis* and *M. tuberculosis*, it is plausible that generating mutants via ORBIT may be less efficient or even impossible in *M. marinum*. However, as the Bxb1 integrase and the RecT annealase gp61 both originate from mycobacteriophage it is unlikely that these would not function in *M. marinum* (Mediavilla et al., 2000; Van Kessel and Hatfull, 2008). Furthermore, the Bxb1 phage integrase acts independently of host machinery, it is not species specific and has been shown to assist in the genetic engineering of numerous bacterial candidates (Wang et al., 2017). Nevertheless, a previous study has also reported difficulties using the RecT annealase for generating *M. marinum* recombinants in which it was hypothesised that the Che9c phage enzymes were genotoxic at high concentrations (Koliwer-Brandl et al., 2019).

We are therefore no clearer on whether the ORBIT method can be adapted for *M. marinum.* We suggest that the method is further optimised and that alternative genes should be targeted for deletion since it is conceivable that the deletion of *MMAR_50430* could result in defective growth making it more difficult to obtain successful transformants. That said, it is possible that ORBIT cannot be adapted for *M. marinum*, and thus it would be interesting to determine whether the ORBIT method can be applied to other mycobacterial species. An adaptation of ORBIT has already been developed for more efficient recombineering in *E. coli*. In this method, a helper plasmid was employed to induce both the oligonucleotide recombination and plasmid

insertion. More importantly, the Che9c gp61 annealase was replaced with the Collinsella stercoris phage RecT annealease, CspRecT, for more efficient oligonucleotide-mediated recombination (Saunders and Ahmed, 2023). Should it be discovered that the original ORBIT method cannot be applied to all mycobacterial species, a similar system could be developed to the *E. coli* ORBIT method and an alternative RecT annealase could be implemented.

Whilst we initially had limited success at generating *M. smegmatis* mutants via ORBIT, once we discovered that we had incorrectly constructed the targeting oligonucleotides and redesigned them, the following ORBIT attempts were all successful. It is remarkable that we were able to obtain any ORBIT mutants using the incorrectly designed oligonucleotides given that we accidentally targeted the leading strand and generated frameshifts. Frameshifts can create a missense in downstream genes some of which may be essential genes or synthetically lethal thus resulting in the inability to generate ORBIT mutants. That said, whilst *MSMEG_6255* contains an upstream essential gene, asd, none of the targeted genes neighbour a downstream essential gene. Though it is possible that the disruption of downstream genes could result in a growth defect thus hindering the ability to isolate successful transformants, as the targeting oligonucletodies designed for *MSMEG_2107* and *MSMEG_4365* do not cause frameshifts, the failure to generate ORBIT mutants cannot be explained by the creation of frameshifts alone. In contrast, we discovered that all of the initially designed oligonucleotides targeted the leading strand and since it has been discovered that HR is most successful when the lagging strand is targeted and results in around a $10^4$-fold

increase in efficiency (Van Kessel and Hatfull, 2008), it is likely that this was the main cause for our failure to obtain ORBIT transformants.

As we demonstrated in Chapter 4 Section 2.1.2, a minor drawback of the ORBIT method is the generation of off-target mutants in which the payload plasmid can integrate into the mycobacterial genome via illegitimate recombination. In addition, Murphy et al. (2018) reported that only 25-50% of transformants were successfully genetically engineered indicating that the majority of transformants were the result of non-specific insertion of the payload plasmid. In comparison, we discovered an even lower transformation efficiency of ~20% and obtained fewer transformants than previously described (Murphy et al., 2018). We therefore suggest that the ORBIT method may be less efficient than first reported. However, given that Murphy et al. (2018) had a much larger sample size, it is possible that the reduced efficiency observed in our study can be attributed to a reduced HR receptivity in the genes that we targeted for deletion. Regardless of the transformation efficiency and rate of illegitimate recombination, our results validate that ORBIT is a significant improvement on the traditional methods for generating mutants in *M. smegmatis*.

As covered in detail in Chapter 1 Section 2.6, CRISPR systems represent an exciting alternative to the ORBIT method which have been adapted for mycobacteria with great success (Bosch et al., 2021). In 2020, Meijers et al. demonstrated that a 55% efficiency for generating successful gene disruptions in *M. marinum* and *M. tuberculosis* could be achieved using the CRISPR method. In addition, an adaptation of the CRISPR method, called CRISPRi, can be used to generate hypomorphs so that

essential genes can also be analysed (Li et al., 2022). In contrast, ORBIT boasts a higher efficiency in generating mutants and can be used to introduce DNA and chromosomally tag genes (Murphy et al., 2018). Whilst methods for introducing DNA via CRISPR have also been described, these are yet to be adapted to mycobacteria (Yan et al., 2017). Therefore, both the ORBIT and CRISPR systems have their advantages and limitations thus having these two unique methods for generating mycobacterial mutants is extremely valuable. Given that we had difficulty applying the ORBIT method to *M. marinum* and, since CRISPR has been previously employed for the generation of *M. marinum* mutants, we may have had more success generating *MMAR_50430* mutants via the CRISPR method. Therefore, whilst we were successful in using the ORBIT method to obtain markerless gene deletions, other systems might be better suited for generating clean KOs for validating chemical-genomic screening hits in certain species of mycobacteria.

## 3.2 LpqF is a D,D-carboxypeptidase

To phenotypically characterise the ORBIT mutants and thus gain further insight into the function of the genes targeted for deletion, we conducted growth assays, a small chemical-genetic screening and microscopy.

As demonstrated in Chapter 4, Section 2.2.1, we were unable to identify any growth differences in the ORBIT strains through spot assay and growth curves. However, chemical-genetic screening appeared to be more sensitive and thus enabled us to detect subtle variations in mutant fitness by measuring colony growth. Through this

method, we discovered that Δ*lpqF*$_{smeg}$ had altered fitness in response to beta-lactams. Under standard growth conditions Δ*lpqF*$_{smeg}$ demonstrated a defective colony growth. However, when grown in the presence of meropenem, amoxicillin with clavulanic acid and imipenem, growth was comparable with WT.

Unfortunately, whilst we were able to reproduce the meropenem phenotype via growth curve, we saw no Δ*lpqF*$_{smeg}$ fitness change in imipenem or amoxicillin with clavulanic acid. It is possible that due to the transition from solid to liquid medium, the drug concentrations were variable and thus the results differed. Alternatively, this result could be explained by the chemical-genetic screening demonstrating increased sensitivity to changes in strain fitness. However, when we cured Δ*lpqF*$_{smeg}$ (Δ*lpqfu*$_{smeg}$) of the ORBIT payload plasmid and repeated the chemical-genetic screening, like the growth curves, only the meropenem phenotype could be reproduced. In addition, Δ*lpqFu*$_{smeg}$ demonstrated a greater growth defect under standard conditions. However, this was discovered to be the result of an accidental bias introduced during colony arraying which lead to severe edge effects. We were also unable to complement this defect, though this could be due to the design of the complement plasmid. Therefore, ideally, the screen would be repeated without a bias and with a new complemented strain. In addition, when we repeated the screening with an *M. marinum lpqF* mutant, Δ*lpqF*$_{mar}$, although Δ*lpqF*$_{mar}$ had defective growth, WT growth was not restored via meropenem. However, as the same concentration of meropenem was applied to *M. marinum*, and as sub-inhibitory concentrations of drugs can vary dramatically between strains, not to mention different species, the concentration of meropenem may not have been high enough to trigger a response. Furthermore, it is

plausible that the *lpqF* homologues may fulfil different roles in each species or have different levels of genetic redundancy.

Interestingly, in the Li et al. (2022) CRISPRi pooled *M. tuberculosis* chemical-genomic screening dataset, in contrast to our results, it was reported that Δ*lpqF* has significantly reduced fitness in meropenem. Like with *M. marinum*, this could be due to differences between the *lpqF* homologues. However, the Li et al. (2022) screen also reported that Δ*lpqF* had significantly reduced growth in rifampicin, whilst Xu et al. (2018) described the opposite. Furthermore, Li et al. (2022), defined *lpqF* as essential, whilst various other studies have marked it as non-essential (DeJesus et al., 2017; Xu et al., 2018 and Bosch et al., 2021). Therefore, we question the accuracy of the Li et al. (2022) Δ*lpqF* screening data. Nevertheless, as meropenem is a carbapenem, which are resistant to many β-lactamases, we were surprised to discover that meropenem caused a recovery of fitness in Δ$lpqF_{smeg}$ as the deletion of a β-lactamases should have little to no effect. It is plausible that meropenem cannot inhibit *lpqF* but in this circumstance a *lpqF* mutant would be expected to exhibit increased susceptibility to meropenem. We therefore concluded that *lpqF* may not function as a β-lactamases and instead fulfils another role. As meropenem targets L,D-transpeptidases and PBPs, we suggested that *lpqF* may have a role in PG synthesis and in particular, PG crosslinking.

Via microscopy, we discovered that Δ$lpqF_{smeg}$ and Δ$lpqF_{mar}$ again had contrasting phenotypes, in which Δ$lpqF_{mar}$ cells were longer and wider than WT whereas Δ$lpqF_{smeg}$ cells were slightly thinner but similar in length. However, both Δ$lpqF_{smeg}$ and Δ$lpqF_{mar}$

did exhibit an increase in chaining. Given the differences between the two species, the microscopy data doesn't really provide any further insight into *lpqF* function though it appears both homologues have some role in regulating cell size.

Although currently unpublished, Clare Thomas of the Patrick Moynihan research group, recently found that, based on sequence analysis, *lpqF* also contains sequence signatures comparable to low-molecular weight PBPs with a conserved serine in the active site of the C-terminal PBP domain and a PG-binding N-terminal NTF2 domain. In addition, through PG binding, bocillin-fluorescence and nitrocefin assays, it was discovered that whilst LpqF bound to both PG and PBPs, though PBP binding was inhibited by carbapenems, it did not exhibit β-lactamase activity. Furthermore, through a PG degradation assay, LpqF was shown to have D,D-carboxypeptidase activity, cleaving the D-Ala-D-Ala terminus of the PG peptide chain. Interestingly, the LpqF crystal structure revealed an additional N-terminal NTF2 domain that could inhibit the active site thus identifying a novel mechanism for the control of PBP activity. Therefore, since we discovered that the deletion of $lpqF_{smeg}$ results in a restoral of fitness in the presence of meropenem, and given that Δ*lpqF* demonstrates a change in the cell size and chaining, our results further support the hypothesis that *lpqF* encodes a PBP rather than a β-lactamase.

Interestingly, likely due to genetic redundancy, the deletion of individual D,D-carboxypeptidases in *E. coli* has no effect on cell morphology and drug susceptibility (Denome et a., 199). However, upon the deletion of multiple D,D-carboxypeptidases severe growth defects can be observed as the cells become spherical (Denome et al.,

1999). Likewise, in *M. smegmatis*, it has been reported that the individual deletion of D,D-carboxypeptidases *MSMEG_1661*, *MSMEG_2433* and *MSMEG_2432* do not result in any differences in growth (Pandey et al., 2018; Ealand et al., 2019). However, Ismael (2016) demonstrated that in strains which harbour deletions in all of these genes, cells exhibited an increase in chaining, size and susceptibility to β-lactams and acidic minimal media. Furthermore, Pandey et al. (2018) also demonstrated that just the double KO of *MSMEG_2433* and *MSMEG_2432* was sufficient to trigger the increased antibiotic susceptibility, cell chaining and cell length phenotypes. Pandey et al. (2018) also discovered that these double KO resulted in increased 4,3-crosslinking, highlighting a compensatory pathway upon the inhibition of D,D-carboxypeptidases. In *M. tuberculosis* similar phenotypes can also be observed in which the deletion of *dacB2* impaired growth in acidic Sauton's medium (Bourai et al., 2012). In addition, the overexpression of *dacB2* resulted in a smooth colony morphology and a reduced growth rate (Bourai et al., 2012). Given that the deletion of *lpqF* in *M. marinum* and *M. smegmatis* resulted in similar phenotypes it is plausible that LpqF also functions as a D,D-carboxypeptidase and, since the individual deletion of other *M. smegmatis* D,D-carboxypeptidases does not result in phenotypic changes, LpqF perhaps plays a more crucial role.

## 3.3 *MSMEG_0740* is an α-mannanase

Unfortunately, no significant changes in fitness were determined for the chemical-genetic screening of the GH183 family mutants, Δ*MSMEG_4352* and Δ*MSMEG_6255*,

or the putative mannosidase mutant, ∆*MSMEG_0740*. As there are five different GH183 family proteins in *M. smegmatis* and as there are at least three suspected mannosidases, it is plausible that these may fulfil similar roles and can therefore compensate each other's function. This would explain why no change in mutant phenotype could be observed. Similar findings were discovered in both the Xu et al. (2017) and the Li et al. (2022) pooled *M. tuberculosis* chemical-genomic screenings as no significant fitness alterations were observed in mutants of *Rv3707c* or *Rv0365c*. The other *M. tuberculosis* GH183 family member, *Rv1754* demonstrated a significantly reduced growth in rifampicin in the Xu et al. (2017) study, but this was not replicated by Li et al. (2022). However, as we did not manage to successful generate a mutant in the *M. smegmatis Rv1754* homologue, we cannot directly compare this phenotype with our own results. Likewise, in the Li et al. (2022) screen, whilst the majority of the putative mannosidases had no significant change in fitness, ∆*Rv0640* was found to be susceptible to vancomycin and resistant to linezolid. However, we cannot validate these results as these conditions were not tested in our chemical-screening and we were unable to generate a knockout of the *Rv0640* homologue. We therefore suggest that increasing the number of conditions used in the chemical-genetic screening could help to generate more significant phenotypes for these mutants. Nevertheless, in general, it appears that the deletion of the GH183 family members or the putative mannosidases does not result in growth defects, likely due to a genetic redundancy.

In contrast, through microscopy (Chapter 4 Section 2.2.4), we discovered that the deletion of *MSMEG_0740*, *MSMEG_4352* and *MSMEG_6255* resulted in significantly altered phenotypes when compared to WT with all mutants demonstrating reduced

cell width and cell length. Given that these genes are thought to function as cell wall hydrolases, we suspected that a change is cell size might be observed. However, since we expected that a decrease in the degradation of AG and LAM would result in a higher abundance of AG and LAM in the cell envelope and thus an increase in thickness, we were surprised to discover that the cell sizes were reduced. Nevertheless, it is possible that AG and LAM hydrolases are important for cell division, growth and repair in order to make space for newly synthesised AG or LAM to be inserted. In addition, the degradation of AG and LAM could be vital for the recycling of key substrates for AG and LAM synthesis or for modifications necessary for the stability of the cell wall. It is also conceivable that a compensatory protein could have an enhanced rate of cleavage, a characteristic that could be beneficial for survival under certain growth conditions, or that in the absence of degradation, synthesis of AG or LAM is downregulated. Under each of these circumstances, a reduction in cell size would be plausible. Lastly, as the presence of arabinomannans and mannans in mycobacterial capsule are largely attributed to the activity of arabinofuranases and mannosidases (Lemassu and Daffe, 1994), it is reasonable that the deletion of the proteins responsible would result in a reduction in the capsule and thus a decrease in cell thickness.

Although our findings do not reveal the function of the *M. smegmatis* GH183 family proteins, this work substantiates the research conducted by Al-Jourani et al. (2023) by providing new phenotypic data for this family of enzymes. Furthermore, our findings also supported a recent publication in which the *M. tuberculosis MSMEG_0740* homologue, *Rv0365c*, renamed *lamH*, was found to exhibit α-1,6 mannanase activity

(Franklin et al., 2023). In this study, both the deletion of *MSMEG_0740* and *Rv0365c* resulted in the loss of mannans from the mycobacterial capsule, thus reaffirming our hypothesis that the decreased cell size of Δ*MSMEG_0740* could be due to a reduction in the capsule. In contrast to our findings, Δ*Rv0365c* exhibited a prolonged lag phase. However, the failure to replicate this phenotype in Δ*MSMEG_0740* could suggest that there are other *M. smegmatis* genes that fulfil a similar role and can compensate for the loss of *MSMEG_0740*. Lastly, it was discovered that *Rv0365c* is important for *M. tuberculosis* replication within macrophages and that deletion of *Rv0365c* resulted in reduced LAM biosynthesis thus indicating that there is communication between the LAM synthesis and LAM hydrolysis pathways. Given the importance of these enzymes for maintenance of the cell envelope, growth and virulence, we conclude that these enzymes are ideal candidates for drug targets and thus the characterisation of similar proteins should be further investigated.

# Part 4 – Conclusion

Due to their thick and hydrophobic cell wall, mycobacteria are inherently resistant to many commonly used antibiotics (Parte et al., 2020). Instead, many of the front-line anti-mycobacterial drugs target this unique cell wall as it is exposed and essential for survival. Nevertheless, multidrug resistance is on the rise and thus there is a demand for new therapeutics (WHO, 2023). Whilst, the synthesis of the cell wall is relatively well characterised, the hydrolysis and recycling of cell wall components as well as the mechanisms for drug resistance are still poorly understood (Fullam et al., 2016; Al-Jourani et al., 2023). By identifying the genes responsible for these biological processes, we may be able to identify new drug targets. A genetic engineering method, called ORBIT, was recently developed by Murphy et al. (2018) for the efficient generation of *M. tuberculosis* and *M. smegmatis* mutants. As ORBIT can be combined with reverse genetics, it represents a powerful tool for the phenotypic characterisation of genes and thus can be used to gain insight into gene function and potentially identify new drug targets. We therefore explored whether ORBIT could be applied to other mycobacterial species and attempted to generate knockouts of putative mycobacterial proteins involved in cell wall degradation and antibiotic resistance. In addition, we explored whether ORBIT was an appropriate tool for generating clean KOs that could be used to validate chemical-genomic screening hits.

Although we were unable to adapt the ORBIT method to *M. marinum*, we successful generated *M. smegmatis lpqF, MSMEG_0740, MSMEG_4352 and MSMEG_6255*

mutants and were able to generate markerless gene deletions thus demonstrating the utility of the ORBIT method for validating chemical-genomic screening hits. Through chemical-genetic screening and growth assay we found that Δ*lpqF* had defective growth that was restored to WT upon the addition of meropenem. In addition, through microscopy we identified that Δ*lpqF* had increased chaining and an altered cell size. Similarly, Δ*MSMEG_0740* had reduced cell length and width compared to WT. These findings support research conducted in collaborative studies in which it was discovered that *MSMEG_0740* (Rv0365c) functions as an α-1,6 mannanase (Franklin et al., 2023) and *lpqF* likely encodes a D,D-carboxypeptidase, rather than a β-lactamase (Thomas et al., Unpublished Data). As these enzymes are suggested to play an important role in the maintenance of the cell wall, they could represent exciting drug target candidates.

# Materials and Methods

## 1.1 Bacterial Strains, Bacteriophage and Media

The bacterial strains used in this study are listed in table 1.2. Unless stated otherwise, for liquid growth, mycobacterial strains were grown from glycerol stocks in 7H9 broth (Middlebrook 7H9 (Sigma) supplemented with 1% glycerol (Thermo Fisher Scientific), 0.05% tween80 (Sigma) and 10% ADC (albumin, dextrose and catalase; Sigma). For solid growth, strains were plated on 7H9 agar (Middlebrook 7H9 supplemented with 1% glycerol, 10% ADC and 2% BD Difco™ agar (Thermo Fisher Scientific). For growth of the *M. abscessus* and *M. bovis* BCG library strains, 7H9 broth or 7H9 agar was further supplemented with 25 µg/ml kanamycin (Bio Basic), 0.2% Bacto™ casamino acids (Thermo Fisher Scientific), 50 µg/ml L-Tryptophan (sigma) and 50 µg/mL Natures Aid multivitamins. For simplicity, these will be referred to as enriched 7H9 broth and enriched 7H9 agar. For the chemical-genomic screening of the *M. bovis* BCG library, Sauton's and Roisin's media were also used as minimal media (table 1.3).

**Table 1.2 The bacterial strains used in our study.**

| **Bacteria** | | **Source** |
|---|---|---|
| ***M. abscessus* subsp. *abscessus* ATCC 19977** | WT | Prof Tracy Palmer (Newcastle University) |
| ***M. abscessus* subsp. *masiliense* CIP 108297** | WT | Lab Stock |

| | | |
|---|---|---|
| ***M. bovis* BCG Danish 1331** | WT | Lab Stock |
| ***M. bovis* BCG Danish 1331** | A library of transposon insertion mutants | Borgers et al. (2018) |
| ***M. marinum 1218R*** | WT | Lab Stock |
| ***M. marinum 1218R*** | Δ*lpqF* | Lab Stock |
| ***M. smegmatis* MC$^2$ 115** | WT | Lab Stock |
| ***M. smegmatis* MC$^2$ 115** | Δ*lpqF* | Lab Stock |
| ***M. smegmatis* MC$^2$ 115** | Δ*MSMEG_0740* | This Study |
| ***M. smegmatis* MC$^2$ 115** | Δ*MSMEG_4352* | This Study |
| ***M. smegmatis* MC$^2$ 115** | Δ*MSMEG_6255* | This Study |
| ***M. smegmatis* MC$^2$ 115** | Δ*MSMEG_4722* | Bhatt et al. (2008) |

**Table 1.3 Minimal growth media.**

| Medium | Components |
|---|---|

| | |
|---|---|
| **Sauton's Medium** | 0.5 g $K_2HPO_4$ (Thermo Fisher Scientific), 0.5 g $MgSO_4*7H_2O$ (Thermo Fisher Scientific), 2.0 g citric acid (Sigma), 0.05 g ferric ammonium citrate (Sigma), 4 g L-asparagine monohydrate (Acros Organics), 0.05% tyloxapol (Sigma) and 1% glycerol. $dH_2O$ is added for a final volume of 1 L. L-asparagine and glycerol can be substituted for alternate nitrogen (4 g/L unless stated otherwise) and carbon (4 mg/mL unless stated otherwise) sources. |
| **Roisin's Medium** | 1.0 g $KH_2PO_4$ (Thermo Fisher Scientific), 2.5 g $Na_2HPO_4$ (VWR), 5.9 g $NH_4Cl$ (Thermo Fisher Scientific), 2.0 g $K_2SO_4$ (VWR), 1 mL trace element solution, 0.5 mL 1M $CaCl_2$ (Thermo Fisher Scientific), 0.5 mL 1M $MgCl_2$ (VWR), 1% glycerol, 0.05% tyloxapol. $dH_2O$ is added to a final volume of 1 L.<br>Trace element solution: 80 mg $ZnCl_2$ (BioServ), 400 mg $FeCl_3*6H_2O$ (VWR), 20 mg $CuCl_2*2H_2O$ (VWR), 20 mg $MnCl_2*4H_2O$ (BioServ), 20 mg $Na_2B_4O_7*10H_2O$ (Thermo Fisher Scientific), 20 mg $(NH_4)_6Mo_7O_{24}*4H_2O$ (VWR), $dH_2O$ is added to a final volume of 1 L. |

The bacteriophages used in this study are documented in table 1.4. Bacteriophages

were stored in MP buffer containing 2 mM $CaCl_2$, 150 mM NaCl (Thermo Fisher Scientific) and 50 mM Tris (Sigma) and adjusted to pH 7.5. The method for phage propagation is described in Chapter 5, Section 1.2.

**Table 1.4 The bacteriophages used in our study.**

| Bacteriophage | Resistance | Source |
|---|---|---|
| **ΦMycoMarT7** | Kanamycin | Prof Christopher Sassetti (University of Massachusetts) |
| **phDB36** | Zeocin™ | Dr Daniel Barkan (Koret School of Veterinary Medicine) |

## 1.2 Propagation and Titration of Bacteriophage

The bacteriophage titration and propagation method was adapted from the protocol outlined in Siegrist and Rubin (2009).

To calculate the bacteriophage titre, a 10 mL culture *of M. smegmatis* at an $OD_{600nm}$ of 1.0 was prepared. The cells were washed by pelleting the cells via centrifugation (10,000 rpm, 3 min), removing the supernatant and resuspending in 2 mL of MP buffer. This process was then repeated for a total of three washes. Following the final wash,

cells were resuspended in 1mL of MP buffer and 100 µL of the resuspension was added to separate sterile glass test tubes containing 4 mL of 0.6% molten agar. The bacteriophage was serially diluted by 10-fold increments and 50 µL of each dilution was added to each tube. The molten agar was then vortexed and poured on top of 7H9 agar plates. Plates were incubated at the lytic permissive temperature of 30°C for 2-3 days and the PFU/mL was calculated by counting the number of plaques on each dilution plate.

To propagate the bacteriophage, plates were prepared via the same method as bacteriophage titration. However, only the bacteriophage dilution that resulted in a lacey plaque formation was added to the test tubes containing 0.6% molten agar. Following a 3 day incubation, 3 mL of MP buffer was added to each plate. Plates were then incubated for 24 h on a shaking platform at 25 rpm. For the plate breathing method, the temperature was alternated between RT and 4°C. Lastly, the phage containing MP buffer was collected and filter sterilised either immediately or after a further 24 h incubation period.

## 1.3 Resazurin Assay

Resazurin assays were conducted in black 96-well plates (CELLSTAR®). Mycobacterial cultures were grown to mid-log, diluted to 1 x $10^6$ cells/mL and 50 µL added to each well. Drugs (supplementary table 1), diluted in 7H9, were added to each well with increasing concentrations to achieve a 150 µL total volume for each well. The

maximum drug concentrations were 1.7 mg/mL for Gibco™ Zeocin™ (Thermo Fisher Scientific) and 1.0 mg/mL for amphotericin B (Cayman Chemicals). The outer rows and columns were filled with blank 7H9 medium to create a liquid barrier to prevent evaporation. Plates were then incubated for 7 days for *M. bovis* BCG and 24 h for *M. abscessus* at 37°C. Following incubation, 30 µL of 0.05% resazurin and 12.25 µL of 20% Tween80 were added to each well and plates were incubated for a further 24 h at 37°C. Fluorescence was measured via a BMG plate reader at an excitation wavelength of 550 nm and detection wavelength of 590 nm.

