Whole-genome transposon mutagenesis to elucidate the genetic requirements for vitamin B_{12} biosynthesis and assimilation in mycobacteria

Rendani Donald Mbau

MBXREN001

Division of Medical Microbiology

Department of Pathology

A dissertation submitted to the Faculty of Health Sciences,

University of Cape Town,

For the Degree of Doctor of Philosophy,

June, 2021.

Declaration

I declare that this thesis is my own work in design and in execution. It is being submitted for the degree of Doctor of Philosophy at the University of Cape Town. It has not been submitted for any degree or examination at any other university.

Rendani D. Mbau

15 June 2021

Date

Abstract

Comparative genomic analyses have identified an altered capacity for cobalamin biosynthesis as a critical step in the evolution of the pathogenic *Mycobacterium tuberculosis* Complex strains from a common environmental ancestor. However, resolving the full gene complement involved in the complex, multi-step pathway for *de novo* cobalamin biosynthesis, assimilation, and salvage in different mycobacterial species is challenging. A genome-scale approach was adopted to yield detailed genetic maps of *de novo* cobalamin biosynthesis in *M. smegmatis*, a non-pathogenic saprophyte. To this end, a combination of whole-genome transposon (Tn) mutagenesis and next generation sequencing (TnSeq) was applied in *M. smegmatis* Δ*metE*, a gene-deletion mutant in which the cobalamin-independent methionine synthase is inactivated, rendering the cobalamin-dependent isoform, MetH, essential for viability. Following growth of the *metE* mutant in rich medium, genomic DNA was extracted, amplified by PCR, and subjected to high-throughput sequencing to quantify all Tn junctions. Thereafter, the library was cultivated in defined minimal medium to enable identification of all conditionally essential genes – including those required for *de novo* cobalamin biosynthesis. A Δ*metE* library comprising 400,000 individual Tn insertion mutants (cfu/ml) was generated. Of the predicted 6,716 genes in the *M. smegmatis* genome, 213 genes were identified as essential for growth on rich agar while 356, 301, and 337 genes were identified as essential in unsupplemented, cyanocobalamin (CNCbl; vitamin B_{12})-supplemented and cobalt-supplemented Sauton's minimal medium, respectively. A total of 424 genes were identified as essential across all conditions tested with only 10, 13 and 24 genes (ES plus GD) uniquely required for growth in unsupplemented, CNCbl-supplemented and cobalt supplemented Sauton's minimal medium, respectively. On average, predicted cobalamin pathway genes were underrepresented in number of Tn insertions and read counts, indicating the likely essentiality of these genes during growth of the *metE* mutant in minimal medium. Notably, elucidation of cobalamin biosynthetic and assimilatory genes required the analysis of libraries exposed to CNCbl-unsupplemented minimal media for extended durations, probably reflecting the need to exhaust the organism's capacity for co-factor storage and recycling. Utilizing targeted silencing of individual genes by CRISPR interference, candidate cobalamin biosynthesis genes were validated, providing functional evidence of their essentiality for *metE* survival in minimal medium, in turn supporting the validity of the cobalamin biosynthetic pathway constructed from the TnSeq results. In addition, the results add further evidence in support of the functionality of the cobalamin riboswitch upstream of *metE*. This is an important observation as it suggests the

potential to apply an analogous approach in *M. tuberculosis*, a major human pathogen whose ability to synthesize cobalamins remains unresolved. Moreover, elucidating the genetic requirements for optimal growth under specific conditions can inform our basic understanding of mycobacterial physiology and pathogenicity, identifying potential vulnerabilities for novel anti-tuberculosis therapeutics.

Acknowledgements

First of all, I would like to thank the Almighty Lord for granting me countless blessings, strength and knowledge that have enabled me to get this far.

I would like to acknowledge the South African Medical Research Council (SAMRC), the Molecular Mycobacteriology Research Unit (MMRU), as well as the University of Cape Town (UCT) for funding and financial support.

To my Mom, Vho-Masakona Mbau, thank you for your love, care, and everything you have done for me. I would not be where I am without your unfailing support. To my late Father, Vho-John Mbau, thank you for being there for me and for being a wonderful mentor. I appreciate everything you have taught me, that have made me a wonderful man that I am today. To elder siblings, Reuben, Freddy, Florah, Oriel and Shonisani, as well as my late sister Merriam, and to my sister in-laws, Grace and Dobsy you have all showed me a great deal of support, encouragement and inspired me to pursue my dreams. Thank you all for believing in me.

To my Supervisor, Prof. Digby Warner, thank you for giving me the opportunity to work under your wonderful supervision. For your guidance and support I needed to succeed within this research field. It has been a privilege to be part of the MMRU where under your supervision and counsel I grew as a scientist. You inspired me to learn more about this challenging research field. Your guidance made everything look easy.

To my co-supervisor, Prof. Valerie Mizrahi, I cannot express enough gratitude for your tremendous support, your encouragement, as well as your insight throughout the course of my studies.

To Dr. Gabriel Mashabela, thank you for believing in me, motivating, and inspiring me. For the insight you instilled in me regarding biochemistry as well as your Vitamin B_{12} expertise. I cannot thank you enough for your great mentorship and amazing friendship. Dr Raju Mukherjee, thank you for guidance and support as well as your invaluable scientific advice. To Dr. Atica Moosa and Dr. Terry Kipkorir, thank you for insights in Vitamin B_{12} metabolism and molecular techniques.

To Dr. Mandy Mason and Irene Gobe, thank you for your insights in the TnSeq technique. Extra gratitude to Dr. Mandy Mason for your guidance, support and instilling the knowledge needed for sequence analyses. To Dr. Melissa Chengalroyen for critical reading and comments for this thesis. To Dr. Antonina Wasuna, Dr. Krupa Naran and Zela Martin, thank you for training me and introducing me to the tricks and techniques of the laboratory.

Finally, to my colleagues at MMRU, thank you for welcoming me and showing me the ropes around the laboratory. Your support and assistance are very much appreciated.

Table of Contents

List of Figures

List of Tables

List of Abbreviations

5-CH_3-THF	5-methyltetrahydrofolate
ABC (buffer)	Ammonium bicarbonate buffer
ABC	ATP-binding cassette
AdoCba	Adenosylcobamide
AdoCbl	Adenosylcobalamin (coenzyme B_{12})
ALA	5-aminolevulinic acid
ATc	Anhydrotetracycline
BCAA	Branched chain amino acids
BCG	Bacille Calmette-Guérin
bp	base pair(s)
CFU	Colony forming unit
cDNA	Complementary DNA
CNCba	Cyanocobamide
CNCbl	Cyanocobalamin (vitamin B_{12})
CRISPRi	Clustered regularly interspaced short palindromic repeats interference
CRISPRi-Seq	CRISPRi sequencing
CTAB	Cetyltrimethylammonium bromide
dCas9Sth1	Deactivated CRISPR-associated protein-9 nuclease from *Streptococcus thermophilus*
DMB	Dimethylbenzimidazole
DMSO	Dimethylsulphoxide
DNA	Deoxyribonucleic acid
dNTP	Deoxyribonucleotide triphosphate
DOTS	Directly-observed therapy, short course
EDTA	Ethylenediaminetetraacetic acid
EMB	Ethambutol
ESX	Early secretory antigenic target secretion system
et al.	et alibi
FDA	US Food and Drug Administration
g	Gravitational force
hr	Hour(s)

H37Rv	Wild-type laboratory strain of *Mycobacterium tuberculosis*
HIV	Human immunodeficiency virus
ICL	Isocitrate lyase
i.e.	id est
IRIS	Immune Reconstitution Inflammatory Syndrome
INH	Isoniazid
kbp	Kilo base pairs
KD	Knockdown
Kan	Kanamycin
Kan^R	Kanamycin resistant
KO	Knockout
LA	Luria-Bertani agar
LB	Luria-Bertani broth
LTBI	Latent tuberculosis infection
Msm	*Mycobacterium smegmatis*
Mtb	*Mycobacterium tuberculosis*
mc^2155	*Mycobacterium smegmatis* strain mc^2155
MDR-TB	Multi-drug resistant tuberculosis
MeCba	Methylcobamide
MeCbl	Methylcobalamin
Mg	Milligram
min	Minute(s)
ml	Millilitre
mM	Millimolar
MMRU	Molecular Mycobacteriology Research Unit
mRNA	Messenger RNA
MTBC	*Mycobacterium tuberculosis* complex
NaCl	Sodium chloride
NADPH	Nicotinamide adenine dinucleotide phosphate
ng	nanogram
nM	nanomolar
nm	nanometres
OADC	Oleic Albumin Dextrose Catalase
OD_{600} nm	Optical density at 600nanometre wavelength)

ORF	Open reading frame
HOCba	Hydroxycobamide
HOCbl	Hydroxycobalamin
Oligo	Oligonucleotide
PABA	para-aminobenzoic acid
PAM	protospacer adjacent motif
PAS	Para-aminosalicylic acid
PBS	Phosphate-buffered saline
PCR	Polymerase chain reaction
PDIM	Phthiocerol dimycocerosate
PEP	Phosphoenolpyruvate
PEPCK	Phosphoenolpyruvate carboxykinase
PRR	pattern recognition receptor
PZA	Pyrazinamide
RD	Region of difference
RIF	Rifampicin
RNA	Ribonucleic acid
RNI	Reactive nitrogen intermediates
ROI	Reactive oxygen intermediates
rpm	revolutions per minute
rRNA	ribosomal Ribonucleic Acid
SAM	S-adenosyl-L-methionine
sdH_20	Sterile distilled water
SDS	Sodium dodecyl sulphate
sec	Seconds
SEM	standard error of mean
sgRNA	Short guide RNA
SMZ	Sulfamethoxazole
STR	Streptomycin
TB	Tuberculosis
TCA	Tricarboxylic acid
THF	Tetrahydrofolate
TMP	Trimethoprim
Tn	Transposon

TNF	Tumor necrosis factor
TnSeq	Transposon sequencing
Tween	Polyoxyethylene sorbitan monooleate
U	Units
UCT	University of Cape Town
Uro' III	Uroporphyrinogen III
UTR	Untranslated region
UV	Ultraviolet
V	Volt
v/v	Volume per volume
WGS	Whole-genome sequencing
WHO	World Health Organization
WT	Wild-type
XDR	Extensively Drug Resistant
µg	Microgram
µl	Microlitre
µM	Micromolar

Chapter 1: Literature Review

1.1 Tuberculosis

Tuberculosis (TB) is an infectious disease mainly caused by *Mycobacterium tuberculosis* (*Mtb*), a facultative intracellular pathogen that infects via the airborne route (Chaoui et al., 2009). TB is one of the top 10 causes of death worldwide and the leading infectious killer ranking above HIV/AIDS (WHO, 2020). It typically affects the lungs (pulmonary TB), where *Mtb* bacilli infect alveolar macrophages and avoid elimination by interfering with innate antimicrobial host defence mechanisms (Ramakrishnan, 2012; Russell et al., 2009; Vrieling et al., 2020). However, TB can also affect other sites, with extrapulmonary TB accounting for at least 8-24% of new infections annually (WHO, 2020). These different disease loci imply differences in metabolic interactions between host and bacilli in different microenvironments. On entry into host alveolar macrophages, *Mtb* and other intracellular pathogens are enclosed in endocytic vacuoles called phagosomes (Clark-Curtiss & Haydel, 2003). If phagosome-lysosome fusion occurs to form a phagolysosome loaded with lysosomal enzymes, the contained bacteria are exposed to a hostile environment characterized by gradual acidification and reactive oxygen (ROI) and nitrogen intermediates (RNI) that target cellular and biochemical structures (Hestvik et al., 2005; Smith, 2003). However, *Mtb* has evolved mechanisms that allow survival, escape, growth, and dissemination within the phagosome. These include prevention of the phagosomal maturation cycle (Jayachandran et al., 2007; Malik et al., 2000; Vergne et al., 2005), resistance to ROI and RNI (Miller et al., 2004) and prevention of apoptosis (Hickman et al., 2002; Sly et al., 2003). As a result, the bacilli avoid eradication and continue to proliferate within phagosomes until phagosomal lysis occurs (Grosset, 2003; Sun et al., 2015).

Despite advances in TB research, *Mtb* infections, exacerbated by the Human Immunodeficiency Virus (HIV) co-pandemic, present a large healthcare problem globally, particularly in developing countries (Dheda et al., 2014; Dheda et al., 2017; Gandhi et al., 2006; Zuniga et al., 2012). It is estimated that about one quarter of the world's population - an estimated 1.7 billion people - is latently infected with *Mtb* and thus at risk of developing active TB disease (WHO, 2020). Moreover, the highest incidence of TB is associated with HIV infection (Corbett et al., 2003; Ducati et al., 2006), and remains a significant public health threat (Bruchfeld et al., 2015; Dheda et al., 2014; Gandhi et al., 2006). This co-morbidity is

contributed by socioeconomic and biological factors. These include poor nutrition, smoking, alcohol consumption, drug abuse, diabetes mellitus and vitamin D deficiency.

In 2019, new cases of TB disease were estimated at 10 million people (WHO, 2020), which is equivalent to 130 new cases per 100 000 population per year. Moreover, most TB cases were in South-East Asia (44%) and the African region (25%). The Western Pacific accounted for 18% of the world's cases, with smaller percentages in the Eastern Mediterranean at 8.2%, the Americas and Europe at 2.9% and 2.5% new cases (WHO, 2020). In 2019, there were an estimated 1.2 million deaths from TB disease and 208 000 deaths among people living with HIV (WHO, 2020). In addition, India, China, Indonesia, Philippines, Pakistan, Nigeria, Bangladesh and South Africa accounted for two thirds of the global TB incidence in 2019 (WHO, 2020). **Figure 1.1** shows the geographic distribution of new TB cases globally.

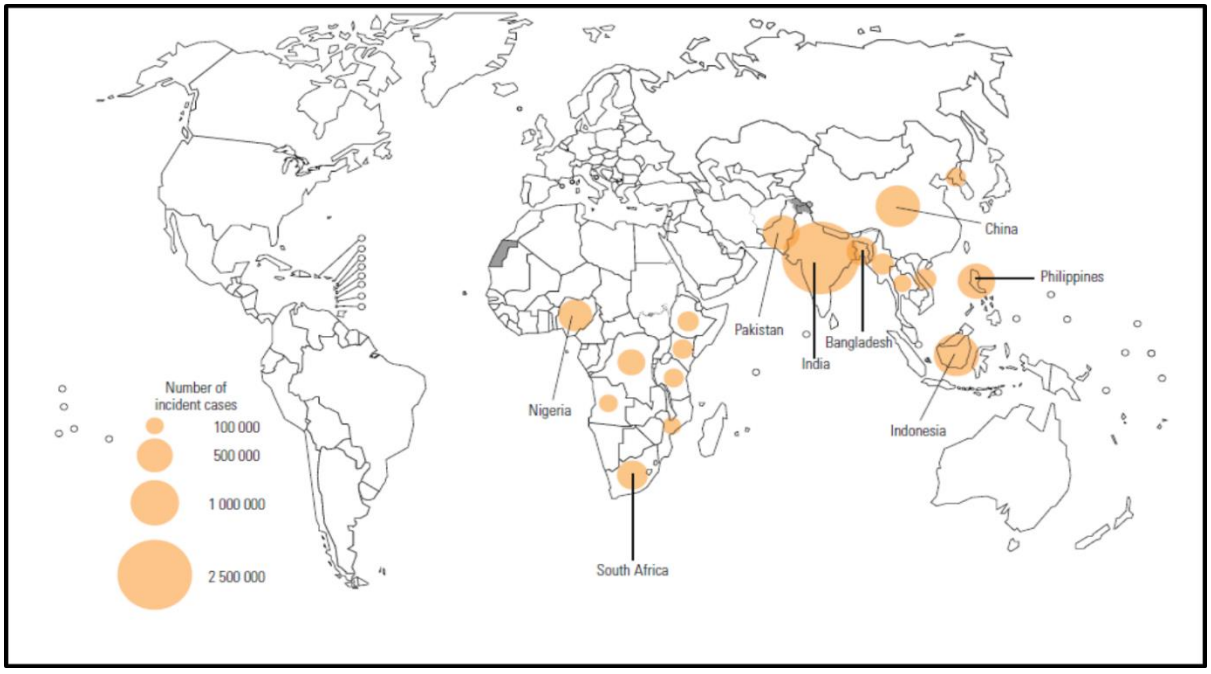


Figure 1.1: Estimated TB incidence in 2019, for countries with at least 100 000 incident cases (WHO, 2020)**.**

1.2 Disease pathology

One-quarter of the world's population is thought to be latently infected with *Mtb*, a condition where individuals are infected with this highly successful intracellular pathogen and rarely leads to active TB disease but are at risk for reactivation (Drain et al., 2018; Lin & Flynn, 2018; VanderVen et al., 2016). *Mtb* is transmitted through inhalation of aerosolized bacilli via the respiratory tract (Ramakrishnan, 2012; Russell et al., 2009; Vrieling et al., 2020). Once inhaled, *Mtb* encounters an innate antimicrobial defence consisting of professional phagocytic cells; macrophages, neutrophils, dendritic cells (DCs) and monocytes (Berry et al., 2010; Blankley et al., 2016; de Martino et al., 2019). Macrophage, neutrophils, natural killer cells (NK), DCs, complement system and mast cells are the key players of innate immune response (de Martino et al., 2019).

Mtb interact with alveolar macrophages through a variety of phagocytic receptors (Ernst, 1998; Villeneuve et al., 2005). These include complement receptors (CR1, CR3, and CR4), Fc receptors (FcR), mannose receptors (MR), surfactant protein A receptor and toll-like receptor (TLR) (Boritsch et al., 2016; Lukácsi et al., 2020; Queval et al., 2017; Satoh & Akira, 2017; Villeneuve et al., 2005). Some studies have suggested that ingestion of *Mtb* through certain receptors determine macrophage response and the outcome of the infection. For instance, ingestion of *Mtb* through FcR triggers the generation of ROI and phagosomal-lysosomal fusion, whereas ingestion through predominant complement receptor 3 (CR3) or MR prevents the oxidative burst and phagosomal maturation (Kang et al., 2005; Lugo-Villarino et al., 2011; Mortaz et al., 2015; Sakamoto, 2012), to promote the survival of *Mtb* bacilli during latent infection. If innate immunity succeeds in eradicating the bacilli rapidly, the infection aborts (Forrellad et al., 2013). However, if *Mtb* avoids eradication, more macrophages are infected and the bacilli reproduce inside the cells, initially causing few, if any, clinical manifestations (de Martino et al., 2019; Forrellad et al., 2013). The development of active TB or latent TB infection (LTBI) and/or reactivation of LTBI to active TB depends on the complex host-pathogen complex interaction (Boom et al., 2021; Cadena et al., 2017; de Martino et al., 2019). HIV and type-2 diabetes can accelerate the establishment of infection as well as the progression to active TB.

Previously, *Mtb* was thought to lack classical virulence factors such as toxins (Forrellad et al., 2013; Gordon et al., 2009; Mukhopadhyay et al., 2012). Instead, mycobacterial pathogenicity was believed to depend only on down-regulation of host immune modulators, modulation of

lipid metabolism, the prevention of phagosome-lysosome fusion, metal ion transporters, proteases, and inhibition of the antimicrobial mechanisms of macrophages such as cytokines-mediated host defences and inhibition of antigen presentation, interfering with immune recognition (de Martino et al., 2019; Forrellad et al., 2013; Harding & Boom, 2010; Hmama et al., 2015). However, the outer membrane protein CpnT of *Mtb* was discovered as the main cytotoxicity factor of the bacilli in macrophages. CpnT protein consists of the N-terminal domain that is used for uptake of nutrients and a secreted toxic C-terminal domain known as the tuberculosis necrotizing toxin (TNT) (Danilchanka et al., 2014; Sun et al., 2015). TNT contain NAD^+-glycohydrolase activity, which depletes the cellular NAD^+ pool and resulting in the activation of the necroptosis pathway, leading to host cell death (Sun et al., 2015; Tak et al., 2019). Among the strategies employed by *Mtb* are the inhibition of intracellular trafficking, inhibition of macrophage inflammatory responses, the inhibition of autophagy, the induction of necroptosis, leading to host cell death, the neutralization of virulent components such as ROI, and the prevention of antigen presentation that drives the recruitment and activation of various immune cells types to the site of infection, including tumor necrosis factor (TNF) and interleukin (IL)-12 (de Martino et al., 2019; Lerner et al., 2015). *Mtb* manages to avoid phagosome-lysosome fusion through the ESX-1 secretion system (Conrad et al., 2017) and, therefore, to block phagosomal maturation via nucleotide diphosphate kinase (Npk), which prevent lysosomal trafficking and NADPH-oxidase activity (Mihret, 2012; Sun et al., 2013). *Mtb* continues to multiply intracellularly until phagosomal lysis occurs and spreads inside the lymph nodes. Chemokine and inflammatory cytokines released by lysed alveolar macrophages stimulate the recruitment of different populations of cells, including more monocytes-derived macrophages and DCs to the infection site (van Crevel et al., 2003). During this time, activated DCs traffic to the lymphoid organs where they prime naïve antigen-specific T cells to initiate the adaptive immunity against mycobacterial antigens (Behar et al., 2011; Tufariello et al., 2003). Furthermore, T lymphocytes undergo a process of activation and expansion of the specific populations for the *Mtb* antigens, to strengthen antimycobacterial mechanisms in macrophages, including the production of ROI and RNI which further slows down *Mtb* replication (Russell, 2011) and recruiting additional immune cells to the infection site (de Martino et al., 2019).

Following the development of adaptive immunity, a complex and well-coordinated mechanism is established between innate and adaptive immunity to form an organized structure called granuloma, a wall of macrophages and highly differentiated immune cells such as

multinucleated giant cells, epithelioid cells and Foamy cells, intended to constrain the infection, which ultimately traps *Mtb* in a hypoxic and lipid-rich environment (de Martino et al., 2014; Gengenbacher & Kaufmann, 2012; Russell, 2007). This process occurs in at least 90% of infected individuals and leads to LTBI. During LTBI, the bacilli can enter a phase of non-replicating persistence (Tailleux et al., 2008). Latent TB progresses to active TB in fewer than 10% of infected individuals when, for various reasons, a condition of immunosuppression develops (de Martino et al., 2019). If the host fails to restrain *Mtb* growth, bacilli can eventually break out of the caseous granuloma resulting in granulomatous lesions with more necrotic macrophage death and increased inflammatory cell recruitment (Liu et al., 2017).

1.3 TB Chemotherapy

A major breakthrough in the fight against TB occurred between 1906 and 1919 when Calmette and Guérin developed the attenuated vaccine strain, Bacille Calmette-Guérin (BCG) (Calmette et al., 1921). Although the vaccine has variable efficacy in adults, it does prevent TB in children (Eum et al., 2010). The other major breakthrough in the management of TB was the discovery of streptomycin as the first anti-tuberculosis drug in 1944, then followed by the discovery of isoniazid (INH), rifampicin (RIF), and pyrazinamide (PZA), in 1952, 1957, and 1980, respectively (Sakamoto, 2012; Van Scoy & Wilkowske, 1999). Notwithstanding these early successes, the lack of a completely protective TB vaccine, the need for prolonged TB treatment, the slow development of new anti-tuberculosis drugs and the emergence of multidrug-resistant (MDR) and extensively drug-resistant (XDR) strains of *Mtb* continues to impose a huge toll on public health programs globally (Sakamoto, 2012).

TB is a curable disease: drug-sensitive TB is effectively controlled using current first-line anti-TB drugs (INH, RIF, PZA, and ethambutol (EMB)) which are given as part of the Directly Observed Therapy, Short Course (DOTS) strategy (Chetty et al., 2017; Evans & Mizrahi, 2018). The DOTS-based regimen was launched in the early 1990s to address the operational problem that effective TB treatment requires a prolonged 6-9-month treatment with a combination of the first line anti-TB drugs. However, despite the global implementation of DOTS-based TB control programmes, the current regimen is inadequate for totally eradicating TB and the limited effectiveness of current TB chemotherapy is evidenced by the emergence and spread of MDR and XDR strains (Dheda et al., 2017). Current treatment regimens for TB require combinations of several drugs (used both to increase efficacy and to prevent the

emergence and the spread of resistant strains) (Laurenzi et al., 2007), ranging from 6 months regimen for drug-susceptible TB to typically 9–20 months for MDR-TB (WHO, 2020) and possibly longer if there is additional drug resistance. Moreover, treatment for MDR-TB and XDR-TB require prolonged administration of second-line anti-TB drugs (Bloom et al., 2017; Esmail et al., 2018).

Second-line drugs, which consists of the fluoroquinolones (ciprofloxacin, levofloxacin moxifloxacin, ofloxacin, and gatifloxacin), the three injectable drugs (amikacin, kanamycin and capreomycin) and others, such as ethionamide, D-cycloserine and p-aminosalicylic acid, are given, as part of the so-called "DOTS plus" programme (WHO, 2017) to individuals who fail to respond to first-line treatment. The DOTS plus programme has mainly focused on the management and treatment of MDR-TB, specifically in developing countries with high HIV co-infections. MDR-TB treatment is considerably more complicated as some drugs were administered by injection, is much more expensive and more toxic with serious side effects; in combination, these factors result in lower cure rates and increased transmission of the TB disease (Espinal & Dye, 2005; Marahatta, 2010). However, recently, the three-oral drug regimen known as BPaL, consisting of bedaquiline, pretomanid and linezolid received FDA approval for the treatment of MDR-TB and XDR-TB cases. The BPaL regimen was shown to be highly effective on XDR-TB and complicated MDR-TB (Conradie et al., 2020). Co-infection with HIV, and the emergence of drug resistance, have amplified the problem in the developing world (Bruchfeld et al., 2015).

The main challenges in TB treatment are the prolonged administrations and complexity of drug regimens, both of which affect patient adherence; toxic side-effects, especially for the drugs used to treat drug-resistant TB, and the lack or limited availability of paediatric drug formulations for second-line treatment, are additional confounding factors (Bloss et al., 2010; Falzon et al., 2013; Warner & Mizrahi, 2006; Zheng & Av-Gay, 2016). TB chemotherapy for people living with HIV is further complicated by drug–drug interactions between anti-TB drugs and antiretroviral treatment, and by cumulative drug toxicities that amplify the risk of immune reconstitution inflammatory syndrome (IRIS) (Bruchfeld et al., 2015; Pepper et al., 2007). There is, therefore, a pressing need for regimens that are more effective, more affordable and nontoxic, and that shorten the duration of treatment. However, other factors such as socioeconomic and demographic factors like income constraints, poor hygiene, overcrowding,

poor ventilation and malnutrition, lack of social support, and stigmatization also contribute to the persistence of TB (Creswell et al., 2011; Duarte et al., 2018; Lönnroth et al., 2010).

1.4 Global epidemiology of drug-resistant TB

About 3.3% of new TB patients and 18% of previously treated cases in 2019 had MDR-TB or rifampicin-resistant TB (RR-TB) (WHO, 2020), with the highest proportions in countries of the former Soviet Union. **Figure 1.2** Shows the geographic distribution of new TB cases with MDR/RR-TB globally. In 2019, there were about half a million new cases of rifampicin-resistant TB (of which 78% had MDR-TB) (WHO, 2020). In addition, India, China and the Russian Federation accounted for the largest share of TB burden globally. Overall, there were an estimated 465 000 incident cases of MDR/RR-TB in 2019 and they were about 182 000 deaths from MDR/RR-TB (WHO, 2020).

The emergence of drug-resistant strains of *Mtb* is a global problem posing a threat to TB-control programmes in both developing and industrialised countries (Dheda et al., 2017; Kaufmann et al., 2014). Drug resistance in *Mtb* occurs through the acquisition of chromosomal mutations within drug target genes or genes encoding proteins involved in drug uptake or activation of prodrugs. The majority of these are detectable as single nucleotide polymorphisms (SNPs) (Sandgren et al., 2009), gene duplications, or chromosomal rearrangements (Farhat et al., 2016; Gevers et al., 2004; Witney et al., 2016; Zhang, 2003) which lead to strain variation and antibiotic resistance. The emergence of clinical drug resistance in TB is categorized as *acquired resistance* – which refers to drug resistance in patients with a prior history of TB treatment and it is often due to poor adherence, poor absorption or limited availability of drugs and the addition of single agents to the regimens of patients who are already failing treatment (Kochi et al., 1993), or as *primary resistance* – which refers to the presence of resistance to one or more anti-TB drugs in a treatment-naïve patient (Johnson et al., 2007). The emergence of MDR-TB, which is defined as disease that is resistant to at least INH and RIF; and XDR-TB, defined as resistance to at least INH, RIF, any fluoroquinolone and to any of the three injectable SLDs further complicates the efficacy of treatment of the disease (Abubakar et al., 2013; Dheda et al., 2017).

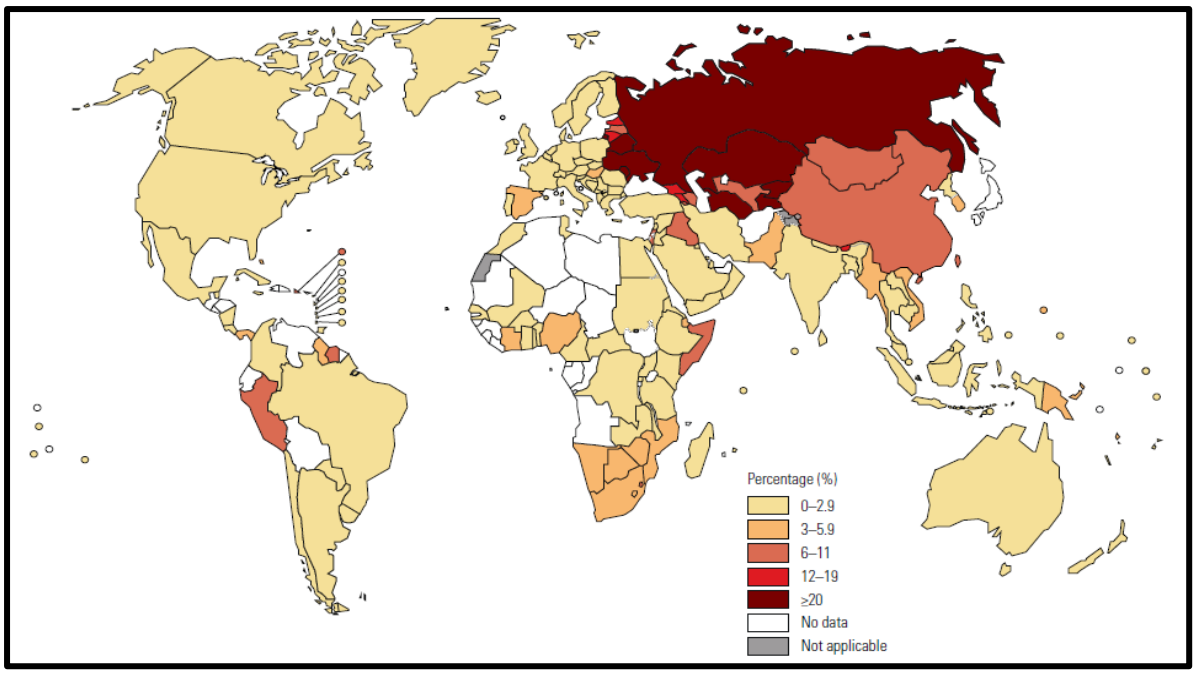


Figure 1.2: Percentage of new TB cases with MDR/RR-TB (WHO, 2020)**.**

1.5 Carbon metabolism in *Mtb*

The success of *Mtb* as a facultative intracellular pathogen depends on the ability of the bacilli to replicate and persist in hostile microenvironments within the human host for long periods of time. To do so, the bacillus possesses a flexible metabolic arsenal, which allows it to acquire and metabolise nutrients from surrounding host tissue (Ehrt et al., 2018; Mashabela et al., 2019; Warner, 2015). *Mtb* has the ability to oxidize a variety of carbon substrates, including carbohydrates, alcohols, fatty acids and tricarboxylic acids (de Carvalho et al., 2010; Ehrt et al., 2018; Muñoz-Elías & McKinney, 2006; Wheeler & Blanchard, 2005). Even though *Mtb* can utilise different carbon sources for growth *in vitro*, lipids have emerged as a major source of carbon for growth during intracellular infection (Cole et al., 1998; Muñoz-Elías & McKinney, 2006; Passemar et al., 2014; Russell et al., 2010; Schnappinger et al., 2003; Timm et al., 2003; Vromman & Subtil, 2014; Williams et al., 2011). Transcriptional profiling studies have demonstrated that genes involved in fatty acid metabolism, and together with those of the glyoxylate cycle, are upregulated during infection of macrophages (Schnappinger et al., 2003) and mice (Dubnau et al., 2005; Timm et al., 2003). However, other studies have revealed that transcriptome signatures of *Mtb* in human granulomas can differ, depending on the site of infection (Fenhalls et al., 2002; Rachman et al., 2006; Teng et al., 2017; Timm et al., 2003).

The β-oxidation cycle and the glyoxylate shunt are the dominant pathway utilized in bacteria for fatty acid catabolism (Marrero et al., 2010; McKinney et al., 2000; Muñoz-Elías & McKinney, 2006; Wilburn et al., 2018). The glyoxylate shunt is essential for carbon anaplerosis in the tricarboxylic acid cycle (TCA) and facilitates oxidation of fatty acids into carbohydrates by bypassing the two oxidative decarboxylation reactions of the TCA cycle during growth on C_2 substrates such as fatty acids, which are the only abundant carbon sources in mammalian tissues (Wheeler & Ratledge, 1994). The glyoxylate shunt consists of two enzymes, isocitrate lyases (ICL) and a malate synthase (MS), encoded by *glcB* and is widespread among prokaryotes and plants, but absent in vertebrates (McKinney et al., 2000). Furthermore, the *Mtb* genome was initially thought to contain a cluster of ~250 genes involved in the five enzymatic reactions of β-oxidation (Cole et al., 1998). However, many studies have demonstrated that many of these putative lipid oxidation enzymes are involved in lipid biosynthesis (Krithika et al., 2006; Trivedi et al., 2004) and function as multiprotein complexes or are not involved in the β-oxidation reactions (Casabon et al., 2013; Thomas & Sampson, 2013; Yang et al., 2014; Yang et al., 2015). However, *Mtb* still possesses a huge number of genes that are involved in fatty acid β-oxidation, although the exact functions of these putative proteins remain uncharacterized (Wilburn et al., 2018). In contrast, *Escherichia coli* only has one enzyme set for each of the β-oxidation pathway reactions under either anaerobic or aerobic conditions (Muñoz-Elías & McKinney, 2006). It has been suggested that the gene redundancy observed in *Mtb* allows the bacillus to shift its metabolism to utilise the available carbon sources encountered in the various microenvironments in the host, an important adaptation strategy to ensure growth and survival during chronic infection (Williams et al., 2011). However, this remains to be proven, together with the identity of specific type of fatty acids available to *Mtb* within the differing host environment (Griffin et al., 2012). Furthermore, *Mtb* genome contains ~80 genes (Van der Geize et al., 2007) dedicated to the complex processes of cholesterol uptake, catabolic pathway and regulations (Capyk et al., 2009; Casabon et al., 2014; Crowe et al., 2017; Frank et al., 2014; Griffin et al., 2011; Griffin et al., 2012; Ho et al., 2016; Lu et al., 2015; Wipperman et al., 2014; Yang et al., 2015). Transposon mutagenesis studies have classified the majority of β-oxidation genes as non-essential *in vitro*, possibly due to extensive gene redundancy (DeJesus et al., 2017; Griffin et al., 2011). Catabolism of saturated fatty acids and lipids through β-oxidation produce acetyl-CoA, which feeds into the TCA cycle. This generates succinate which can be converted into glucose (Muñoz-Elías & McKinney, 2006; zu Bentrup & Russell, 2001). The *Mtb* genome encodes two enzymes that are required to utilize acetyl-CoA derived from fatty acids: isocitrate lyase (ICL) (encoded by

two genes, *icl1* and *icl2*) and the gluconeogenic rate-limiting enzyme, phosphoenolpyruvate carboxykinase (PEPCK) (encoded by *pckA*) (Eoh & Rhee, 2014; Marrero et al., 2010; Muñoz-Elías et al., 2006).

In addition to acetyl-coA, β-oxidation of cholesterol and odd-chain fatty acids produces propionyl-CoA (Griffin et al., 2012; Thomas et al., 2011; Yang et al., 2009), which is metabolized via methylmalonyl-CoA and methylcitrate pathways (Muñoz-Elías et al., 2006; Savvi et al., 2008; Upton & McKinney, 2007) or through incorporation into methyl-branched lipids in the cell wall (Lee et al., 2013). The methylmalonyl coenzyme A route requires the cobalamin-dependent enzyme, (*R*)-methylmalonyl-CoA mutase (Savvi et al., 2008). The methylcitrate cycle appears to be the dominant route for propionyl-CoA metabolism in bacteria, that consists of two specific enzymes, methylcitrate synthase and methylcitrate dehydrogenase, encoded by *prpC* and *prpD*, respectively (Claes et al., 2002; Horswill & Escalante-Semerena, 1999; Textor et al., 1997). The methylcitrate cycle involves the cleavage of methylisocitrate to succinate and pyruvate by methylisocitrate lyase, as end-products of catabolic pathways, which enter the TCA cycle via the glyoxylate shunt (Eoh & Rhee, 2014; McKinney et al., 2000), analogous to the reaction catalysed by ICL in the glyoxylate cycle (**Figure 1.3**).

Alternatively, propionyl-CoA may undergo α-carboxylation via the methylmalonyl pathway, which is converted to methylmalonyl-CoA, which is further converted to succinyl-CoA by the cobalamin-dependent methylmalonyl-CoA mutase (encoded by *mutAB*) (Savvi et al., 2008). Additionally, the methylmalonyl pathway is a major source of methylmanonyl-CoA precursor for the synthesis of methyl-branched lipids (thus, propionyl-CoA products can be incorporated into cell wall lipids and detoxified) such as the surface-exposed lipids, phthiocerol dimycocerosates (PDIMs), sulfolipid (SL)-1 and polyacylated trehaloses (PATS) (Yang et al., 2009) (**Figure 1.3**), which are associated with mycobacterial virulence (Lee et al., 2013; Lovewell et al., 2016; Russell et al., 2010), highlighting the potential importance of cobamides during host infection (Sokolovskaya et al., 2020).