## 1.4 Pooled Transposon Mutagenesis Library Generation

We followed a previously described method for the generation of the pooled *M. abcessus* transposon insertion library (Rifat et al., 2021). Briefly, a 30 mL culture of *M. abscessus* was grown to an $OD_{600nm}$ between 1.0 and 2.0. Cells were pelleted via centrifugation at 4000rpm for 10 min, the supernatant was removed and the cells were resuspended in 30 mL of prewarmed (37°C) MP buffer. This wash step was repeated before resuspending in 5 mL of MP buffer. At a multiplicity of infection (MOI) of 20, high titre bacteriophage (> $5x10^{10}$ PFU/mL) was incubated with the washed cells for 4 h at the lysogenic permissive temperature of 37°C and 180 rpm. Following incubation, to eliminate the bacteriophage, cells were pelleted via centrifugation for 10 min at 4000 rpm, the supernatant was removed, and the cells were resuspended in 5 mL of 7H9. 50 µL was then spread onto enriched 7H9 agar plates supplemented with 25 µg/mL of kanamycin or 50 µg/mL of Zeocin™ via sterile glass beads. Plates were incubated at 37°C for 4-5 days until visible colonies were formed.

## 1.5 PCR

PCR was conducted using 25 µL reactions containing 12.5 µL of Q5 High Fidelity 2X Master Mix (New England Biolabs), 1.25 µL of a 10µM forward and reverse primer mix (Sigma), 11.25 µL of Nanopure $H_2O$ and a single colony. The Eppendorf Mastercycler Nexus was used to conduct the thermocycling steps. The following settings were applied: 5 min 98°C heating step; followed by 35 cycles of: 10 s at 98°C, 30 s of the primer annealing temperature, and, depending on the size of the fragment, between 1-3 min at the polymerisation temperature of 72°C; lastly, a final 72°C polymerisation step for twice the length of time of the within cycle polymerisation step. Gel electrophoresis was conducted with a 1 kb DNA ladder (New England Biolabs) on a 1% agarose gel (Meridian Bioscience; Scientific Laboratory Supplies) diluted with 1X TRIS-acetate-EDTA (TAE) (Fisher) buffer, dyed with MIDORI Green (Nippon Genetics) at 140mA and 140V for 40 min and imaged via the Bio-Rad Gel DocTM XR+ Imager. Primers used for PCR are listed in table 1.5.

**Table 1.5 List of primers used for PCR.**

| Primer target | Primer sequences |
|---|---|
| **pKM461** | catgcttaattaagaagg |
| | accaccgcgtcggaattc |
| **pKM464** | ttaagtcgcgttgcgctc |
| | tacgcccggtagtgatct |

| | |
|---|---|
| ***MSMEG_0740*** | acatgccgacatgtcgaa |
| | cactgaccttctgtcact |
| ***MSMEG_1130*** | tttctcatcgcccttcag |
| | ataccaggcagtctgtga |
| ***MSMEG_1361*** | gtcgacatcgcggaatga |
| | tttcccgctcgtcgatct |
| ***MSMEG_2107*** | cggcgcgcgcaagacgta |
| | ggcggggcgttactggca |
| ***MSMEG_4352*** | cggcagttgggcgcccat |
| | gaacgtcgacggcgggta |
| ***MSMEG_4360*** | ctcggcagcgtcgcggga |
| | tgcccacgcgggtgcgtc |
| ***MSMEG_4365*** | ttttttcgatcgcccacg |
| | gaggagtacggctcgaaa |
| ***MSMEG_6255*** | tcggatggtcagacggtg |
| | gccgacctctgacctgaa |
| ***lpqF*** | aagcctcctggagttcga |
| | ccagatccttctggcgca |
| **Kanamcyin Cassette** | aacaagatggattgcacg |
| | gaactcgtcaagaaggcg |
| **Coordinate** | gccttcttgacgagttcttctgag |
| **Sequencing PCR 1** | gactggagttcagacgtgtgctcttccgatc |

| | |
|---|---|
| **Coordinated Sequencing PCR 2** | NEBNext® Multiplex Oligos for Illumina® (Index Primers Set 1) |

## 1.6 Robotic Cherry-Picking

The BM3-BC robot (S&P Robotics) equipped with a cherry-picking tool was used for the robotic ordering of the library. First, successful transductants were selected on Single Well Cell Culture Plates (VWR®) containing enriched 7H9 agar plus 25 µg/mL of kanamycin. Bead-beating was employed to ensure that the colonies grew separately. Plates were imaged via the BM3-BC robot and through colony recognition software, 96 colonies were selected and located for transfer to a 96-well round-bottom microwell plate (Nunc™) containing 150 µL of enriched 7H9 plus 25 µg/ml kanamycin. The plates were then covered with Breathe-EASIER™ (Diversified Biotech) seals and grown at 37°C for 10 days.

## 1.7 Manual *M. abscessus* mutant library arraying

Using sterile pipette tips, successful transductants generated in Chapter 5, Section 1.3, were picked into 96-well round-bottom microwell plates containing 150 µL of enriched 7H9 plus 25 µg/mL of kanamycin. Clusters of mutants were ignored to avoid cross-contamination and colonies were picked at random, across a range of morphotypes to prevent bias. In total, 18,432 colonies were picked and thus arrayed

into 192 96-well plates. The plates were then sealed with Breathe-EASIER™ seals and grown at 37°C for 10 days.

## 1.8 *M. abscessus* library replication

Using the Biomek i5 robot (Beckman Coulter) fitted with a Multidrop™ Combi Reagent Dispenser (Thermo Scientific™), 110 µL of 35% glycerol and 40 µL of each well from each library source plate was added to corresponding wells in 2 replicate 96-well flat-bottom microwell plates (Nunc™). For storage, the replicate library plates were covered with AlumaSeal® II film seals and placed at -80°C.

## 1.9 Cartesian Pooling

The Cartesian Pooling technique was adapted from a previously described method (Vandewalle et al., 2015) outlined in Chapter 3, Section 1.3. Briefly, the library was split into two subsets, each containing 96 plates, and XY and Z master plates were generated for each set. The Biomek i5 robot was used for XY pooling in which 15 µL of culture was transferred from each XY coordinate across all library source plates to a corresponding XY coordinate in the master plate. In contrast, an OT-2 Liquid Handling Robot (Opentrons) was employed for the Z pooling, in which 15 µL of culture from every well in a library source plate was transferred into a single well of the master plate. To generate the sequencing pools, row and column pools were generated from the master plates using the OT-2 robot. For the row pools, 400 µL of culture was

transferred from each well in a row of the master plates into a 15 mL falcon tube. For the column pools, 600 µL of culture was transferred from each well in a column of the master plates into a 15mL falcon tube. This resulted in a total of 80 sequencing pools comprising either 768 or 1152 mutants. Lastly, each pool was centrifuged for 5 min at 4000 rpm and the supernatant was removed, leaving just the pellet, which was then stored at -20°C until gDNA was extracted as described in Chapter 5, Section 1.8.

## 1.10 Genomic DNA Extraction

To extract gDNA, the Monarch® Genomic DNA Purification Kit was used. The greatest concentration of DNA was achieved through both mechanical and chemical lysis in which samples were exposed to bead-beating immediately followed by lysozyme (Sigma) degradation. The Qubit™ Broad-Range Assay Kit was used to measure gDNA concentration using the Qubit™ 4 Fluorometer (Invitrogen).

## 1.11 Coordinate Sequencing

500 µL of nuclease-free deionised water was added to 1 µg of gDNA from each cartesian sequencing pool and sonicated using a Diagenode Bioruptor® Plus at the following settings: 13 cycles of: 30s low intensity sonication followed by 90s of a rest period. Sonicated gDNA was concentrated to around 50 µL using a vacuum concentrator (Eppendorf). The NEBNext® Ultra™ DNA Library Prep Kit for Illumina® was then used to prepare the sonicated DNA for sequencing. In this method, the DNA

fragments were blunted and 5' phosphorylated to facilitate 3' dA-tailing. NEBNext Adapters were ligated and activated, fragments sizes were selected using SPRI beads and through PCR using primers that target both the transposon and the adapter (table 1.5), the fragments were amplified. Size selection was repeated and then an additional PCR was performed using primers expressing flanking P5 and P7 sequences necessary for Illumina® sequencing (table 1.5). The PCR parameters were as follows: A 48 s initial denaturation step at 98°C; 10 cycles (first PCR step) or 20 cycles (second PCR step) of: 15 s at 98°C (denaturation), 30 s at 65°C (annealing) and 30 s at 72°C (extension); followed by a 60 s final extension step at 72°C. Lastly, size selection was performed for a third and final time and samples were sent for 150 bp paired-end Illumina® sequencing (Novogene).

## 1.12 *M. bovis* BCG Library Condensing and Replication

The condensed *M. bovis* BCG library was selected by filtering for the highest quality copy of each gene deletion. First, mutants without a library location or gene assignment were removed from the library dataset. Next, mutants that harboured an insertion in the latter 20% of the gene were also removed. Then, mutants were selected based on the reliability of their location assignment with a 'unique' assignment having the highest priority, followed by XY/Zok_unique, XY/Zok_nonunique and lastly, heuristic, which was least favoured. Finally, selection of any remaining duplicates was achieved by choosing the KO duplicate which harboured the insertion closest to the start codon.

The Biomek i5 robot was used to cherry-pick 5 µL of culture from library plates into 96-well deep-well plates (Agilent) containing 1.4 mL of enriched 7H9 medium with 25 µg/mL of kanamycin (unless stated otherwise). WT *M. bovis* BCG was added to well E5 for each plate and thus E5 was not supplemented with kanamycin. The deep-well plates were sealed with BreatheEasier™ seals and incubated at 37°C for 5 weeks. Following incubation, using the OT-2 robot, 100 µL from each well of the condensed library was transferred into 96-well plates containing 100 µL of 50% glycerol and stored at -80°C. Any contaminated wells were replaced with WT *M. bovis* BCG.

## 1.13 Chemical-Genetic Screening

For chemical-genetic screening of the *M. abscessus* and *M. bovis* BCG mutant libraries, using the Biomek i5 robot, 5 µL from each well of the selected library plates was transferred into 96-well deep-well plates (Agilent) containing enriched 7H9 medium supplemented with 25 µg/mL of kanamycin and, for *M. bovis* BCG, 25 µg/mL of amphotericin B. For wells containing WT, kanamycin was omitted. The deep-well plates were then incubated for 5 weeks at 37°C for *M. bovis* BCG, and 4 days for *M. abscessus*. To generate the chemical screening source plates, 200 µL of culture was transferred into 96-well deep-well round bottom microwell plates (Nunc™) using the Biomek i5 robot. In contrast, for chemical-genetic screening of *M. marinum* and *M. smegmatis*, selected mutants were grown to an $OD_{600nm}$ of around 0.6 and manually arranged into 96-well deep-well round bottom microwell plates (Nunc™). Chemical

screening replicate plates were prepared by adding 42 mL of 7H9 2% agar supplemented with sub-inhibitory concentrations of various drugs (supplementary table 1) into Single Well Cell Culture Plates (VWR®). Using the BM3-BC robot, the source plates were replicated onto the chemical screening replicate plates in a 384-colony format. These replicate plates were then incubated for 2 weeks for *M. bovis* BCG, 4 days for *M. marinum* and 40 h for both *M. abscessus* and *M. smegmatis*. The plates were imaged at various time-points using the BM3-BC robot and these images were then analysed via IRIS and ChemGAPP to generated mutant fitness S scores. For clustering analysis an uncentred Pearson Correlation was applied using TreeView3.

# 1.14 ORBIT

## 1.14.1 ORBIT Plasmids

The plasmids used in our study, outlined in table 1.6, were obtained from Dr. Kenan Murphy, the author of the ORBIT paper (Murphy et al., 2018).

**Table 1.6 Plasmids required for ORBIT.**

| Plasmid | Drug resistance | Addgene ID |
|---|---|---|
| **pKM461** | Kanamycin | 108320 |

| | | |
|---|---|---|
| **pKM464** | Hygromycin (Toku E; Cambridge Bioscience) | 108322 |
| **pKM512** | Zeocin™ | NA |

## 1.14.2 Oligonucleotides

Oligonucleotides were designed as described by Murphy et al. (2018) though initial attempts at designing the oligonucleotides were incorrect as described in Chapter 4, Section 2.1.2. Oligonucleotides were ordered via IDT as desalted, 100 µM Ultramers and diluted to 1 µg/µL using nuclease free water (Omega Bio-Tek). A list of the oligonucleotides can be found in table 1.7.

**Table 1.7 ORBIT oligonucleotide sequences**

| Oligonucleotide Gene Target | Sequence |
|---|---|
| ***MSMEG_0740*** | ggcaggatcgttgccatggatcagctatgggccaaccgggcggc cagtgccgaagctgcgatcaccgaacggtttgtctggtcaaccac cgcggtctcagtggtgtacggtacaaacctgatgccggagacca aaacggcgccggagacggcgaagtgggacacgacgagaac gaggatgaggagtagcca |
| ***MSMEG_1361*** | caccaccacctggaccgcgggcgagtagagcagccggcccgc acagccgacccggatcagtgcccaccatggtttgtctggtcaacc |

| | |
|---|---|
| | accgcggtctcagtggtgtacggtacaaacctgatcctcggccgg<br>tccggtgtaccgctccgtcgactccgcggatgtgacgtgcatctca<br>aaaccctacggcg |
| ***MSMEG_2107*** | ccgcggtggaattgccgcgcgcgttctgcatatcgtcgagagcca<br>ccctagcccaccaacgctgtgccagggtttgtctggtcaaccacc<br>gcggtctcagtggtgtacggtacaaacctgaggcaaggacgtgg<br>gcaccagctgcccattccttgacatgaagttggagagcaacacc<br>agccggatgttgc |
| ***MSMEG_4352*** | tgggcgcggggtgacgctgccgctgctcaacggcaaggccgcc<br>gacatcacggggccgggccgcaccgacggtttgtctggtcaacc<br>accgcggtctcagtggtqtacqqtacaaacctqagccctaccgc<br>gcaatgcaattcaaqacacqgctgaaggacaccaccaccqcg<br>ccgcccagcgtactatga |
| ***MSMEG_4360*** | tcgccgggcacgtctcggaccgatcacqgatccgtaccggatcg<br>tgaccggcaaaccacacctgtctctgggtttgtctggtcaaccacc<br>gcggtctcagtggtgtacggtacaaacctgaacctgtacttcacgc<br>tgtcacgttggtcggactacagcgtgatgctgatgcgcacgtccct<br>gtcgaccag |
| ***MSMEG_4365*** | cggagcactggctgtggcactgggcatcggcgcagcactgacc<br>accggaaccggtgtggcatgggccgacggtttgtctggtcaacca<br>ccgoggtctcagtggtgtacggtacaaacctgactgtggggcga<br>ctacaacgtggtgctcatgcagaccgacctgtcgggcctgaaaa<br>ccgttgcggtgtgaa |

| | |
|---|---|
| ***MSMEG_6255*** | cctgaaccccgttgcgccgaaaccgcattccagcaggcgtcgtct ccatctttctctgctggaatgccatggtttgtctggtcaaccaccgcg gtctcagtgatatacqatacaaacctaagacacccattcataacc aattcatagccaattcttgatcaccacctaaacaacagacacata ctcatgac |
| ***MSMEG_6255*** **(Corrected)** | gcgtctcggcaggcgtgccgaccggagtggtcggatggtcagac ggtgtttcactcatgccacgagcatgggtttgtaccgtacaccactg agaccgcggtggttgaccagacaaaccggggttcaggtcagag gtcggcggccaacagttctgcgatctggatggtgttgagcgcggc gcccttgcg |
| ***MSMEG_1130*** | ttttcgacacgtgtcagccgtcacccgggtgtgacgttgcggccct gggaattacagccgggactcaccaggtttgtaccgtacaccactg agaccgcggtggttgaccagacaaaccgtccacgtggtcacctc attcgtgttgtcggaccattcgaggaaggacggcacccccatcgg actcccat |

## 1.14.3 ORBIT method

The method for ORBIT was adapted from the protocol outlined by Murphy et al. (2018). For *M. smegmatis*, 20 mL of the pKM461 containing strains were grown at 37°C in a shaking incubator using kanamycin containing media. At an $OD_{600nm}$ of 0.5 - 1.0, ATc was added at a concentration of 500 ng/mL and incubated for a further 3 h at 37°C,

though, ATc concentrations of 250 ng/mL, 1000 ng/mL and 2000 ng/mL were also used for the optimisation of the method. Electrocompetent pKM461 containing *M. smegmatis* cells were produced via two wash steps. Washing steps were conducted at 4°C in which the cultures were centrifuged for 10 min at 4000rpm followed by the removal of the supernatant and subsequent resuspension of the bacterial pellet in 20 mL of 10% cold glycerol. Following the second wash, cells were collected via further centrifugation and resuspended in 2 mL of 10% cold glycerol. On ice, 380 µL of the electrocompetent cells are added to 2 mm gap width electroporation cuvettes along with 200 ng of the pKM464 payload plasmid and 1 µg of the targeting oligonucleotide. A control electroporation was also conducted without the addition of the targeting oligonucleotide. In addition, The cells were then electroporated at 2.5 kV before overnight incubation (37°C) in recovery media. The next day, 0.5 mL of culture was spread onto solid 50 µg/mL hygromycin plates and incubated for 3-4 days at 37°C for the selection of transformants.

For *M. marinum*, the method differed only slightly. Cultures were again grown to an $OD_{600nm}$ of 0.5 - 1.0, but in a 30°C incubator. Upon induction via ATc, cells were instead incubated for 6h before a further overnight incubation following the addition of 2 mL of 2 M glycine. For optimisation, the method was also attempted without the additional incubation step with glycine. All washing and electroporation steps were the same except these steps were conducted at room temperature (RT). However, for optimisation of the method, electroporation was also attempted at 4°C. The cells were recovered overnight at 30°C before 0.5 mL of culture was spread onto solid 50 µg/mL

hygromycin plates and incubated for 1-2 weeks at 30°C for the selection of transformants.

### 1.14.4 Markerless Gene Deletions

The full method for generating markerless gene deletions is reported in Murphy et al. (2018). In brief: ORBIT mutants were transformed with pKM512 and transformants were selected via growth on 50 µg/mL Zeocin™. Successful transformants were grown to mid-log phase and then induced via the addition of ATc. Bxb1 integrase along with gp47 instigate the restoring of the attP site and thus remove pKM464. Through counter-selection on 10% sucrose (Fisher), mutants cured of all three plasmids can then be isolated.

## 1.15 Growth Assay

### 1.15.1 Spot Assay

*M. smegmatis* strains were grown to mid-log phase and back-diluted to an $OD_{600nm}$ of 0.05. Each stain was then serially back-diluted (10-fold) and 3 µL of each dilution was spotted onto LB agar. Plates were incubated for 3 days at 37°C.

### 1.15.2 Growth Curve

Mycobacterial strains were grown to mid-log phase and back-diluted to an $OD_{600nm}$ of 0.05. 5 mL of each strain was added to a 18mm outer diameter Hungate tube (Chemglass Life Sciences) containing 7H9 medium supplemented with sub-inhibitory concentrations of various drugs and conditions. Each strain was then incubated at 37°C for *M. smegmatis* and 30°C for *M. marinum* in a MicrobeMeter (Humane Technologies). $OD_{600nm}$ readings were taken every 5 min until the bacterial density reached a plateau.

## 1.16 Microscopy

Mycobacterial strains were grown to mid-log phase and back-diluted to an $OD_{600nm}$ of 0.05. Cells were then centrifuged at 1000 rpm for 1 min and 3 µL of supernatant just above the pellet was spotted onto a 1% agarose mounted microscope slides and sealed with a cover slip. For each sample over 300 cells were imaged via phase contrast on the Zeiss Axio Observer microscope at 100x magnification. Cell size, width and chaining was then calculated using the MicrobeJ plugin for ImageJ (Ducret et al., 2016).

For HADA (3-[7-hydroxycoumarin]-carboxamide-D-alanine; TOCRIS) staining, the mycobacterial strains are grown to mid-log phase and back-diluted to an $OD_{600nm}$ of

0.3. 0.5 mL of cells are aliquoted into Eppendorf tubes and HADA is added at a concentration of 500 μM. Cells are then incubated at 180rpm for one generation time (6 h at 30°C for *M. marinum* and 3 h at 37°C for *M. smegmatis*). Cells were washed 3 times via centrifugation at 12,000 rpm in 0.5 mL of 7H9 medium and then centrifuged at 1000 rpm for 1 min. 3 μL of supernatant just above the pellet was spotted onto a 1% agarose mounted microscope slides and sealed with a cover slip. Lastly, cells were imaged via epifluorescence microscopy ($\lambda_{abs}$ $405_{nm}$) using the Zeiss Axio Observer microscope.

# General Discussion

## 6.1 Introduction

The *Mycobacterium* genus encompasses numerous important environmental and animal pathogens including *M. tuberculosis* which is the causative agent of TB and one of the leading causes of death from a single infectious agent (WHO, 2023). Mycobacteria can be characterised by their thick and waxy cell wall which acts as an effective protective barrier and thus these bacteria are intrinsically resistant to many antibiotics (Parte et al., 2020). Furthermore, although the anti-mycobacterial arsenal is limited, mycobacterial diseases are curable. However, the treatments are often extensive and there are growing concerns over antibiotic resistance (Zhou et al., 2020). Therefore, there is great demand for the discovery of new therapeutic options.

Although there is an abundance of mycobacterial sequence information readily available, functional annotation of the genome is lagging and thus these bacteria are still poorly understood (National Institutes for Health, 2023). For example, although widely studied, in *M. tuberculosis*, 25% of the genome remains functionally uncharacterised and thus new methods for deciphering mycobacterial gene function are needed (Modlin et al., 2021).

Genetic engineering is a useful tool for determining the function of genes in which, through reverse genetics, genes can be characterised by linking mutant genotypes to phenotypic observations. Unfortunately, due to reduced HR and transformation

efficiency, traditional genetic engineering methods have limited success in mycobacteria (Van Kessel and Hatfull, 2008). However, CRISPR and ORBIT represent exciting new tools for mycobacterial genetic engineering (Murphy et al., 2018; Meijers et al., 2020). Whilst these methods offer an extremely targeted approach, given that large portions of mycobacterial genomes remain functionally unannotated, employing a gene-by-gene approach is not suitable for interrogating gene function on a global scale. In contrast, functional genomics permits the characterisation of entire bacterial genomes (Typas et al., 2010; Nichols et al., 2011)

In recent years, due to the extremely high-throughput, functional genomics has been dominated by TnSeq. TnSeq involves the generation of pooled libraries of mutants via transposon insertion mutagenesis followed by conditional essentiality analysis in which mutant fitness is determined based on relative abundance via sequencing analysis (Griffin et al., 2011). Although the method is fast and has proven an effective tool for deciphering gene function, the output is limited to sequencing data and thus nuanced alterations in mutant fitness could be overlooked. Instead, the functional genomic screening of ordered mutant libraries enables a greater diversity of data output facilitating a more in-depth investigation of mutant fitness such as the analysis of mutant growth kinetics, colony morphology and cellular conformation (Nichols et al., 2011; Kritikos et al., 2017). Ordered libraries also represent valuable resources, boasting entire collections of mutants which can be individually selected for further study.

Nevertheless, generating ordered mutant libraries is a challenging task. The first ever ordered bacterial mutant library used a gene-by-gene approach in which each gene was targeted for deletion via gene substitution with an antibiotic resistance cassette (Baba et al., 2006). This method therefore required the generation of individual DNA constructs and thus proved a slow and laborious process (Borgers et al., 2020). Instead, transposon mutagenesis represents a much more efficient method in which transposons are introduced into bacterial cells and insert within the genome at random thus disrupting genes through the insertion of premature stop codons (Vandewalle et al., 2015). Therefore, a pooled library of mutants can be easily generated en masse and in tandem. These mutants can then be arrayed into a library and sequencing methods can be employed to determine the mutant identities (Vandewalle et al., 2015; Baym et al., 2016).

## 6.2 Constructing an Ordered *M. abscessus* Transposon Insertion Library

Although *M. abscessus* is an emerging clinical pathogen often associated with nosocomial infections and respiratory diseases, it remains poorly understood and genetic resources are limited (Johansen et al., 2020). We therefore adapted a method for generating ordered mycobacterial transposon insertion libraries (Borgers et al., 2020) in order to construct a *M. abscessus* mutant library which could be combined with reverse genetics and functional genomics to help functionally characterise the *M. abscessus* genome.

To construct the library, we first investigated the suitability of two different phage systems, ΦMycoMarT7 and phDB36, and two different *M. abscessus* subspecies, *M. a. abscessus* and *M. a. massiliense*. Due to a substantially increased transduction efficiency we determined that ΦMycoMarT7 was the most optimal phage system though a different antibiotic background would have been more favourable since many existing mycobacterial genetic tools already use kanamycin. In addition, since we discovered that *M. a. massiliense* exhibited intrinsic Zeocin™ resistance we reasoned that *M. a. abscessus* was the most appropriate strain for library construction as this would avoid the generation of mutants with multidrug resistance.

Next, we investigated a method for ordering mutant libraries using colony-picking robotics. However, we encountered a selection bias and whilst it would be possible to optimise this procedure by adjusting the robot colony selection criteria, we instead opted for a manual colony picking method for the ordering of the library.

Therefore, to generate the library we incubated *M. a. abscessus* cells with ΦMycoMarT7 and selected for successful transductions of selective agar plates. 18,432 colonies were then manually picked into 96-well plates.

To determine the identity of each library mutant we could sequence each individual well of the library, however, this would be extremely cost and time consuming (Borgers et al., 2020). Instead, we employed a library deconvolution method coined CP-CSeq (Vandewalle et al., 2015). In this method, the library mutants are systematically

ordered into 80 pools and sequenced. Following sequencing a deconvolution algorithm can be applied to determine the location of each mutant within the library.

Unfortunately, we were unable to finish the CP-CSeq method and thus the library currently exists as an ordered collection of unknown mutants. However once finished, this library will represent the largest collection of *M. abscessus* mutants in existence and can be combined with functional genomics to help decipher the function of the *M. abscessus* genome and thus better our understanding of these bacteria.

## 6.3 Chemical-Genomic Screening of *M. bovis* BCG

Nevertheless, since we had access to an ordered *M. bovis* BCG mutant library (Borgers et al., 2020), we were able to optimise a method for conducting functional genomics on ordered mycobacterial libraries. *M. bovis* is an important animal pathogen, responsible for significant economic costs to the farming industry and given that *M. bovis* BCG is a biosafety level 2 pathogen which is closely related to *M. tuberculosis* it can be also used as a model organism (Kanipe and Palmer, 2020). Like *M. tuberculosis* and *M. abscessus* much of the *M. bovis* genome remains uncharacterised and thus the functional genomic screening of these bacteria could elucidate the function of various previously uncharacterised genes.

Given that IRIS, a high-throughput colony morphology analysis tool compatible with functional genomics, has been recently developed (Kritikos et al., 2017), we determined that analysing colony growth would result in the best quality functional

genomics dataset by providing the greatest diversity of data output whilst maintaining experimental throughput. Furthermore, Doherty et al. (2023) recently developed a chemical-genomics analysis tool called ChemGAPP which can generate mutant fitness scores from IRIS data as well as quality control and normalise the dataset thus further streamlining the method.

The *M. bovis* BCG library is too large to conduct a chemical-genomic screening and many of the mutants are redundant (Borgers et al., 2020). Therefore, we first condensed the library to contain only the best quality gene knockout duplicate by only selecting mutants with the largest gene deletion and removing any mutants that resulted in less than 20% of disruption to the gene.

We next optimised the screening method and determined that using high-throughput colony arraying and imaging robotics *M. bovis* mutants could be ordered into a 384-colony format and grown on agar plates containing subinhibitory concentrations of various drugs. To validate the method, we also conducted a pilot screen with 2 library plates across 22 conditions. We identified 76 mutants that had a significant alteration in fitness in response to at least one condition and we also discovered that various mutants exhibited expected drug responses thus verifying that our method could accurately identify mutant phenotypes. In addition, we phenotypically profiled numerous mutants of genes which had not been previously characterised. As genes of similar function are likely to cluster together based on their mutant chemical response profiles, through clustering analysis we were also able to infer function and thus highlighted various genes that could be worth researching further.

Unfortunately, we were unable to complete a full screen. However, given the success of the pilot screen which only represents around 1% of the final dataset. It is likely that this method will be extremely effective at deciphering the function of genes and thus could lead to the discovery of novel drug and vaccine targets as well as important antibiotic resistance and virulence pathways.

## 6.4 Validation of Chemical-Genomic Hits and Functional Characterisation of Cell Wall Associated Enzymes

Although we controlled for poor quality mutants by removing any mutants from our condensed library that did not cause a gene disruption of at least 20%, it is possible that a functional protein could be produced. In addition, transposon insertions can cause polar effects and thus disruptions to downstream genes (Borgers et al., 2020). It is therefore important that any chemical-genomic screening hits are validated by analysing the phenotype of a clean KO. ORBIT represents an efficient method for generating mycobacterial mutants but at current has only been applied to *M. smegmatis* and *M. tuberculosis* (Murphy et al., 2018). We therefore investigated the adaptability of this method to other mycobacterial species whilst also conducting a phenotypic analysis of various mutants thought to play a role in the mycobacterial cell envelope and cell-wall associated drug resistance.

Given that the mycobacterial cell envelope is unique, exposed and essential, it is the perfect target for anti-mycobacterial drugs (Abrahams and Besra, 2018). However, whilst the mycobacterial cell envelope synthesis pathways are well characterised, cell wall recycling and maintenance is poorly understood and thus, by elucidating the genes involved in these processes, new drug targets could be discovered (Abrahams and Besra, 2021). In addition, there are rising concerns over increasing resistance to our current arsenal of anti-tuberculosis drugs, therefore, by understanding the mechanisms for drug resistance we can help to inform clinicians to determine the most effective treatments options (Zhou et al., 2020).

Unfortunately, we were unable to adapt the ORBIT method to *M. marinum*. However, we were able to successfully generate *M. smegmatis lpqF, MSMEG_0740, MSMEG_4352 and MSMEG_6255* deletion mutants, all thought to play a key role in either the recycling of cell envelope components or cell-wall associated antibiotic resistance. Through growth kinetics and chemical-genetic screening we discovered that Δ*lpqF* had defective growth under standard conditions but that WT growth was restored upon the addition of meropenem. Furthermore, through microscopy, we found that Δ*lpqF* exhibited altered chain length and size. However, *M. smegmatis* and *M. marinum* Δ*lpqF* strains demonstrated contrasting phenotypes. Similarly, we identified that Δ*MSMEG_0740* had reduced cell width and length when compared to WT. However, no significant phenotypes were discovered via chemical-genetics or growth kinetics. These findings have helped to support research in collaborative studies in which it was discovered that LpqF functions as a D,D-carboxypeptidase (Thomas et al., Unpublished Data) and *MSMEG_0740* (Rv0365c) acts as an α-1,6

mannanase (Franklin et al., 2023). Therefore, both of these enzymes play critical roles in the cell wall and thus have potential as drug targets.

## 6.5 Conclusion

Given that mycobacteria are clinically relevant and yet remain poorly understood, we have developed a pipeline for characterising mycobacterial genomes. We first outlined a method for generating mycobacterial ordered mutant libraries and in the process, constructed the largest collection of *M. abscessus* mutants. We also optimised a method for mycobacterial chemical-genomic screening and conducted a pilot screen in *M. bovis* BCG, facilitating the phenotypic analysis of various mutants, many of which were knockouts of previously uncharacterised genes. Lastly, we established a method for validating screening hits in which clean knockouts are generated via ORBIT and can be analysed through chemical-genetic screening, growth kinetics and microscopy. Furthermore, during this process we helped to characterise two cell wall associated proteins, a D,D-carboxypeptidate, LpqF and an α-1,6 mannanase, *MSMEG_0740* (Rv0365c).

# List of References


Abrahams, K.A. and Besra, G.S. (2018) Mycobacterial cell wall biosynthesis: a multifaceted antibiotic target. *Parasitology*, *145*(2), pp.116-133.

Abrahams, K.A. and Besra, G.S. (2021) Synthesis and recycling of the mycobacterial cell envelope. *Current Opinion in Microbiology*, *60*, pp.58-65.

Agarwal, N. (2020) Construction of a novel CRISPRi-based tool for silencing of multiple genes in *Mycobacterium tuberculosis*. *Plasmid* [online], *110*, p.102515, doi: https://doi.org/10.1016/j.plasmid.2020.102515

Ahangar, M.S., Furze, C.M., Guy, C.S., Cooper, C., Maskew, K.S., Graham, B., Cameron, A.D. and Fullam, E. (2018) Structural and functional determination of homologs of the *Mycobacterium tuberculosis* N-acetylglucosamine-6-phosphate deacetylase (NagA), *Journal of Biological Chemistry*, *293*(25), pp.9770-9783.

Akusobi, C., Benghomari, B.S., Choudhery, S., Singhvi, S., Hicks, A.L., Ioerger, T.R. and Rubin, E.J. (2023) Transposon-sequencing across multiple *Mycobacterium abscessus* isolates reveals significant functional genomic diversity among strains. *bioRxiv* [online], doi: https://doi.org/10.1101/2023.03.29.534805.

Alderwick, L.J., Dover, L.G., Veerapen, N., Gurcha, S.S., Kremer, L., Roper, D.L., Pathak, A.K., Reynolds, R.C. and Besra, G.S. (2008) Expression, purification and characterisation of soluble GlfT and the identification of a novel galactofuranosyltransferase Rv3782 involved in priming GlfT-mediated galactan polymerisation in *Mycobacterium tuberculosis*. *Protein expression and purification*, *58*(2), pp.332-341.

Alderwick, L.J., Radmacher, E., Seidel, M., Gande, R., Hitchen, P.G., Morris, H.R., Dell, A., Sahm, H., Eggeling, L. and Besra, G.S. (2005) Deletion of Cg-emb in corynebacterianeae leads to a novel truncated cell wall arabinogalactan, whereas inactivation of Cg-ubiA results in an arabinan-deficient mutant with a cell wall galactan core. *Journal of Biological Chemistry*, *280*(37), pp.32362-32371.

Alderwick, L.J., Seidel, M., Sahm, H., Besra, G.S. and Eggeling, L. (2006) Identification of a novel arabinofuranosyltransferase (AftA) involved in cell wall arabinan biosynthesis in *Mycobacterium tuberculosis*. *Journal of Biological Chemistry*, *281*(23), pp.15653-15661.

Aldovini, A., Husson, R.N. and Young, R.A. (1993) The uraA locus and homologous recombination in *Mycobacterium bovis* BCG. *Journal of bacteriology*, *175*(22), pp.7282-7289.

Al-Jourani, O., Benedict, S.T., Ross, J., Layton, A.J., van der Peet, P., Marando, V.M., Bailey, N.P., Heunis, T., Manion, J., Mensitieri, F. and Franklin, A. (2023) Identification of d-arabinan-degrading enzymes in mycobacteria. *Nature Communications* [online], *14*(1), doi: https://doi.org/10.1038/s41467-023-37839-5

Anglada-Girotto, M., Handschin, G., Ortmayr, K., Campos, A.I., Gillet, L., Manfredi, P., Mulholland, C.V., Berney, M., Jenal, U., Picotti, P. and Zampieri, M. (2022) Combining CRISPRi and metabolomics for functional annotation of compound libraries. *Nature chemical biology*, *18*(5), pp.482-491.

Baba, T., Ara, T., Hasegawa, M., Takai, Y., Okumura, Y., Baba, M., Datsenko, K.A., Tomita, M., Wanner, B.L. and Mori, H. (2006) Construction of *Escherichia coli* K-12 in-frame, single-gene knockout mutants: the Keio collection. *Molecular Systems Biology* [online], *2*, doi: https://doi.org/10.1038/msb4100050.

Balganesh, M., Kuruppath, S., Marcel, N., Sharma, S., Nair, A. and Sharma, U. (2010) Rv1218c, an ABC transporter of *Mycobacterium tuberculosis* with implications in drug discovery. *Antimicrobial agents and chemotherapy*, *54*(12), pp.5167-5172.

Banaei, N., Kincaid, E.Z., Lin, S.Y.G., Desmond, E., Jacobs Jr, W.R. and Ernst, J.D. (2009) Lipoprotein processing is essential for resistance of *Mycobacterium tuberculosis* to malachite green. *Antimicrobial agents and chemotherapy*, *53*(9), pp.3799-3802.

Bannantine, J.P., Zinniel, D.K. and Barletta, R.G. (2019) 'Transposon mutagenesis in *Mycobacterium avium* subspecies *paratuberculosis*' in *Microbial Transposon Mutagenesis: Protocols and Applications*, New York: Humana Press, pp.117-125.

Bansal-Mutalik, R. and Nikaido, H. (2014) Mycobacterial outer membrane is a lipid bilayer and the inner membrane is unusually rich in diacyl phosphatidylinositol dimannosides. *Proceedings of the National Academy of Sciences*, *111*(13), pp.4958-4963.

Baranowski, C., Welsh, M.A., Sham, L.T., Eskandarian, H.A., Lim, H.C., Kieser, K.J., Wagner, J.C., McKinney, J.D., Fantner, G.E., Ioerger, T.R. and Walker, S. (2018) Maturing *Mycobacterium smegmatis* peptidoglycan requires non-canonical crosslinks to maintain shape. *Elife* [online], *7*, p.e37516, doi: https://doi.org/10.7554/eLife.37516.

Bardarov, S., Kriakov, J., Carriere, C., Yu, S., Vaamonde, C., McAdam, R.A., Bloom, B.R., Hatfull, G.F. and Jacobs Jr, W.R. (1997) Conditionally replicating mycobacteriophages: a system for transposon delivery to *Mycobacterium tuberculosis*. *Proceedings of the National Academy of Sciences*, *94*(20), pp.10961-10966.

Barrangou, R., Fremaux, C., Deveau, H., Richards, M., Boyaval, P., Moineau, S., Romero, D.A. and Horvath, P. (2007) CRISPR provides acquired resistance against viruses in prokaryotes. *Science*, *315*(5819), pp.1709-1712.

Barratt, A.S., Arnoult, M.H., Ahmadi, B.V., Rich, K.M., Gunn, G.J. and Stott, A.W. (2018) A framework for estimating society's economic welfare following the introduction of an animal disease: the case of Johne's disease. *PLoS One* [online], *13*(6), doi: https://doi.org/10.1371/journal.pone.0198436

Barreteau, H., Kovač, A., Boniface, A., Sova, M., Gobec, S. and Blanot, D. (2008) Cytoplasmic steps of peptidoglycan biosynthesis. *FEMS microbiology reviews*, *32*(2), pp.168-207.

Basavannacharya, C., Robertson, G., Munshi, T., Keep, N.H. and Bhakta, S. (2010) ATP-dependent MurE ligase in *Mycobacterium tuberculosis*: biochemical and structural characterisation. *Tuberculosis*, *90*(1), pp.16-24.

Batt, S.M., Minnikin, D.E. and Besra, G.S. (2020) The thick waxy coat of mycobacteria, a protective layer against antibiotics and the host's immune system. *Biochemical Journal*, *477*(10), pp.1983-2006.

Baumgart, M., Schubert, K., Bramkamp, M. and Frunzke, J. (2016) Impact of LytR-CpsA-Psr proteins on cell wall biosynthesis in *Corynebacterium glutamicum*. *Journal of bacteriology*, *198*(22), pp.3045-3059.

Behura, A., Das, M., Kumar, A., Naik, L., Patel, S., Nayak, D.K., Mishra, A., Mishra, A. and Dhiman, R. (2023) 'Mycobacterial biofilm: Structure and its functional relevance in the pathogenesis'. In *Understanding Microbial Biofilms*. India: Academic Press, pp. 461-474.

Bender, J., Kuo, J. and Kleckner, N. (1991) Genetic evidence against intramolecular rejoining of the donor DNA molecule following IS10 transposition. *Genetics*, *128*(4), pp.687-694.

Benson, T.E., Marquardt, J.L., Marquardt, A.C., Etzkorn, F.A. and Walsh, C.T. (1993) Overexpression, purification, and mechanistic study of UDP-N-acetylenolpyruvylglucosamine reductase. *Biochemistry*, *32*(8), pp.2024-2030.

Bernheim, A., Calvo-Villamañán, A., Basier, C., Cui, L., Rocha, E.P., Touchon, M. and Bikard, D. (2017) Inhibition of NHEJ repair by type II-A CRISPR-Cas systems in bacteria. *Nature communications* [online], *8*(1), doi: https://doi.org/10.1038/s41467-017-02350-1.

Besra, G.S., Khoo, K.H., McNeil, M.R., Dell, A., Morris, H.R. and Brennan, P.J. (1995) A new interpretation of the structure of the mycolyl-arabinogalactan complex of *Mycobacterium tuberculosis* as revealed through characterization of oligoglycosylalditol fragments by fast-atom bombardment mass spectrometry and 1H nuclear magnetic resonance spectroscopy. *Biochemistry*, *34*(13), pp.4257-4266.

Beste, D.J., Espasa, M., Bonde, B., Kierzek, A.M., Stewart, G.R. and McFadden, J. (2009) The genetic requirements for fast and slow growth in mycobacteria. *PloS one* [online], *4*(4), p.e5349, doi: https://doi.org/10.1371/journal.pone.0005349.

Bhamidi, S., Scherman, M.S., Rithner, C.D., Prenni, J.E., Chatterjee, D., Khoo, K.H. and McNeil, M.R. (2008) The identification and location of succinyl residues and the characterization of the interior arabinan region allow for a model of the complete primary structure of *Mycobacterium tuberculosis* mycolyl arabinogalactan. *Journal of Biological Chemistry*, *283*(19), pp.12992-13000.

Bhatt, A., Brown, A.K., Singh, A., Minnikin, D.E. and Besra, G.S. (2008) Loss of a mycobacterial gene encoding a reductase leads to an altered cell wall containing β-oxo-mycolic acid analogs and accumulation of ketones. *Chemistry & biology*, *15*(9), pp.930-939.

Bianco, M.V., Blanco, F.C., Imperiale, B., Forrellad, M.A., Rocha, R.V., Klepp, L.I., Cataldi, A.A., Morcillo, N. and Bigi, F. (2011) Role of P27-P55 operon from *Mycobacterium tuberculosis* in the resistance to toxic compounds. *BMC infectious diseases*, *11*(1), pp.1-9.

Billman-Jacobe, H., Haites, R.E. and Coppel, R.L. (1999) Characterization of a Mycobacterium smegmatis mutant lacking penicillin binding protein 1. *Antimicrobial agents and chemotherapy*, *43*(12), pp.3011-3013.

Birch, H.L., Alderwick, L.J., Bhatt, A., Rittmann, D., Krumbach, K., Singh, A., Bai, Y., Lowary, T.L., Eggeling, L. and Besra, G.S. (2008) Biosynthesis of mycobacterial arabinogalactan: identification of a novel α (1→ 3) arabinofuranosyltransferase. *Molecular microbiology*, *69*(5), pp.1191-1206.

Boeck, L., Burbaud, S., Skwark, M., Pearson, W.H., Sangen, J., Wuest, A.W., Marshall, E.K., Weimann, A., Everall, I., Bryant, J.M. and Malhotra, S. (2022) *Mycobacterium abscessus* pathogenesis identified by phenogenomic analyses. *Nature Microbiology*, *7*(9), pp.1431-1441.

Boehme, C.C., Nabeta, P., Hillemann, D., Nicol, M.P., Shenai, S., Krapp, F., Allen, J., Tahirli, R., Blakemore, R., Rustomjee, R. and Milovic, A. (2010) Rapid molecular detection of tuberculosis and rifampin resistance. *New England Journal of Medicine*, *363*(11), pp.1005-1015.

Borgers, K., Ou, J.Y., Zheng, P.X., Tiels, P., Van Hecke, A., Plets, E., Michielsen, G., Festjens, N., Callewaert, N. and Lin, Y.C. (2019) Reference genome and comparative genome analysis for the WHO reference strain for *Mycobacterium bovis* BCG Danish, the present tuberculosis vaccine. *BMC genomics*, *20*(1), pp.1-14.

Borgers, K., Vandewalle, K., Van Hecke, A., Michielsen, G., Plets, E., van Schie, L., Vanmarcke, S., Schindfessel, L., Festjens, N. and Callewaert, N. (2020) Development of a counterselectable transposon to create markerless knockouts from an 18,432-

clone ordered *Mycobacterium bovis* bacillus Calmette-Guerin mutant resource. *Msystems* [online], *5*(4), doi: 10.1128/mSystems.00180-20.

Bosch, B., DeJesus, M.A., Poulton, N.C., Zhang, W., Engelhart, C.A., Zaveri, A., Lavalette, S., Ruecker, N., Trujillo, C., Wallach, J.B. and Li, S. (2021) Genome-wide gene expression tuning reveals diverse vulnerabilities of M. tuberculosis. *Cell*, *184*(17), pp.4579-4592.

Bourai, N., Jacobs Jr, W.R. and Narayanan, S. (2012) Deletion and overexpression studies on DacB2, a putative low molecular mass penicillin binding protein from *Mycobacterium tuberculosis* H37Rv. *Microbial pathogenesis*, *52*(2), pp.109-116.

Boutte, C.C., Baer, C.E., Papavinasasundaram, K., Liu, W., Chase, M.R., Meniche, X., Fortune, S.M., Sassetti, C.M., Ioerger, T.R. and Rubin, E.J. (2016) A cytoplasmic peptidoglycan amidase homologue controls mycobacterial cell wall synthesis. *Elife* [online], *5*(1), doi: https://doi.org/10.7554/eLife.14590.001.

Brouns, S.J., Jore, M.M., Lundgren, M., Westra, E.R., Slijkhuis, R.J., Snijders, A.P., Dickman, M.J., Makarova, K.S., Koonin, E.V. and Van Der Oost, J. (2008) Small CRISPR RNAs guide antiviral defense in prokaryotes. *Science*, *321*(5891), pp.960-964.

Camacho, L.R., Ensergueix, D., Perez, E., Gicquel, B. and Guilhot, C. (1999) Identification of a virulence gene cluster of *Mycobacterium tuberculosis* by signature-tagged transposon mutagenesis. *Molecular microbiology*, *34*(2), pp.257-267.

Campos, M., Govers, S.K., Irnov, I., Dobihal, G.S., Cornet, F. and Jacobs-Wagner, C. (2018) Genomewide phenotypic analysis of growth, cell morphogenesis, and cell cycle events in *Escherichia coli*. *Molecular systems biology* [online], *14*(6), p.e7573, doi: https://doi.org/10.15252/msb.20177573.

Cantrell, S.A., Leavell, M.D., Marjanovic, O., Iavarone, A.T., Leary, J.A. and Riley, L.W. (2013) Free mycolic acid accumulation in the cell wall of the mce1 operon mutant strain of *Mycobacterium tuberculosis*. *Journal of microbiology*, *51*(5), pp.619-626.

Carey, A.F., Rock, J.M., Krieger, I.V., Chase, M.R., Fernandez-Suarez, M., Gagneux, S., Sacchettini, J.C., Ioerger, T.R. and Fortune, S.M. (2018) TnSeq of *Mycobacterium tuberculosis* clinical isolates reveals strain-specific antibiotic liabilities. *PLoS pathogens* [online], *14*(3), p.e1006939, doi: https://doi.org/10.1371/journal.ppat.1007846.

Cascioferro, A., Boldrin, F., Serafini, A., Provvedi, R., Palù, G. and Manganelli, R. (2010) Xer site-specific recombination, an efficient tool to introduce unmarked deletions into mycobacteria. *Applied and Environmental Microbiology* [online], *76*(15), pp.5312-5316, doi: https://doi.org/10.1128/AEM.00382-10.

Centers for Disease Control and Prevention (CDC). (2023) Treatment for TB Disease [online].

Available at: https://www.cdc.gov/tb/topic/treatment/tbdisease.htm
[Accessed 29th July, 2024]

Choudhary, E., Thakur, P., Pareek, M. and Agarwal, N. (2015) Gene silencing by CRISPR interference in mycobacteria. *Nature communications* [online], *6*(1), p.6267, doi: https://doi.org/10.1038/ncomms7267.

Chung, B.C., Zhao, J., Gillespie, R.A., Kwon, D.Y., Guan, Z., Hong, J., Zhou, P. and Lee, S.Y. (2013) Crystal structure of MraY, an essential membrane enzyme for bacterial cell wall synthesis. *Science*, *341*(6149), pp.1012-1016.

Cilloni, L., Fu, H., Vesga, J.F., Dowdy, D., Pretorius, C., Ahmedov, S., Nair, S.A., Mosneaga, A., Masini, E., Sahu, S. and Arinaminpathy, N. (2020) The potential impact of the COVID-19 pandemic on the tuberculosis epidemic a modelling analysis. *EClinicalMedicine* [online], *28*, doi: https://doi.org/10.1016/j.eclinm.2020.100603.

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), pp.10-1128.

Cox, J.A., Abrahams, K.A., Alemparte, C., Ghidelli-Disse, S., Rullas, J., Angulo-Barturen, I., Singh, A., Gurcha, S.S., Nataraj, V., Bethell, S. and Remuiñán, M.J. (2016) THPP target assignment reveals EchA6 as an essential fatty acid shuttle in mycobacteria. *Nature microbiology*, *1*(2), pp.1-10.

Crellin, P.K., Kovacevic, S., Martin, K.L., Brammananth, R., Morita, Y.S., Billman-Jacobe, H., McConville, M.J. and Coppel, R.L. (2008) Mutations in pimE restore lipoarabinomannan synthesis and growth in a *Mycobacterium smegmatis* lpqW mutant. *Journal of bacteriology*, *190*(10), pp.3690-3699.

Daffe, M. and Etienne, G. (1999) The capsule of *Mycobacterium tuberculosis* and its implications for pathogenicity.*Tubercle and lung disease*, *79*(3), pp.153-169.

Daffe, M., Brennan, P.J. and McNeil, M. (1990) Predominant structural features of the cell wall arabinogalactan of *Mycobacterium tuberculosis* as revealed through characterization of oligoglycosyl alditol fragments by gas chromatography/mass spectrometry and by 1H and 13C NMR analyses. *Journal of Biological Chemistry*, *265*(12), pp.6734-6743.

Dao, D.N., Kremer, L., Guérardel, Y., Molano, A., Jacobs, W.R., Porcelli, S.A. and Briken, V. (2004) *Mycobacterium tuberculosis* lipomannan induces apoptosis and interleukin-12 production in macrophages. *Infection and immunity*, *72*(4), pp.2067-2074.

Datsenko, K.A. and Wanner, B.L. (2000) One-step inactivation of chromosomal genes in *Escherichia coli* K-12 using PCR products. *Proceedings of the National Academy of Sciences*, *97*(12), pp.6640-6645.

De Berardinis, V., Vallenet, D., Castelli, V., Besnard, M., Pinet, A., Cruaud, C., Samair, S., Lechaplais, C., Gyapay, G., Richez, C. and Durot, M. (2008) A complete collection of single-gene deletion mutants of *Acinetobacter baylyi* ADP1. *Molecular systems biology* [online], *4*(1), doi: https://doi.org/10.1038/msb.2008.10.