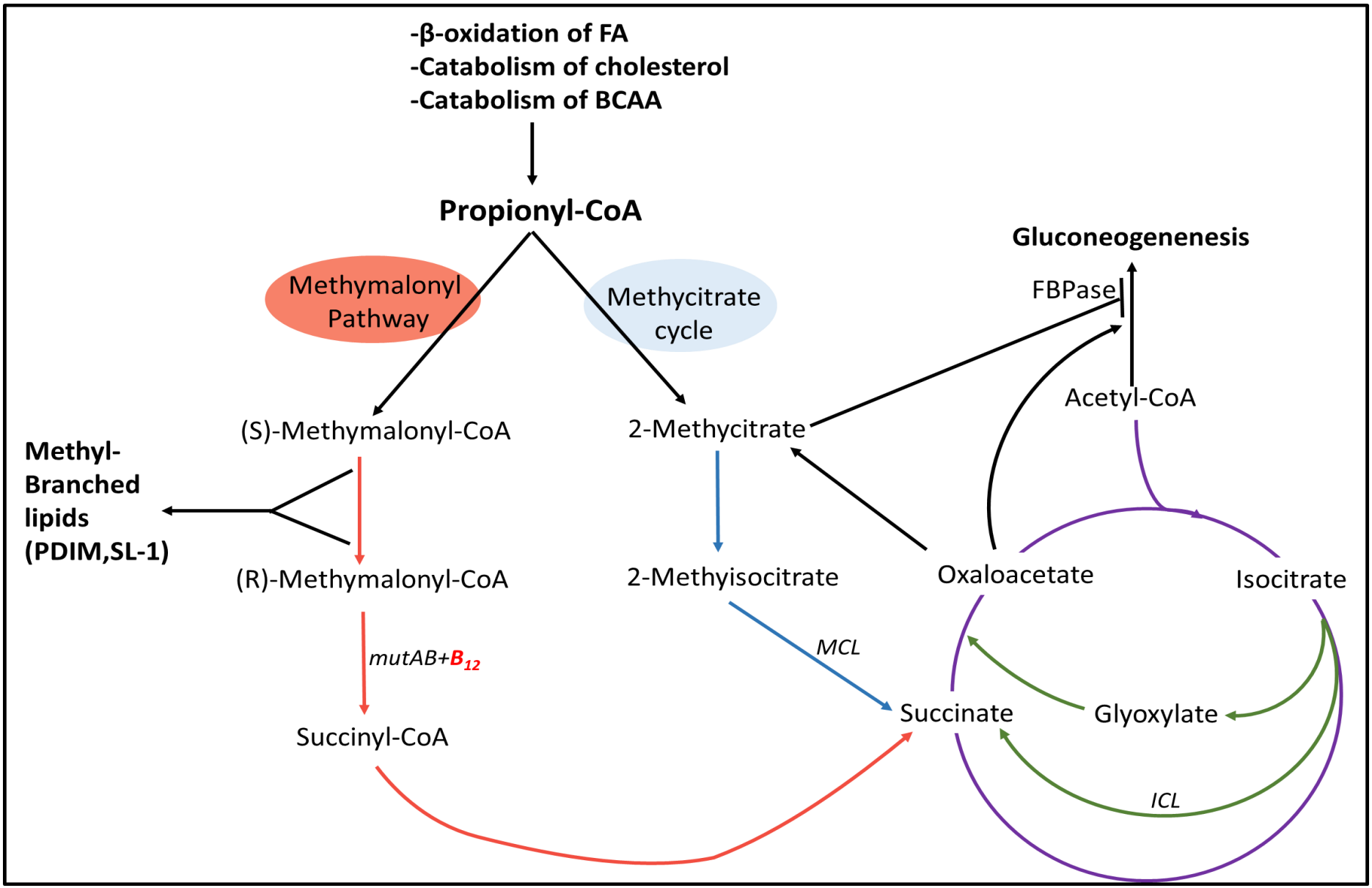

-β-oxidation of FA
-Catabolism of cholesterol
-Catabolism of BCAA
Propionyl-CoA
Methymalonyl Pathway
Methycitrate cycle
Gluconeogenenesis
FBPase
Acetyl-CoA
(S)-Methymalonyl-CoA
2-Methycitrate
Methyl-Branched lipids (PDIM,SL-1)
(R)-Methymalonyl-CoA
2-Methyisocitrate
Oxaloacetate
Isocitrate
mutAB+B12
MCL
Succinate
Glyoxylate
Succinyl-CoA
ICL

Figure 1.3: Propionate metabolism in *Mtb*. Catabolism of alternative carbon sources including odd- and branched-chain FAs, branched chain amino acids (BCAAs), and cholesterol generates propionyl-CoA as a three-carbon (C3) terminal product. Propionyl-CoA can be oxidised to succinate through either the methylcitrate and glyoxylate cycles or the methylmalonyl pathway. *Mtb* ICLs encode a bifunctional enzyme: isocitrate lyase activity (green line) and methylisocitrate lyase activity (blue line) (Gould et al., 2006; Muñoz-Elías et al., 2006). TCA cycle (purple line); glyoxylate shunt (green line); methylcitrate cycle (blue line); B_{12}-dependent methylmalonyl-CoA pathway (red line); gluconeogenic pathway (black line). Secondary accumulation of 2-Methlycitrate allosterically inhibits the gluconeogenic enzyme, fructose 1,6-bisphosphatase (FBPase).

1.6 Cobamide

Cobamides are structurally diverse cobalt-containing cofactors, the most familiar of which is vitamin B_{12} (also known as cyanocobalamin), initially discovered in 1926 as anti-pernicious anaemia factor by Minot and Murphy (1926). Twenty years later, it was purified from liver and kidney by Rickes et al. (1948) and Smith (1948) as a dark-red crystalline compound and was Vitamin B_{12}. Its unique corrin ligand structure was revealed by Dorothy Hodgkin using X-ray crystallography (Hodgkin et al., 1956; Raux et al., 2000). Together with Lenhert, she later revealed one of the biologically active forms structure of cobalamin, adenosylcobalamin (AdoCbl) involved in rearrangement or reductase reactions (Lenhert & Hodgkin, 1961). Although most widely appreciated for their role in human health, cobalamin and other cobamides also fulfil critical functions in microbial communities such as in the catabolism of various carbon sources, nucleotide biosynthesis, and natural product biosynthesis (Banerjee & Ragsdale, 2003; Bridwell-Rabb & Drennan, 2017; Roth et al., 1996; Shelton et al., 2019; Sokolovskaya et al., 2020).

Cobamides are cofactors in various important enzymatic reactions and microbes are the only natural sources of cobamide-derivatives, even though organisms across all domains of life depend on these cobalt corrinoids (Kräutler, 2005; Nielsen et al., 2012). Given that many organisms depend on surrounding species to synthesize this nutrient, cobamide-dependent interactions are well described (Sokolovskaya et al., 2020). Cobamides are corrinoids that contain an upper and lower ligand (Crofts et al., 2013) and are grouped into three chemical classes according to the structure of their lower ligand: benzimidazolyl, phenolyl and purinyl cobamides – containing benzimidazoles, phenolics and purines, respectively (**Figure 1.4**). Different cobamides with structural variability in the lower ligand have been described (Crofts et al., 2013; Renz, 1999). The structural diversity of lower ligands is important because, in most cases, cobamide-dependent metabolism varies on the basis of lower ligand structure, and

the sets of cobamides that support growth or enzymatic activity in different organisms are distinct (Helliwell et al., 2016; Shelton et al., 2019; Sokolovskaya et al., 2019; Yan et al., 2016).

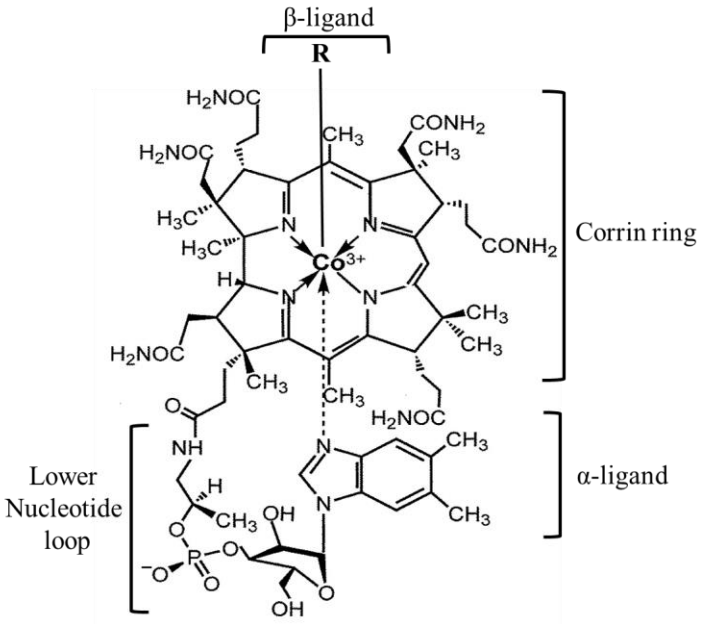
1.6.1 Structure of Cobamide

Cobamides constitute one of the largest cofactor families in biology, and are derived from the common precursor for all naturally occurring tetrapyrroles, uroporphyrinogen III (UroIII) together with other prosthetic groups such as chlorophyll, heme, siroheme, and coenzyme F_{430} (Randaccio et al., 2010; Raux et al., 1998; Rodionov et al., 2003; Warren et al., 2002). Cobalamins are structurally complex, consisting of a modified tetrapyrrole with centrally chelated cobalt ion coordinated equatorially by the nitrogen atoms of a cyclic tetrapyrrole known as the corrin ring, which is held in place by the lower base 5,6-dimethylbezimidazole (DMB) and an upper alkyl group comprising a cyano ligand (cyanocobalamin (CNCbl)) which can be replaced by either an adenosyl group, to give AdoCbl, a methyl group, to give methylcobalamin (MeCbl), or a hydroxyl group (OHCbl) in aquacobalamin (Raux et al., 1999) (**Figure 1.4A**). CNCbl is industrially produced form of vitamin B_{12}, that contain cyano group occupying the upper ligand as a result of the extraction and purification procedure (Martens et al., 2002). The lower axial ligand of cobalamin is the purine analogue base which is attached to the corrin ring by a phosphodiester bond between an aminopropanol substituent of the ring and the phosphoryl moiety of the DMB-riboside monophosphate (Lenhert & Hodgkin, 1961). DMB can either coordinate via the imidazole N atom to the lower (R)-axial site of Cobalamin ("base-on" configuration) or remain uncoordinated ("base-off" configuration). It was discovered that during the MeCbl-dependent methionine synthase reaction, DMB is replaced by the imidazole group of a histidine (His) residue from a side chain of the metalloenzyme, the so-called "base off/His-on" conformation (Stupperich et al., 1990). It is now known that this conformation occurs in several cobalamin-dependent enzymes, including methylmalonyl Co-A mutase and glutamate mutase (Drennan et al., 1994; Zelder et al., 1995) (**Figure 1.4C**).

In organisms that utilize a different lower base (Gruber et al., 2011), DMB is replaced with adenine, substituted benzimidazoles that include hydroxy or methoxybenzimidazole, or phenolic compounds (Crofts et al., 2013; Smith et al., 2018) (**Figure 1.4B**). Naturally occurring cobalamin analogues have either an methyl or adenosyl group occupying the upper β-ligand: the adenosyl group is found in cobalamin analogues involved in radical reactions such as carbon rearrangement or reductase reactions and deaminations (Halpern, 1985; Marsh, 1999),

while the methyl group is found in the cofactor that is involved in cobalamin-dependent methionine synthesis and dehalogenation reactions (Dassanayake et al., 2013; Dowling et al., 2016; Koutmos et al., 2009; Payne et al., 2015).

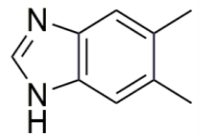
A



R-groups:	
CH_3 (Methly)	= methylcobamide
5'-deoxyadenosine	= adenosylcobamide
OH (Hydroxy)	= hyroxocobamide
CN (Cyano)	= cyanocobamide

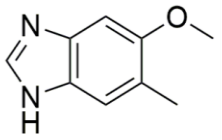
B

(i)



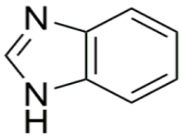
5,6-dimethylbenzimidazole
Cobalamin, B_{12}

(ii)



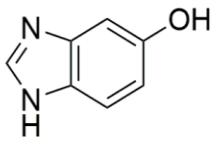
5-methoxy-6-methylbenzimidazole
[5-OMe-6-MeBza]Cba

(iii)



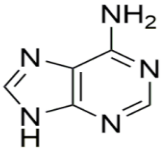
Benzimidazole
[Bza]Cba

(iv)



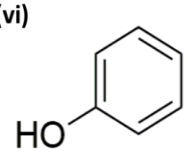
5-hydroxybenzimidazole
[5-OHBza]Cba

(v)



Adenine
[Ade]Cba pseudocobalamin,
pseudo-B12

(vi)



Phenol
[Phe]Cba

C

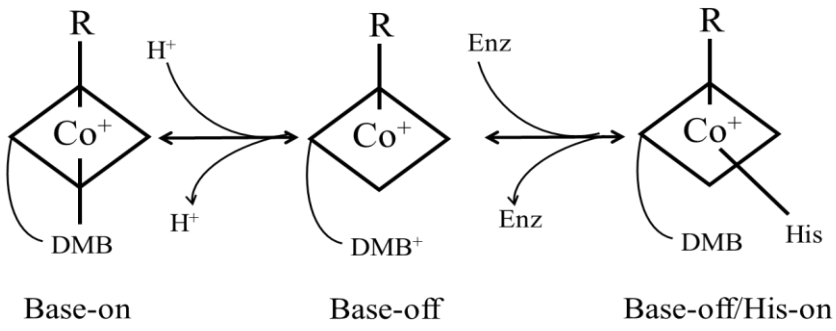


Figure 1.4: The chemical structure of a cobalamin. (**A**) Cobalamin structure comprises central corrin ring, which houses a centrally-chelated cobalt ion, coordinated by an alkyl upper ligand (R-group) and a lower dimethlybenzimidazole (DMB) ligand. The R-group can be occupied by cyano- (cyanocobalamin), adenosyl- (adenosylcobalamin), methyl- (methylcobalamin) or hydroxyl- (hydroxycobalamin) groups (Martens et al., 2002; Moore et al., 2013; Warren et al., 2002). (**B**) Structures of cobamide lower ligands. The name of each compound and the abbreviation for the corresponding cobamide are given below each structure. (i) DMB (5,6-dimethylbezimidazole) (Campbell et al., 2006; Lamm et al., 1982; Martens et al., 2002; Rickes et al., 1948); (ii) [5-OMe-6-MeBza]Cba (5-methoxy-6-methylbenzimidazole), (Stupperich & Kräutler, 1988); (iii) [Bza]Cba (benzimidazole), (Crofts et al 2013); (iv) [5-OHBza]Cba (5-hydroxybenzimidazole), (Krautler et al., 1987; Lezius & Barker, 1965; Pol et al., 1982; Stupperich & Kräutler, 1988); (v) [Ade]Cba (adenine) aka pseudocobalamin, pseudo-B_{12}, (Barker et al., 1958; Hoffmann et al., 2000; Keck & Renz, 2000; Santos et al., 2007; Stadtman, 1960; Stupperich & Kräutler, 1988; Tanioka et al., 2009); (vi)[Phe]Cba (phenol), (Stupperich, 1989), (**C**) Illustration of three conformations of cobalamin: Base-on (DMB coordinated), Base-off (DMB not coordinated), Base-off/His-on (enzyme histidine coordinated).

1.7 Cobamide biosynthesis

Cobamide is an essential nutrient that is required by many organisms and humans (Nielsen et al., 2012; Roth et al., 1993). However, cobamide biosynthesis is limited to a certain group of bacteria and archaea species (Doxey et al., 2015; Martens et al., 2002; Roth et al., 1996; Zhang et al., 2009). These include the major human pathogen, *Mtb*, which has the ability to synthesize and also to take up cobamide *in vitro* (Gopinath, Venclovas, et al., 2013). *De novo* cobalamin biosynthesis from UroIII, a common precursor of haem, sirohaem, and chlorophyll synthesis requires more than 30 enzymatic reactions (Lawrence & Roth, 1995) .These tetrapyrole compounds including cobamide, heme, and chlorophyll, are derived from δ-aminolevulinate (ALA) (Yin & Bauer, 2013). Both aerobic and anaerobic pathways have been described for biosynthesis of cobamide and the distinctive features of these pathways are the mechanism of ring contraction, insertion of the cobalt (which is inserted at early stage in anaerobic biosynthesis, prior to ring contraction and late in aerobic biosynthesis) and in their

requirements for molecular oxygen (Warren et al., 2002). This has resulted in the classification of cobamide biosynthesis into two major stages (Rodionov et al., 2003): the first stage, corrin ring contraction steps, differs for aerobic and anaerobic pathways, whereas the second stage comprises adenosylation, aminopropanol attachment, and the assembly of nucleotide loop, is common for both aerobic and anaerobic pathways (Warren et al., 2002). The two distinct pathways for cobamide biosynthesis were characterized in *Pseudomonas denitrificans* (for oxygen-dependent pathway) (Stamford, 1994) and in *Salmonella typhimurium,* (Roth et al., 1993) and *Bacillus megaterium* (Raux et al., 1998) (for oxygen-independent pathway) (Martens et al., 2002; Moore et al., 2012; Roman et al., 2001). The cobamide biosynthesis genes in *P. denitrificans* are prefixed "*cob*", whereas in *S. typhimurium* the prefix "*cbi*" is used, thus distinguishing genes involved in the oxygen-dependent and oxygen-independent pathways (Blanche et al., 1995; Mattes et al., 2017; Raux et al., 1998; Warren et al., 2002). Both routes differ in terms of corrin ring contraction, oxygen requirements (the aerobic pathway requires oxygen to promote ring-contraction, while the anaerobic pathway does not require oxygen in this step) and cobalt chelation – which occurs via precorrin-2 with CbiK in anaerobic pathway, *versus* hydrogenobyrinic a,c-diamide with the CobNST complex in the aerobic pathway (Fang et al., 2017; Kang et al., 2012).

The first committed precursor of the tetrapyrrole synthesis pathway is ALA. ALA is formed by either the C4 or the C5 pathway. In the C4 pathway, the enzyme ALA synthase catalyses the condensation of glycine and succinyl-CoA to produce ALA. In the C5 pathway, ALA is synthesized from glutamate through three enzymatic reactions (Fang et al., 2017; Kang et al., 2012; Oh-Hama et al., 1988). ALA undergoes condensation to form monopyrrole porphobilinogen (via porphobilinogen synthase), then polymerization and cyclization to form UroIII (Fang et al., 2017). UroIII is methylated at C-2 and C-7 to form precorrin-2 with *S*-adenosyl-L-methionine (SAM) as a methyl donor (Kang et al., 2012; Warren et al., 2002). Precorrin-2 is the last common precursor of cobalamin, siroheme and coenzyme F_{430} (Fang et al., 2017; Warren et al., 2002). While the cobamide biosynthetic pathways diverge at precorrin-2, they re-converge at adenosyl-cobyric acid (Raux et al., 1999; Warren et al., 2002). The second part of the cobamide biosynthesis, adenosylation, aminopropanol attachment, and the assembly of nucleotide loop and DMB, can occur either aerobically or anaerobically (Hazra et al., 2015; Keck et al., 1998; Renz, 1970; Warren et al., 2002).

1.8 Mycobacterial cobamide biosynthesis

The largest component of cobamide is the corrinoid ring, synthesized from UroIII, (Roth et al., 1996). Homologues of nearly all proteins necessary for the aerobic synthesis of cobamide have been identified in mycobacteria, except for members of the MTBC which are missing *cobF* – encoding a precorrin-6a synthase. CobF functions as methyltransferase that catalyses the conversion of precorrin-5 to precorrin-6x (Min et al., 1993) and, in *Mtb*, possible replacements for this protein have been suggested (Gopinath, Moosa, et al., 2013). A nonorthologous CobF is predicted to be encoded either by *Rv2067c*, which contain a C-terminal methyltransferase domain, or *Rv0391*(designated as *metZ*) (Rodionov et al., 2003). Although *Mtb* possesses the remaining genes necessary for cobalamin biosynthesis, genetic evidence suggests that bacilli cannot produce cobalamin *in vitro* (Gopinath, Moosa, et al., 2013; Savvi et al., 2008; Warner et al., 2007). The *cobF* gene was lost from the *Mycobacterium canettii*-like ancestor, along with two further genes and N- and C-terminal portions of flanking genes Rv0943c and Rv0944 (Supply et al., 2013).