De Lorenzo, V., Herrero, M., Jakubzik, U. and Timmis, K.N. (1990) Mini-Tn5 transposon derivatives for insertion mutagenesis, promoter probing, and chromosomal insertion of cloned DNA in gram-negative eubacteria. *Journal of bacteriology*, *172*(11), pp.6568-6572.

de Wet, T.J., Winkler, K.R., Mhlanga, M., Mizrahi, V. and Warner, D.F. (2020) Arrayed CRISPRi and quantitative imaging describe the morphotypic landscape of essential mycobacterial genes. *Elife* [online], *9*, p.e60083, doi: https://doi.org/10.7554/eLife.60083.

Dianišková, P., Korduláková, J., Škovierová, H., Kaur, D., Jackson, M., Brennan, P.J. and Mikušová, K. (2011) Investigation of ABC transporter from mycobacterial arabinogalactan biosynthetic cluster. *General physiology and biophysics*, *30*(3), p.239.

DeJesus, M.A., Gerrick, E.R., Xu, W., Park, S.W., Long, J.E., Boutte, C.C., Rubin, E.J., Schnappinger, D., Ehrt, S., Fortune, S.M. and Sassetti, C.M. (2017) Comprehensive essentiality analysis of the *Mycobacterium tuberculosis* genome via saturating transposon mutagenesis. *MBio* [online], *8*(1), p.e02133-16, doi: https://doi.org/10.1128/mBio.02133-16.

Denome, S.A., Elf, P.K., Henderson, T.A., Nelson, D.E. and Young, K.D. (1999) *Escherichia coli* mutants lacking all possible combinations of eight penicillin binding proteins: viability, characteristics, and implications for peptidoglycan synthesis. *Journal of bacteriology*, *181*(13), pp.3981-3993.

Dong, C., Fontana, J., Patel, A., Carothers, J.M. and Zalatan, J.G. (2018) Synthetic CRISPR-Cas gene activators for transcriptional reprogramming in bacteria. *Nature communications* [online], *9*(1), p.2489. doi: https://doi.org/10.1038/s41467-018-04901-6.

Dubée, V., Triboulet, S., Mainardi, J.L., Ethève-Quelquejeu, M., Gutmann, L., Marie, A., Dubost, L., Hugonnet, J.E. and Arthur, M. (2012) Inactivation of *Mycobacterium tuberculosis* L, D-transpeptidase LdtMt1 by carbapenems and cephalosporins. *Antimicrobial agents and chemotherapy*, *56*(8), pp.4189-4195.

Ducret, A., Quardokus, E.M. and Brun, Y.V. (2016) MicrobeJ, a tool for high throughput bacterial cell detection and quantitative analysis. *Nature microbiology*, *1*(7), pp.1-7.

Dulberger, C.L., Rubin, E.J. and Boutte, C.C. (2020) The mycobacterial cell envelope – a moving target. *Nature Reviews Microbiology*, *18*(1), pp.47-59.

Ealand, C.S., Asmal, R., Mashigo, L., Campbell, L. and Kana, B.D. (2019) Characterization of putative DD-carboxypeptidase-encoding genes in *Mycobacterium smegmatis*. *Scientific Reports*, *9*(1), p.5194, doi: https://doi.org/10.1038/s41598-019-41001-x.

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 [online]*, *33*(2). Available at: https://doi.org/10.1093/nar/gni013.

El Zoeiby, A., Sanschagrin, F. and Levesque, R.C. (2003) Structure and function of the Mur enzymes: development of novel inhibitors. *Molecular microbiology*, *47*, pp.1-12.

Fay, A., Czudnochowski, N., Rock, J.M., Johnson, J.R., Krogan, N.J., Rosenberg, O. and Glickman, M.S. (2019) Two accessory proteins govern MmpL3 mycolic acid transport in mycobacteria. *MBio* [online], *10*(3), doi: https://doi.org/10.1128/mBio.00850-19.

Fernandes, N.D. and Kolattukudy, P.E. (1996) Cloning, sequencing and characterization of a fatty acid synthase-encoding gene from *Mycobacterium tuberculosis* var. *bovis* BCG. *Gene*, *170*(1), pp.95-99.

Fey, P.D., Endres, J.L., Yajjala, V.K., Widhelm, T.J., Boissy, R.J., Bose, J.L. and Bayles, K.W. (2013) A genetic resource for rapid and comprehensive phenotype screening of nonessential *Staphylococcus aureus* genes. *Mbio* [online], p.e00537-12, doi: https://doi.org/10.1128/mbio.00537-12.

Filteau, M., Hamel, V., Pouliot, M.C., Gagnon-Arsenault, I., Dubé, A.K. and Landry, C.R. (2015) Evolutionary rescue by compensatory mutations is constrained by genomic and environmental backgrounds. *Molecular Systems Biology* [online], *11*(10), p.832, doi: https://doi.org/10.15252/msb.20156444.

Flores, A.R., Parsons, L.M. and Pavelka Jr, M.S. (2005) Characterization of novel *Mycobacterium tuberculosis* and *Mycobacterium smegmatis* mutants hypersusceptible to β-lactam antibiotics. *Journal of bacteriology*, *187*(6), pp.1892-1900.

Foreman, M., Gershoni, M. and Barkan, D. (2020) A simplified and efficient method for Himar-1 transposon sequencing in bacteria, demonstrated by creation and analysis of a saturated transposon-mutant library in *Mycobacterium abscessus*. *MSystems* [online], *5*(5), p.e00976-20, doi: https://doi.org/10.1128/msystems.00976-20.

Franklin, A., Layton, A.J., Mize, T., Salgueiro, V.C., Sullivan, R., Benedict, S.T., Gurcha, S.S., Anso, I., Besra, G.S., Banzhaf, M. and Lovering, A.L. (2023) The mycobacterial glycoside hydrolase LamH enables capsular arabinomannan release and stimulates growth. *bioRxiv* [online].

Freed, E.F., Winkler, J.D., Weiss, S.J., Garst, A.D., Mutalik, V.K., Arkin, A.P., Knight, R. and Gill, R.T. (2015) Genome-wide tuning of protein expression levels to rapidly engineer microbial traits. *ACS Synthetic Biology*, *4*(11), pp.1244-1253.

Frehel, C., Rastogi, N., Benichou, J.C. and Ryter, A. (1988) Do test tube-grown pathogenic mycobacteria possess a protective capsule? *FEMS microbiology letters*, *56*(2), pp.225-229.

Fujiwara, M., Kawasaki, M., Hariguchi, N., Liu, Y. and Matsumoto, M. (2018) Mechanisms of resistance to delamanid, a drug for *Mycobacterium tuberculosis*. *Tuberculosis*, *108*, pp.186-194.

Fukuda, T., Matsumura, T., Ato, M., Hamasaki, M., Nishiuchi, Y., Murakami, Y., Maeda, Y., Yoshimori, T., Matsumoto, S., Kobayashi, K. and Kinoshita, T. (2013) Critical roles for lipomannan and lipoarabinomannan in cell wall integrity of mycobacteria and pathogenesis of tuberculosis. *MBio* [online], *4*(1), doi: https://doi.org/10.1128/mBio.00472-12.

Fullam, E., Prokes, I., Fütterer, K. and Besra, G.S. (2016) Structural and functional analysis of the solute-binding protein UspC from *Mycobacterium tuberculosis* that is specific for amino sugars. *Open biology*, *6*.

Giaever, G., Chu, A.M., Ni, L., Connelly, C., Riles, L., Véronneau, S., Dow, S., Lucau-Danila, A., Anderson, K., André, B. and Arkin, A.P. (2002) Functional profiling of the *Saccharomyces cerevisiae* genome. *nature*, *418*(6896), pp.387-391.

Garnier, T., Eiglmeier, K., Camus, J.C., Medina, N., Mansoor, H., Pryor, M., Duthoy, S., Grondin, S., Lacroix, C., Monsempe, C. and Simon, S. (2003) The complete genome sequence of *Mycobacterium bovis*. *Proceedings of the National Academy of Sciences*, *100*(13), pp.7877-7882.

Gawronski, J.D., Wong, S.M., Giannoukos, G., Ward, D.V. and Akerley, B.J. (2009) Tracking insertion mutants within libraries by deep sequencing and a genome-wide screen for Haemophilus genes required in the lung. *Proceedings of the National Academy of Sciences*, *106*(38), pp.16422-16427.

Gilbert, L.A., Larson, M.H., Morsut, L., Liu, Z., Brar, G.A., Torres, S.E., Stern-Ginossar, N., Brandman, O., Whitehead, E.H., Doudna, J.A. and Lim, W.A. (2013) CRISPR-mediated modular RNA-guided regulation of transcription in eukaryotes. *Cell*, *154*(2), pp.442-451.

Gilleron, M., Nigou, J., Cahuzac, B. and Puzo, G. (1999) Structural study of the lipomannans from Mycobacterium bovis BCG: characterisation of multiacylated forms of the phosphatidyl-myo-inositol anchor. *Journal of molecular biology*, *285*(5), pp.2147-2160.

Gilleron, M., Quesniaux, V.F. and Puzo, G. (2003) Acylation state of the phosphatidylinositol hexamannosides from *Mycobacterium bovis* Bacillus Calmette Guerin and *Mycobacterium tuberculosis* H37Rv and its implication in Toll-like receptor response. *Journal of Biological Chemistry*, *278*(32), pp.29880-29889.

Goffin, C. and Ghuysen, J.M. (1998) Multimodular penicillin-binding proteins: an enigmatic family of orthologs and paralogs. *Microbiology and molecular biology reviews*, *62*(4), pp.1079-1093.

Goodall, E.C., Morris, F.C., McKeand, S.A., Sullivan, R., Warner, I.A., Sheehan, E., Boelter, G., Icke, C., Cunningham, A.F., Cole, J.A. and Banzhaf, M. (2022) LI-Detector: a Method for Curating Ordered Gene-Replacement Libraries. *Microbiology Spectrum* [online], *10*(4), pp.e00833-22, doi: https://doi.org/10.1128/spectrum.00833-22.

Goodall, E.C., Robinson, A., Johnston, I.G., Jabbari, S., Turner, K.A., Cunningham, A.F., Lund, P.A., Cole, J.A. and Henderson, I.R. (2018) The essential genome of *Escherichia coli* K-12. *MBio* [online], *9*(1), doi: https://doi.org/10.1128/mbio.02096-17.

Goodfellow, M., Kämpfer, P., Busse, H.-J., Trujillo, M. E., Suzuki, K.I., Ludwig, W. and Whitman, W. B. (2012) *Bergey's Manual® of Systematic Bacteriology*. The Actinobacteria, Part A and B, Vol 5. 2nd ed. New York: Springer. pp.171-206.

Goodman, A.L., McNulty, N.P., Zhao, Y., Leip, D., Mitra, R.D., Lozupone, C.A., Knight, R. and Gordon, J.I. (2009) Identifying genetic determinants needed to establish a human gut symbiont in its habitat. *Cell host & microbe*, *6*(3), pp.279-289.

Gray, A.N., Koo, B.M., Shiver, A.L., Peters, J.M., Osadnik, H. and Gross, C.A. (2015) High-throughput bacterial functional genomics in the sequencing era. *Current opinion in microbiology*, *27*, pp.86-95.

Grzegorzewicz, A.E., De Sousa-D'Auria, C., McNeil, M.R., Huc-Claustre, E., Jones, V., Petit, C., kumar Angala, S., Zemanová, J., Wang, Q., Belardinelli, J.M. and Gao, Q. (2016) Assembling of the *Mycobacterium tuberculosis* cell wall core. *Journal of Biological Chemistry*, *291*(36), pp.18867-18879.

Griffin, J.E., Gawronski, J.D., DeJesus, M.A., Ioerger, T.R., Akerley, B.J. and Sassetti, C.M. (2011) High-resolution phenotypic profiling defines genes essential for mycobacterial growth and cholesterol catabolism. *PLoS pathogens* [online], *7*(9), p.e1002251, doi: https://doi.org/10.1371/journal.ppat.1002251.

Guerin, M.E., Kaur, D., Somashekar, B.S., Gibbs, S., Gest, P., Chatterjee, D., Brennan, P.J. and Jackson, M. (2009) New insights into the early steps of phosphatidylinositol mannoside biosynthesis in mycobacteria: PimB′ is an essential enzyme of *Mycobacterium smegmatis*. *Journal of Biological Chemistry*, *284*(38), pp.25687-25696.

Guilhot, C., Otal, I., Van Rompaey, I., Martin, C. and Gicquel, B. (1994) Efficient transposition in mycobacteria: construction of *Mycobacterium smegmatis* insertional mutant libraries. *Journal of bacteriology*, *176*(2), pp.535-539.

Gupta, R., Barkan, D., Redelman-Sidi, G., Shuman, S. and Glickman, M.S. (2011) Mycobacteria exploit three genetically distinct DNA double-strand break repair pathways. *Molecular microbiology*, *79*(2), pp.316-330.

Gupta, R., Lavollay, M., Mainardi, J.L., Arthur, M., Bishai, W.R. and Lamichhane, G. (2010) The *Mycobacterium tuberculosis* protein LdtMt2 is a nonclassical transpeptidase required for virulence and resistance to amoxicillin. *Nature medicine*, *16*(4), pp.466-469.

Harrison, J., Lloyd, G., Joe, M., Lowary, T.L., Reynolds, E., Walters-Morgan, H., Bhatt, A., Lovering, A., Besra, G.S. and Alderwick, L.J. (2016) Lcp1 is a phosphotransferase responsible for ligating arabinogalactan to peptidoglycan in *Mycobacterium tuberculosis*. *MBio* [online], *7*(4), doi: https://doi.org/10.1128/mBio.00972-16.

Hashish, E., Merwad, A., Elgaml, S., Amer, A., Kamal, H., Elsadek, A., Marei, A. and Sitohy, M. (2018) *Mycobacterium marinum* infection in fish and man: epidemiology, pathophysiology and management; a review. *Veterinary Quarterly*, *38*(1), pp.35-46.

Helena-Bueno, K., Rybak, M.Y., Ekemezie, C.L., Sullivan, R., Brown, C.R., Dingwall, C., Baslé, A., Schneider, C., Connolly, J.P., Blaza, J.N. and Csörgő, B. (2024) A new family of bacterial ribosome hibernation factors. *Nature*, pp.1-8.

Hensel, M., Shea, J.E., Gleeson, C., Jones, M.D., Dalton, E. and Holden, D.W. (1995) Simultaneous identification of bacterial virulence genes by negative selection. *Science*, *269*(5222), pp.400-403.

Heo, J., Cho, K., Kim, U., Cho, D.H., Ko, S., Tran, Q.G., Lee, Y.J., Ryu, C.M. and Kim, H.S. (2020) Genome-wide high-throughput screening of interactive bacterial metabolite in the algal population using *Escherichia coli* K-12 Keio collection. *Scientific Reports* [online], *10*(1), doi: https://doi.org/10.1038/s41598-020-67322-w.

Hermans, J., Boschloo, J.G. and De Bont, J.A.M. (1990) Transformation of *Mycobacterium aurum* by electroporation: the use of glycine, lysozyme and isonicotinic acid hydrazide in enhancing transformation efficiency. *FEMS Microbiology letters*, *72*(1-2), pp.221-224.

Hett, E.C., Chao, M.C., Deng, L.L. and Rubin, E.J. (2008) A mycobacterial enzyme essential for cell division synergizes with resuscitation-promoting factor. *PLoS pathogens* [online], *4*(2), doi: https://doi.org/10.1371/journal.ppat.1000001.

Hlavsa, M.C., Moonan, P.K., Cowan, L.S., Navin, T.R., Kammerer, J.S., Morlock, G.P., Crawford, J.T. and LoBue, P.A. (2008) Human tuberculosis due to *Mycobacterium bovis* in the United States, 1995-2005. *Clinical Infectious Diseases*, *47*(2), pp.168-175.

Hoefsloot, W., Van Ingen, J., Andrejak, C., Ängeby, K., Bauriaud, R., Bemer, P., Beylis, N., Boeree, M.J., Cacho, J., Chihota, V. and Chimara, E. (2013) The geographic diversity of nontuberculous mycobacteria isolated from pulmonary samples: an NTM-NET collaborative study. *European Respiratory Journal*, *42*(6), pp.1604-1613.

Husson, R.N., James, B.E. and Young, R. (1990) Gene replacement and expression of foreign DNA in mycobacteria. *Journal of bacteriology*, *172*(2), pp.519-524.

Ismail, Z.S. (2016) Characterization of DD-Carboxypeptidase function in mycobacteria: Genetic knockout and recombinant protein production (Doctoral dissertation: University of Witwatersrand).

Jackson, M., Raynaud, C., Lanéelle, M.A., Guilhot, C., Laurent-Winter, C., Ensergueix, D., Gicquel, B. and Daffé, M. (1999) Inactivation of the antigen 85C gene profoundly affects the mycolate content and alters the permeability of the *Mycobacterium tuberculosis* cell envelope. *Molecular microbiology*, *31*(5), pp.1573-1587.

Jacobs, W.R., Tuckman, M. and Bloom, B.R. (1987) Introduction of foreign DNA into mycobacteria using a shuttle phasmid. *Nature*, *327*(6122), pp.532-535.

Jankute, M., Alderwick, L.J., Noack, S., Veerapen, N., Nigou, J. and Besra, G.S. (2017) Disruption of mycobacterial AftB results in complete loss of terminal β (1→ 2) arabinofuranose residues of lipoarabinomannan. *ACS chemical biology*, *12*(1), pp.183-190.

Jin, Y., Xin, Y., Zhang, W. and Ma, Y. (2010) *Mycobacterium tuberculosis* Rv1302 and *Mycobacterium smegmatis* MSMEG_4947 have WecA function and MSMEG_4947 is required for the growth of *M. smegmatis*. *FEMS microbiology letters*, *310*(1), pp.54-61.

Jinek, M., Chylinski, K., Fonfara, I., Hauer, M., Doudna, J.A. and Charpentier, E. (2012) A programmable dual-RNA–guided DNA endonuclease in adaptive bacterial immunity. *science*, *337*(6096), pp.816-821.

Jinek, M., Jiang, F., Taylor, D.W., Sternberg, S.H., Kaya, E., Ma, E., Anders, C., Hauer, M., Zhou, K., Lin, S. and Kaplan, M. (2014) Structures of Cas9 endonucleases reveal RNA-mediated conformational activation. *Science* [online], *343*(6176), doi: 10.1126/science.1247997.

Johansen, M.D., Herrmann, J.L. and Kremer, L. (2020) Non-tuberculous mycobacteria and the rise of *Mycobacterium abscessus*. *Nature Reviews Microbiology*, *18*(7), pp.392-407.

Johnes Information Centre (2018) Assessment of surveillance and control of Johne's disease in farm animals in GB [online]. Available at: https://johnes.org/wp-content/uploads/2018/11/Scottish_Report_JD-245-pages_2002.pdf
[Accessed 30th April 2023]

Johnson, E.O., LaVerriere, E., Office, E., Stanley, M., Meyer, E., Kawate, T., Gomez, J.E., Audette, R.E., Bandyopadhyay, N., Betancourt, N. and Delano, K. (2019) Large-scale chemical–genetics yields new *M. tuberculosis* inhibitor classes. *Nature*, *571*(7763), pp.72-78.

Judd, J.A., Canestrari, J., Clark, R., Joseph, A., Lapierre, P., Lasek-Nesselquist, E., Mir, M., Palumbo, M., Smith, C., Stone, M. and Upadhyay, A. (2021) A mycobacterial systems resource for the research community. *MBio* [online], *12*(2), p.e02401-20, doi: https://doi.org/10.1128/mbio.02401-20.

Kalpana, G.V., Bloom, B.R. and Jacobs Jr, W.R. (1991) Insertional mutagenesis and illegitimate recombination in mycobacteria. *Proceedings of the National Academy of Sciences*, *88*(12), pp.5433-5437.

Kalscheuer, R., Weinrick, B., Veeraraghavan, U., Besra, G.S. and Jacobs, W.R. (2010) Trehalose-recycling ABC transporter LpqY-SugA-SugB-SugC is essential for virulence of *Mycobacterium tuberculosis*. *Proceedings of the National Academy of Sciences*, *107*(50), pp.21761-21766.

Kandasamy, S. and Narayanan, S. (2015) Phenotypic characterization of a novel double knockout PknI/DacB2 from *Mycobacterium tuberculosis*. *Microbiological Research*, *170*, pp.255-262.

Kanipe, C. and Palmer, M.V. (2020) *Mycobacterium bovis* and you: a comprehensive look at the bacteria, its similarities to *Mycobacterium tuberculosis*, and its relationship with human disease. *Tuberculosis* [online], *125*, p.102006, doi: https://doi.org/10.1016/j.tube.2020.102006.

Karakousis, P.C., Bishai, W.R. and Dorman, S.E. (2004) *Mycobacterium tuberculosis* cell envelope lipids and the host immune response. *Cellular microbiology*, *6*(2), pp.105-116.

Kaur, D., Obregón-Henao, A., Pham, H., Chatterjee, D., Brennan, P.J. and Jackson, M. (2008) Lipoarabinomannan of *Mycobacterium*: mannose capping by a multifunctional terminal mannosyltransferase. *Proceedings of the National Academy of Sciences*, *105*(46), pp.17973-17977.

Kieser, K.J., Baranowski, C., Chao, M.C., Long, J.E., Sassetti, C.M., Waldor, M.K., Sacchettini, J.C., Ioerger, T.R. and Rubin, E.J. (2015) Peptidoglycan synthesis in *Mycobacterium tuberculosis* is organized into networks with varying drug susceptibility. *Proceedings of the National Academy of Sciences*, *112*(42), pp.13087-13092.

Kim, S.Y., Kim, C.K., Bae, I.K., Jeong, S.H., Yim, J.J., Jung, J.Y., Park, M.S., Kim, Y.S., Kim, S.K., Chang, J. and Kang, Y.A. (2015) The drug susceptibility profile and inducible resistance to macrolides of *Mycobacterium abscessus* and *Mycobacterium massiliense* in Korea. *Diagnostic microbiology and infectious disease*, *81*(2), pp.107-111.

Koliwer-Brandl, H., Knobloch, P., Barisch, C., Welin, A., Hanna, N., Soldati, T. and Hilbi, H. (2019) Distinct *Mycobacterium marinum* phosphatases determine pathogen vacuole phosphoinositide pattern, phagosome maturation, and escape to the cytosol. *Cellular microbiology* [online], *21*(6), doi: https://doi.org/10.1111/cmi.13008.

Koo, B.M., Kritikos, G., Farelli, J.D., Todor, H., Tong, K., Kimsey, H., Wapinski, I., Galardini, M., Cabal, A., Peters, J.M. and Hachmann, A.B. (2017) Construction and analysis of two genome-scale deletion libraries for *Bacillus subtilis*. *Cell systems*, *4*(3), pp.291-305.

Kleckner, N., Roth, J. and Botstein, D. (1977) Genetic engineering in vivo using translocatable drug-resistance elements: new methods in bacterial genetics. *Journal of molecular biology*, *116*(1), pp.125-159.

Kremer, L., Nampoothiri, K.M., Lesjean, S., Dover, L.G., Graham, S., Betts, J., Brennan, P.J., Minnikin, D.E., Locht, C. and Besra, G.S. (2001) Biochemical Characterization of Acyl Carrier Protein (AcpM) and Malonyl-CoA: AcpM Transacylase (mtFabD), Two Major Components of *Mycobacterium tuberculosis* Fatty Acid Synthase II. *Journal of Biological Chemistry*, *276*(30), pp.27967-27974.

Kritikos, G., Banzhaf, M., Herrera-Dominguez, L., Koumoutsi, A., Wartel, M., Zietek, M. and Typas, A. (2017) A tool named Iris for versatile high-throughput phenotyping in microorganisms. *Nature microbiology*, *2*(5), pp.1-10.

Kumar, G., Galanis, C., Batchelder, H.R., Townsend, C.A. and Lamichhane, G. (2022) Penicillin binding proteins and β-lactamases of *Mycobacterium tuberculosis*: reexamination of the historical paradigm. *MSphere* [online], *7*(1), pp.e00039-22, doi: https://doi.org/10.1128/msphere.00039-22.

Kumar, P., Arora, K., Lloyd, J.R., Lee, I.Y., Nair, V., Fischer, E., Boshoff, H.I. and Barry III, C.E. (2012) Meropenem inhibits D, D-carboxypeptidase activity in *Mycobacterium tuberculosis*. *Molecular microbiology*, *86*(2), pp.367-38.

Lane, J.M. and Rubin, E.J. (2006) Scaling down: a PCR-based method to efficiently screen for desired knockouts in a hig density *Mycobacterium tuberculosis* picked mutant library. *Tuberculosis*, *86*(3), pp.310-313.

Langridge, G.C., Phan, M.D., Turner, D.J., Perkins, T.T., Parts, L., Haase, J., Charles, I., Maskell, D.J., Peters, S.E., Dougan, G. and Wain, J. (2009) Simultaneous assay of every *Salmonella Typhi* gene using one million transposon mutants. *Genome research*, *19*(12), pp.2308-2316.

Lampe, D.J., Akerley, B.J., Rubin, E.J., Mekalanos, J.J. and Robertson, H.M. (1999) Hyperactive transposase mutants of the Himar 1 mariner transposon. *Proceedings of the National Academy of Sciences*, *96*(20), pp.11428-11433.

Lavollay, M., Arthur, M., Fourgeaud, M., Dubost, L., Marie, A., Veziris, N., Blanot, D., Gutmann, L. and Mainardi, J.L. (2008) The peptidoglycan of stationary-phase *Mycobacterium tuberculosis* predominantly contains cross-links generated by L, D-transpeptidation. *Journal of bacteriology*, *190*(12), pp.4360-4366.

Lechartier, B., Hartkoorn, R.C. and Cole, S.T. (2012) In vitro combination studies of benzothiazinone lead compound BTZ043 against *Mycobacterium tuberculosis*. *Antimicrobial agents and chemotherapy*, *56*(11), pp.5790-5793.

Lee, M.H., Pascopella, L., Jacobs Jr, W.R. and Hatfull, G.F. (1991) Site-specific integration of mycobacteriophage L5: integration-proficient vectors for *Mycobacterium smegmatis, Mycobacterium tuberculosis*, and bacille Calmette-Guérin. *Proceedings of the National Academy of Sciences*, *88*(8), pp.3111-3115.

Lefrancois, L.H., Nitschke, J., Panis, G., Prados, J., Butler, R.E., Mendum, T.A., Hanna, N., Stewart, G.R. and Soldati, T. (2023) Temporal genome-wide fitness analysis of *Mycobacterium marinum* during infection reveals genetic requirement for virulence and survival in amoebae and microglial cells. *bioRxiv* [online], doi: https://doi.org/10.1101/2023.03.22.533734.

Lemassu, A. and Daffé, M. (1994) Structural features of the exocellular polysaccharides of *Mycobacterium tuberculosis*. *Biochemical Journal*, *297*(2), pp.351-357.

Lewis, A.H. and Falkinham, J.O. (2015) Microaerobic growth and anaerobic survival of *Mycobacterium avium*, *Mycobacterium intracellulare* and *Mycobacterium scrofulaceum*. *International Journal of Mycobacteriology*, *4*(1), pp.25-30.

Li, F., Feng, L., Jin, C., Wu, X., Fan, L., Xiong, S. and Dong, Y. (2018) LpqT improves mycobacteria survival in macrophages by inhibiting TLR2 mediated inflammatory cytokine expression and cell apoptosis. *Tuberculosis*, *111*, pp.57-66.

Li, S., Poulton, N.C., Chang, J.S., Azadian, Z.A., DeJesus, M.A., Ruecker, N., Zimmerman, M.D., Eckartt, K.A., Bosch, B., Engelhart, C.A. and Sullivan, D.F. (2022) CRISPRi chemical genetics and comparative genomics identify genes mediating drug potency in *Mycobacterium tuberculosis*. *Nature Microbiology*, *7*(6), pp.766-779.

Li, Y.Y., Liu, H.M., Wang, D., Lu, Y., Ding, C., Zhou, L.S., Wu, X.Y., Zhou, Z.W., Xu, S.Q., Lin, C. and Qin, L.H. (2022) Arabinogalactan enhances *Mycobacterium marinum* virulence by suppressing host innate immune responses. *Frontiers in Immunology*, *13*, p.879775, doi: https://doi.org/10.3389/fimmu.2022.879775.

Lin, T.Y., Lo, Y.H., Tseng, P.W., Chang, S.F., Lin, Y.T. and Chen, T.S. (2012) A T3 and T7 recombinant phage acquires efficient adsorption and a broader host range. *PloS one* [online], *7*(2), p.e30954, doi: https://doi.org/10.1371/journal.pone.0030954.

Lipworth, S., Hough, N., Weston, N., Muller-Pebody, B., Phin, N., Myers, R., Chapman, S., Flight, W., Alexander, E., Smith, E.G. and Robinson, E. (2021) Epidemiology of *Mycobacterium abscessus* in England: an observational study. *The Lancet Microbe* [online], *2*(10), pp.e498-e507, doi: https://doi.org/10.1016/S2666-5247(21)00128-2.

Liu, J., Tran, V., Leung, A.S., Alexander, D.C. and Zhu, B. (2009) BCG vaccines: their mechanisms of attenuation and impact on safety and protective efficacy. *Human vaccines*, *5*(2), pp.70-78.

Mainardi, J.L., Fourgeaud, M., Hugonnet, J.E., Dubost, L., Brouard, J.P., Ouazzani, J., Rice, L.B., Gutmann, L. and Arthur, M. (2005) A novel peptidoglycan cross-linking enzyme for a β-lactam-resistant transpeptidation pathway. *Journal of Biological Chemistry*, *280*(46), pp.38146-38152.

Majumdar, G., Mbau, R., Singh, V., Warner, D.F., Dragset, M.S. and Mukherjee, R. (2017) 'Genome-wide transposon mutagenesis in *Mycobacterium tuberculosis* and *Mycobacterium smegmatis*' in *In Vitro Mutagenesis: Methods and Protocols,* New York: Humana Press, pp.321-335.

Makarova, K.S., Aravind, L. and Koonin, E.V. (2002) SWIM, a novel Zn-chelating domain present in bacteria, archaea and eukaryotes. *Trends in biochemical sciences*, *27*(8), pp.384-386.

Malaga, W., Perez, E. and Guilhot, C. (2003) Production of unmarked mutations in mycobacteria using site-specific recombination. *FEMS microbiology letters*, *219*(2), pp.261-268.

Manjunatha, U.H., Boshoff, H., Dowd, C.S., Zhang, L., Albert, T.J., Norton, J.E., Daniels, L., Dick, T., Pang, S.S. and Barry III, C.E. (2006) Identification of a nitroimidazo-oxazine-specific protein involved in PA-824 resistance in *Mycobacterium tuberculosis*. *Proceedings of the National Academy of Sciences*, *103*(2), pp.431-436.

Marquardt, J.L., Brown, E.D., Walsh, C.T. and Anderson, K.S. (1993) Isolation and structural elucidation of a tetrahedral intermediate in the UDP-N-acetylglucosamine enolpyruvoyl transferase enzymic pathway. *Journal of the American Chemical Society*, *115*(22), pp.10398-10399.

Marrakchi, H., Ducasse, S., Labesse, G., Montrozier, H., Margeat, E., Emorine, L., Charpentier, X., Daffé, M. and Quémard, A. (2002) MabA (FabG1), a *Mycobacterium tuberculosis* protein involved in the long-chain fatty acid elongation system FAS-II. *Microbiology*, *148*(4), pp.951-960.

Marrakchi, H., Lanéelle, M.A. and Daffé, M. (2014) Mycolic acids: structures, biosynthesis, and beyond. *Chemistry & biology*, *21*(1), pp.67-85.

Martin, C., Timm, J., Rauzier, J., Gomez-Lus, R., Davies, J. and Gicquel, B. (1990) Transposition of an antibiotic resistance element in mycobacteria. *Nature*, *345*(6277), pp.739-743.

Mazurkiewicz, P., Tang, C.M., Boone, C. and Holden, D.W. (2006) Signature-tagged mutagenesis: barcoding mutants for genome-wide screens. *Nature Reviews Genetics*, *7*(12), pp.929-939.

McAdam, R.A., Weisbrod, T.R., Martin, J., Scuderi, J.D., Brown, A.M., Cirillo, J.D., Bloom, B.R. and Jacobs Jr, W.R. (1995) In vivo growth characteristics of leucine and methionine auxotrophic mutants of *Mycobacterium bovis* BCG generated by transposon mutagenesis. *Infection and immunity*, *63*(3), pp.1004-1012.

McClintock, B. (1950) The origin and behavior of mutable loci in maize. *Proceedings of the National Academy of Sciences*, *36*(6), pp.344-355.

McCready, A.R., Paczkowski, J.E., Cong, J.P. and Bassler, B.L. (2019) An autoinducer-independent RhlR quorum-sensing receptor enables analysis of RhlR regulation. *PLoS pathogens* [online], *15*(6), p.e1007820, doi: https://doi.org/10.1371/journal.ppat.1007820.

McDonough, J.A., Hacker, K.E., Flores, A.R., Pavelka Jr, M.S. and Braunstein, M. (2005) The twin-arginine translocation pathway of *Mycobacterium smegmatis* is functional and required for the export of mycobacterial β-lactamases. *Journal of bacteriology*, *187*(22), pp.7667-7679.

McNeil, M.B., Keighley, L.M., Cook, J.R., Cheung, C.Y. and Cook, G.M. (2021) CRISPR interference identifies vulnerable cellular pathways with bactericidal phenotypes in *Mycobacterium tuberculosis*. *Molecular microbiology*, *116*(4), pp.1033-1043.

McNeil, M., Daffe, M. and Brennan, P.J. (1991) Location of the mycolyl ester substituents in the cell walls of Mycobacteria. *Journal of Biological Chemistry*, *266*(20), pp.13217-13223.

McNeil, M.R., Robuck, K.G., Harter, M. and Brennan, P.J. (1994) Enzymatic evidence for the presence of a critical terminal hexa-arabinoside in the cell walls of *Mycobacterium tuberculosis*. *Glycobiology*, *4*(2), pp.165-174.

Mediavilla, J., Jain, S., Kriakov, J., Ford, M.E., Duda, R.L., Jacobs Jr, W.R., Hendrix, R.W. and Hatfull, G.F. (2000) Genome organization and characterization of mycobacteriophage Bxb1. *Molecular microbiology*, *38*(5), pp.955-970.

Meijers, A.S., Troost, R., Ummels, R., Maaskant, J., Speer, A., Nejentsev, S., Bitter, W. and Kuijl, C.P. (2020) Efficient genome editing in pathogenic mycobacteria using *Streptococcus thermophilus* CRISPR1-Cas9. *Tuberculosis* [online], *124*, doi: https://doi.org/10.1016/j.tube.2020.101983.

Melief, E., Kokoczka, R., Files, M., Bailey, M.A., Alling, T., Li, H., Ahn, J., Misquith, A., Korkegian, A., Roberts, D. and Sacchettini, J. (2018) Construction of an overexpression library for *Mycobacterium tuberculosis*. *Biology Methods and Protocols* [online], *3*(1), doi: https://doi.org/10.1093/biomethods/bpy009.

Melly, G.C., Stokas, H., Dunaj, J.L., Hsu, F.F., Rajavel, M., Su, C.C., Edward, W.Y. and Purdy, G.E. (2019) Structural and functional evidence that lipoprotein LpqN supports cell envelope biogenesis in *Mycobacterium tuberculosis*. *Journal of Biological Chemistry*, *294*(43), pp.15711-15723.

Mengin-Lecreulx, D., Texier, L., Rousseau, M. and van Heijenoort, J. (1991) The murG gene of *Escherichia coli* codes for the UDP-N-acetylglucosamine: N-acetylmuramyl-(pentapeptide) pyrophosphoryl-undecaprenol N-acetylglucosamine transferase involved in the membrane steps of peptidoglycan synthesis. *Journal of Bacteriology*, *173*(15), pp.4625-4636.

Mills, J.A., Motichka, K., Jucker, M., Wu, H.P., Uhlik, B.C., Stern, R.J., Scherman, M.S., Vissa, V.D., Pan, F., Kundu, M. and Ma, Y.F. (2004) Inactivation of the mycobacterial rhamnosyltransferase, which is needed for the formation of the arabinogalactan-peptidoglycan linker, leads to irreversible loss of viability. *Journal of Biological Chemistry*, *279*(42), pp.43540-43546.

Milstien, J.B. and Gibson, J.J. (1990) Quality control of BCG vaccine by WHO: a review of factors that may influence vaccine effectiveness and safety. *Bulletin of the World Health organization*, *68*(1), pp.93-108.

Minnikin, D.E. and Brennan, P.J. (2020) Lipids of clinically significant mycobacteria. *Health Consequences of Microbial Interactions with Hydrocarbons, Oils, and Lipids*, pp.33-108.

Mirsaeidi, M., Farnia, P., Sadikot, R., Hsueh, P.R. and Aliberti, S. (2015) Nontuberculous Mycobacteria: epidemiologic, mycobacteriologic, and clinical aspects. *BioMed Research International* [online]. doi: https://doi.org/10.1155/2015/523697.

Mishra, A.K., Alderwick, L.J., Rittmann, D., Wang, C., Bhatt, A., Jacobs Jr, W.R., Takayama, K., Eggeling, L. and Besra, G.S. (2008) Identification of a novel α (1→ 6) mannopyranosyltransferase MptB from *Corynebacterium glutamicum* by deletion of a conserved gene, NCgl1505, affords a lipomannan-and lipoarabinomannan-deficient mutant. *Molecular microbiology*, *68*(6), pp.1595-1613.

Modlin, S.J., Elghraoui, A., Gunasekaran, D., Zlotnicki, A.M., Dillon, N.A., Dhillon, N., Kuo, N., Robinhold, C., Chan, C.K., Baughn, A.D. and Valafar, F. (2021) Structure-aware *Mycobacterium tuberculosis* functional annotation uncloaks resistance, metabolic, and virulence genes. *MSystems* [online]. *6*(6), doi: 10.1128/mSystems.00673-21.

Mojica, F.J., Díez-Villaseñor, C., García-Martínez, J. and Almendros, C. (2009) Short motif sequences determine the targets of the prokaryotic CRISPR defence system. *Microbiology*, *155*(3), pp.733-740.

Moule, M.G. and Cirillo, J.D. (2020) Mycobacterium tuberculosis dissemination plays a critical role in pathogenesis. *Frontiers in cellular and infection microbiology* [online], *10*, p.65, doi: https://doi.org/10.3389/fcimb.2020.00065

Moynihan, P.J., Cadby, I.T., Veerapen, N., Jankute, M., Crosatti, M., Mukamolova, G.V., Lovering, A.L. and Besra, G.S. (2019) The hydrolase LpqI primes mycobacterial peptidoglycan recycling. *Nature communications* [online], *10*, p.2647, doi: https://doi.org/10.1038/s41467-019-10586-2.

Murphy, K.C. (1998) Use of bacteriophage λ recombination functions to promote gene replacement in *Escherichia coli*. *Journal of bacteriology*, *180*(8), pp.2063-2071.

Murphy, K.C., Nelson, S.J., Nambi, S., Papavinasasundaram, K., Baer, C.E. and Sassetti, C.M. (2018) ORBIT: a new paradigm for genetic engineering of mycobacterial chromosomes. *MBio* [online]. *9*(6), doi: https://doi.org/10.1128/mBio.01467-18.

Murry, J.P., Pandey, A.K., Sassetti, C.M. and Rubin, E.J. (2009) Phthiocerol dimycocerosate transport is required for resisting interferon-γ–independent immunity. *The Journal of infectious diseases*, *200*(5), pp.774-782.

Musayev, F., Sachdeva, S., Scarsdale, J.N., Reynolds, K.A. and Wright, H.T. (2005) Crystal structure of a substrate complex of *Mycobacterium tuberculosis* β-ketoacyl-acyl carrier protein synthase III (FabH) with lauroyl-coenzyme A. *Journal of molecular biology*, *346*(5), pp.1313-1321.

Mutalik, V.K., Novichkov, P.S., Price, M.N., Owens, T.K., Callaghan, M., Carim, S., Deutschbauer, A.M. and Arkin, A.P. (2019) Dual-barcoded shotgun expression library sequencing for high-throughput characterization of functional traits in bacteria. *Nature communications* [online], *10*(1), p.308, doi: https://doi.org/10.1038/s41467-018-08177-8.

Nambi, S., Long, J.E., Mishra, B.B., Baker, R., Murphy, K.C., Olive, A.J., Nguyen, H.P., Shaffer, S.A. and Sassetti, C.M. (2015) The oxidative stress network of *Mycobacterium tuberculosis* reveals coordination between radical detoxification systems. *Cell host & microbe*, *17*(6), pp.829-837.

National Statistics (2021) Quarterly TB in cattle in Great Britain statistics notice: December 2021 [online]. Available at: https://www.gov.uk/government/statistics/incidence-of-tuberculosis-tb-in-cattle-in-great-britain/quarterly-tb-in-cattle-in-great-britain-statistics-notice-december-2021
[Accessed 22nd May 2021]

National Institutes of Health (2023) Genome List [online]
Available at: https://www.ncbi.nlm.nih.gov/genome/browse/#!/overview/
[Accessed 23rd June 2023]

Newton-Foot, M. and van Pittius, N.C.G. (2013) The complex architecture of mycobacterial promoters. *Tuberculosis*, *93*(1), pp.60-74.

Nichols, R.J., Sen, S., Choo, Y.J., Beltrao, P., Zietek, M., Chaba, R., Lee, S., Kazmierczak, K.M., Lee, K.J., Wong, A. and Shales, M. (2011) Phenotypic landscape of a bacterial cell. *Cell*, *144*(1), pp.143-156.

Ortalo-Magne, A., Lemassu, A., Laneelle, M.A., Bardou, F., Silve, G., Gounon, P., Marchal, G. and Daffé, M. (1996) Identification of the surface-exposed lipids on the cell envelopes of *Mycobacterium tuberculosis* and other mycobacterial species. *Journal of bacteriology*, *178*(2), pp.456-461.

Pandey, S.D., Pal, S., Kumar N, G., Bansal, A., Mallick, S. and Ghosh, A.S. (2018) Two DD-carboxypeptidases from *Mycobacterium smegmatis* affect cell surface properties through regulation of peptidoglycan cross-linking and glycopeptidolipids. *Journal of Bacteriology* [online], *200*(14), p.10.1128, https://doi.org/10.1128/jb.00760-17.

Paradis-Bleau, C., Kritikos, G., Orlova, K., Typas, A. and Bernhardt, T.G. (2014) A genome-wide screen for bacterial envelope biogenesis mutants identifies a novel factor involved in cell wall precursor metabolism. *PLoS genetics* [online], *10*(1), p.e1004056, doi: https://doi.org/10.1371/journal.pgen.1004056.

Parish, T., Mahenthiralingam, E., Draper, P., Davis, E.O. and Colston, E.O. (1997) Regulation of the inducible acetamidase gene of *Mycobacterium smegmatis*. *Microbiology*, *143*(7), pp.2267-2276.

Parte, A.C., Sardà Carbasse, J., Meier-Kolthoff, J.P., Reimer, L.C. and Göker, M. (2020) List of Prokaryotic names with Standing in Nomenclature (LPSN) moves to the DSMZ. *International Journal of Systematic and Evolutionary Microbiology* [online], *70*(11), pp.5607-5612, doi: https://doi.org/10.1093/nar/gkt1111

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

Pelicic, V., Reyrat, J.M. and Gicquel, B. (1996) Generation of unmarked directed mutations in mycobacteria, using sucrose counter-selectable suicide vectors. *Molecular microbiology*, *20*(5), pp.919-925.

Penn, B.H., Netter, Z., Johnson, J.R., Von Dollen, J., Jang, G.M., Johnson, T., Ohol, Y.M., Maher, C., Bell, S.L., Geiger, K. and Golovkine, G. (2018) An Mtb-human protein-protein interaction map identifies a switch between host antiviral and antibacterial responses. *Molecular cell*, *71*(4), pp.637-648.

Peters, J.M., Colavin, A., Shi, H., Czarny, T.L., Larson, M.H., Wong, S., Hawkins, J.S., Lu, C.H., Koo, B.M., Marta, E. and Shiver, A.L. (2016) A comprehensive, CRISPR-based functional analysis of essential genes in bacteria. *Cell*, *165*(6), pp.1493-1506.

Peterson, D.O. and Bloch, K. (1977) *Mycobacterium smegmatis* fatty acid synthetase. Long chain transacylase chain length specificity. *Journal of Biological Chemistry*, *252*(16), pp.5735-5739.

Portevin, D., de Sousa-D'Auria, C., Montrozier, H., Houssin, C., Stella, A., Lanéelle, M.A., Bardou, F., Guilhot, C. and Daffé, M. (2005) The acyl-AMP ligase FadD32 and AccD4-containing acyl-CoA carboxylase are required for the synthesis of mycolic acids and essential for mycobacterial growth: identification of the carboxylation product and determination of the acyl-CoA carboxylase components. *Journal of Biological Chemistry*, *280*(10), pp.8862-8874.

Porwollik, S., Santiviago, C.A., Cheng, P., Long, F., Desai, P., Fredlund, J., Srikumar, S., Silva, C.A., Chu, W., Chen, X. and Canals, R. (2014) Defined single-gene and multi-gene deletion mutant collections in Salmonella enterica sv Typhimurium. *PloS one* [online], *9*(7), p.e99820, doi: https://doi.org/10.1371/journal.pone.0099820.

Power, D.A. and Hanks, J.H. (1965) The effect of organic acids, serum albumin, and wetting agents on lag phase, dispersed growth, and pH stabilization in mycobacterial cultures. *American Review of Respiratory Disease*, *92*(1), pp.83-93.

Price, M.N., Wetmore, K.M., Waters, R.J., Callaghan, M., Ray, J., Liu, H., Kuehl, J.V., Melnyk, R.A., Lamson, J.S., Suh, Y. and Carlson, H.K. (2018) Mutant phenotypes for thousands of bacterial genes of unknown function. *Nature*, *557*(7706), pp.503-509.

Próchnicki, T., Mangan, M.S. and Latz, E. (2016) Recent insights into the molecular mechanisms of the NLRP3 inflammasome activation. *F1000Research* [online], *5*, doi: 10.12688/f1000research.8614.1.

Prouty, M.G., Correa, N.E., Barker, L.P., Jagadeeswaran, P. and Klose, K.E. (2003) Zebrafish-*Mycobacterium marinum* model for mycobacterial pathogenesis. *FEMS Microbiology Letters*, *225*(2), pp.177-182.

Qi, L.S., Larson, M.H., Gilbert, L.A., Doudna, J.A., Weissman, J.S., Arkin, A.P. and Lim, W.A. (2013) Repurposing CRISPR as an RNA-guided platform for sequence-specific control of gene expression. *Cell*, *152*(5), pp.1173-1183.

Queiroz, A., Medina-Cleghorn, D., Marjanovic, O., Nomura, D.K. and Riley, L.W. (2015) Comparative metabolic profiling of mce1 operon mutant vs wild-type *Mycobacterium tuberculosis* strains. *Pathogens and disease* [online], *73*(8), doi: https://doi.org/10.1093/femspd/ftv066

Raghuvanshi, S., Sharma, P., Singh, S., Van Kaer, L. and Das, G. (2010) *Mycobacterium tuberculosis* evades host immunity by recruiting mesenchymal stem cells. *Proceedings of the National Academy of Sciences*, *107*(50), pp.21653-21658.

Rastogi, S., Singh, A.K., Chandra, G., Kushwaha, P., Pant, G., Singh, K., Mitra, K., Sashidhara, K.V. and Krishnan, M.Y. (2017) The diacylglycerol acyltransferase Rv3371 of *Mycobacterium tuberculosis* is required for growth arrest and involved in stress-induced cell wall alterations. *Tuberculosis*, *104*, pp.8-19

Rathnaiah, G., Bannantine, J.P., Bayles, D.O., Zinniel, D.K., Stabel, J.R., Gröhn, Y.T. and Barletta, R.G. (2016) Analysis of *Mycobacterium avium* subsp. *paratuberculosis* mutant libraries reveals loci-dependent transposition biases and strategies for novel mutant discovery. *Microbiology*, *162*(4), pp.633-641.

Ratnatunga, C.N., Lutzky, V.P., Kupz, A., Doolan, D.L., Reid, D.W., Field, M., Bell, S.C., Thomson, R.M. and Miles, J.J. (2020) The rise of non-tuberculosis mycobacterial lung disease. *Frontiers in immunology* [online], *11*(303), doi: https://doi.org/10.3389/fimmu.2020.00303

Raymond, J.B., Mahapatra, S., Crick, D.C. and Pavelka, M.S. (2005) Identification of the namH gene, encoding the hydroxylase responsible for the N-glycolylation of the mycobacterial peptidoglycan. *Journal of Biological Chemistry*, *280*(1), pp.326-333.

Rego, E.H., Audette, R.E. and Rubin, E.J. (2017) Deletion of a mycobacterial divisome factor collapses single-cell phenotypic heterogeneity. *Nature*, *546*(7656), pp.153-157.

Rengarajan, J., Bloom, B.R. and Rubin, E.J. (2005) Genome-wide requirements for *Mycobacterium tuberculosis* adaptation and survival in macrophages. *Proceedings of the National Academy of Sciences*, *102*(23), pp.8327-8332.

Rifat, D., Chen, L., Kreiswirth, B.N. and Nuermberger, E.L. (2021) Genome-wide essentiality analysis of *Mycobacterium abscessus* by saturated transposon mutagenesis and deep sequencing. *Mbio* [online], *12*(3), doi: https://doi.org/10.1128/mbio.01049-21.

Rivera-Marrero, C.A., Ritzenthaler, J.D., Roman, J. and Moremen, K.W. (2001) Molecular cloning and expression of an α-mannosidase gene in *Mycobacterium tuberculosis*. *Microbial pathogenesis*, *30*(1), pp.9-18.

Rock, J.M., Hopkins, F.F., Chavez, A., Diallo, M., Chase, M.R., Gerrick, E.R., Pritchard, J.R., Church, G.M., Rubin, E.J., Sassetti, C.M. and Schnappinger, D. (2017) Programmable transcriptional repression in mycobacteria using an orthogonal CRISPR interference platform. *Nature microbiology*, *2*(4), pp.1-9.

Rominski, A., Selchow, P., Becker, K., Brülle, J.K., Dal Molin, M. and Sander, P. (2017) Elucidation of *Mycobacterium abscessus* aminoglycoside and capreomycin resistance by targeted deletion of three putative resistance genes. *Journal of Antimicrobial Chemotherapy*, *72*(8), pp.2191-2200.

Rose, N.L., Completo, G.C., Lin, S.J., McNeil, M., Palcic, M.M. and Lowary, T.L. (2006) Expression, purification, and characterization of a galactofuranosyltransferase involved in *Mycobacterium tuberculosis* arabinogalactan biosynthesis. *Journal of the American Chemical Society*, *128*(20), pp.6721-6729.

Rubin, E.J., Akerley, B.J., Novik, V.N., Lampe, D.J., Husson, R.N. and Mekalanos, J.J. (1999) In vivo transposition of mariner-based elements in enteric bacteria and mycobacteria. *Proceedings of the National Academy of Sciences*, *96*(4), pp.1645-1650.

Rustad, T.R., Minch, K.J., Ma, S., Winkler, J.K., Hobbs, S., Hickey, M., Brabant, W., Turkarslan, S., Price, N.D., Baliga, N.S. and Sherman, D.R. (2014) Mapping and manipulating the *Mycobacterium tuberculosis* transcriptome using a transcription factor overexpression-derived regulatory network. *Genome biology*, *15*, pp.1-11.