The RD9 deletion in RD9-deleted branch of MTBC (*Mycobacterium africanum, Mycobacterium bovis* and *Mycobacterium microti*) causes further attenuation of the cobamide pathway, deleting the amino terminus of CobL, which occurs in an operon with *cobM* and *cobK* (Brosch et al., 2002), and potentially disrupting the functions of all three genes. It has been suggested, therefore, that MTBC members may have come to rely on the host environment as a source of vitamin. The obligate pathogen *Mycobacterium leprae* has a severely reduced genome (Cole et al., 2001) lacking all of the cobamide pathway with the exception of a set of genes required for scavenging exogenous cobamide precursors (Gopinath, Moosa, et al., 2013; Rodionov et al., 2003; Young et al., 2015; Zhang et al., 2009). However, the adaptation of the MTBC members to an intracellular lifestyle might have gradually abrogated the need for cobamide biosynthesis because of its availability from the host. This scenario is further supported by the finding of various nonsynonymous single-nucleotide polymorphisms (SNPs) and frameshift mutations in other cobalamin biosynthesis genes, which were discovered in a genomic screen including more than 200 clinical *Mtb* isolates (Boritsch et al., 2016; Comas et al., 2013; Young et al., 2015).

1.9 Alternate cobamides

A great variety of cobamide analogues exist in nature whose functions are not yet fully understood. These have historically been referred to as pseudo-cobamides (Barker et al., 1958), though current terminology prefers "alternate cobamides". Some organisms synthesize cobamide analogues that contain bases in the α-ligand other than DMB. They share a structural architecture consisting of a corrin ring with a cobalt ion chelated at the core and the β-axial ligand consisting of an organometallic 5'-deoxy-5'adenosyl ligand, or a methyl group, but with a different lower nucleotide loop (referred to α-axial ligand) (Crofts et al., 2013). Anderson and colleagues (2008) demonstrated that *Salmonella enterica* is able to synthesize cobamide *de novo* (Roth et al., 1996), inserting DMB or adenine (to form pseudo-coenzyme cobamide [Ado-pseudo-B_{12}]) or 2-methyl-adenine (to form adenosyl-factor A)) as the axial ligand (Keck & Renz, 2000). However, *S. enterica* requires incomplete corrinoid such as cobinamide (the cobyric acid plus aminopropanol side chain) and DMB supplementation in order to synthesize cobamide during aerobic growth (Anderson et al., 2008). Furthermore, the authors demonstrated that the synthesis of pseudo-cobamide from cobinamide precursor require the same set of genes (*cobC, cobU, cobS* and *cobT*) that are involved in biosynthesis of cobamide which houses DMB base as a lower ligand for cobalt (Anderson et al., 2008). This result indicated that enzymes that are involved in the synthesis of the lower nucleotide loop and attachment can accommodate different substrates. Pseudo-cobamide cofactors can be used by all the cobamide-dependent enzymes of *S. enterica* and may be the only corrinoid types made under periods of high demand unless DMB is provided exogenously (Anderson et al., 2008). This variant form of cobamide was found to be produced by a number of bacteria such as *Clostridium sticklandii* (Stadtman, 1960), *Propionibacterium shermanii* (Friedmann & Fyfe, 1969), and *Lactobacillus reuteri* (Santos et al., 2007). Moreover, a similar response was observed when *E.coli* mutants responded to pseudo-cobamide supplementation (Davis, 1952). Subsequently, it was demonstrated that *Synechocystis sp.* PCC6803 is capable of synthesizing pseudo-cobamide *de novo* and this cofactor supports the activity of the cobamide-dependent methionine synthase (*metH*) in this organism (Tanioka et al., 2009).

Whether mycobacteria possess a similar ability to synthesize alternate cobamides remains to be determined. However, the ability of *Mtb* to adapt to various *in vivo* microenvironments supports the possibility that the bacilli might be able to utilize different variant forms of cobamide cofactors. Moreover, comparative genomic analyses have identified purine

biosynthetic and scavenging pathways (including adenine) in *M. lepare* (Dawes & Mizrahi, 2001; Wheeler, 1987), which suggests that alternate forms of cobamide cofactors might be relevant in different mycobacterial pathogens.

1.10 Cobamide as a Coenzyme

In prokaryotes, cobamides act as a cofactor ("helper molecules" that do not bind to the enzyme), coenzyme (bind to active site of an enzyme) or gene regulator in different essential biochemical processes (Banerjee & Ragsdale, 2003; Gruber et al., 2011; Kräutler, 2005; Osman et al., 2020; Smith et al., 2018). In contrast, in eukaryotes including mammals, cobamide has a limited function in assisting only two enzymes. Cobamide-dependent enzymes fall into four broad classes that involve reductive dehalogenation, isomerization, methylation, and radical SAM processes (Bridwell-Rabb & Drennan, 2017; Vey & Drennan, 2011). The main active biological forms of cobalamin are where the upper ligand is either a adenosyl or an methyl group, which are present in MeCbl and AdoCbl, respectively (Banerjee & Ragsdale, 2003; Gruber et al., 2011). MeCbl acts as a cofactor in a number of methyl transferase reactions, including in methionine synthesis. AdoCbl acts as a coenzyme in rearrangement/isomerase reactions such as with methylmalonyl CoA mutase. AdoCbl and MeCbl catalyse probably the two best-known cobamide-dependent reactions which also represent the only two known cobamide-dependent processes in mammals (Osman et al., 2020).

Prokaryotes possess a variety of other cobamide-dependent enzymes including the ethanolamine ammonia lyases, diol dehydratases, reductive dehalogenases, ribonucleotide reductases (RNRs) and a host of radical SAM enzymes (Bridwell-Rabb & Drennan, 2017; Costa, 2020). Cobamides can also influence levels of transcription and translation through interacting with riboswitches, where binding of specific cobamide forms to regulatory regions within the mRNA controls the production of encoded proteins (Rodionov et al., 2006; Rodionov et al., 2003). Most cobamide riboswitches regulate genes associated with cobamide metabolism and cobalt transport (Osman et al., 2020). All these cobamide-mediated reactions are facilitated by the ability of the cobalt ion to change its state of oxidation (Smith et al., 2018) from Co(I), Co(II), and Co(III). However, as the oxidation state of cobalt changes, so the coordination of the metal changes from four to six, respectively (Smith et al., 2018). For example, in the methyltransfer reaction associated with methionine synthase, the cobalt in methylcobamide is in the Co(III) state and the methyl group is transferred after heterolytic

cleavage of the Co-CH_3 bond to generate the nucleophilic Co(I) species, which rapidly acquires another methyl group from methyltetrahydrofolate (CH_3-THF) (Banerjee & Ragsdale, 2003; Gruber et al., 2011). In this process, the cobamide cycles between Co(III) and Co(I), allowing for efficient transfer of the methyl group from CH_3-THF to homocysteine in the synthesis of methionine with the cobamide acting as a recycling cofactor. With isomerisation reactions, the Co-carbon bond of adenosylcobamide (AdoCba) undergoes homolytic cleavage, converting the Co(III) of AdoCba to a Co(II) species and generating an adenosyl radical. The adenosyl radical is then able to abstract a hydrogen from the substrate, thereby inducing substrate radical formation and rearrangement, prior to hydrogen abstraction from the adenosyl group and reformation of the Co(III) AdoCba (Banerjee & Ragsdale, 2003; Gruber et al., 2011).

1.10.1 Cobamide-dependent enzymes in *Mtb*

The *Mtb* genome encodes three putative cobamide-dependent enzymes which are involved in various aspects of metabolic processes: a class II RNRs (encoded by *nrdZ*) (Dawes et al., 2003), methylmalonyl CoA mutase (encoded by *mutAB*) (Savvi et al., 2008) and methionine synthase (encoded by *metH*) (Warner et al., 2007). The RNRs function in the conversion of ribonucleosides to deoxyribonucleosides which fulfil an essential role in nucleotide cycling and the maintenance of deoxyribonucleotide triphosphate pools for DNA repair and replication (Högbom et al., 2004; Jordan & Reichard, 1998; Nordlund & Reichard, 2006). There are three different class of RNRs: class I, II and III which differ in oxygen requirement. Class I enzymes can be further sub-classified into classes Ia, Ib and Ic and are dependent on oxygen for catalytic function; the class II enzymes (such as NrdZ) require a cobamide cofactor for activity but do not require oxygen, while class III enzymes are inactivated by oxygen and so function under anaerobic conditions only (Jordan & Reichard, 1998; Nordlund & Reichard, 2006).

The methylmalonyl-CoA mutase catalyzes the conversion of (R)-methylmalonyl-CoA to succinyl-CoA, a critical step in propionyl-CoA detoxification (Muñoz-Elías et al., 2006; Savvi et al., 2008). The intermediates from this pathway contribute to the biosynthesis of cell wall lipids (Yang et al., 2009), which are thought to pay crucial role in the survival of *Mtb* during host infection. The *metH*-encoded methionine synthase generates methionine from homocysteine. MetH (5-methyltetrahydrofolate-homocysteine methyltransferase) transfers a methyl group from N^5 methyl-tetrahydrofolate to the thiolate of homocysteine as the final step in methionine synthesis (Banerjee & Matthews, 1990). Methionine can also be synthesized

both aerobically and anaerobically by a cobamide-independent methionine synthase (MetE) (Roth et al., 1993). In addition to MetH, the *Mtb* genome encodes cobamide-independent MetE that catalyzes the same methyl transfer reaction but uses N^5 methyl-tetrahydrofolate directly as the methyl donor (Gonzalez et al., 1992; Matthews et al., 2003; Pejchal & Ludwig, 2004). Both cobamide-independent and cobamide-dependent isoforms require zinc to activate homocysteine (Matthews & Goulding, 1997). The cobamide-dependent enzymes found in mammals are methylmalonyl-CoA mutase and methionine synthase (Ludwig et al., 1996). In humans, inhibition of methionine synthase, either as a consequence of cobamide deficiency or following exposure to nitrous oxide, results in the development of megaloblastic anaemia (Allen et al., 1993; Banerjee & Matthews, 1990; Ludwig et al., 1996).

1.11 Cobamide transport in bacteria

The uptake of scarce nutrients in Gram-negative bacteria requires a series of transporters that can actively transfer nutrients across the outer and inner membranes. Components in the periplasmic space are also sometimes required for transport of these essential nutrients into the cytosol (Pérez, Rodionov, et al., 2016; Wiener, 2005). Molecules that require active transport are usually unable to cross the bacterial membrane owing to their molecular weight, charge, or polarity. Examples include some amino acids, siderophores, enzymes, polymers, and vitamins (Seth & Taga, 2014). In these cases, these molecules need to be imported into the intracellular environment, via energy-dependent transport systems such as the ATP binding cassette (ABC) transporter family (Rees et al., 2009) or the phosphotransferase system. Corrinoids, which consist of four pyrrole rings, fall into this category (Seth & Taga, 2014). Cobalamin is a corrinoid which is actively transported through the bacterial membrane via ABC transporters (D'Souza et al., 2018). ABC transporters are universally distributed across all domains of life and are involved in many various aspects of microbial cellular metabolism, including the regulation of several metabolic processes and bacterial pathogenesis (Thomas et al., 2020). These transporters plays a significant role in mediating the uptake of essential nutrients and the export of toxic compounds in prokaryotes, but they can also facilitate the transport of other physiological substrates (Davidson et al., 2008). ABC transporters are one of many different types of transporters found in bacteria and other living organisms and other structurally and functionally different of transporters have been identified in living organisms (Davidson et al., 2008).

ABC transporters can be subdivided in three functional groups. First as importers that facilitate the uptake of a wide range of substrates in prokaryotes, such as ions, metals, amino acids, siderophores, peptides, sugars or complex organic molecules and vitamins. Also as exporters that mediate the secretion of various substrates, including toxic compounds (such as antibiotics and hemolysin), lipids, proteins, hydrophobic drugs, polysaccharides and peptides. The third category of systems is apparently not involved in the import or export of any physiological substrates, with some members being involved mainly in DNA metabolism, including translation of mRNA and in DNA repair (Davidson et al., 2008).

Cobamide uptake is essential for cobamide-utilizing organisms which lack the ability to synthesize the coenzyme *de novo*, and the only known transport systems for cobamides in prokaryotes are BtuBFCD (DeVeaux & Kadner, 1985; Korkhov et al., 2012) and the energy coupling factor (ECF)-type ABC transporter, ECF-CbrT (Rodionov et al., 2009; Santos et al., 2018). BtuBFCD ABC transporter consists of BtuB, a TonB-dependent outer membrane transporter (TBDT) (Shultis et al., 2006) and an inner membrane transporter (BtuCDF) of the ABC family (Borths et al., 2005; Pérez, Rodionov, et al., 2016; Zhang et al., 2009). BtuB is found only in Gram-negative bacteria, while the periplasmic binding protein BtuF and ABC transporter BtuCD are found across bacterial taxa (Degnan et al., 2014). In Gram-negative bacteria, exogenous cobamides are transported into the cell via an ABC transport system, consisting of BtuC, BtuD, and BtuF, which are membrane permease, ATPase, and periplasmic-binding protein components, respectively. BtuB is coupled to an inner membrane complex consisting of TonB, ExbB, and ExbD and shares the canonical 22-strand, β-barrel architecture of other TBDTs together with an N-terminal, globular periplasmic domain that occludes a channel through the transporter and thus prevents passive diffusion into the periplasm (Shultis et al., 2006). The inner-membrane ABC transporter consists of a high-affinity, cobalamin-binding periplasmic protein (BtuF), which binds cobalamin in the periplasmic space and delivers it to an inner membrane ABC transporter (BtuCD). This complex consists of a pair of membrane-spanning permease subunits (BtuC) and a pair of ATP binding and hydrolysis subunits (BtuD) (Locher, 2004). The inner membrane ABC transporter system utilizes ATP hydrolysis-driven conformational changes for translocation of cobalamin into the cytoplasm, but BtuB relies on TonB, ExbB and ExbD to couple proton motive force to cobalamin uptake (Locher, 2004; Shultis et al., 2006). TonB protein is also required for siderophore-mediated iron uptake. These iron uptake systems also employ specific receptors in the outer membrane as well as other gene products that mediate transport steps after receptor binding (Noinaj et al.,

2010; Wiener, 2005). ECF-type ABC transporter, ECF-CbrT, was recently identified to be cobalamin-specific transporter in bacteria (Rodionov et al., 2009; Santos et al., 2018). ECF-type ABC transporter consists of two cytoplasmic ATPases domains (EcfA and EcfA'), similar to the ATPases of conical ABC transporters, and two transmembrane proteins, EcfT (T-component)- for energy coupling, and EcfS (S-component) -for substrate binding. The two ATPases subunits and one of the transmembrane proteins, EcfT, forms a ternary complex (the ECF module). The S-component interact with the energizing ECF module to allow for substrate translocation (Eitinger et al., 2011; Rodionov et al., 2009; Santos et al., 2018).

The uptake of cobamide has been linked to pathogenesis of TB (Gopinath, Venclovas, et al., 2013). *Mtb* does not encode any characterized cobamide transporter (Korkhov et al., 2012); however, BacA, an ABC transporter protein encoded by Rv1819c was found to be essential for the uptake of cobamide (Gopinath, Venclovas, et al., 2013; Rempel et al., 2020). Furthermore, BacA was initially annoatated as a protein involved in the transport of antimicrobial peptides (Domenech et al., 2009). This suggests that cobamide uptake in mycobacteria is different from that of Gram-negative bacteria which depend on a well-defined system.

1.12. Cobalt transport in bacteria

Cobalt is an essential trace element for many living organisms, as it plays a significant biological role as the centrally coordinated ion in macrocyclic tetrapyrroles known as cobyric acid and it requirement is generally distributed among microorganisms that can produce cobamide *de novo* (Cheng et al., 2011; Escalante-Semerena, 2007; Martens et al., 2002). Cobalt can also be associated directly with cobalt-dependent enzymes (noncorrin enzymes) (Kobayashi & Shimizu, 1999). To acquire sufficient cobalt for metabolism, bacteria have high-affinity uptake systems to scavenge Co^{2+} from the environment, where it is often available only in trace amounts (Eitinger et al., 2005; Rodionov et al., 2006; Vitreschak et al., 2003). CbiMNQO (found in *Rhodobacter capsulatus*) and the NikMNQO transporter are the most widespread transport system for Ni and Co uptake in bacteria (Rodionov et al., 2006; Zhang et al., 2009), which are cofactors of a number of essential metabolic enzymes (Bao et al., 2017; Okamoto & Eltis, 2011). Similar to the transport of cobamide, cobalt ion must be traffic across the outer membrane into the periplasmic space, where it is captured and delivered to the ABC-type importer, which transport cobalt into the bacterial cytoplasm. The uptake of cobalt ion can also occur via non-specific metal uptake system; for example, in most bacteria, the uptake of

cobalt and nickel is mediated by CorA proteins. CorA proteins are generally associated with uptake of magnesium ion (Maguire, 2006; Niegowski & Eshaghi, 2007; Zhang et al., 2009). When external metal concentrations are very high, Co^{2+} accumulation may become toxic, and excess Co^{2+} can be removed from cells by efflux systems (Cheng et al., 2011; Nies, 2003).

1.13 Microbial cobamide-dependent regulation

Cobamide metabolism in bacteria is largely regulated by RNA regulatory elements known as riboswitches (Choudhary et al., 2013). Cobamide is known to regulate the expression of genes required for its own biosynthesis and transport; this regulation occurs through the activity of cobamide-dependent riboswitches (Vitreschak et al., 2003) – a messenger RNA (mRNA) structural elements that serve as ligand-responsive genetic control elements (Blount & Breaker, 2006; Nahvi et al., 2004). In a pioneering bioinformatic study, Rodionov et al. (2003) demonstrated that cobamide elements regulate not only genes related to cobamide biosynthesis and transport but also several genes from cobamide-dependent pathways. They observed that, in many cases, the cobamide-independent isozymes of methionine synthase and RNRs are regulated by cobamide elements in organisms possessing both cobamide-dependent and cobamide-independent isozymes. Moreover, cobamide regulons of various bacteria are thought to include enzymes from known cobamide-dependent or alternative pathways (Rodionov et al., 2003).

Riboswitches are a class of noncoding RNA (ncRNA) regulatory elements that specifically bind small-molecular ligands with high specificity and affinity (Li et al., 2020; Mellin et al., 2014; Mironov et al., 2002; Sudarsan et al., 2003; Winkler & Breaker, 2005). These RNA elements controls gene expression through transcriptional or translational modifications (Li et al., 2020; Mellin et al., 2014; Polaski et al., 2017), by binding metabolites such as cobalamin, tRNAs, sugars and amino acids, (Ignatov & Johansson, 2017). Binding induces conformational changes in the RNA that result in attenuation at either the transcriptional or the translational level (Nahvi et al., 2004; Pérez, Liu, et al., 2016). Riboswitches are highly structured and are embedded within in the 5′ untranslated leader regions (5′ UTR) of mRNAs. They consist of two functionally distinct domains (Nahvi et al., 2002; Santillán & Mackey, 2005), an evolutionarily conserved aptamer domain and a variable expression platform, which is located downstream of the aptamer domain (Haller et al., 2011; Mandal & Breaker, 2004; Roth & Breaker, 2009). Ligand binding to the aptamer brings about structural changes in both the

aptamer and the downstream expression platform (Kubodera et al., 2003; Loh et al., 2009; Soukup & Soukup, 2004; Winkler et al., 2004; Winkler et al., 2003). Riboswitches sensing corrinoids (cobamide derivatives), commonly named cobalamin riboswitches, are prevalent throughout the bacterial domain and constitute the most complex class of riboswitches discovered thus far (Nahvi et al., 2004; Vitreschak et al., 2003).