Rybniker, J., Wolke, M., Haefs, C. and Plum, G. (2003) Transposition of Tn 5367 in *Mycobacterium marinum*, using a conditionally recombinant mycobacteriophage. *Journal of bacteriology*, *185*(5), pp.1745-1748.

Sacco, E., Covarrubias, A.S., O'Hare, H.M., Carroll, P., Eynard, N., Jones, T.A., Parish, T., Daffé, M., Bäckbro, K. and Quémard, A. (2007) The missing piece of the type II fatty acid synthase system from *Mycobacterium tuberculosis*. *Proceedings of the National Academy of Sciences*, *104*(37), pp.14628-14633.

Sancho-Vaello, E., Albesa-Jové, D., Rodrigo-Unzueta, A. and Guerin, M.E. (2017) Structural basis of phosphatidyl-myo-inositol mannosides biosynthesis in mycobacteria. *Biochimica et Biophysica Acta (BBA)-Molecular and Cell Biology of Lipids*, *1862*(11), pp.1355-1367.

Sani, M., Houben, E.N., Geurtsen, J., Pierson, J., De Punder, K., van Zon, M., Wever, B., Piersma, S.R., Jiménez, C.R., Daffé, M. and Appelmelk, B.J. (2010) Direct visualization by cryo-EM of the mycobacterial capsular layer: a labile structure containing ESX-1-secreted proteins. *PLoS pathogens* [online], *6*(3), doi: https://doi.org/10.1371/journal.ppat.1000794

Sassetti, C.M. and Rubin, E.J. (2003) Genetic requirements for mycobacterial survival during infection. *Proceedings of the National Academy of Sciences*, *100*(22), pp.12989-12994.

Sassetti, C.M., Boyd, D.H. and Rubin, E.J. (2001) Comprehensive identification of conditionally essential genes in mycobacteria. *Proceedings of the National Academy of Sciences*, *98*(22), pp.12712-12717.

Sassetti, C.M., Boyd, D.H. and Rubin, E.J. (2003) Genes required for mycobacterial growth defined by high density mutagenesis. *Molecular microbiology*, *48*(1), pp.77-84.

Saunders, S.H. and Ahmed, A.M. (2024) ORBIT for *E. coli*: Kilobase-scale oligonucleotide recombineering at high throughput and high efficiency. *Nucleic Acids Research*, p.gkae227, doi: https://doi.org/10.1093/nar/gkae227.

Sauvage, E., Kerff, F., Terrak, M., Ayala, J.A. and Charlier, P. (2008) The penicillin-binding proteins: structure and role in peptidoglycan biosynthesis. *FEMS microbiology reviews*, *32*(2), pp.234-258.

Seidel, M., Alderwick, L.J., Birch, H.L., Sahm, H., Eggeling, L. and Besra, G.S. (2007) Identification of a novel arabinofuranosyltransferase AftB involved in a terminal step of cell wall arabinan biosynthesis in Corynebacterianeae, such as *Corynebacterium glutamicum* and *Mycobacterium tuberculosis*. *Journal of Biological Chemistry*, *282*(20), pp.14729-14740.

Shapiro, J.A. (1979) Molecular model for the transposition and replication of bacteriophage Mu and other transposable elements. *Proceedings of the National Academy of Sciences*, *76*(4), pp.1933-1937.

Sharma, K., Gupta, M., Pathak, M., Gupta, N., Koul, A., Sarangi, S., Baweja, R. and Singh, Y. (2006) Transcriptional control of the mycobacterial embCAB operon by PknH through a regulatory protein, EmbR, in vivo. *Journal of bacteriology*, *188*(8), pp.2936-2944.

Shen, L., Viljoen, A., Villaume, S., Joe, M., Halloum, I., Chêne, L., Méry, A., Fabre, E., Takegawa, K., Lowary, T.L. and Vincent, S.P. (2020) The endogenous galactofuranosidase GlfH1 hydrolyzes mycobacterial arabinogalactan. *Journal of Biological Chemistry*, *295*(15), pp.5110-5123.

Sher, J.W., Lim, H.C. and Bernhardt, T.G. (2020) Global phenotypic profiling identifies a conserved actinobacterial cofactor for a bifunctional PBP-type cell wall synthase. *Elife* [online], *9*, p.e54761, doi: https://doi.org/10.7554/eLife.54761.

Shi, L., Berg, S., Lee, A., Spencer, J.S., Zhang, J., Vissa, V., McNeil, M.R., Khoo, K.H. and Chatterjee, D. (2006) The carboxy terminus of EmbC from *Mycobacterium smegmatis* mediates chain length extension of the arabinan in lipoarabinomannan. *Journal of Biological Chemistry*, *281*(28), pp.19512-19526.

Shiver, A.L., Culver, R., Deutschbauer, A.M. and Huang, K.C. (2021) Rapid ordering of barcoded transposon insertion libraries of anaerobic bacteria. *Nature protocols*, *16*(6), pp.3049-3071.

Shiver, A.L., Osadnik, H., Kritikos, G., Li, B., Krogan, N., Typas, A. and Gross, C.A. (2016) A chemical-genomic screen of neglected antibiotics reveals illicit transport of kasugamycin and blasticidin S. *PLoS genetics* [online], *12*(6), p.e1006124, doi: https://doi.org/10.1371/journal.pgen.1006902.

Siegl, T., Petzke, L., Welle, E. and Luzhetskyy, A. (2010) I-Sce I endonuclease: a new tool for DNA repair studies and genetic manipulations in streptomycetes. *Applied microbiology and biotechnology*, *87*, pp.1525-1532.

Siegrist, M.S. and Rubin, E.J. (2009) 'Phage transposon mutagenesis' in *Mycobacteria Protocols*, 2nd Ed, London: Humana Press, pp.311-323.

Silvis, M.R., Rajendram, M., Shi, H., Osadnik, H., Gray, A.N., Cesar, S., Peters, J.M., Hearne, C.C., Kumar, P., Todor, H. and Huang, K.C. (2021) Morphological and transcriptional responses to CRISPRi knockdown of essential genes in Escherichia coli. *Mbio* [online], *12*(5), doi: https://doi.org/10.1128/mbio.02561-21.

Simner, P.J., Stenger, S., Richter, E., Brown-Elliott, B.A., Wallace Jr, R.J. and Wengenack, N.L. (2015) 'Mycobacterium: laboratory characteristics of slowly growing mycobacteria'. *Manual of clinical microbiology*, In Jorgensen J.H., Pfaller M.A., Carroll K.C., Funke G., Landry M.L., Richter S.S., Warnock D.W. (eds.), *Manual of clinical microbiology*, 11th Ed, Washington, DC: ASM Press, pp.570-594.

Singh, A.K., Carette, X., Potluri, L.P., Sharp, J.D., Xu, R., Prisic, S. and Husson, R.N. (2016) Investigating essential gene function in *Mycobacterium tuberculosis* using an efficient CRISPR interference system. *Nucleic acids research* [online], *44*(18), doi:

https://doi.org/10.1093/nar/gkw625.

Singh, V., Brecik, M., Mukherjee, R., Evans, J.C., Svetlíková, Z., Blaško, J., Surade, S., Blackburn, J., Warner, D.F., Mikušová, K. and Mizrahi, V. (2015) The complex mechanism of antimycobacterial action of 5-fluorouracil. *Chemistry & biology*, *22*(1), pp.63-75.

Škovierová, H., Larrouy-Maumus, G., Zhang, J., Kaur, D., Barilone, N., Korduláková, J., Gilleron, M., Guadagnini, S., Belanová, M., Prevost, M.C. and Gicquel, B. (2009) AftD, a novel essential arabinofuranosyltransferase from mycobacteria. *Glycobiology*, *19*(11), pp.1235-1247.

Smith, L.M., Jackson, S.A., Gardner, P.P. and Fineran, P.C. (2021) SorTn-seq: a high-throughput functional genomics approach to discovering regulators of bacterial gene expression. *Nature Protocols*, *16*(9), pp.4382-4418.

Smith, V., Botstein, D. and Brown, P.O. (1995) Genetic footprinting: a genomic strategy for determining a gene's function given its sequence. *Proceedings of the National Academy of Sciences*, *92*(14), pp.6479-6483.

Snapper, S.B., LuGosI, L.A.S.Z.L., Jekkel, A., Melton, R.E., Kieser, T., Bloom, B.R. and Jacobs Jr, W.R. (1988) Lysogeny and transformation in mycobacteria: stable expression of foreign genes. *Proceedings of the National Academy of Sciences*, *85*(18), pp.6987-6991.

Soroka, D., Ourghanlian, C., Compain, F., Fichini, M., Dubée, V., Mainardi, J.L., Hugonnet, J.E. and Arthur, M. (2017) Inhibition of β-lactamases of mycobacteria by avibactam and clavulanate. *Journal of Antimicrobial Chemotherapy*, *72*(4), pp.1081-1088.

Squeglia, F., Moreira, M., Ruggiero, A. and Berisio, R. (2019) The cell wall hydrolytic NlpC/P60 endopeptidases in mycobacterial cytokinesis: a structural perspective. *Cells*, *8*(6), p.609.

Sreevatsan, S., Pan, X., Stockbauer, K.E., Williams, D.L., Kreiswirth, B.N. and Musser, J.M. (1996) Characterization of rpsL and rrs mutations in streptomycin-resistant *Mycobacterium tuberculosis* isolates from diverse geographic localities. *Antimicrobial agents and chemotherapy*, *40*(4), pp.1024-1026.

Steele, J.H. (1995) 'Regional and country status repor't. In: Thoen, C.O., Steele, J.H. and Gilsdorf, M.J. (eds.) (2006) *Mycobacterium bovis infection in animals and humans.* 2nd Ed. Iowa State University Press, Ames IA: Blackwell Publishing, pp.169-172.

Stephan, J., Stemmer, V. and Niederweis, M. (2004) Consecutive gene deletions in *Mycobacterium smegmatis* using the yeast FLP recombinase. *Gene*, *343*(1), pp.181-190.

Sun, H., Yang, J. and Song, H. (2020) Engineering mycobacteria artificial promoters and ribosomal binding sites for enhanced sterol production. *Biochemical Engineering Journal* [online], *162*, p.107739. doi: https://doi.org/10.1016/j.bej.2020.107739.

Sun, B., Yang, J., Yang, S., Ye, R.D., Chen, D. and Jiang, Y. (2018) A CRISPR-Cpf1-assisted non-homologous end joining genome editing system of *Mycobacterium smegmatis*. *Biotechnology journal* [online], *13*(9), doi: https://doi.org/10.1002/biot.201700588.

Takayama, K., Wang, C. and Besra, G.S. (2005) Pathway to synthesis and processing of mycolic acids in *Mycobacterium tuberculosis*. *Clinical microbiology reviews*, *18*(1), pp.81-101.

Talaat, A.M. and Trucksis, M. (2000) Transformation and transposition of the genome of *Mycobacterium marinum*. *American journal of veterinary research*, *61*(2), pp.125-128.

Thibault, D., Jensen, P.A., Wood, S., Qabar, C., Clark, S., Shainheit, M.G., Isberg, R.R. and van Opijnen, T. (2019) Droplet Tn-Seq combines microfluidics with Tn-Seq for identifying complex single-cell phenotypes. *Nature Communications* [online], *10*(1), doi: https://doi.org/10.1038/s41467-019-13719-9.

Turapov, O., Forti, F., Kadhim, B., Ghisotti, D., Sassine, J., Straatman-Iwanowska, A., Bottrill, A.R., Moynihan, P.J., Wallis, R., Barthe, P. and Cohen-Gonsaud, M. (2018) Two faces of CwlM, an essential PknB substrate, in *Mycobacterium tuberculosis*. *Cell reports*, *25*(1), pp.57-67.

Turnbull, W.B., Shimizu, K.H., Chatterjee, D., Homans, S.W. and Treumann, A. (2004) Identification of the 5-methylthiopentosyl substituent in *Mycobacterium tuberculosis* lipoarabinomannan. *Angewandte Chemie International Edition*, *43*(30), pp.3918-3922.

Typas, A., Banzhaf, M., van Saparoea, B.V.D.B., Verheul, J., Biboy, J., Nichols, R.J., Zietek, M., Beilharz, K., Kannenberg, K., von Rechenberg, M. and Breukink, E. (2010) Regulation of peptidoglycan synthesis by outer-membrane proteins. *Cell*, *143*(7), pp.1097-1109.

Typas, A., Nichols, R.J., Siegele, D.A., Shales, M., Collins, S.R., Lim, B., Braberg, H., Yamamoto, N., Takeuchi, R., Wanner, B.L. and Mori, H. (2008) High-throughput, quantitative analyses of genetic interactions in *E. coli*. *Nature methods*, *5*(9), pp.781-787.

Ustinova, V.V., Smirnova, T.G., Sochivko, D.G., Varlamov, D.A., Larionova, E.E., Andreevskaya, S.N., Andrievskaya, I.Y., Kiseleva, E.A., Chernousova, L.N. and Ergeshov, A. (2019) New assay to diagnose and differentiate between Mycobacterium tuberculosis complex and nontuberculous mycobacteria. *Tuberculosis*, 114, pp.17-23.

Van der Oost, J., Jore, M.M., Westra, E.R., Lundgren, M. and Brouns, S.J. (2009) CRISPR-based adaptive and heritable immunity in prokaryotes. *Trends in biochemical sciences*, *34*(8), pp.401-407.

Van Kessel, J.C. and Hatfull, G.F. (2007) Recombineering in *Mycobacterium tuberculosis*. *Nature methods*, *4*(2), pp.147-152.

Van Kessel, J.C. and Hatfull, G.F. (2008) Efficient point mutagenesis in mycobacteria using single-stranded DNA recombineering: characterization of antimycobacterial drug targets. *Molecular microbiology*, *67*(5), pp.1094-1107.

Van Opijnen, T., Bodi, K.L. and Camilli, A. (2009) Tn-seq: high-throughput parallel sequencing for fitness and genetic interaction studies in microorganisms. *Nature methods*, *6*(10), pp.767-772.

Van Wyk, N., Drancourt, M., Henrissat, B. and Kremer, L. (2017) Current perspectives on the families of glycoside hydrolases of *Mycobacterium tuberculosis*: their importance and prospects for assigning function to unknowns. *Glycobiology*, *27*(2), pp.112-122.

Vandewalle, K., Festjens, N., Plets, E., Vuylsteke, M., Saeys, Y. and Callewaert, N. (2015) Characterization of genome-wide ordered sequence-tagged *Mycobacterium* mutant libraries by Cartesian Pooling-Coordinate Sequencing. *Nature communications* [online], *6*(1), p.7106. doi: https://doi.org/10.1038/ncomms8106.

Vatlin, A.A., Bekker, O.B., Shur, K.V., Ilyasov, R.A., Shatrov, P.A., Maslov, D.A. and Danilenko, V.N. (2023) Kanamycin and Ofloxacin Activate the Intrinsic Resistance to Multiple Antibiotics in *Mycobacterium smegmatis*. *Biology* [online], *12*(4), p.506, doi: https://doi.org/10.3390/biology12040506.

Villeneuve, M., Kawai, M., Kanashima, H., Watanabe, M., Minnikin, D.E. and Nakahara, H. (2005) Temperature dependence of the Langmuir monolayer packing of mycolic acids from *Mycobacterium tuberculosis*. *Biochimica et Biophysica Acta (BBA)-Biomembranes*, *1715*(2), pp.71-80.

Wang, J.Y., Lee, M.C., Shu, C.C., Lee, C.H., Lee, L.N., Chao, K.M. and Chang, F.Y. (2015) Optimal duration of anti-TB treatment in patients with diabetes: nine or six months? *Chest*, *147*(2), pp.520-528.

Wang, R., Preamplume, G., Terns, M.P., Terns, R.M. and Li, H. (2011) Interaction of the Cas6 riboendonuclease with CRISPR RNAs: recognition and cleavage. *Structure*, *19*(2), pp.257-264.

Wang, X., Tang, B., Ye, Y., Mao, Y., Lei, X., Zhao, G. and Ding, X. (2017) Bxb1 integrase serves as a highly efficient DNA recombinase in rapid metabolite pathway assembly. *Acta Biochimica et Biophysica Sinica*, *49*(1), pp.44-50.

Wang, Y., Cui, T., Zhang, C., Yang, M., Huang, Y., Li, W., Zhang, L., Gao, C., He, Y., Li, Y. and Huang, F. (2010) Global protein−protein interaction network in the human pathogen *Mycobacterium tuberculosis* H37Rv. *Journal of proteome research*, *9*(12), pp.6665-6677.

Wei, J.R., Krishnamoorthy, V., Murphy, K., Kim, J.H., Schnappinger, D., Alber, T., Sassetti, C.M., Rhee, K.Y. and Rubin, E.J. (2011) Depletion of antibiotic targets has widely varying effects on growth. *Proceedings of the National Academy of Sciences*, *108*(10), pp.4176-4181.

Wei, W., Zhang, S., Fleming, J., Chen, Y., Li, Z., Fan, S., Liu, Y., Wang, W., Wang, T., Liu, Y. and Ren, B. (2019) *Mycobacterium tuberculosis* type III-A CRISPR/Cas system crRNA and its maturation have atypical features. *The FASEB Journal*, *33*(1), pp.1496-1509

Weld, R.J., Butts, C. and Heinemann, J.A. (2004) Models of phage growth and their applicability to phage therapy. *Journal of Theoretical Biology*, *227*(1), pp.1-11.

Wetmore, K.M., Price, M.N., Waters, R.J., Lamson, J.S., He, J., Hoover, C.A., Blow, M.J., Bristow, J., Butland, G., Arkin, A.P. and Deutschbauer, A. (2015) Rapid quantification of mutant fitness in diverse bacteria by sequencing randomly bar-coded transposons. *MBio* [online], *6*(3), p.e00306-15, doi: 10.1128/mBio.00306-15.

Winkler, K.R., Mizrahi, V., Warner, D.F. and De Wet, T.J. (2023) High-throughput functional genomics: A (myco) bacterial perspective. *Molecular Microbiology*, *120*, pp.141-158.

Wipperman, M.F., Yang, M., Thomas, S.T. and Sampson, N.S. (2013) Shrinking the FadE proteome of *Mycobacterium tuberculosis*: insights into cholesterol metabolism through identification of an α2β2 heterotetrameric acyl coenzyme A dehydrogenase family. *Journal of bacteriology*, *195*(19), pp.4331-4341.

World Health Organisation. (2021) Global tuberculosis report 2021 [online].
Available at: https://www.who.int/publications/i/item/9789240037021
[Accessed 12th May, 2022]

World Health Organisation. (2023) Global tuberculosis report 2023 [online].
Available at: https://www.who.int/teams/global-tuberculosis-programme/tb-reports/global-tuberculosis-report-2023
[Accessed 18th Dec, 2023]

Wolucka, B.A., McNeil, M.R., de Hoffmann, E., Chojnacki, T. and Brennan, P.J. (1994) Recognition of the lipid intermediate for arabinogalactan/arabinomannan biosynthesis and its relation to the mode of action of ethambutol on *Mycobacteria*. *Journal of Biological Chemistry*, *269*(37), pp.23328-23335.

Xin, Y., Huang, Y. and McNeil, M.R. (1999) The presence of an endogenous endo-D-arabinase in *Mycobacterium smegmatis* and characterization of its oligoarabinoside product. *Biochimica et Biophysica Acta (BBA)-General Subjects*, *1473*(2-3), pp.267-271.

Xu, L., Wu, D., Liu, L., Zheng, Q., Song, Y., Ye, L., Sha, S., Kang, J., Xin, Y. and Ma, Y. (2014) Characterization of mycobacterial UDP-N-acetylglucosamine enolpyruvyle transferase (MurA), *Research in microbiology*, *165*(2), pp.91-101.

Xu, W., DeJesus, M.A., Rücker, N., Engelhart, C.A., Wright, M.G., Healy, C., Lin, K., Wang, R., Park, S.W., Ioerger, T.R. and Schnappinger, D. (2017) Chemical genetic interaction profiling reveals determinants of intrinsic antibiotic resistance in *Mycobacterium tuberculosis*. *Antimicrobial agents and chemotherapy* [online], *61*(12), doi: https://doi.org/10.1128/aac.01334-17.

Yan, M.Y., Yan, H.Q., Ren, G.X., Zhao, J.P., Guo, X.P. and Sun, Y.C. (2017) CRISPR-Cas12a-assisted recombineering in bacteria. *Applied and environmental microbiology* [online], *83*(17), doi: https://doi.org/10.1128/AEM.00947-17.

Yang, Y., Kulka, K., Montelaro, R.C., Reinhart, T.A., Sissons, J., Aderem, A. and Ojha, A.K. (2014) A hydrolase of trehalose dimycolate induces nutrient influx and stress sensitivity to balance intracellular growth of *Mycobacterium tuberculosis*. *Cell host & microbe*, *15*(2), pp.153-163.

Yasir, M., Turner, A.K., Bastkowski, S., Baker, D., Page, A.J., Telatin, A., Phan, M.D., Monahan, L., Savva, G.M., Darling, A. and Webber, M.A. (2020) TraDIS-Xpress: a high-resolution whole-genome assay identifies novel mechanisms of triclosan action and resistance. *Genome research*, *30*(2), pp.239-249.

Zahir, T., Camacho, R., Vitale, R., Ruckebusch, C., Hofkens, J., Fauvart, M. and Michiels, J. (2019) High-throughput time-resolved morphology screening in bacteria reveals phenotypic responses to antibiotics. *Communications biology* [online], *2*(1), doi: https://doi.org/10.1038/s42003-019-0480-9.

Zhang, Y.J., Ioerger, T.R., Huttenhower, C., Long, J.E., Sassetti, C.M., Sacchettini, J.C. and Rubin, E.J. (2012) Global assessment of genomic regions required for growth in *Mycobacterium tuberculosis*. *PLOS Pathogens* [online], *9*(7),

Zhou, L., Xu, D., Liu, H., Wan, K., Wang, R. and Yang, Z. (2020) Trends in the prevalence and antibiotic resistance of non-tuberculous mycobacteria in Mainland China, 2000–2019: systematic review and meta-analysis. *Frontiers in Public Health*, *8*, p.295.

Zhu, X., Hong, A., Sun, X., Wang, W., He, G., Luo, H., Wu, Z., Xu, Q., Hu, Z., Wu, X. and Huang, D. (2022) Nigericin is effective against multidrug resistant gram-positive bacteria, persisters, and biofilms. *Frontiers in Cellular and Infection Microbiology* [online], *12*, p.1055929, doi: https://doi.org/10.3389/fcimb.2022.1055929.