These riboswitches mainly regulate the gene products associated with corrinoid biosynthesis and transport, porphyrin and cobalt transport, and cobamide-independent RNRs, as well as enzymes involved in glutamate and succinate fermentation (Winkler & Breaker, 2005). Regulation of these gene products can take place at either the translational or the transcriptional level (Nahvi et al., 2002; Nou & Kadner, 2000; Perdrizet et al., 2012). The aptamer of cobalamin riboswitches consists of a well-organized secondary structure divided into distinct stems and loops. All of the cobalamin binding aptamers consist of a central four-way junction that mostly forms the core ligand-binding pocket (Johnson Jr et al., 2012; Nahvi et al., 2004; Peselis & Serganov, 2012). Cobalamin riboswitches were recently divided into two classes: (i) aquacobalamin (AqCbl) and (ii) AdoCbl riboswitches depending on their cognate metabolite (Johnson Jr et al., 2012). Cobalamin is also known to regulate the *btu* genes responsible for its active transport (Nou & Kadner, 2000) and the *cob* operon responsible for its *de novo* biosynthesis (Ravnum & Andersson, 1997). In bacteria such as *E. coli* and *S. typhimurium*, cobamide interacts with the 5′ UTR of the cobamide biosynthesis operon to repress translation of the corresponding genes, including the *cob* and *btuB* operons (Li et al., 2020; Lundrigan et al., 1991; Richter-Dahlfors et al., 1994). Cobalamin riboswitches that control the transcription of the Class Ia RNRs genes, *nrdABS,* are also known (*e.g*., in *Streptomyces coelicolor)* (Borovok et al., 2006).

The *Mtb* genome encodes two putative cobamide riboswitch motifs, one of which is located upstream of *metE* and has been experimentally shown to regulate transcription of the cobamide-independent methionine synthase in cobamide-replete conditions (Warner et al., 2007). The second riboswitch is situated upstream of the putative *PPE2- cobQ1-cobU* operon (Rodionov et al., 2003; Vitreschak et al., 2003), but the functionality of this riboswitch has not been confirmed. The fact that *Mtb* has the capacity to regulate core metabolic functions according to cobamide availability, whether acquired via endogenous synthesis or through uptake from the host environment, implies that there is a role for cobamide in pathogenesis, though this remains unproven (Gopinath, Moosa, et al., 2013).

1.14 Folate-Cobamide Interrelationship

The folate biosynthetic pathway is essential in the production of tetrahydrofolate (THF) and derivatives required as cofactors in the biosynthesis of purines, thymidylate, serine, and methionine (Zheng et al., 2013). Many microorganisms and plants possess the ability to synthesize folic acid derivatives *de novo*, initially forming dihydrofolate (Green & Matthews, 2007). Reduction of dihydrofolate to THF is carried out in most bacteria and eukaryotes by dihydrofolate reductase (DHFR) (Zheng et al., 2013). All the folic acid derivatives that serve as recipients and donors of one-carbon units are derivatives of THF, which is formed from dihydrofolate by an NADPH-dependent reduction catalyzed by dihydrofolate reductase (FolA) (Green & Matthews, 2007). THF is best known for its role as a donor for one-carbon units in various oxidation states that are attached to N-5 or N-10 in a variety of biosynthetic processes, including the formation of methionine, purines and thymine (Bermingham & Derrick, 2002; Green & Matthews, 2007). However, THF can also act as an acceptor of one-carbon units in degradative reactions. While the cellular requirement for folates is universal, methods for obtaining them differ between prokaryotes and eukaryotes (Bermingham & Derrick, 2002). Folate derivatives also participate in other metabolic processes that do not involve one-carbon transfers (Green & Matthews, 2007).

In bacteria, THF is also required for the synthesis of formylmethionyl $tRNA^{fMet}$, which is essential for the initiation of protein synthesis (Zheng et al., 2013). DHFR is currently a target for antibacterial drugs such as Trimethoprim (TMP), a bacteriostatic antimicrobial agent that potently inhibits DHFR in various bacterial species (Minato et al., 2015; Zheng et al., 2013). The metabolism of cobamide and folate is interrelated and these pathways are joined metabolically at the methionine cycle (Froese et al., 2019). Methylenetetrahydrofolate (CH_2-THF) can be reduced to form 5-methlytetrahydrofolate (5-CH_3-THF), the donor of the methyl group used for conversion of homocysteine to methionine in the final step of methionine biosynthesises by methionine synthase (MetH- (using Co(I)cobalamin as a co-factor) or MetE) (Blanco et al., 1998). Folate metabolism is also tied to the methyl cycle, which is involved in the biosynthesis of SAM. SAM-dependent methyltransferases are essential for many cellular functions in *Mtb*, including DNA methylation, biotin synthesis, modification of mycolic acids, and methylation of rRNA (Boissier et al., 2006; Minato et al., 2015; Parveen & Cornell, 2011).

In humans, cobalamin deficiency and/or lack of MetH enzymatic activity lead to accumulation of cellular THF in the methylated form (5-CH_3-THF), thus interrupting the normal flow of the

one-carbon metabolic reactions, leading to a phenomenon called the "methyl folate trap." (Guzzo et al., 2016; Palmer et al., 2017). The interruption of the normal flow of the one-carbon metabolic reactions was recently described in prokaryotes, including mycobacteria (Guzzo et al., 2016).

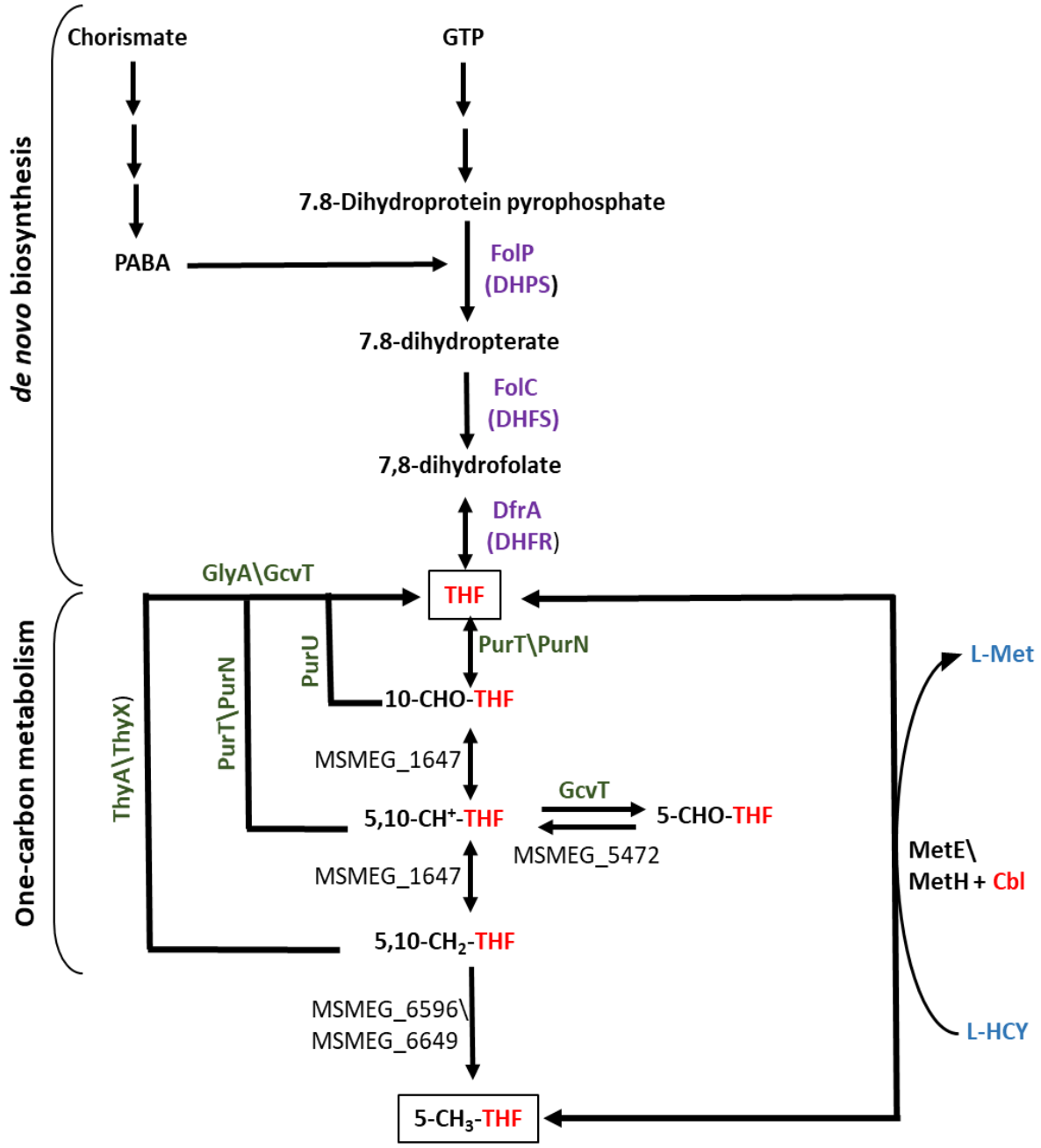


Figure 1.5: Schematic representation of the link between folate metabolism, methionine biosynthesis and the one carbon metabolism in *Msm*. Enzymes of folate biosynthesis are puple, and enzymes of one-carbon metabolism are in green, Cobalamin in red. Pathway intermediates are connected by black arrows. Abbreviations: GTP, guanosine-5'-triphosphate; PABA, *para*-aminobenzoic acid; THF, tetrahydrofolate; CHO-THF, 10-formyltetrahydrofolate; CH^{+}-THF,5,10-methenyltetrahydrofolate; CHO-THF,5-formyltetrahydrofolate; CH_2-THF, 5,10-methylenetetrahydrofolate; CH_3-THF, 5-methyltetrahydrofolate; DHPS, dihydropteroate synthase; DHFS, dihydrofolate synthase; DHFR, dihydrofolate reductase; L-HCY, homocysteine; L-Met, methionine; Cbl, cobalamin with upper axial ligand (eg, cyano-, hydroxo-).

1.15 Transposon mutagenesis

Significant advances have been made in the development of genetic tools to disrupt gene function in mycobacteria utilizing both targeted and random approaches (Gopinath et al., 2015; Rock et al., 2017). Whole-genome transposon (Tn) mutagenesis is a powerful method for categorizing gene function in a variety of microorganisms (Sassetti et al., 2001). These genes include those that can be assigned to essential metabolic pathways as well as many of unknown function. Different genetic approaches have been performed to determine the set of genes that essential for survival in different microorganisms. These approaches include antisense-mediated gene inactivation in *Haemophilus influenzae* (Akerley et al., 2002), Tn mutagenesis in *Pseudomonas aerguinosa* (Jacobs et al., 2003), plasmid-insertion mutagenesis in *Bacillus subtilis* (Kobayashi et al., 2003) and insertion-duplication mutagenesis in *S. enterica* (Knuth et al., 2004). These different techniques have provided estimates for essential gene sets ranging from fewer than 300 to 600 genes, depending on the microorganism. This major difference is more likely due to a variety of factors, such as experimental techniques producing false positives and negatives, the changes in Tn library size during the experiment, the growth conditions and genetic properties of the cell being manipulated (Barquist et al., 2013; Barquist et al., 2016; Sassetti et al., 2001).

Tns, also known as "jumping genes", are biological elements consisting of a DNA sequence (Reznikoff, 1993) that randomly moves from one location to another in the genome. These transposable elements (TEs) are widely distributed in all domains of life and have played a significant role in the evolution of many genomes (Biémont, 2010; Feschotte & Pritham, 2007; Jangam et al., 2017). The TEs were first discovered in 1948 by Barbara McClintock in maize as controlling elements. However, over the years, they have been identified in all forms of life, prokaryotes and eukaryotes (Feschotte & Pritham, 2007; Kidwell & Lisch, 2001). There are

two classes of TEs, categorised based on their mechanism of transposition. Class I transposons, or retrotransposons move through the copy and paste mechanism, using RNA intermediates and reverse transcriptase to copy themselves. Class II transposons are DNA transposons that transpose by the cut and paste mechanism in which the transposon is excised from one DNA sequence and inserted into a new target DNA sequence (Hamer et al., 2001; Muñoz-López & García-Pérez, 2010). The Transposition of DNA transposons is catalysed by transposase, an enzyme that cuts out the transposon sequence and inserts into a new genomic location (Hamer et al., 2001; Muñoz-López & García-Pérez, 2010). In this study, the *Himar1* Mariner transposon - isolated from the horn fly, *Haematobia irritans* - was used to yield detailed genetic maps of *de novo* cobalamin biosynthesis and salvage in *Msm*. The *Himar1* Mariner transposon is part of the class II elements that randomly inserts into genomic DNA sequences resulting in the formation of mutants (Akerley et al., 1998; Reznikoff & Winterberg, 2008; Van Opijnen & Camilli, 2013). A common approach to identifying genomic requirement for survival and growth under a particular set of conditions is to screen large pools of mutants simultaneously. This can be done with defined mutants (Baba et al., 2006; Hobbs et al., 2010), but this is labour-intensive and requires accurate genomic annotation, which can be particularly difficult to define for non-coding regions. The most frequently used Tns in bacterial genetics are based on the Tn *5* and Tn *10* platforms (Jacobs et al., 2003; Ruvkun et al., 1982; Way et al., 1984). Both Tn *5* and *Himar1* Mariner Tns have been used to generate non-redundant mutant libraries in *P. aeruginosa* (Held et al., 2012; Jacobs et al., 2003). The *Himar1* Mariner transposon, isolated from the horn fly *Haematobia irritans*, is a relatively simple Tn resembling an IS element and consisting of the *Himar1* transposase bounded by 29 bp inverted repeats (Rubin et al., 1999). *Himar1* transposon inserts into the dinucleotide -TA- in a mechanism similar to that of the Tn *5* and Tn *10* Tns, in which the donor Tn DNA is cut and then pasted into the target site by the action of the transposase (Kulasekara, 2014).

1.15.1 Tn insertion sequencing (TIS) methods

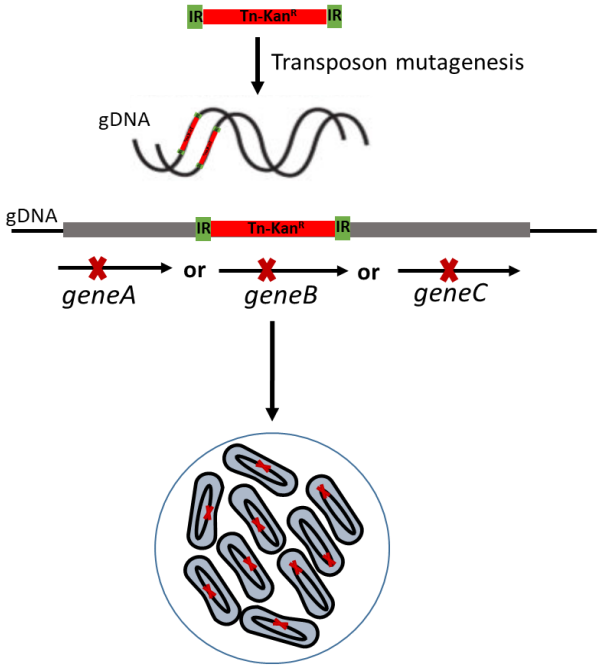
Several methods were developed for high-throughput sequencing of Tn insertion sites: high-throughput insertion tracking by deep sequencing (HITS) (Gawronski et al., 2009), Tn sequencing (TnSeq) (Van Opijnen et al., 2009), insertion sequencing (INSeq) (Goodman et al., 2009), and Tn-directed insertion site sequencing (TraDIS) (Langridge et al., 2009), making it possible to reveal genotype-phenotype relationships in a high-throughput manner. However, they all have the same fundamental methodology. First, Tn mutagenesis is used to generate a

high-density Tn insertion mutant in which ideally all genomic loci have been disrupted at multiple sites. In principle, the starting Tn mutants containing inactivating insertions in any of the genomic regions that are required for growth (*i.e*., 'essential' genes/regions) will be eliminated. This library can then be grown under a specific condition of interest, so that mutants that are attenuated under a specific experimental condition are outcompeted and mutants with increased growth and survival become overrepresented in the population. High-throughput sequencing is used to quantify all Tn junctions, to determine the relative abundance of mutants containing a Tn at particular insertion site (Chao et al., 2016) (**Figure 1.6**).

High-throughput sequencing technology enables routine monitoring of as many as one million Tn insertion mutants simultaneously in any genetically tractable organism (Van Opijnen et al., 2009). The majority of TIS studies utilize a library of pooled Tn mutants, although defined Tn libraries can also be used and are preferred in certain circumstances (for example insertions in small genes, where a single mutant can be retrieved and characterize in any condition). High-density Tn libraries are commonly created using either mariner-based transposons, which target -TA- dinucleotides (Akerley et al., 1998; Chiang & Rubin, 2002; Rubin et al., 1999) or Tn *5*-based vectors, which insert at random sites (Goryshin & Reznikoff, 1998). For most bacteria, -TA- dinucleotides are distributed relatively evenly across the genome, although total number of -TA- sites will vary according to genomic GC content, and the frequency of -TA- sites can differ locally within the genome; *e.g*., in horizontally acquired DNA regions (Chao et al., 2016).

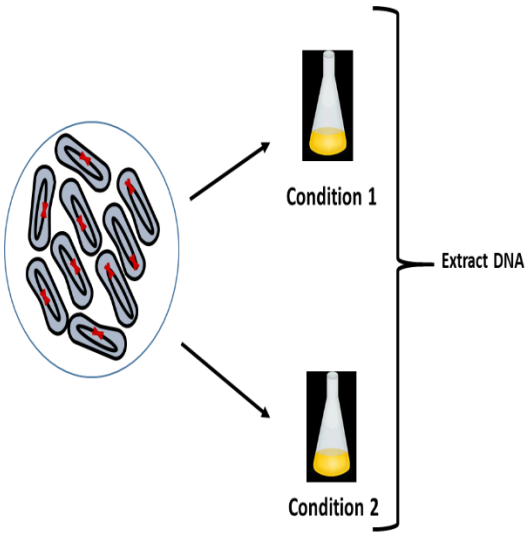
Mariner family transposons have been widely used in fish pathogen *Mycobacterium. marinum*, and also human pathogens such as *P. aeruginosa*, *Campylobacter jejuni*, *Leptospira interrogans*, and *Rickettsia prowazekii* (Gao et al., 2003; Hendrixson et al., 2001; Liberati et al., 2006; Liu et al., 2007; Murray et al., 2009). Tn-insertion sequencing (TIS) allows rapid identification of the complete set of genes required for growth under different conditions (Sassetti et al., 2001). The identification of a set of essential genes for *Mtb* has greatly enhanced the ability to elucidate metabolic pathways and processes, and to define new, potential drug targets of the pathogen (Chevalier et al., 2014). Essential genes, defined as genes indispensable for growth and/or survival, are potential targets for new types of antibacterial drugs (DeJesus et al., 2017; DeJesus & Ioerger, 2013). Gene essentiality can be assessed by targeted gene disruptions, where genes that cannot be disrupted are typically categorized as being essential (Minato et al., 2019).

A Creating random transposon mutants



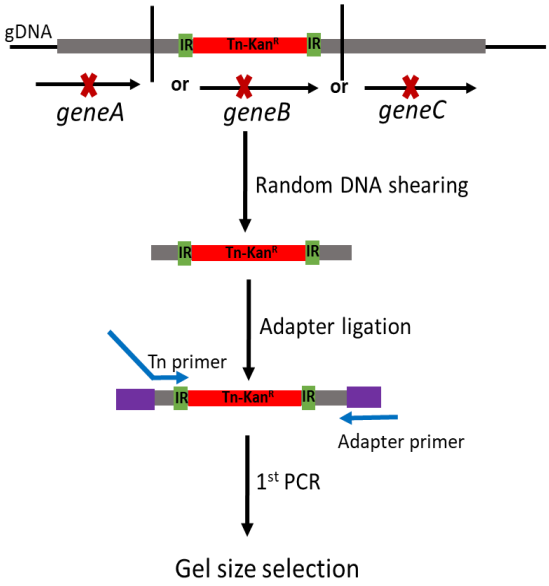
Pool high-density transposon mutant library

B Selection



Grow transposon library under desired conditions

C DNA library construction



Fragment, attach adaptor sequences and Amplify transposon junctions

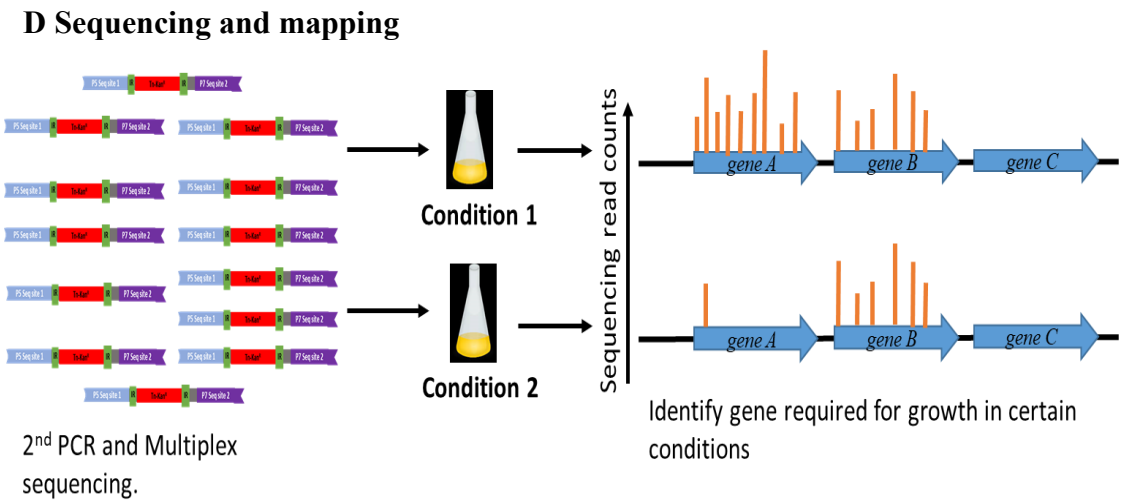


Figure 1.6: Tn insertion sequencing (TIS) workflow. The TIS method has four steps. (**A**) Creation of a high-density Tn insertion library containing multiple insertions in every non-essential genomic locus. (**B**) Tn library is grown under different conditions (*e.g.*, selective and non-selective), and mutants that are viable under each condition are recovered and the genomic DNA is extracted. (**C**) The Tn junctions in both selected pools are attached to sequencing adaptors and amplified. (**D**) Therefore, high-throughput sequencing is used to quantify all Tn junctions, then the sequences are mapped to the genome and the read counts at each insertion site are subjected to statistical analysis to define genomic loci that appear significantly underrepresented in the selective growth condition.

1.16 Aims and Objectives

The overall goal of this project was to exploit the loosely related and non-pathogenic saprophyte, *Mycobacterium smegmatis* strain mc^{2}155, as a model mycobacterium, to yield detailed genetic maps of *de novo* cobalamin biosynthesis in mycobacteria. To address this question, a combination of a genome-scale approach and molecular techniques, including TnSeq, were adopted.

1.16.1 Main Aim

The main aim of this research was to elucidate the genetic requirements for vitamin B_{12} biosynthesis and assimilation in the non-pathogenic *Msm* mc^{2}155.

1.16.2 Objectives

The specific objectives of this study were:

(i) To construct high-coverage transposon (Tn) libraries of *Msm* mc^2155 and a derivative Δ*metE* mutant containing a targeted deletion of the gene encoding the B_{12}-independent methionine synthase.

(ii) To utilize the libraries generated in (i) above to identify genes required for the biosynthesis and/or assimilation of vitamin B_{12} during growth in defined media *in vitro*.

(iii) To utilize CRISPRi (Rock et al., 2017) to validate the role of selected gene(s) identified as essential for cobalamin biosynthesis, including putative cobalt transporters.

Chapter 2: Materials and Methods

2.1 Mycobacterial strain and culture conditions

Msm mc^{2}155 strains were grown in Middlebrook 7H9 broth medium (Difco) supplemented with 0.2% glycerol, 10% OADC/ADC and 0.05% Tween 80 or on Middlebrook 7H10 solid medium supplemented with 0.05% glycerol and 10% OADC. Where appropriate, growth media were supplemented with cobalt at concentrations of 0.65 µg/ml and/or CNCbl at concentrations of 10 µg/ml, as per published methods (Gopinath, Venclovas, et al., 2013). Where relevant, media were supplemented with 25 µg/ml kanamycin (Kan). *E. coli* strains were grown in Luria-Bertani (LB) broth or on agar (LA) with 25 µg/ml Kan. All cultures were incubated at 37°C unless indicated otherwise. Growth was monitored by measuring the optical density at 600 nm (OD_{600}) of cultures using at least three independent biological replicates. Bacterial strains used in this study are listed in **Table 2.1**.

Table 2.1: Strains used in this study

Strain's name	**Description**	**References**
Escherichia coli		
DH5α	*supE44 ΔlacU169 (F80 lacZΔM15) hsdR17 recA1 endA1 gyrA96 thi-1 relA1*	Promega
***Mycobacterium smegmatis* (*Msm*)**		
mc^{2}155	High-frequency transformation mutant of *Msm* ATCC 706	(Snapper et al., 1990)
Δ*metE*	*metE* knockout in *Msm* mc^{2}155	Dawes, unpublished
Δ*metE cobA*-KD	*cobA* knockdown in Δ*metE*; Kan^R	This study
Δ*metE cobIJ*-KD	*cobIJ* knockdown in Δ*metE*; Kan^R	This study
Δ*metE cobN*-KD	*cobN* knockdown in Δ*metE*; Kan^R	This study
Δ*metE cobO*-KD	*cobO* knockdown in Δ*metE*; Kan^R	This study
Δ*metE cobU*-KD	*cobU* knockdown in Δ*metE*; Kan^R	This study
Δ*metE cobT*-KD	*cobT* knockdown in Δ*metE*; Kan^R	This study
Δ*metE bluB*-KD	*bluB* knockdown in Δ*metE*; Kan^R	This study
Δ*metE mmpL3*-KD	*mmpL3* knockdown in Δ*metE*; Kan^R	This study

Table 2.2: Plasmids used in this study

Plasmid name	Description	References
pLJR962	CRISPRi backbone for *Msm*	(Rock et al., 2017)
pLJR962_*cobA*	*Msm cobA* knockdown construct	This study
pLJR962_*cobIJ*	*Msm cobIJ* knockdown construct	This study
pLJR962_*cobN*	*Msm cobN* knockdown construct	This study
pLJR962_*cobO*	*Msm cobO* knockdown construct	This study
pLJR962_*cobU*	*Msm cobU* knockdown construct	This study
pLJR962_*cobT*	*Msm cobT* knockdown construct	This study
pLJR962_*bluB*	*Msm blub* knockdown construct	This study
pLJR962_*mmpL3*	*Msm mmpL3* knockdown construct	This study

2.2 DNA Extraction

2.2.1 Plasmid DNA extraction and purification

2.2.1.1 Small-scale extraction

Cultures were grown overnight at 37°C or over 48 hr at 30°C and were harvested by centrifugation in an Eppendorf 5415D microcentrifuge (13 000 rpm for 1 min) at room temperature. The supernatant was discarded, and the pellet resuspended in 100 µl lysis solution I (0.5 M glucose, 50 mM Tris-HCl pH 8.0, 10 mM ethylenediaminetetraacetic acid (EDTA, pH 8.0)); thereafter, 200 µl of solution II (0.2 M sodium hydroxide (NaOH), 1% sodium dodecyl sulphate (SDS)) was added and the suspension was mixed gently. After 5 min of incubation, 150 µl of neutralization solution III (3 M potassium acetate, pH 5.5) was added and gently mixed, followed by centrifugation at 13 000 rpm for 5 min at room temperature. The supernatant was transferred to fresh microfuge tubes and treated with 1 µl of 10 µg/ml RNase A (Sigma Aldrich) for 10 min at 42°C. Plasmid DNA was precipitated through the addition of 350 µl isopropanol followed by centrifugation at 13 000 rpm for 10 min following 10 min incubation at room temperature. The pelleted DNA was washed with 70% ice-cold ethanol by centrifugation (13 000 rpm for 5 min) and dried at 40°C in a vacuum centrifuge (MiVac DNA concentrator, GeneVac). The plasmid DNA was resuspended in 20-30 µl of sterile distilled water (sdH2O) and heated to 42°C. The final DNA plasmid solution was either stored at 4°C for short time or -20°C for extended periods of time.

2.2.1.2 Large-scale extraction

Fifty ml *E. coli* cultures were grown overnight at 37ºC or at 30ºC for 48 hr and cells were harvested by centrifugation at (4500 × *g*, 20ºC, for 10 min) (Beckmann Allegra X-22R). Following centrifugation, the supernatant was discarded, and pellets were re-suspended in 1 ml of Solution I, followed by 2 ml of Solution II and inverted gently to mix. This was followed by addition of 1.5 ml of Solution III and vigorous shaking. The suspension was aliquoted into 1.5 ml Eppendorf tubes and centrifuged for 10 min at 13 000 rpm at room temperature. The supernatant was transferred into a fresh 1.5 ml Eppendorf tube and 10 µl RNaseA (10 mg/ml stock) was added to each tube, followed by incubation at 42°C for 30 min. After incubation, 700 µl of isopropanol was added to each tube to precipitate the DNA and centrifuged for 10 min at 13 000 rpm. The pelleted DNA was washed with 500 µl of cold 70% ethanol and dried in the vacuum-dryer to remove excess ethanol. The DNA pellet from each tube was re-suspended in 100 µl of sterile autoclaved water (dH_2O) and pooled to a volume of 500 µl. Thereafter, 50 µl of 3 M sodium acetate (pH 5.2) and 700 µl of phenol: chloroform (1:1) was added to the DNA mixture. The mixture was vortexed before centrifuging at 13 000 rpm for 10 min. The aqueous layer was transferred into a clean 1.5 ml Eppendorf tube, 350 µl of chloroform: isoamyl alcohol (24:1) was added to the aqueous solution and the mixture was vortexed and centrifuged at 13 000 rpm for 10 min. Following this, the aqueous layer was transferred into a clean 1.5 ml Eppendorf tube and 1ml of cold 100% ethanol was added to the mixture. To precipitate the DNA, the tube containing the mixture was gently inverted several times and stored at -20°C for 2 hr or overnight before centrifugation at room temperature (13 000 rpm, 20 min). The DNA pellet was washed with 70% cold ethanol, vacuum-dried and re-suspended 50-100 µl of sterile RNase free water. The DNA was quantified using a NanoDrop 2000c Spectrophotometer (Thermo Scientific) according to the manufacturer's instructions.

2.2.2 *Msm* genomic DNA extraction

Genomic DNA was isolated using a modified cetyltrimethylammonium bromide (CTAB; ICN Biomedicals, Aurora, Ohio) method (Larsen, 2000; van Helden et al., 2001). Briefly, 20 ml cells were grown to log-phase and harvested by centrifugation (4500 × *g*, 20ºC, 10 min) (Beckmann Allegra X-22R). Thereafter, cells were heat-killed in microfuge tubes for 20 min at 65°C, then harvested by centrifugation at 13 000 rpm for 10 min. After resuspending the

pellet in 500 µl TE buffer (1 mM Tris·HCl pH 8.0, and 1 mM EDTA), 50 µl of lysozyme (10 mg/ml) and 10 µl RNase (10 mg/ml) was added and the reaction incubated overnight at 37°C. The following morning, 70 µl of 10% SDS and 50 µl of proteinase K (10 mg/ml) was added and the mixture incubated at 65°C, shaking at 400 rpm in a Eppendorf thermomixer (Merck) for 2 hr. Next, 100 µl of pre-warmed CTAB/NaCl mix (10% CTAB made in 0.7 M NaCl) and 100 µl of 5 M NaCl was added and the mixture was further incubated for 10 min at 65°C. To harvest the DNA, 700 µl of 24:1 (v/v) chloroform: isoamyl alcohol was added and the tubes left to stand for 10 min at room temperature, followed by centrifugation at 13 000 rpm for 10 min. The aqueous phase was added to an equal volume of isopropanol and incubated on ice for at least 30 min. The DNA was recovered by centrifugation (13 000 rpm, 20 min), the pellet was washed with ice cold 70% ethanol and dried in a vacuum centrifuge before resuspending in 100 µl of sdH2O. The extracted DNA was run on a 1% gel to check for quality and quantified using the NanoDrop ND-1000 Spectrophotometer (Thermo Scientific).

2.3 DNA manipulations

All DNA manipulations and molecular biology techniques were performed according to standard protocols (Russell & Sambrook, 2001; Sambrook et al., 1989).

2.3.1 Agarose gel electrophoresis

Standard electrophoretic techniques were used for separation of DNA fragments. High molecular weight DNA fragments were separated on 1% agarose gels while low molecular weight fragments were separated on 2% gels. Agarose gels were prepared in 1×TAE buffer (1 mM EDTA, 40 mM Tris-acetic acid, pH 8.0), agarose powder (Sigma-Aldrich) and contained 0.5 µg/ml ethidium bromide for DNA detection. DNA samples were loaded with tracking dye (0.025% bromophenol blue in 30% glycerol). DNA molecular weight markers III, IV and V (Roche Applied Science) were used to estimate fragment sizes. Agarose gels were electrophoresed in a Mini-Sub Cell GT mini gel horizontal submarine unit (Bio-Rad) at 80-100 volts and visualized under UV-light using the Gel Doc (WealTeach Keta Imaging System). Fragments used for cloning were visualised using the blue-light Dark Reader DR88M transilluminator (Inqaba Biotec) to avoid UV-induced DNA damage.

2.3.2 Gel extraction and PCR purification

DNA fragments of the desired size was excised from the agarose gel using a sterile scalpel blade and purified using Qiagen Gel Extraction kit (WhiteSci), as per the manufacturer's instructions. For purification of PCR reactions, the amplicons were cleaned using the QIAquick PCR purification kit according to the manufacturer's instructions. The DNA was then eluted using prewarmed sterile RNase-free water. The DNA was quantified either on agarose gels or using the NanoDrop ND-1000 Spectrophotometer (Thermo Scientific).

2.3.3 Restriction digests

Restriction enzymes were from New England Biolabs Inc. (Inqaba Biotech, South Africa) or Fermentas (Thermo Scientific) and used as per manufacturer's instructions. All restriction enzyme digests were performed at optimal temperatures of the respective enzymes, mostly 37ºC. Plasmid DNA (1µg) or mycobacterial chromosomal DNA (5 µg) was digested for at least 3 hr or overnight with the appropriate reaction buffer(s). Digested DNA fragments were separated and analysed on agarose gels (as described previously).

2.3.4 Ligation

DNA ligations were performed using either the Fast-Link™ ligation kit (Epicentre® Biotechnologies) or T4 DNA Ligase (Roche Applied Science), as per manufacturer's instructions. Blunt-end ligations were performed at 4°C and sticky-end ligations were performed at room temperature, followed by heat inactivation of the ligase at 70°C for 15 min.

2.3.5 Polymerase Chain Reaction (PCR)

For screening and preliminary PCRs, the lower-fidelity Faststart Taq (Roche Applied Science) polymerase was used, as per the manufacturer's instructions. Briefly, the standard 25 µl reactions containing between 50-100 ng of plasmid or genomic DNA were set up with 1× reaction buffer, 200 µM of each dNTP, 0.5 µM of each primer, 1.5 mM $MgCl_2$, and 2U/50 µl of DNA polymerase. DNA amplification conditions were as follows: initial denaturation at 95°C for 4 min followed by 30 cycles of denaturation at 95°C for 30 sec, annealing (appropriate temperature for 30 sec) extension at 72°C for 1 min/kbp and final extension at 72°C for 7 min.

All PCR reactions of were performed in the MyCycler™ thermal cycler (Bio-Rad) with oligonucleotide primers purchased from UCT (Molecular and Cell Biology, University of Cape Town).

2.4 Bacterial transformation

2.4.1 Chemical transformation of *E. coli* cells

2.4.1.1 Preparation of competent cells

Rubidium chloride-treated *E. coli* DH5α cells were used for transformation of plasmid DNA into *E. coli* cells. Briefly, 1 ml of an overnight culture of *E. coli* DH5α was inoculated into 100 ml LB and grown to an OD_{600} of 0.48-0.55. Cells were kept on ice for 15 min and harvested by centrifugation at 3901 × *g* for 5 min at 4°C. The pellet was resuspended in 20 ml transformation buffer I (30 mM potassium acetate, 100 mM rubidium chloride, 50 mM manganese chloride and 15% (v/v) glycerol, pH 5.8) before storing on ice for a further 15 min and then harvesting by centrifugation at 3901 × g for 5 min at 4°C. The pellet was resuspended in 2 ml TfbII (10 mM MOPS, 75 mM calcium chloride, 10 mM rubidium chloride and 15% v/v glycerol: pH 6.5). Aliquots (500 µl) were flash-frozen in ethanol and stored at -80°C until required.

2.4.1.2 Transformation of *E. coli* with plasmid DNA

Competent *E. coli* DH5α cells were thawed on ice and 100 µl aliquots were incubated with up to 1µg plasmid DNA and kept on ice for 20 min. After 20 min, the cells were heat-shocked at 42°C for 90 sec then chilled on ice for 1-2 min; thereafter, 4 volumes of 2TY (Tryptone Yeast broth) was added to rescue the cells at 37°C for 1 hr. Cells were plated on LA plates containing the appropriate antibiotics, and incubated overnight at 37°C (Russell & Sambrook, 2001; Sambrook et al., 1989) or for two days at 30°C, for plasmids larger than 8kbp (Parish & Stoker, 2000).