# Appendix

**Supplementary Table 1** Drugs used in chemical-genomic screening with MICs and test concentrations for agar plates. Pretest concentration refers to the concentrations tested for pretesting, these concentrations were chosen based on values from literature. Pretest MICs refers to the concentration discovered during pretesting that inhibited bacterial growth. Pilot screen concentration refers to the concentration of drug used in the pilot screen. MICs and appropriate screening concentrations could not be determined for all conditions and thus further pretesting is necessary.

| Name | Source | Class | MOA | Pretest Concentrations | Pretest MICs | Pilot Screen Concentration |
|---|---|---|---|---|---|---|
| **2-Chloroadenosine** | Sigma | Others | Adenosine receptor agonist | 250 ng/mL, 500 ng/mL, 1 µg/mL | >1 µg/mL | |
| **4-aminosalicylic acid** | VWR | Drugs against mycobacteria | Folate synthesis inhibition | 250 ng/mL, 500 ng/mL, 1 µg/mL, | <250 ng/mL | |
| **5-fluorouracil** | Thermo Fisher Scientific | multipurpose drugs | DNA pyrimidine, Nucleic acid synthesis | | | |
| **Acetate** | Thermo Fisher Scientific | Carbon Source | Carbon Source | | | |
| **Acidic pH** | NA | pH | pH | pH 5.4, pH 5.8 | pH 5.8 | |
| **Actinomycin D** | Stratech Scientific | multipurpose drugs | pantothenate synthase | 250 ng/mL, 500 ng/mL, 1 µg/mL | 1 µg/mL | |
| **Adriamycin (Doxorubicin)** | Stratech Scientific | multipurpose drugs | Transcription, DNAG | | | |

| | | | | | | |
|---|---|---|---|---|---|---|
| **Alanine** | Sigma | Nitrogen Source | Nitrogen Source | | | |
| **Amikacin** | Sigma | Aminoglycoside | Translation, 16sRNA, 30S subunit | | | |
| **Amoxicillin (clavulanate)** | Cambridge Bioscience | Penicillin | PBPs | 8 µg/mL, 16 µg/mL, 32 µg/mL | 32µg/mL | 32 µg/mL |
| **Asparagine** | Sigma | Nitrogen Source | Nitrogen Source | | | |
| **Aspartate** | Sigma | Nitrogen Source | Nitrogen Source | | | |
| **Basic pH** | NA | pH | pH | pH 7.5, pH 8.0 | pH 8.0 | |
| **Bedaquiline** | Thermo Fisher Scientific | Drugs against mycobacteria | ATP/F-ATP synthase | 5 ng/mL, 10 ng/mL, 20 ng/mL | >20 µg/mL | |
| **Bleach (biocleanse)** | Thermo Fisher Scientific | Disinfectant | | | | |
| **Bleomycin** | Stratech Scientific | multipurpose drugs | DNA radical damage | | | |
| **BTZ043/Macrozinone** | Stratech Scientific | Drugs against mycobacteria | AG | 500 pg/mL, 1 ng/mL, 2 ng/mL | 500pg/mL | 500pg/mL |
| **Butyrate** | Sigma | Carbon Source | Carbon Source | | | |
| **CaCl2** | Sigma | Metal | Metal Stress | | | |
| **Capreomycin** | Cambridge Bioscience | Drugs against mycobacteria | Translation, 70S | 100 ng/mL, 200 ng/mL, 400 ng/mL | 200 ng/mL | |
| **CCCP (Carbonyl cyanide 3-chlorophenylhydrazone)** | Sigma | Others | PMF | | | |
| **Cefadroxil (clavulanate)** | Sigma | Cephalosporins (1) | PBPs | | | |

| | | | | | | |
|---|---|---|---|---|---|---|
| **Cefotaxime (clavulanate)** | BioServ UK | Cephalosporins (3) | PBPs | | | |
| **Cefoxitin (clavulanate)** | Sigma | Cephalosporins (2) | PBPs | | | |
| **Cerulenin** | Sigma | Others | MA/FAS | 50 ng/mL, 100 ng/mL, 200 ng/mL | | |
| **Chloramphenicol** | Thermo Fisher Scientific | Others | Translation, 50s | 2 µg/mL, 4 µg/mL, 8 µg/mL | 2 µg/mL | 2 µg/mL |
| **Chlorhexidine dihydrochloride** | Sigma | Disinfectant | | | | |
| **Cholesterol** | Sigma | Carbon Source | Carbon Source | | | |
| **Ciprofloxacin** | Sigma | Quinolones/Fluoroquinolones | Transcription, type II topoisomerase (DNA gyrase), topoisomerase IV | | | |
| **Clofazimine** | Thermo Fisher Scientific | Drugs against mycobacteria | DNA Guanine, accumulation of lysophospholipids | 50 ng/mL, 100 ng/mL, 200 ng/mL | > 200 ng/mL | 500 ng/mL |
| **CoCl2** | Sigma | Metal | Metal Stress | | | |
| **Cold Shock** | NA | Environment | Environment | 30°C | 30°C | |
| **CuSO4** | Sigma | Metal | Metal Stress | 200 µM, 400 µM, 1 mM | 200 µM | 200 µM |
| **Cycloserine** | Sigma | Drugs against mycobacteria | PG biosynthesis, D-Ala analogue | | | |
| **D-Trehalose** | Thermo Fisher Scientific | Carbon Source | | | | |
| **Dapsone** | Cambridge Bioscience | Drugs against mycobacteria | Folate synthesis inhibition | 8 µg/mL, 16 µg/mL, 32 µg/mL | 16 µg/mL | 16 µg/mL |

| | | | | | | |
|---|---|---|---|---|---|---|
| **Delamanid** | Cambridge Bioscience | Drugs against mycobacteria | MA | 500ng/mL, 1 µg/mL, 2 µg/mL | 1 µg/mL | 1 µg/mL |
| **Dequalinium chloride** | Sigma | | Permeability + enzyme acitivty | | | |
| **Diphenyleneiodonium chloride** | Cambridge Bioscience | Other | NADPH/NADH oxidase inhibitor | 5 ng/mL, 10 ng/mL, 20 ng/mL | <5 ng/mL | |
| **Disulfiram** | Stratech Scientific | Drugs against mycobacteria | | 500 ng/mL, 1 µg/mL, 2 µg/mL | >2 µg/mL | |
| **Doxycycline** | BioServ UK | Tetracyclines | Translation, 30S | 125 ng/mL, 250 ng/mL, 500 ng/mL | 125 ng/mL | |
| **EDTA** | Sigma | Chealator | Membrane | 50 µM, 100 µM, 200 µM | > 200 µM | 1 mM |
| **EGTA** | Sigma | Chealator | Membrane | 500 µM, 1 mM, 2 mM | 1 mM | |
| **Erythromycin** | Sigma | Macrolides | Translation, 50S | 8 µg/mL, 16 µg/mL, 32 µg/mL | 8 µg/mL | 8 µg/mL |
| **Ethambutol** | SLS | Drugs against mycobacteria | AG | 500 ng/mL, 1 µg/mL, 2 µg/mL | 500 ng/mL | 500 ng/mL |
| **Ethanol** | VWR | Alcohol | Protein degredation | 5% | 5% | |
| **Ethidium Bromide** | Sigma | DNA intercalating agent | DNA damage | | | |
| **Ethionamide** | SLS | Drugs against mycobacteria | MA | 10 µg/mL, 20 µg/mL, 40 µg/mL | | |
| **Ferric ammonium citrate (excess)** | Sigma | Metal | Metal Stress | 4 mM, 8 mM | <4 mM | |

| | | | | | | |
|---|---|---|---|---|---|---|
| **ferric ammonium citrate (starvation)** | Sigma | Metal | Metal Stress | 5 µM, 10 µM, 20 µM | >20 µM | |
| **Fusidic acid** | Cambridge Bioscience | Others | Protein synthesis EF-G | 2 µg/mL, 4 µg/mL, 8 µg/mL | 2 µg/mL | 2 µg/mL |
| **Gliotoxin** | Cambridge Bioscience | Others | Protein degredation/NA | 5 ng/mL, 10 ng/mL, 25 ng/mL | 5 ng/mL | |
| **Glutamate** | Sigma | Nitrogen Source | Nitrogen Source | | | |
| **Glutamine** | Sigma | Nitrogen Source | Nitrogen Source | | | |
| **Glycerol** | Thermo Fisher Scientific | Carbon Source | Carbon Source | | | |
| **GSK-286** | NA | Drugs against mycobacteria | Cholesterol metabolism | | | |
| **GSK3036656** | NA | Drugs against mycobacteria | Transcription, LeuRS | | | |
| **H2O2** | Thermo Fisher Scientific | | Oxidative Stress | 5 mM, 10 mM, 20 mM | 5 mM | 5 mM |
| **Heat Shock** | NA | Environment | Environment | 40°C | 40°C | |
| **Hydroxyurea** | VWR | | DNA damage | | | |
| **Hypoxia** | NA | Environment | Environment | | | |
| **Imipenem (clavulanate)** | Thermo Fisher Scientific | Carbapenems | PBPs | 500 ng/mL, 1 µg/mL, 2 µg/mL | >8 µg/mL (no clavulanate) | |
| **Isoniazid** | SLS | Drugs against mycobacteria | MA | 25 ng/mL, 50 ng/mL, 100 ng/mL | >100 ng/mL | |
| **Itaconic acid** | Sigma | Enzyme | inhibitor of isocitrate lyase | 125 µg/mL, 250 µg/mL, 500 µg/mL | 125 µg/mL | 125 µg/mL |

| | | | | | | |
|---|---|---|---|---|---|---|
| **Lactate (Lactic Acid)** | Sigma | Carbon Source | Carbon Source | | | |
| **Leucine** | Sigma | Nitrogen Source | Nitrogen Source | | | |
| **Linezolid** | Stratech Scientific | Oxazolidinones | Translation, Protein Sythesis initiation | 100 ng/mL, 250 ng/mL, 500 ng/mL | >500 ng/mL | |
| **Loperamide** | SLS | multipurpose drugs | MA | 8 µg/mL, 16 µg/mL, 32 µg/mL | >32 µg/mL | |
| **Lysozyme (0.05% tween80 (40ug/mL))** | Sigma | multipurpose drugs | PG hydrolysis | | | |
| **Meropenem** | Sigma | Carbapenems | PBPs | 250 ng/mL, 500 ng/mL, 1 µg/mL | 250 ng/mL | 250 ng/mL |
| **Metformin** | Thermo Fisher Scientific | multipurpose drugs | | | | |
| **Methoctramine** | Cambridge Bioscience | multipurpose drugs | | 1 µg/mL, 2 µg/mL, 4 µg/mL | >4 µg/mL | |
| **MgCl2** | Sigma | Salt | Osmotic stress | 10 mM, 20 mM, 40 mM | 20 mM | |
| **MMS** | VWR | Carcinogen | DNA damage | | | |
| **NaCl** | Thermo Fisher Scientific | Salt | Osmotic stress | 20 µM, 40 µM, 80 µM | 100 mM (high salt) | |
| **NH4Cl** | Sigma | Nitrogen Source | Nitrogen Source | | | |
| **Niclosamide** | Thermo Fisher Scientific | multipurpose drugs | | | | |
| **Nigericin** | Stratech Scientific | multipurpose drugs | Membrane potential | 125 ng/mL, 250 ng/mL, 500 ng/mL | 250 ng/mL | 250 ng/mL |

| | | | | | | |
|---|---|---|---|---|---|---|
| **NiSO4** | Sigma | Metal | Metal Stress | 400 µM, 800 µM, 2 mM | >2 mM | |
| **Ofloxacin** | Sigma | Quinolones/Fluoro quinolones | Transcription, type II topoisomerase (DNA gyrase), topoisomerase IV | 125 ng/mL, 250 ng/mL, 500 ng/mL | >500 ng/mL | |
| **Oleic Acid** | Sigma | Carbon Source | Carbon Source | | | |
| **Palmitate** | Sigma | Carbon Source | Carbon Source | | | |
| **Phenazine methosulfate (PMS)** | SLS | Other | DNA damage | 5 µg/mL, 10 µg/mL, 20 µg/mL | 5 µg/mL | 5 µg/mL |
| **Phenothiazine** | SLS | Drugs against mycobacteria | Ndh | 500 ng/mL, 1 µg/mL, 2 µg/mL | >2 µg/mL | |
| **Piperacillin (clavulanate)** | Sigma | Penicillin | PBPs | | | |
| **Polymyxin B** | Stratech Scientific | Polypeptides | Membrane | 2.5 µg/mL, 5 µg/mL, 10 µg/mL | >10 µg/mL | |
| **Propanol** | VWR | Alcohol | Protein degredation | | | |
| **Propidium iodide** | VWR | Others | DNA damage | 25 µg/mL, 50 µg/mL, 100 µg/mL | >50 µg/mL | |
| **Propionate** | Sigma | Carbon Source | Carbon Source | | | |
| **Pyrazinamide (pH 5.5)** | Thermo Fisher Scientific | Drugs against mycobacteria | | 2 µg/mL, 4 µg/mL, 8 µg/mL | 4 µg/mL | |
| **Pyruvate** | Sigma | Carbon Source | Carbon Source | | | |
| **Reserpine** | SLS | multipurpose drugs | Efflux pumps, biofilm formation | 4 µg/mL, 8 µg/mL, 16 µg/mL | 16 µg/mL | 16 µg/mL |
| **Rifampicin** | BioServ UK | Drugs against mycobacteria | Transcription, RNAP | 25 ng/mL, 50 ng/mL, 100 ng/mL | 25 ng/mL | 25 ng/mL |

| | | | | | | |
|---|---|---|---|---|---|---|
| **SDS** | Thermo Fisher Scientific | Detergent | Membrane | 0.05%, 0.1%, 0.2% | <0.05% | |
| **Silver nitrate** | Sigma | Metal | Metal Stress | 5 µM, 10 µM, 20 µM | >10 µM | |
| **Sodium fluoride** | SLS | Others | Phosphotase inhibition | 25 mM, 50 mM, 100 mM | >50 mM | |
| **SQ109** | Sigma | Drugs against mycobacteria | MA | 25 ng/mL, 50 ng/mL, 100 ng/mL | 50 ng/mL | |
| **SQ609** | Sigma | Drugs against mycobacteria | | | | |
| **Stearate** | Thermo Fisher Scientific | Carbon Source | Carbon Source | | | |
| **Streptomycin** | Thermo Fisher Scientific | Aminoglycoside | Translation, 16sRNA, 30S subunit | 2 ng/mL, 10 ng/mL, 50 ng/mL | >50 ng/mL | |
| **Sulfamethoxazole** | Thermo Fisher Scientific | Sulfonamide | Folic Acid Biosyntheis | 10 µg/mL, 20 µg/mL, 30 µg/mL | 10 µg/mL | 10 µg/mL |
| **t-Butanol** | Thermo Fisher Scientific | Alcohol | Protein degredation | 1%, 2%, 3% | 2% | |
| **Telacebec (Q203)** | Cambridge Bioscience | Drugs against mycobacteria | ATP synthesis, Electron transport chain, cytochrome bc-1 | 250 pg/mL, 500 pg/mL, 1 ng/mL | 250 pg/mL | 250 pg/mL |
| **Thiacetazone** | Cambridge Bioscience | Acetamide | MA/cyclopropane synthesis | 50 ng/mL, 100 ng/mL, 200 ng/mL | 200ng/mL | |

| | | | | | | |
|---|---|---|---|---|---|---|
| **Tigecycline(Bs)** | Cambridge Bioscience | Others | Translation, 30S, 70S | 250 ng/mL, 500 ng/mL, 1 µg/mL | >1 µg/mL | |
| **Tobramycin** | Thermo Fisher Scientific | Aminoglycoside | Translation, 30S subunit | 1 µg/mL, 2 µg/mL, 4 µg/mL | 2 µg/mL | |
| **TP053** | SLS | Drugs against mycobacteria | Oxidative stress | 50 ng/mL, 100 ng/mL, 200 ng/mL | 200 ng/mL | 200 ng/mL |
| **Triton X-100** | Sigma | Detergent | Membrane | 0.05%, 0.1%, 0.2% | <0.05% | |
| **Tween80** | Sigma | Detergent | Membrane | 0.2%, 0.4%, 0.8% | | |
| **Tween80** | Sigma | Carbon Source | Carbon Source | | | |
| **UV** | NA | Environment | Environment | 1min, 2min, 4min | > 4min | |
| **Vancomycin** | Stratech Scientific | Glycopeptides | PG crosslinking | 4 µg/mL, 8 µg/mL, 16 µg/mL | 8 µg/mL | |
| **Vanoxerine** | Stratech Scientific | Others | efflux pumps | 5 µg/mL, 7.5 µg/mL, 10 µg/mL | >10 µg/mL | |
| **Verapamil** | Stratech Scientific | multipurpose drugs | Efflux pumps | 50 µg/mL, 100 µg/mL, 200 µg/mL | 200 µg/mL | |
| **ZnSO4** | Sigma | Metal | Metal Stress | 200 µM, 400 µM, 1 mM | >1 mM | |

**Supplementary Table 2** Pilot screen mutant fitness S-scores in response to different chemical stresses. The colony morphology data was analysed via IRIS and 6 of the original 22 conditions were removed due to incompatibility with the IRIS software. Fitness S-scores were generated via ChemGAPP and a normality threshold of 40%, a Mann-Whitney plate threshold of 0.005846 and an average variance threshold of 6954000.0 was applied as a data quality control measure. This resulted in the removal of 4 additional conditions. The fitness S-scores for the remaining 12 conditions are represented. An S-score < -3 or > 2 signifies significantly reduced or significantly enhanced mutant fitness respectively.

| Gene | 7H9 | BTZ | CUSO4 | DLM | EGTA | EMB | IA | MER | NIG | PMS | RES | TP053 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **BCGDan_0025** | 0.241 | 1.804 | 0.890 | 1.442 | -0.666 | 1.119 | -0.537 | 0.268 | -0.803 | | 0.077 | -0.491 |
| **BCGDan_0054** | 0.612 | | -1.073 | 1.144 | 0.105 | 0.697 | -1.304 | -1.491 | -0.526 | -0.165 | -1.014 | 0.865 |
| **BCGDan_0076** | -0.002 | | 0.556 | -0.420 | 0.999 | -0.109 | -0.384 | -0.081 | 0.261 | -0.077 | 0.008 | -0.019 |
| **BCGDan_0079** | -0.615 | | 0.511 | -0.217 | 0.573 | -0.189 | -0.063 | 0.051 | -0.229 | -0.100 | -0.052 | 0.569 |
| **BCGDan_0082** | 0.570 | 0.245 | -0.516 | 0.267 | -0.250 | 0.202 | -0.867 | -0.282 | -0.162 | | 1.340 | 0.829 |
| **BCGDan_0087** | -3.263 | | 0.970 | -1.172 | -1.056 | -2.450 | -0.268 | 1.925 | 2.255 | 1.258 | 1.485 | -2.659 |
| **BCGDan_0105** | 0.938 | -1.208 | 0.705 | 3.251 | -11.403 | -1.038 | 0.027 | -0.842 | -1.696 | | 1.126 | 1.206 |
| **BCGDan_0110** | -1.729 | 1.482 | 2.481 | 0.104 | -0.971 | -0.698 | 1.506 | -0.215 | 3.125 | | -2.090 | -1.390 |
| **BCGDan_0123** | -0.119 | 1.098 | 0.265 | -0.114 | -1.055 | -0.088 | 0.027 | -0.404 | -1.193 | | -0.258 | -0.268 |
| **BCGDan_0133** | 0.141 | | -0.249 | 2.092 | 0.384 | -0.464 | -1.906 | -0.466 | 0.893 | 1.523 | 0.288 | -0.255 |
| **BCGDan_0140** | 2.276 | | 0.241 | 0.826 | 0.467 | 0.842 | -0.659 | -0.432 | 4.163 | -1.430 | -1.300 | -1.051 |
| **BCGDan_0142** | 0.142 | | 0.137 | 0.356 | 0.426 | -0.205 | -1.345 | 0.247 | 1.963 | -1.809 | -1.836 | -0.330 |
| **BCGDan_0152** | 0.104 | -0.023 | 0.716 | 2.747 | -0.907 | -1.139 | -0.567 | 1.419 | -0.206 | | -0.674 | 0.594 |
| **BCGDan_0174** | -1.578 | | 0.784 | -0.802 | -1.301 | -0.951 | 0.424 | 0.674 | 0.573 | 2.488 | 0.809 | -1.703 |
| **BCGDan_0179** | -2.083 | 2.577 | 0.801 | 1.126 | 0.498 | -1.219 | -0.463 | -1.137 | 1.135 | | 1.210 | -1.468 |
| **BCGDan_0231** | 0.669 | | 0.161 | 0.425 | 0.660 | 0.246 | -0.625 | -0.469 | 0.791 | -1.190 | -0.950 | 0.111 |
| **BCGDan_0252** | -1.787 | | 1.082 | 0.074 | -1.222 | -2.386 | 0.751 | 1.659 | 0.486 | -0.113 | 0.470 | -0.712 |
| **BCGDan_0256** | -1.291 | -0.209 | -0.144 | 0.442 | -0.297 | -1.403 | -0.900 | 1.966 | 0.918 | | 1.163 | 4.245 |

| | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **BCGDan_0260** | -0.457 | | 0.368 | 0.267 | -0.089 | -1.040 | -0.404 | -0.283 | 0.294 | -0.329 | 0.334 | 0.566 |
| **BCGDan_0278** | 0.180 | | 0.980 | 1.046 | -0.115 | 0.836 | -0.652 | -0.950 | 0.143 | -1.302 | 0.062 | 0.610 |
| **BCGDan_0327** | -0.900 | 0.040 | 0.834 | -0.570 | 0.448 | -0.874 | 0.056 | 0.469 | 0.294 | | -1.611 | 0.532 |
| **BCGDan_0341** | 1.218 | 0.585 | -0.570 | -0.331 | -0.766 | 0.701 | -0.875 | -1.095 | 0.166 | | 0.730 | -0.805 |
| **BCGDan_0404** | -1.717 | 0.463 | 0.664 | 1.095 | -1.795 | -1.915 | 0.573 | 0.556 | 1.793 | | -0.480 | 0.068 |
| **BCGDan_0419** | 0.565 | | -0.553 | 0.102 | 0.412 | 0.814 | -3.018 | 0.211 | -0.297 | -1.053 | -0.397 | 1.794 |
| **BCGDan_0421** | -2.285 | 0.949 | 2.452 | 5.142 | 0.261 | -1.497 | -1.281 | 1.336 | 7.010 | | -0.462 | -4.867 |
| **BCGDan_0423** | 0.000 | | 0.000 | 0.000 | 0.293 | 5.246 | 0.000 | 0.000 | 0.066 | 0.000 | 0.149 | 0.000 |
| **BCGDan_0470** | 0.506 | 0.803 | -0.679 | -1.215 | -0.538 | 0.409 | -0.800 | -1.643 | -0.052 | | 1.765 | -0.088 |
| **BCGDan_0486** | -0.409 | 2.310 | 0.126 | 0.644 | -2.004 | 0.457 | -0.480 | 0.569 | 3.348 | | 0.441 | -1.703 |
| **BCGDan_0558** | -0.362 | | 0.096 | -0.206 | 0.334 | -0.443 | -0.780 | 0.599 | 0.081 | -0.332 | 0.426 | 0.779 |
| **BCGDan_0572** | 0.838 | -0.332 | -0.541 | -0.608 | 1.773 | 0.227 | -2.554 | -1.738 | -0.093 | | 1.351 | 0.991 |
| **BCGDan_0601** | -0.655 | 2.548 | 0.767 | -0.437 | -2.462 | 0.546 | -0.242 | 0.072 | 0.486 | | 0.304 | -1.608 |
| **BCGDan_0620** | 0.063 | 1.685 | -0.183 | 0.532 | -0.260 | -1.120 | 0.392 | -0.560 | -0.294 | | 1.898 | -0.194 |
| **BCGDan_0699** | -0.604 | 0.128 | 0.439 | 2.824 | 0.156 | 0.583 | -1.050 | 0.247 | 0.933 | | -0.488 | -0.819 |
| **BCGDan_0723** | 1.000 | | -0.528 | 0.481 | 0.118 | 0.059 | 0.294 | -0.916 | 5.077 | -2.052 | -1.263 | 0.119 |
| **BCGDan_0727** | -1.627 | 1.265 | 0.828 | 0.786 | -1.908 | -2.347 | 0.389 | 1.455 | -0.132 | | 0.092 | -1.148 |
| **BCGDan_0785** | -0.111 | 0.564 | -0.757 | 0.382 | 0.993 | -0.911 | -2.466 | -0.860 | -0.113 | | 0.700 | 1.752 |
| **BCGDan_0789** | -0.871 | | 2.479 | -0.583 | -0.744 | -0.190 | -0.640 | 2.228 | 0.135 | 4.577 | 6.321 | -0.699 |
| **BCGDan_0815** | -0.075 | | 0.299 | -0.100 | 0.067 | -1.585 | -0.317 | 0.652 | 0.778 | -0.456 | 0.950 | -0.500 |
| **BCGDan_0832** | -0.467 | | -0.131 | 0.826 | -0.221 | -0.855 | 0.172 | 0.801 | 0.027 | 0.974 | 1.371 | -0.655 |
| **BCGDan_0874** | -0.056 | 1.709 | -0.605 | 0.927 | -1.193 | -0.638 | 0.309 | -1.280 | 1.149 | | 0.959 | -0.731 |
| **BCGDan_0926** | -1.301 | | 1.041 | -0.866 | -1.128 | -1.750 | 0.240 | 0.931 | 1.204 | 0.997 | 1.083 | -1.410 |
| **BCGDan_0965** | -1.546 | | 1.738 | -1.029 | -1.845 | -0.008 | 0.108 | 1.277 | -0.351 | 2.357 | 1.349 | -0.767 |
| **BCGDan_0984** | -2.426 | 1.684 | 1.053 | 0.944 | -1.042 | -0.516 | 0.718 | -0.025 | -0.422 | | 0.930 | 0.112 |
| **BCGDan_1005** | -2.094 | | 1.107 | -1.275 | -1.325 | -1.687 | -0.032 | 1.441 | 6.552 | 1.508 | 1.125 | -2.811 |
| **BCGDan_1016** | 0.228 | | -1.119 | 0.963 | 0.323 | 0.427 | -1.855 | -1.734 | -0.232 | 1.115 | -0.032 | 1.189 |

| | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **BCGDan_1028** | -0.095 | 0.705 | -0.780 | 0.505 | 0.775 | -1.880 | -2.063 | -0.526 | -0.256 | | 1.344 | 1.616 |
| **BCGDan_1040** | -0.776 | | -0.483 | -0.547 | -0.242 | -1.763 | 0.145 | 0.799 | 1.006 | 2.009 | 3.022 | -0.176 |
| **BCGDan_1061** | 0.138 | -4.108 | -0.399 | 0.914 | 0.589 | 0.549 | -0.463 | -0.020 | -0.042 | | 0.939 | 0.922 |
| **BCGDan_1080** | 3.345 | | -0.026 | -1.128 | 1.408 | 2.643 | 1.322 | -0.054 | -3.890 | 0.148 | -8.082 | -1.483 |
| **BCGDan_1087** | -1.256 | -1.837 | 1.997 | 1.513 | 1.026 | -0.128 | -0.446 | -1.393 | 3.154 | | 0.473 | 0.709 |
| **BCGDan_1113** | 0.140 | | -0.474 | 1.449 | -0.333 | -0.826 | -1.275 | -0.863 | 0.157 | 1.415 | 0.422 | -0.907 |
| **BCGDan_1121** | -0.104 | 1.802 | -0.529 | -0.429 | -0.370 | -0.188 | -1.235 | -0.470 | 1.730 | | 1.045 | -1.629 |
| **BCGDan_1140** | -0.558 | -0.054 | 0.337 | 1.044 | 0.904 | -0.075 | -0.423 | 0.702 | -0.342 | | -0.988 | 0.342 |
| **BCGDan_1198** | 1.981 | | -0.481 | -0.037 | -0.355 | 0.764 | -0.369 | -2.114 | -0.704 | 0.876 | -0.008 | 0.678 |
| **BCGDan_1225** | 2.436 | | -0.464 | 0.150 | 0.974 | 0.613 | -0.340 | -2.203 | -0.518 | 0.114 | -0.100 | 1.998 |
| **BCGDan_1227** | 0.524 | | -0.670 | 2.479 | 0.662 | 1.442 | -0.312 | -1.076 | -1.479 | -6.360 | 1.036 | 1.271 |
| **BCGDan_1230** | 1.310 | -0.351 | 0.027 | -3.090 | -1.836 | 1.180 | 3.112 | -2.154 | 9.345 | | -3.505 | -3.392 |
| **BCGDan_1238** | -0.902 | 0.081 | 0.546 | 0.023 | -0.678 | -0.871 | 1.889 | 0.076 | 0.221 | | 0.463 | -1.109 |
| **BCGDan_1242** | -0.519 | | 0.532 | -0.831 | -1.008 | 0.387 | -1.100 | 1.939 | 0.705 | 0.216 | 0.145 | -0.257 |
| **BCGDan_1247** | -0.544 | | 0.451 | -0.083 | -1.270 | -0.863 | -0.012 | 2.476 | -0.850 | 2.568 | 0.792 | 0.031 |
| **BCGDan_1258** | -0.291 | | 1.432 | -0.502 | -0.359 | -0.006 | -1.802 | 1.381 | 0.118 | 1.831 | 0.903 | -0.444 |
| **BCGDan_1271** | -1.062 | | 0.703 | 1.006 | 0.341 | -1.116 | 0.524 | 1.593 | -4.051 | 1.832 | -0.428 | -0.775 |
| **BCGDan_1280** | 0.441 | | 0.339 | -0.319 | -0.188 | -0.404 | -0.927 | 1.221 | 0.491 | -1.146 | -0.460 | 0.112 |
| **BCGDan_1285** | 0.241 | 0.892 | -0.624 | 2.237 | 1.395 | -3.006 | -1.453 | -0.761 | -0.371 | | 0.493 | 0.197 |
| **BCGDan_1292** | 0.077 | -0.905 | 0.463 | 2.040 | -2.293 | -0.572 | 0.375 | 0.614 | 0.594 | | -0.539 | -0.366 |
| **BCGDan_1294** | -0.058 | | -0.372 | 0.215 | -0.973 | 0.608 | 0.009 | -0.161 | -0.005 | -0.326 | 0.245 | 0.145 |
| **BCGDan_1306** | 3.786 | -3.840 | -1.410 | -0.266 | 0.055 | 0.336 | 4.734 | 1.390 | 4.144 | | -6.371 | -1.966 |
| **BCGDan_1426** | -1.692 | | 1.870 | -0.996 | -1.269 | -1.076 | 0.074 | 0.069 | 4.514 | 0.686 | 0.932 | -0.556 |
| **BCGDan_1429** | 0.693 | 1.298 | 0.980 | 1.793 | -0.848 | -0.156 | -0.896 | 0.154 | -1.122 | | -0.481 | -0.351 |
| **BCGDan_1432** | 1.758 | | -1.231 | 1.117 | 0.247 | 0.664 | -2.381 | -0.707 | -0.486 | 0.539 | -0.373 | 1.016 |
| **BCGDan_1477** | 0.193 | | -0.388 | 0.719 | -0.192 | -0.276 | -0.678 | 1.037 | -0.726 | 0.179 | -0.355 | 0.754 |
| **BCGDan_1478** | -2.214 | | 1.712 | -0.264 | -1.209 | -2.066 | -0.561 | 2.034 | 2.512 | 3.258 | 0.579 | -1.809 |

| | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **BCGDan_1541** | 0.000 | 0.000 | 0.000 | 0.000 | 0.562 | 0.000 | 0.329 | 0.000 | 0.268 | | 0.000 | 0.000 |
| **BCGDan_1561** | 0.081 | 1.898 | -0.663 | 1.256 | 1.364 | -2.410 | -1.127 | -0.969 | -0.305 | | 1.155 | 1.401 |
| **BCGDan_1562** | -0.492 | | -0.407 | 1.026 | -0.583 | -0.371 | -1.032 | -0.021 | 1.537 | 1.235 | 1.537 | -0.210 |
| **BCGDan_1600** | -1.074 | -0.513 | 1.022 | -0.299 | -0.940 | 0.379 | 0.126 | 0.181 | 0.984 | | 0.270 | -0.050 |
| **BCGDan_1602** | -0.291 | | -0.342 | 0.172 | -0.090 | -1.506 | -0.837 | 0.585 | 0.186 | 0.587 | 0.724 | -0.201 |
| **BCGDan_1655** | -0.887 | -0.337 | -0.456 | -0.899 | 0.140 | -1.745 | 0.873 | -0.491 | 1.197 | | 1.926 | 0.817 |
| **BCGDan_1682** | 0.461 | 0.749 | -0.788 | -1.453 | -0.422 | 0.085 | -0.627 | -0.579 | 3.819 | | 1.863 | -1.474 |
| **BCGDan_1688** | 0.463 | -0.080 | -0.614 | -0.827 | 1.149 | -0.388 | 0.094 | -0.229 | -0.078 | | 0.666 | 0.826 |
| **BCGDan_1746** | -1.022 | | 0.566 | -1.023 | 0.003 | -1.973 | 0.247 | 4.046 | 1.930 | -0.353 | 1.847 | -0.294 |
| **BCGDan_1758** | -1.286 | | 0.521 | 0.495 | -0.534 | -2.117 | -1.008 | 3.632 | 3.064 | 0.623 | 0.004 | -1.637 |
| **BCGDan_1763** | -0.775 | -0.132 | -0.255 | -0.373 | 0.516 | 0.141 | 0.358 | -0.897 | -0.164 | | 1.254 | 0.135 |
| **BCGDan_1785** | -0.814 | | 0.796 | -0.938 | 0.291 | -0.102 | 0.270 | 1.147 | 3.520 | -0.039 | -1.107 | 0.415 |
| **BCGDan_1792** | -0.218 | -2.104 | 1.768 | 2.407 | 0.184 | 1.368 | 1.071 | -1.373 | -5.184 | | -0.705 | 0.189 |
| **BCGDan_1849** | -2.020 | | 1.039 | -0.682 | -1.398 | -1.268 | 0.102 | 1.027 | 0.234 | 0.718 | 1.264 | -0.910 |
| **BCGDan_1899** | -1.050 | | 1.351 | -0.644 | -1.545 | 0.190 | -0.948 | 0.980 | -1.694 | 0.572 | 1.646 | 0.572 |
| **BCGDan_1900** | 0.046 | | -0.553 | 0.626 | -0.398 | -0.669 | -1.035 | -0.084 | -0.861 | 0.295 | -0.057 | 0.882 |
| **BCGDan_1903** | 0.411 | 0.589 | -0.480 | -0.370 | 1.110 | -0.028 | 0.239 | -0.887 | -0.100 | | 1.833 | 1.715 |
| **BCGDan_1910** | -0.505 | 0.262 | 0.466 | 3.013 | -1.040 | -0.473 | -0.229 | 0.717 | 0.031 | | -0.852 | 0.064 |
| **BCGDan_1911** | 0.385 | 1.404 | -1.604 | 0.657 | -0.594 | -0.983 | -0.843 | 0.230 | -0.238 | | 1.419 | 1.454 |
| **BCGDan_1913** | -1.594 | | 1.336 | -0.197 | -1.663 | -1.093 | 0.214 | 0.755 | 0.511 | 1.125 | 1.489 | -0.640 |
| **BCGDan_1940** | -0.079 | 0.928 | -0.514 | -0.251 | -0.097 | -1.034 | -0.329 | -0.350 | 0.169 | | 1.551 | 0.987 |
| **BCGDan_1949** | -0.361 | | -0.522 | 0.503 | 0.050 | 0.251 | -1.220 | -0.994 | -0.579 | 0.715 | 0.276 | 0.984 |
| **BCGDan_1965** | -0.725 | 0.875 | 0.440 | 0.175 | -0.775 | 0.684 | 1.170 | -1.832 | -1.388 | | -0.856 | -0.274 |
| **BCGDan_1998** | -1.068 | | -0.215 | 0.274 | -0.260 | -0.760 | 0.360 | 0.261 | -0.734 | 1.827 | 1.478 | -0.149 |
| **BCGDan_2006** | -1.760 | | 0.520 | -0.241 | -1.641 | -0.912 | 0.264 | 0.926 | 0.241 | 0.475 | 1.312 | -0.585 |
| **BCGDan_2043** | -2.127 | | 1.099 | -1.671 | -2.010 | 0.023 | -0.479 | 1.006 | 0.060 | 2.707 | 2.538 | -0.445 |
| **BCGDan_2048** | 0.551 | -3.925 | -0.383 | 0.512 | 0.270 | -0.747 | -0.173 | -0.615 | 0.001 | | 1.535 | 0.162 |

| | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **BCGDan_2073** | 1.472 | | -0.877 | 1.362 | 0.271 | 1.576 | -0.484 | -0.214 | 8.184 | -8.221 | -5.278 | 0.450 |
| **BCGDan_2085** | -0.485 | 0.423 | -0.345 | -0.094 | 1.044 | -1.127 | 0.332 | -0.257 | -0.245 | | 1.225 | 0.551 |
| **BCGDan_2096** | 0.374 | | 0.231 | 0.082 | 0.203 | -0.283 | 0.392 | -0.607 | -1.003 | 0.068 | 0.320 | -0.287 |
| **BCGDan_2105** | 0.620 | | -0.478 | 1.191 | -0.306 | 0.013 | -1.631 | -1.897 | -0.521 | 0.506 | 0.038 | 1.115 |
| **BCGDan_2111** | -2.449 | -2.719 | 1.104 | -0.354 | -1.373 | -0.101 | 1.490 | -1.646 | 1.818 | | 0.936 | -0.871 |
| **BCGDan_2122** | -1.031 | | 0.649 | 0.251 | -1.370 | -0.949 | 0.913 | 0.749 | 0.296 | -0.318 | 1.042 | -3.410 |
| **BCGDan_2128** | -0.691 | -3.913 | -0.169 | 0.135 | 1.491 | -0.292 | 0.271 | 0.103 | 0.588 | | 1.989 | 0.359 |
| **BCGDan_2132** | -0.453 | | 0.131 | -4.980 | -0.024 | 0.026 | 0.655 | 2.187 | -5.942 | 0.476 | -5.394 | -0.824 |
| **BCGDan_2138** | -3.313 | -1.898 | 0.602 | 0.451 | -2.276 | 0.670 | 0.483 | -0.896 | 0.773 | | -0.212 | 0.433 |
| **BCGDan_2148** | 0.159 | | -0.432 | 0.330 | -0.092 | 0.432 | -2.237 | 0.036 | 0.189 | -0.038 | -0.245 | 1.380 |
| **BCGDan_2174** | 0.762 | | -1.079 | 0.935 | -0.142 | 1.874 | -1.313 | -1.003 | -0.208 | -9.947 | 0.202 | 1.476 |
| **BCGDan_2230** | -0.162 | 0.416 | 0.981 | 0.488 | 0.133 | -1.136 | -0.470 | 1.172 | -0.376 | | -0.221 | 0.347 |
| **BCGDan_2296** | -0.989 | | 0.616 | -0.299 | -0.380 | 0.272 | -0.880 | 0.888 | -4.168 | 1.148 | 1.083 | 0.086 |
| **BCGDan_2297** | 0.433 | | 0.467 | -0.276 | -0.675 | 0.530 | 0.560 | -0.236 | 0.134 | -0.748 | 0.532 | -0.118 |
| **BCGDan_2310** | 0.558 | 1.011 | -0.927 | 0.222 | -1.129 | -1.592 | -2.153 | -0.788 | -0.175 | | 0.740 | 0.690 |
| **BCGDan_2318** | -0.049 | | -0.573 | 0.855 | -0.177 | 0.232 | -1.774 | -0.695 | 0.942 | 0.690 | -0.385 | 0.682 |
| **BCGDan_2323** | -2.183 | | 0.548 | -1.309 | -1.549 | -2.869 | 0.065 | 0.822 | 1.780 | 0.838 | 1.074 | -2.282 |
| **BCGDan_2329** | -0.272 | | -0.637 | -0.478 | 0.652 | 0.125 | 2.325 | -0.220 | 0.862 | -0.015 | 0.036 | 1.022 |
| **BCGDan_2348** | -0.002 | 0.713 | 0.581 | -2.050 | -4.974 | -1.209 | 1.033 | 0.572 | 2.030 | | -3.862 | -3.054 |
| **BCGDan_2372** | -0.420 | | -0.229 | 0.357 | -0.009 | 0.013 | 0.242 | -0.172 | -0.173 | 2.894 | 2.058 | 0.106 |
| **BCGDan_2432** | -0.346 | | 3.005 | -0.433 | 0.038 | 0.465 | 0.507 | 2.992 | -1.441 | -0.841 | -1.566 | 0.684 |
| **BCGDan_2463** | 0.121 | | -0.108 | 1.921 | 0.101 | -0.488 | -0.065 | -0.867 | -0.137 | 1.300 | 0.688 | -0.445 |
| **BCGDan_2485** | 0.001 | | -0.457 | 0.989 | -0.393 | 0.141 | -1.405 | -0.028 | -0.350 | 0.810 | 0.105 | 0.503 |
| **BCGDan_2495** | -1.114 | | 1.185 | -0.863 | -2.241 | -0.714 | 0.247 | 1.062 | 0.389 | 2.438 | 1.804 | -1.212 |
| **BCGDan_2505** | 0.179 | 1.440 | 0.984 | 0.686 | -0.960 | -1.150 | -0.782 | 0.355 | 0.147 | | 0.384 | -1.377 |
| **BCGDan_2506** | 0.241 | | 1.003 | 0.352 | 0.317 | -0.082 | 0.302 | -0.648 | 1.536 | -1.031 | -0.394 | -0.301 |
| **BCGDan_2550** | -1.387 | | 1.246 | -0.445 | -2.152 | -1.777 | 0.246 | 0.828 | 0.210 | 0.830 | 1.061 | -1.463 |

| | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **BCGDan_2586** | -1.601 | 2.666 | 1.340 | 0.377 | -0.767 | -1.052 | 1.446 | -0.386 | -0.766 | | 0.975 | -0.428 |
| **BCGDan_2621** | -0.183 | | 0.752 | 1.666 | 0.451 | -0.831 | -0.186 | -0.161 | 0.604 | -0.539 | -0.197 | -0.637 |
| **BCGDan_2624** | 1.161 | | -0.370 | 0.689 | 1.589 | 0.897 | -1.592 | -2.130 | -0.481 | -1.853 | -0.575 | 0.979 |
| **BCGDan_2629** | -1.590 | 0.229 | 1.222 | 0.235 | -0.815 | -1.358 | 0.888 | 0.953 | 0.556 | | 0.325 | -0.234 |
| **BCGDan_2642** | -1.249 | | 1.223 | -2.583 | -1.602 | -1.504 | -0.895 | 2.408 | 7.069 | 1.383 | 0.593 | 0.310 |
| **BCGDan_2707** | -0.569 | 2.100 | 0.939 | -0.357 | 0.022 | 1.032 | 0.443 | 0.177 | -0.476 | | -1.262 | -0.117 |
| **BCGDan_2710** | -0.436 | | -0.504 | 3.005 | 0.788 | -0.284 | 0.804 | -0.605 | -0.019 | 0.433 | -0.264 | -0.685 |
| **BCGDan_2718** | -0.345 | 0.545 | 0.611 | 1.213 | -1.464 | -0.524 | 0.311 | 1.127 | 0.387 | | -0.392 | -0.842 |
| **BCGDan_2720** | -2.235 | | 0.217 | -1.044 | -1.551 | -3.280 | -0.864 | 0.962 | 2.494 | 0.540 | 1.617 | 0.049 |
| **BCGDan_2740** | -3.163 | -0.658 | 0.892 | 0.423 | -1.404 | -0.585 | 0.844 | 0.293 | 0.998 | | 0.806 | -0.235 |
| **BCGDan_2779** | 0.896 | 0.700 | -0.614 | 0.084 | -0.497 | 0.916 | -1.146 | -1.317 | -0.027 | | 1.350 | 0.426 |
| **BCGDan_2823** | -0.936 | 0.974 | 0.501 | 1.881 | -0.048 | -1.299 | -0.288 | 1.027 | -0.001 | | -0.716 | -0.194 |
| **BCGDan_2838** | -1.351 | | 0.764 | 0.087 | -0.479 | -1.571 | -1.293 | 1.728 | 1.152 | 0.350 | 0.469 | -1.194 |
| **BCGDan_2854** | 2.736 | 1.759 | -0.589 | -0.436 | 1.997 | 0.471 | -0.472 | -0.381 | -0.308 | | 1.737 | -0.103 |
| **BCGDan_2870** | -0.074 | | 1.406 | -1.256 | -1.515 | -0.793 | -1.479 | 1.316 | 1.306 | 0.806 | 0.582 | -0.572 |
| **BCGDan_2890** | -1.431 | 2.159 | 0.829 | 0.958 | -2.621 | -1.444 | 1.158 | 1.059 | -0.355 | | 0.437 | -1.007 |
| **BCGDan_2900** | -0.181 | 0.077 | -0.446 | 0.799 | 0.125 | -2.054 | 0.209 | 0.264 | -0.282 | | 1.787 | 0.457 |
| **BCGDan_2932** | 0.384 | 1.135 | -0.915 | 0.600 | -0.695 | -1.335 | -2.120 | -1.178 | 0.708 | | 0.240 | 0.573 |
| **BCGDan_2935** | -0.180 | | -0.052 | 0.445 | -0.370 | -0.246 | 1.128 | -0.191 | 0.001 | 2.831 | 2.222 | -0.346 |
| **BCGDan_2985** | 1.745 | | -0.523 | 0.789 | 0.519 | 0.411 | -2.066 | -0.164 | 0.136 | 0.036 | -0.266 | 1.511 |
| **BCGDan_2996** | -0.953 | -0.156 | 0.227 | 1.761 | -0.682 | -0.156 | -0.091 | 1.147 | -0.134 | | -0.233 | -0.378 |
| **BCGDan_3010** | -0.153 | | 0.275 | 0.237 | -0.210 | -0.977 | -1.727 | 0.952 | 0.842 | 0.457 | -0.534 | 0.051 |
| **BCGDan_3048** | 0.000 | | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.223 | 0.000 | 0.000 | 0.000 |
| **BCGDan_3162** | -1.357 | -1.416 | -0.877 | 2.071 | -0.809 | -1.935 | -0.927 | 1.438 | 1.714 | | 0.980 | 3.180 |
| **BCGDan_3195** | 0.060 | -0.132 | 1.348 | 0.556 | -0.487 | -1.665 | 2.984 | 1.937 | -3.229 | | -0.197 | -0.523 |
| **BCGDan_3200** | 0.221 | 1.368 | -0.642 | 0.475 | -0.597 | -0.455 | -1.714 | -1.805 | 0.470 | | 0.961 | 1.383 |
| **BCGDan_3210** | -0.929 | 0.568 | 0.560 | -0.785 | -1.462 | -1.475 | -0.084 | 1.005 | -1.148 | | 0.355 | 0.947 |

| | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **BCGDan_3229** | -1.434 | 0.595 | 1.039 | 0.946 | -3.115 | -0.208 | 0.409 | 1.190 | 0.532 | | -0.676 | -0.556 |
| **BCGDan_3246** | -2.063 | | 0.724 | -0.443 | -0.208 | -2.991 | -0.087 | -0.074 | 1.132 | 0.794 | 1.183 | -1.462 |
| **BCGDan_3254** | -2.434 | | 1.236 | -1.629 | -0.993 | -2.421 | -0.211 | 0.787 | 0.998 | 1.057 | 3.342 | -2.010 |
| **BCGDan_3271** | -0.981 | 0.953 | 0.593 | 0.781 | -1.523 | -1.167 | 0.556 | 0.797 | 3.626 | | -0.252 | -0.851 |
| **BCGDan_3335** | -0.092 | 0.770 | 0.732 | 1.782 | -2.277 | -0.577 | 0.331 | 0.382 | 0.365 | | -0.502 | -1.437 |
| **BCGDan_3346** | -1.215 | -0.974 | 1.051 | -1.596 | 0.571 | -0.113 | 0.424 | -0.426 | 1.475 | | 0.280 | -0.069 |
| **BCGDan_3349** | 0.501 | 2.230 | -0.600 | 0.683 | 0.158 | 0.339 | -0.908 | -1.411 | -1.018 | | 0.776 | 0.508 |
| **BCGDan_3369** | -0.820 | 0.323 | 1.564 | -0.059 | 1.620 | 0.574 | -0.242 | -0.718 | -0.676 | | -0.148 | 0.471 |
| **BCGDan_3370** | -2.083 | -0.173 | 1.063 | -0.555 | -0.083 | -0.222 | 0.628 | 0.624 | 0.758 | | 0.033 | 0.636 |
| **BCGDan_3383** | -1.395 | 1.847 | 0.536 | 0.729 | -1.784 | -0.845 | 0.014 | 0.153 | 3.762 | | 0.372 | -0.722 |
| **BCGDan_3476** | -2.143 | | 1.639 | -1.350 | -2.165 | -0.187 | -0.303 | 1.185 | 0.810 | 1.298 | 1.104 | -1.462 |
| **BCGDan_3516** | 0.540 | -4.614 | -0.663 | 1.306 | -0.013 | -0.976 | -1.103 | -0.641 | 1.263 | | 0.720 | 0.474 |
| **BCGDan_3523** | -2.402 | -0.366 | 1.151 | -0.497 | -0.746 | 0.276 | 1.870 | -1.135 | 0.198 | | 1.059 | -1.376 |
| **BCGDan_3531** | 1.673 | | -0.519 | 0.442 | 0.053 | 0.938 | -2.216 | -2.195 | -1.089 | -0.455 | -0.674 | 1.042 |
| **BCGDan_3543** | 0.164 | | 0.202 | -3.018 | -1.736 | 0.243 | 4.330 | -0.280 | 0.251 | 1.834 | -6.067 | -5.230 |
| **BCGDan_3560** | -1.414 | -1.160 | 0.834 | 2.538 | -0.429 | 0.282 | 0.539 | -0.457 | 2.038 | | -0.612 | -0.361 |
| **BCGDan_3560** | -1.781 | 0.569 | 1.050 | 4.189 | 0.016 | -2.602 | -1.860 | 0.757 | 1.896 | | -2.941 | 0.054 |
| **BCGDan_3587** | -0.995 | 0.345 | 0.480 | 1.467 | -0.339 | -0.355 | -0.231 | 0.792 | 1.273 | | -1.151 | -0.360 |
| **BCGDan_3624** | 0.944 | 0.934 | -0.514 | 0.263 | 0.164 | -0.857 | -1.315 | -0.344 | -0.128 | | 1.615 | 1.988 |
| **BCGDan_3636** | -1.928 | | 1.984 | -1.100 | -1.170 | -1.369 | -0.333 | 1.472 | 1.771 | 1.822 | 1.501 | -2.051 |
| **BCGDan_3679** | 0.662 | | -1.139 | 0.938 | -0.395 | 0.423 | -1.280 | -0.379 | -0.345 | 0.807 | 0.004 | 0.856 |
| **BCGDan_3711** | 1.652 | | 0.659 | -0.625 | 0.790 | 0.241 | 5.437 | -0.096 | -3.325 | 4.163 | -1.399 | -2.006 |
| **BCGDan_3755** | 2.032 | 0.740 | -0.389 | 0.397 | -1.116 | 1.090 | -0.787 | -0.634 | -0.285 | | 0.785 | 0.083 |
| **BCGDan_3784** | 0.413 | -0.005 | -0.553 | 0.397 | 0.650 | -0.977 | -1.936 | -0.158 | -0.211 | | 1.295 | 0.659 |
| **BCGDan_3786** | -1.687 | 1.571 | 0.590 | -1.215 | -1.114 | -0.367 | 0.563 | -2.125 | 1.152 | | 0.343 | -1.156 |
| **BCGDan_3796** | -1.254 | | 0.925 | -1.688 | -1.139 | 1.189 | -1.061 | 1.769 | -0.279 | 1.057 | 0.327 | 0.208 |
| **BCGDan_3832** | -1.945 | -1.216 | 0.649 | 0.511 | -2.472 | 0.329 | 0.103 | 0.804 | 1.092 | | -0.209 | 0.618 |

| | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **BCGDan_3858** | -0.230 | | 0.062 | 0.932 | -0.786 | 0.201 | -0.774 | -0.273 | -1.049 | 0.745 | 0.941 | 0.810 |
| **BCGDan_3862** | -0.620 | 3.244 | -0.067 | 1.761 | 0.224 | -1.092 | 0.431 | -0.689 | -0.339 | | 0.007 | -1.314 |
| **BCGDan_3897** | -0.921 | | -0.358 | 0.222 | -0.674 | -1.841 | -1.263 | 0.707 | 0.033 | 0.284 | 0.590 | 0.065 |
| **BCGDan_3918** | -0.197 | -1.066 | -0.516 | -1.126 | 1.702 | -0.136 | 2.717 | -0.553 | 0.798 | | 0.434 | 0.411 |
| **BCGDan_3961** | 0.000 | 0.000 | 0.000 | 0.000 | 0.899 | 0.000 | 0.488 | 0.336 | 0.000 | | 0.000 | 0.000 |
| **BCGDan_3996** | 0.507 | 0.601 | -0.861 | 1.095 | -0.967 | -1.478 | -1.884 | -0.844 | 0.844 | | -0.323 | 1.049 |
| **BCGDan_3999** | -1.345 | | 1.928 | -3.390 | -0.682 | -2.719 | -0.420 | 2.599 | 1.601 | 2.845 | 0.820 | -0.720 |
| **BCGDan_4013** | -0.819 | -0.301 | -0.358 | 0.079 | 0.309 | -0.126 | 0.641 | -0.548 | 0.659 | | 1.471 | -0.145 |
| **BCGDan_4020** | 1.398 | | -0.535 | 1.269 | 0.302 | 1.022 | -2.069 | -0.502 | -0.430 | 0.463 | -0.296 | 1.282 |
| **BCGDan_4070** | -0.566 | 1.702 | 0.995 | 0.782 | -1.847 | -1.550 | 1.363 | 0.519 | 0.281 | | 0.346 | -0.931 |
| **WT** | -0.705 | -0.029 | -0.234 | 1.030 | 2.327 | 1.558 | -1.290 | 0.723 | 3.178 | | -0.011 | -0.761 |
| **WT** | -0.139 | | -0.391 | 0.027 | 0.319 | 0.256 | -2.134 | 0.159 | 3.984 | 1.243 | -0.344 | -0.950 |