2.4.2 Electroporation of *Msm*

Briefly, a single *Msm* colony was used to inoculate a pre-culture of 10 ml 7H9 OADC and grown overnight at 37°C. A volume of 1ml of the pre-culture was inoculated into 100 ml 7H9/OADC media and grown to log phase (OD_{600} of ~0.6 - 0.8) at 37°C with shaking. Cells were harvested by centrifugation (3901 × *g*, 10 min, 4°C) and washed three times in pre-chilled 10% glycerol (v/v). The cells were harvested by centrifugation between each washing step. After washing, the cell pellet was resuspended in 2 ml ice-cold 10% glycerol. A minimum of 1 µg to 4 µg plasmid DNA was added to a pre-chilled 2 mm electroporation cuvette (0.2 cm electrode gap, Bio-Rad). 400 µl electrocompetent cells were added to the electroporation cuvettes containing the plasmid DNA and pulsed once in a GenePulser™ (Bio-Rad). The pulsing conditions were as follows; 2 500 V, 1000 Ω, and 25 µF. Cells were rescued immediately with 800 µl LB and then transferred to a fresh Eppendorf tube. The tubes were incubated at 37°C for 4 hr and the cells were plated on 7H10/OADC supplemented with the appropriate antibiotic and incubated for 3-5 days at 37°C.

2.5 Southern blot analysis.

2.5.1 Electroblotting

Genomic DNA (5 µg) was digested overnight at 37°C with appropriate restriction enzymes, after which the DNA was separated by electrophoresis on a 1% agarose gel at 80 V and viewed and photographed with a fluorescent ruler with the Gel Doc (WealTeach Keta Imaging System). The DNA was depurinated by immersing the agarose gels in 0.25 M HCl for 15 min, followed by denaturation buffer (1.5 M NaCl/0.5 M NaOH) for 25 min, and then neutralized (1.5 M NaCl, 5 M Tris·HCl, pH 7.5) for 30 min. Thereafter, the DNA was transferred to a nitrocellulose membrane (Hybond™-N+ membrane, Amersham) overnight capillary transfer. Once transferred, the DNA was cross-linked to the membrane by irradiation in a UV Stratalinker 1800 (Stratagene) at 1200 mJ/cm^2, and membranes were hybridized immediately.

2.5.2 Synthesis and labelling of probes

All probes used for Southern blot analyses were synthesized by PCR using the oligonucleotides in **Table 2.3**. The ECL Direct Nucleic Acid Labelling and Detection System protocol (Amersham) was used to label probes. Briefly, a maximum of 100 ng of probe DNA in 10 µl of dH2O was denatured by boiling for 5 min at 95°C, and immediately cooled on ice for 5 min. Equal volumes of DNA labelling agent and gluteraldehyde (Amersham) were added to the probe, mixed gently and incubated for 15 min at 37°C. Following incubation, the labelled probe was used immediately in hybridization experiments.

Table 2.3: Oligonucleotides used to synthesize Southern blot probes

Name	**Sequence (5'-3')**
metE F	GACGAGGTCACCGAGTCA
metE R	CTCGCGCTTAGACCTA

2.5.3 Hybridisation

The hybridization buffer was prepared and used according to the ECL Direct Nucleic Acid Labelling and Detection System protocol (Amersham). Briefly, for membranes (10 cm^2 in size), 20 ml of hybridization buffer, 5% w/v blocking agent and 0.5 M NaCl was combined and stirred at room temperature for 1 hr and then at 42°C for 1 hr. Subsequent to cross-linking, the membrane was pre-hybridized in roller bottles in the Hybridisation oven/shaker SI3OH (Stuart) for 1 hr at 42°C, thereafter, followed by the addition of the labelled probe and the membrane was hybridized overnight at 42°C.

2.5.4 Detection

Following overnight hybridization, the membrane was washed twice in primary wash buffer (6 M Urea, 0.4% SDS, 0.5 × SSC) for 20 min at 42°C, and then twice for 5 min at room temperature in secondary wash buffer (0.5 × SSC). Thereafter, detection reagents 1 and 2 (Amersham) were mixed in equal quantities, transferred to the membrane and incubated for 1 min at room temperature. The membrane was then drained, saran wrapped and exposed to X-ray film (Amersham Biosciences) in a cassette for time periods ranging from 1 to 10 min at room temperature before developing.

2.5.5 Development of X-ray film

Development of the X-ray film was carried out in a dark room with red-light facilities. Immediately after exposure, for the appropriate time, the X-ray film was submerged in developer solution (GBX Developer, Carestream® Kodak®, Sigma-Aldrich) for 3 min, rinsed briefly in H_2O and submerged in fixer solution (GBX Fixer, Carestream® Kodak®, Sigma-Aldrich) for 3 min. X-ray films were air-dried completely before photographing.

2.6 Constructions of transposon mutant libraries

Tn libraries were constructed in WT *Msm* mc^2155 and Δ*metE* using the MycoMarT7 phage as described (Murry et al., 2008).

2.6.1 Preparation of high-titre phage stock

Briefly, 100 ml culture of *Msm* was grown to an OD_{600} ~ 3-4. Cells were washed twice with 7H9-ADC (no Tween80) and re-suspended in 100 µl of the same medium. 3.5 ml top agar (cooled to 42°C) was added into the washed cells and thereafter the mixture was poured into LA or 7H10 agar plates, the plates were left to dry at room temperature for a few hours. A 10-fold serial dilution of the φMycoMarT7 phage was performed in 50 µl MP buffer and 100 µl of cells were added in the respective dilutions of phage. The cells and phage were mixed with the top agar and poured on the LA plates. The plate was incubated at 30°C for 48 hr until small, distinct plaques appeared. Thereafter, phage from several individual plaques were picked using a sterile micro-tip and resuspended in 50 µl MP buffer. Aliquots of 10 µl were spotted onto two replicate LA plates containing *Msm* in top agar; plates were left to dry at room temperature and then incubated at different temperatures (one plate at 30°C and another at 37°C) for 48 hr. Clones which formed plaques only at 30°C and not at 37°C were selected and resuspended into MP buffer. The phages were recovered by centrifugation at 4000 × *g* for 1 min. The phages were titred to determine the dilution at which confluent plaques appeared on 7H10 plates. Then, 10 ml *Msm* were cultured in standard 7H9/OADC to an OD600 of 3-4 at 37ºC. 500 µl cells were washed twice with 7H9-ADC and re-suspended in 500 µl of the same medium. 100 µl of phage was added to 500 µL of washed cells. Aliquots of 100 µL of this mixture were added to 3.5 ml top agar (cooled to 42°C), mixed thoroughly, and then poured on five 7H10 plates. The plates were incubated at 30°C for about two days until confluent plaques appeared. After

incubation, plates were flooded with 2 ml MP buffer, and mixed gently at 20 rpm at room temperature for 5 hr, or overnight at 4°C. The remaining MP buffer (approximately 1 ml) from all plates was pooled and passed through a 0.22 µm syringe filter to prepare the final phage stock and stored at 4°C.

2.6.2 Transduction of *Msm* with φMycoMarT7

Msm strains were grown in Middlebrook 7H9 liquid medium containing glycerol, ADC and 0.05% Tween80 for 3 days, then the cells were diluted 1:100 in 7H9 medium and cultured at 37°C until they reached an OD_{600} of 2.0. The cells from a 50 ml culture were harvested by centrifugation at 4500 rpm for 15 min and washed three times with 50 ml of MP buffer warmed to 37°C. The supernatant was discarded, and the cells were resuspended in 4 ml of MP buffer. Both the bacterial culture and the phage stock were pre-heated at 39°C for 30 min. 2 ml of 1x 10^{12} warm phage were added into the warm bacterial cells and incubated for 18 hr at 37°C shaking incubator. After incubation 500 ul of the transduced cells were plated on 15 cm LA plates containing 20 µg/ml Kan and 0.05% Tween80. The plates were incubated at 37°C for three days. A small amount of the transduced cells was tittered on 10 cm plates with 20 µg/ml and 0.05% Tween80 to estimate the library size. Approximately 400000 of the transduced cells were harvested by scraping the colonies from the plates into 10 ml 7H9 with 15% glycerol and stored at -80^{o}C.

2.7 DNA Library preparation for sequencing

Genomic DNA was extracted from the Tn libraries, and mutant composition was determined by sequencing amplicons of the Tn-genome junctions, as previously described (Long et al., 2015). Briefly, genomic DNA was sheared into ~500 bp fragments by enzymatic shearing, and fragments were subjected to end repair and A-tailing with *Taq* polymerase and ligated to T-tailed adapters bearing random 7-nucleotide barcodes to distinguish between unique fragments before subsequent PCR amplification. Fragments containing Tn genome junctions were selectively enriched in a first PCR amplification, size selected in the 400 to 600 bp range and amplified in a second heminested PCR to add adapter sequences for Illumina sequencing. PCR amplicons were subjected to 75 to 100 bp paired-end sequencing on an Illumina HiSeq platform, and raw sequence data were exported to fastq files for further analysis (DeJesus et al., 2015; Long et al., 2015) (**Figure 2.1**).

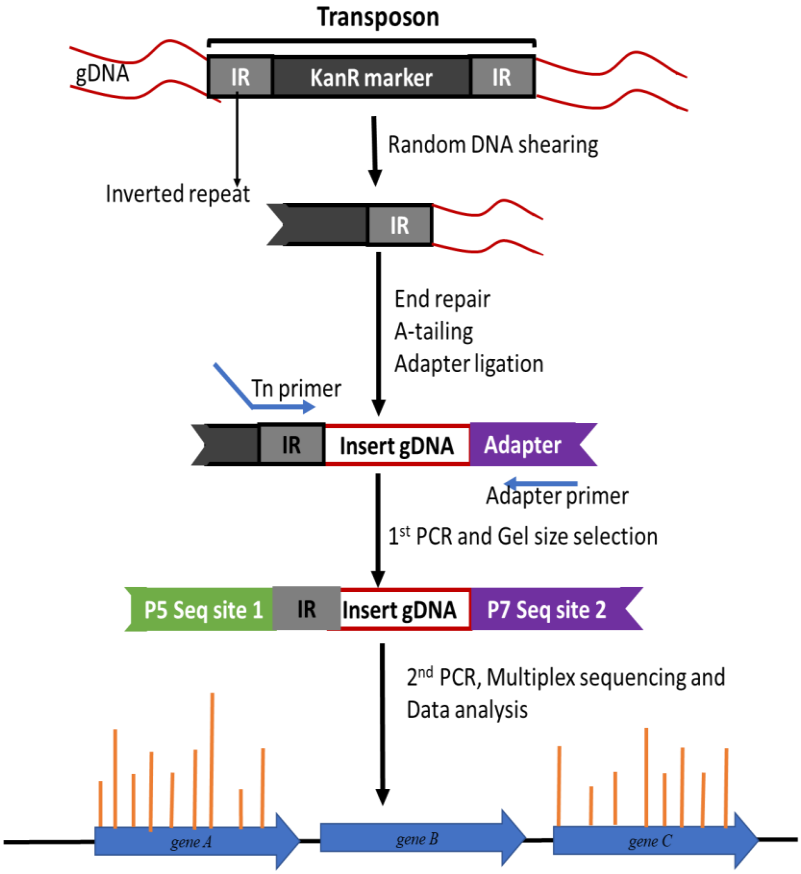


Figure 2.1: Preparation of DNA library for sequencing. An illustration of the workflow of DNA library preparation protocols. Genomic DNA is extracted and randomly fragmented by a restriction digest or mechanical shearing into ~500 bp fragments, followed by end repair and A-tailing with *Taq* polymerase and overnight adapter ligation. PCR amplification is performed to amplify the fragments containing transposon sequences after which sequencing adaptors are ligated. Lastly high-throughput sequencing is used to determine the location of the transposons in the genome and to determine the relative abundance of mutants containing a transposon at the insertion site, thus allowing fitness to be calculated.

2.7.1 DNA Fragmentation

Genomic DNA was sheared into ~50-200 bp fragments by enzymatic shearing. For each reaction: 5 µg of genomic DNA and 3 µl of 10× "next dsFragmentase" reaction buffer was added to dH_2O to up to 28 µl volume. Then, 2µl "next dsFragmentase" enzyme was added into the mixture and incubated for 30 min at 37°C to generate DNA fragments of sizes ranging from 50-200 bp. The fragmented DNA was purified using the QIAquick PCR purification kit as per the manufacturer's instructions and eluted into 50 µl dH_2O.

2.7.2 End repair

The sheared DNA was blunt-ended using Epicentre's "End-It" DNA End Repair kit according to the manufacturer's instructions. Briefly, 1.8 µg DNA, 5 µl of 10× "End-It" DNA repair buffer, 5 µl dNTPs, 5 µl ATP were mixed and water was added up to final volume of 50 µl. Thereafter, 1 µl of enzyme was added, mixed with a micropipette and incubated at 37°C for 45 min. The reaction was purified using a QIAquick PCR purification column and eluted with 65 µl water to perform A-tailing of the end-repaired DNA.

2.7.3 A-tailing

An adenosine nucleotide was added to the 3' end of each strand of blunt-ended DNA by adding 10 µl 10× Taq DNA polymerase buffer, 20 µl 10 mM dATP and 5 µl Taq DNA polymerase, the reaction was mixed with a micropipette and incubated at 72°C in a thermocycler for 45 min. The DNA was then purified using the Qiagen QIAquick PCR purification kit following the manufacturer's instructions and eluted with 50 µl distilled water.

2.7.4 Adaptor ligation

Barcoded adapters were prepared for ligation to the genomic DNA fragments by mixing an equal volume of the 100 µM adapter mix of adapter 1 and 2 **(Table 2.4)**. Briefly, 24 µl from a 100 µM stock of each adapter and 2 µl of 50 mM $MgCl_2$ were mixed into a PCR tube and heated to 95ºC for 10 min in a thermocycler. The temperature was slowly reduced to 20°C over a period of 2 hr with ramping at 1%. For each ligation reaction, 4 µl of the barcoded oligo were added to the 50 µl A-tailed DNA solution, 10 µl 10x ligase buffer and 8 µl T4 DNA ligase, and water was added up to 100 µl final volume. The reaction was incubated at 16°C overnight. The ligation reaction was purified using a QIAquick PCR purification kit as per the manufacturer's instructions with 4 additional washes of the DNA with PE buffer before elution into 100 µl of water.

2.7.5 PCR amplification of Tn-DNA junctions

Genomic DNA fragments were amplified to enrich Tn junction sites by PCR amplification using a primer with Tn homology and a primer with homology to the adapter sequence **(see Table 2.4)**. For the PCR, 8 reactions were prepared for each sample. Briefly, for each PCR, 100 ng of DNA was added into 2.5 µl of DMSO, 5 µl of 10× PCR buffer, 5 µl of 2.5 mM dNTPs, 0.5 µl Taq DNA polymerase, 1.5 µl of 10 µM adapter primer and 1.5 µl of 10 µM Tn primer. Water was added to each reaction to bring it up to a volume of 30 µl. PCR parameters were 95ºC for 10 min; 20 cycles of 95°C for 30 sec, 58°C for 30 sec and 72°C for 45 sec; and 72°C for 5 min. After PCR, 8 reactions were pooled for each sample. The amplified DNA was visualized on a 2% agarose gel run at 60 V for 140 min. Smears of amplified DNA in the 400-600 bp range were excised, placed in the -20°C freezer for 30 min and then gel extracted with the Qiagen Gel Extraction kit and eluted into 50 µl water.

2.7.6 Hemi-nested PCR

The staggered primers were mixed together in equal concentration prior to being added to the PCR mix **(Table 2.4)**. For hemi-nested PCR, 4 µl of amplified DNA product from the first round of PCR was mixed with 2 µl of 1 µM Sol-mar mix, 2 µl of 1 µM Sol-ap-tag mix, 2µl of 10× PCR buffer, 2 µl of 2.5 mM dNTPs, 0.2 µl of Taq polymerase and water was added up to final volume of 20 µl. PCR parameters were 95°C for 5 min; 10 cycles of 95°C for 30 sec, 58°C for 30 sec and 72°C for 45 sec; and 72ºC for 5 min. Amplified DNA was purified using the QIAquick PCR purification kit as per the manufacturer's instructions and and eluted into 50 µl EB buffer. The DNA concentration was measured by Qubit.

2.8 DNA sequencing

All sequencing of plasmid constructs or PCR products was performed as a pay-for-service by the Central Analytical Sequencing Facility at Stellenbosch University. Massive parallel sequencing (MPS) of Tn mutant libraries was performed by Oklahoma Medical Research Foundation Genomics Core (Oklahoma City, USA).

Table 2.4: Primers and adapters used for TnSeq. The long Sol_AP1 primers recognize the adapter portion of the sequence whereas Sol_Mar recognizes the Tn (Long et al., 2015).

Primer name		Sequence	Function
Adapter 1		5′-TACCACGACCA-**NH2**	Barcoded adapter that ligate to genomic DNA after fragmentation
Adapter 2		5′ATGATGGCCGGTGGATTTGTGNNANNANNNTGGTCGTGGTAT	
JEL_AP1 (Adapter primer)		5′-ATGATGGCCGGTGGATTTGTG	Amplifies the sequence between barcoded genomic DNA fragment and inserted Tn.
T7 (Tn primer)		5′-TAATACGACTCACTATAGGGTCTAGAG	
1.Sol_AP1_ tagged 930	**A**	CAAGCAGAAGACGGCATACGAGAT***TGTTCCGAG***TGACTGGAGTTCAGACGTGTGCTCTTCCGATCTGTCAATGATGGCCGGTGGATTTGTG	Multiplex sequencing adapter. Long primers which inserts specific sequence (in bold) allowing identification of samples after multiplex sequencing.
	B	CAAGCAGAAGACGGCATACGAGAT***TGTTCCGAG***TGACTGGAGTTCAGACGTGTGCTCTTCCGATCTCGTCCATGATGGCCGGTGGATTTGTG	
	C	CAAGCAGAAGACGGCATACGAGAT***TGTTCCGAG***TGACTGGAGTTCAGACGTGTGCTCTTCCGATCTACAGTCCCATGATGGCCGGTGGATTTGTG	
	D	CAAGCAGAAGACGGCATACGAGAT***TGTTCCGAG***TGACTGGAGTTCAGACGTGTGCTCTTCCGATCTTAGTGGATGATGGCCGGTGGATTTGTG	
2.Sol_AP1_tagged_949	**A**	CAA GCA GAA GAC GGC ATA CGA GAT ***TTC CGG AG***G TGA CTG GAG TTC AGA CGT GTG CTC TTC CGA TCT GTC AAT GAT GGC CGG TGG ATT TGT G	
	B	CAA GCA GAA GAC GGC ATA CGA GAT ***TTC CGG AG***G TGA CTG GAG TTC AGA CGT GTG CTC TTC CGA TCT CGT CCA TGA TGG CCG GTG GAT TTG TG	
	C	CAA GCA GAA GAC GGC ATA CGA GAT ***TTC CGG AG***G TGA CTG GAG TTC AGA CGT GTG CTC TTC CGA TCT ACA GTC CCA TGA TGG CCG GTG GAT TTG TG	
	D	CAA GCA GAA GAC GGC ATA CGA GAT ***TTC CGG AG***G TGA CTG GAG TTC AGA CGT GTG CTC TTC	

		CGA TCT TAG TGG ATG ATG GCC GGT GGA TTT GTG	
3.Sol_AP1_tagged_651	**A**	CAA GCA GAA GAC GGC ATA CGA GAT ***GCC GAT GT***G TGA CTG GAG TTC AGA CGT GTG CTC TTC CGA TCT GTC AAT GAT GGC CGG TGG ATT TGT G	
	B	CAA GCA GAA GAC GGC ATA CGA GAT ***GCC GAT GT***G TGA CTG GAG TTC AGA CGT GTG CTC TTC CGA TCT CGT CCA TGA TGG CCG GTG GAT TTG TG	
	C	CAA GCA GAA GAC GGC ATA CGA GAT ***GCC GAT GT***G TGA CTG GAG TTC AGA CGT GTG CTC TTC CGA TCT ACA GTC CCA TGA TGG CCG GTG GAT TTG TG	
	D	CAA GCA GAA GAC GGC ATA CGA GAT ***GCC GAT GT***G TGA CTG GAG TTC AGA CGT GTG CTC TTC CGA TCT TAG TGG ATG ATG GCC GGT GGA TTT GTG	
4.Sol_AP1_tagged_373	**A**	CAA GCA GAA GAC GGC ATA CGA GAT ***CAT GAT CG***G TGA CTG GAG TTC AGA CGT GTG CTC TTC CGA TCT GTC AAT GAT GGC CGG TGG ATT TGT G	
	B	CAA GCA GAA GAC GGC ATA CGA GAT ***CAT GAT CG***G TGA CTG GAG TTC AGA CGT GTG CTC TTC CGA TCT CGT CCA TGA TGG CCG GTG GAT TTG TG	
	C	CAA GCA GAA GAC GGC ATA CGA GAT ***CAT GAT CG***G TGA CTG GAG TTC AGA CGT GTG CTC TTC CGA TCT ACA GTC CCA TGA TGG CCG GTG GAT TTG TG	
	D	CAA GCA GAA GAC GGC ATA CGA GAT ***CAT GAT CG***G TGA CTG GAG TTC AGA CGT GTG CTC TTC CGA TCT TAG TGG ATG ATG GCC GGT GGA TTT GTG	
5.Sol_AP1_tagged_473	**A**	CAA GCA GAA GAC GGC ATA CGA GAT **CGC GCG GT**G TGA CTG GAG TTC AGA CGT GTG CTC TTC CGA TCT GTC AAT GAT GGC CGG TGG ATT TGT G	
	B	CAA GCA GAA GAC GGC ATA CGA GAT **CGC GCG GT**G TGA CTG GAG TTC AGA CGT GTG CTC TTC CGA TCT CGT CCA TGA TGG CCG GTG GAT TTG TG	

	C	CAA GCA GAA GAC GGC ATA CGA GAT **CGC GCG GT**G TGA CTG GAG TTC AGA CGT GTG CTC TTC CGA TCT ACA GTC CCA TGA TGG CCG GTG GAT TTG TG	
	D	CAA GCA GAA GAC GGC ATA CGA GAT **CGC GCG GT**G TGA CTG GAG TTC AGA CGT GTG CTC TTC CGA TCT TAG TGG ATG ATG GCC GGT GGA TTT GTG	
6.Sol_AP1_tagged_100	**A**	CAA GCA GAA GAC GGC ATA CGA GAT **ACA CGA TC**G TGA CTG GAG TTC AGA CGT GTG CTC TTC CGA TCT GTC AAT GAT GGC CGG TGG ATT TGT G	
	B	CAA GCA GAA GAC GGC ATA CGA GAT **ACA CGA TC**G TGA CTG GAG TTC AGA CGT GTG CTC TTC CGA TCT CGT CCA TGA TGG CCG GTG GAT TTG TG	
	C	CAA GCA GAA GAC GGC ATA CGA GAT **ACA CGA TC**G TGA CTG GAG TTC AGA CGT GTG CTC TTC CGA TCT ACA GTC CCA TGA TGG CCG GTG GAT TTG TG	
	D	CAA GCA GAA GAC GGC ATA CGA GAT **ACA CGA TC**G TGA CTG GAG TTC AGA CGT GTG CTC TTC CGA TCT TAG TGG ATG ATG GCC GGT GGA TTT GTG	
7.Sol_AP1_tagged_57	**A**	CAA GCA GAA GAC GGC ATA CGA GAT AAG TAG AGG TGA CTG GAG TTC AGA CGT GTG CTC TTC CGA TCT GTC AAT GAT GGC CGG TGG ATT TGT G	
	B	CAA GCA GAA GAC GGC ATA CGA GAT AAG TAG AGG TGA CTG GAG TTC AGA CGT GTG CTC TTC CGA TCT CGT CCA TGA TGG CCG GTG GAT TTG TG	
	C	CAA GCA GAA GAC GGC ATA CGA GAT AAG TAG AGG TGA CTG GAG TTC AGA CGT GTG CTC TTC CGA TCT ACA GTC CCA TGA TGG CCG GTG GAT TTG TG	
	D	CAA GCA GAA GAC GGC ATA CGA GAT AAG TAG AGG TGA CTG GAG TTC AGA CGT GTG CTC TTC CGA TCT TAG TGG ATG ATG GCC GGT GGA TTT GTG	
8.Sol_AP1_tagged_598	**A**	CAA GCA GAA GAC GGC ATA CGA GAT GAG ATC TTG TGA CTG GAG TTC AGA CGT GTG CTC TTC CGA TCT GTC AAT GAT GGC CGG TGG ATT TGT G	

	B	CAA GCA GAA GAC GGC ATA CGA GAT GAG ATC TTG TGA CTG GAG TTC AGA CGT GTG CTC TTC CGA TCT CGT CCA TGA TGG CCG GTG GAT TTG TG	
	C	CAA GCA GAA GAC GGC ATA CGA GAT GAG ATC TTG TGA CTG GAG TTC AGA CGT GTG CTC TTC CGA TCT ACA GTC CCA TGA TGG CCG GTG GAT TTG TG	
	D	CAA GCA GAA GAC GGC ATA CGA GAT GAG ATC TTG TGA CTG GAG TTC AGA CGT GTG CTC TTC CGA TCT TAG TGG ATG ATG GCC GGT GGA TTT GTG	
9.Sol_ Mar	**A**	AATGATACGGCGACCACCGAGATCTACACTCTTT CCCTACACGACGCTCTTCCGATCTCGGGGACTTAT CAGCCAACC	Common sequencing primer which recognizes the Tn.
	B	AATGATACGGCGACCACCGAGATCTACACTCTTT CCCTACACGACGCTCTTCCGATCTTCGGGGACTTA TCAGCCAACC	
	C	AATGATACGGCGACCACCGAGATCTACACTCTTT CCCTACACGACGCTCTTCCGATCTGATACGGGGA CTTATCAGCCAACC	
	D	AATGATACGGCGACCACCGAGATCTACACTCTTT CCCTACACGACGCTCTTCCGATCTATCTACGGGGA CTTATCAGCCAACC	

2.9 Data processing

Sequence data were processed using the TPP tool included with TRANSIT (DeJesus et al., 2015). Reads were mapped to the genome using the Burroughs-Wheeler Aligner. The set of reads in "read1" with a prefix matching the end of the *Himar1* transposon were mapped to the corresponding -TA- site in the genome (stripping off the Tn prefix). The read counts were reduced to template counts by discarding duplicates with the same barcode in "read2." The final template counts were normalized across all data sets using Trimmed Total Reads (TTR) normalization. TTR normalizes data sets so that they have the same mean template count, while ignoring ("trimming") the top and bottom 5% of read counts to reduce the influence of outliers.

2.10 Gene silencing of *Msm* cobalamin genes by CRISPRi

Transcriptional knockdown (KD) of target genes using the clustered randomly interspaced short palindromic repeats interference (CRISPRi) approach was carried out as described (Rock et al., 2017). Short guide RNAs (sgRNAs) were designed as described (de wet et al., 2018), and 8 sgRNA oligo pairs – top oligo, forward primer, and bottom oligo, reverse primer – targeting each "cobalamin gene" were synthesised (**Table 2.5**). The top and bottom oligos were annealed to generate an sgRNA which was then cloned into the pLJR962 plasmid (Rock et al., 2017) using *Bsm*BI restriction. The ligation reaction was performed overnight in a 10 µl reaction volume using T4 DNA ligase (NEB), and the entire ligation mix from each reaction was transformed into 100 µl electrocompetent *E. coli* DH5α. Transformants were selected on LA plates with 50 µg/ml Kan. For each transformation, plasmid DNA was extracted from single colonies and sequenced with primer 1834 (**Table 2.5**) to verify the cobalamin KD constructs. As a positive control, a *mmpL3* KD construct was also generated. Electroporation competent *Msm* Δ*metE* cells were transformed by electroporation with 100ng of the cobalamin KD or *mmpL3* KD constructs. Each independent electroporation was selected on 7H10-OADC containing 25 µg/ml Kan. To induce gene silencing by CRISPRi, 100 ng/ml anhydrotetracycline (ATc) was included in the selection plates. For uninduced controls, plates did not include ATc.

Table 2.5: Oligos used to generate sgRNAs

Oligos ID		**Sequence (5'-3')**
cobA	F1	GGGAGCGCGCGATGTCGGCGCGCAC
	R1	AAACGTGCGCGCCGACATCGCGCGC
cobIJ	F2	GGGAGCCGCCGGGCACCATGGTCAG
	R2	AAACCTGACCATGGTGCCCGGCGGC
cobN	F3	GGGAACGGCCTTGGGATCCACCGA
	R3	AAACTCGGTGGATCCCAAGGCCGT
cobO	F4	GGGAATAGGTGAACTCGTCGAGGAC
	R4	AAACGTCCTCGACGAGTTCACCTAT
cobU	F5	GGGAGCTCACCATGACCAGGGGTGC
	R5	AAACGCACCCCTGGTCATGGTGAGC
cobT	F6	GGGAGCGTCGCACGGCGGCCTGCGC

	R6	AAACGCGCAGGCCGCCGTGCGACGC
bluB	F7	GGGAGATCTCCAGGACCGGCCGGT
	R7	AAACACCGGCCGGTCCTGGAGATC
mmpL3	F8	GGGAACAGACTGGCTGCCCTCGTC
	R8	AAACGACGAGGGCAGCCAGTCTGT
Primer 1834	P1834	TTCCTGTGAAGAGCCATTGATAATG

2.11 Phenotypic characterization of *Msm* B_{12} mutants

Where relevant, selected *Msm* mutants were assessed for their potential to utilize substrates such as vitamin B_{12} (CNCbl) and L-methionine to complement growth. Briefly, a 10ml culture was grown in Sauton's medium (Allen, 1998) until mid-log phase (OD_{600}=0.5-1). Thereafter, cultures were inoculated at an OD_{600} of 0.1 in 25 ml fresh Sauton's and incubated at 37°C with shaking until OD_{600} = 0.6. Cultures were serially diluted in un-supplemented Sauton's medium and spotted on 7H10 OADC agar plates containing the desired supplements and incubated for 3 days. CNCbl and L-methionine were utilized at a final concentration of 10 μg/ml and 1 μM, respectively.

Chapter 3: Results

3 Investigating the functionality of the cobalamin biosynthetic pathway in *Msm* utilizing a high-throughput genetic approach

Comparative genomic analyses have identified an altered capacity for cobalamin biosynthesis as a critical step in the evolution of the pathogenic MTBC strains from a common environmental ancestor (Boritsch et al., 2014; Supply et al., 2013; Young et al., 2015). This was inferred using the presence of a full-length *cobF* gene – encoding a methyltransferase which catalyses the conversion of precorrin 5 to precorrin 6x – as a proxy for a functional *de novo* cobalamin biosynthetic pathway. While compelling, a fundamental limitation of this deduction is that the current "model" pathway for cobalamin biosynthesis is largely based on gene annotation and is not supported by exhaustive functional analyses in all organisms. One consequence is that there is no certainty that *cobF* encodes a non-redundant enzymatic function; stated alternatively, the pathway for cobalamin biosynthesis in *Mtb* remains theoretical. This owes a lot to the complexity inherent in validating the functions of approximately 30 genes individually: *de novo* cobalamin biosynthesis comprises at least 30 enzyme-catalysed steps (Smith et al., 2018; Warren et al., 2002). Therefore, resolving the full gene complement involved in the complex, multi-step pathway for cobalamin biosynthesis, assimilation, and salvage in different species is challenging.

3.1 Constituents of the cobalamin biosynthetic pathway in *Mtb*

The whole-genome sequences of *Mtb*, together with the related mycobacterial pathogens, *M. marinum*, *M. kansasii* and *M. leprae*, revealed diversity amongst genes associated with cobalamin-related metabolism as a consequence of gene deletion events, differential acquisition of genes by horizontal transfer, and single nucleotide polymorphisms with predicted impact on protein function and transcriptional regulation (Young et al., 2015). Differences in cobalamin synthesis, methionine biosynthesis, fatty acid catabolism, and DNA repair and replication are consistent with adaptations to different environmental niches and pathogenic lifestyles (Young et al., 2015). Multiple comparative genomics studies have predicted that mycobacteria can assimilate or utilize cobalamin (Gopinath, Venclovas, et al., 2013; Minias et al., 2018; Rodionov et al., 2003; Shelton et al., 2019; Vitreschak et al., 2003; Zhang et al., 2009). *M. leprae* (like *Mtb*, an obligate pathogen) lacks a complete set of genes

for cobalamin biosynthesis but retains those genes required for transporting exogenous cobalamin precursors (Rodionov et al., 2003). *Mtb*, on the other hand, possesses a near complete cobalamin biosynthetic pathway (Gopinath, Moosa, et al., 2013; Rodionov et al., 2003), whereas *M. marinum* and *M. kansasii* possesses complete cobalamin biosynthetic pathways (Young et al., 2015).

Previous studies in the Molecular Mycobacteriology Research Unit have identified putative homologues of most cobalamin biosynthetic enzymes in *Mtb* and *Msm*, enabling the construction of a predicted cobalamin biosynthetic pathway (Gopinath, Moosa, et al., 2013). The proposed pathway is suggestive of the aerobic type because the CobG mono-oxygenase, which contains an iron-sulphur centre and is responsible for converting precorrin-3A into precorrin-3B, is of the class that requires molecular oxygen for activity (Debussche et al., 1993). Also, *Mtb* possesses aerobic-type *cobN*-Rv2850c-encoded subunits. In addition, it is predicted that cobalt insertion occurs late in the *Mtb* cobalamin biosynthetic pathway, and the *Mtb* CobK and CobJ orthologues contain residues commonly found in aerobic bacteria (Warren et al., 2002). Notably, *Mtb* also appears to possess some features that are characteristic of anaerobic biosynthesis. For example, Rv0259c exhibits homology to CbiX, a cobalt chelatase (Raux et al., 1998), that was identified in *Bacillus megaterium* – possibly implying that mycobacteria might possess partial capacity to use the anaerobic pathway (Gopinath, Moosa, et al., 2013).

As noted above, a distinctive feature of *Mtb* is that the *cobF*-encoded precorrin-6a synthase was lost during evolution from the *M. canettii*-like ancestor (Supply et al., 2013; Young et al., 2015). Instead, it has been suggested that CobF function in *Mtb* is provided by an alternative methyltransferase: either Rv2067c, which contains a C-terminal methyltransferase domain, or Rv0391 (designated as *metZ*) (Gopinath, Moosa, et al., 2013; Rodionov et al., 2003). In addition, it has been postulated that *Mtb* may rely on the host environment as a source of cobamides (Young et al., 2015).

Cobalt insertion into hydrogenobyrinic-acid a,c-diamide is predicted to be catalyzed by a CobN-CobS-CobT fusion protein (Debussche et al., 1992; Rodionov et al., 2003). CobS and CobT are thought to form a complex that interacts with CobN to generate cob(II)yrinic acid a,c-diamide (Moosa, 2013). The gene encoding the cob(II)yrinic acid a,c-diamide reductase is yet to be identified in the mycobacterial genome. In *Mtb*, BluB, which also catalyses the formation of DMB, was predicted to encode the cobalt reductase function (Gopinath, Moosa,

et al., 2013; Rodionov et al., 2003) based on sequence homology to CobR, a predicted cob(II)yrinate a,c diamide reductase (Lawrence et al., 2008).

In contrast to *Mtb*, *Msm* appears to encode all enzymes required for *de novo* cobalamin biosynthesis and the predicted pathway is also suggestive of the aerobic type. Additionally, *Msm* encodes five other putative cobalamin biosynthesis proteins (MSMEG_1123, MSMEG_2607, MSMEG_4934, MSMEG_6048 and MSMEG_6069). Although their enzymatic functions have not been confirmed, their predicted gene annotations are suggestive of roles in cobalamin metabolism. Furthermore, *Msm* encodes putative homologues of additional cobalamin-dependent enzymes that are not encoded in *Mtb*: specifically, glutamate mutase (MSMEG_0969), ethanolamine ammonium lyase made up of small and large subunits (MSMEG_1553-1554), and three glycerol dehydratases made up of small and large subunits (MSMEG_0496-0497, MSMEG_6320-6321 and MSMEG_1547-1548). Therefore, while the genome of *Msm* does not encode the same cobalamin-dependent enzymes that characterizes the genome of *Mtb*, this non-pathogenic mycobacterium appears to contain the full machinery required for *de novo* cobalamin biosynthesis. Moreover, *in vitro* biosynthesis of cobalamin was recently confirmed microbiologically and biochemically (Kipkorir et al., 2021).

In this study, a genome-scale approach was adopted to yield detailed genetic maps of *de novo* cobalamin biosynthesis and salvage in *Msm*. *Msm* is convenient model for the general study of mycobacteria because it has a relatively fast doubling time of approximately 3 hr and requires a biosafety level 2 laboratory. On the other hand, *Mtb* is a slow grower, has doubling time of 24 hr in standard broth medium and requires biosafety level 3 facilities (James et al., 2000). Moreover, different aspects of mycobacterial physiology and metabolism have been studied using non-pathogenic saprophyte, *Msm* strain mc^2155 as a model mycobacterium (Barry, 2001; Reyrat & Kahn, 2001). However, the comparison between an environmental mycobacterium (*Msm*) and an obligate human pathogen (*Mtb*) might be more useful than simply regarding *Msm* as a "model". Differences observed in cobalamin-dependent metabolism between *Msm* and *Mtb* might be more informative in understanding the evolution of *Mtb* into a human pathogen, as well as understanding the role of cobalamin in mycobacterial pathogenesis (Kipkorir et al., 2021). The approach involved selection of conditionally essential genes using whole-genome transposon (Tn) mutagenesis combined with next-generation sequencing, called transposon insertion sequencing (TnSeq) (Cain et al., 2020; Chao et al., 2016). In this study, the frequencies of mutants under-represented or over-represented in pools of an *Msm*

Δ*metE* Tn mutant library we investigated during growth in defined medium with or without CNCbl or cobalt supplementation. Following negative selection under nutrient limited or nutrient-starved conditions, mutants with insertions in genes that are conditionally essential should fail to survive the selection pressure and will, therefore, be underrepresented. In this study, TnSeq was applied to elucidate the full complement of *de novo* cobalamin biosynthetic genes in *Msm*. It was predicted that disruption of any gene required for cobalamin biosynthesis would result in growth impairment of the Δ*metE* mutant – a strain dependent on the alternative, cobamide-dependent methionine synthase, MetH, for methionine production (Kipkorir et al., 2021) – when grown in medium without CNCbl supplementation.

3.2 Phenotypic assessments of the *Msm* Δ*metE* mutant in minimal defined media with or/ without CNCbl and cobalt supplementation

Previous work in my host laboratory demonstrated that an *Mtb* Δ*metE* mutant lacking the cobalamin-independent methionine synthase was unable to grow *in vitro* unless supplemented exogenously with either cyanocobalamin (CNCbl; vitamin B_{12}) (Warner et al., 2007) or cobalt (Moosa, 2013). Since the current study aimed to exploit the non-pathogen, *Msm*, as model mycobacterium, it was necessary first to determine the ability of CNCbl and cobalt to complement growth of the corresponding *Msm* Δ*metE* mutant. The genotype of the Δ*metE* mutant was validated by Southern blot, confirming an in-frame deletion of the entire *metE* open reading frame (ORF) (**Figure 3.1A**). To examine the growth phenotype of the Δ*metE* knockout, cells were grown overnight in Middlebrook 7H9-OADC broth until they reached log-phase (OD_{600} ~0.8). Thereafter, the cells were re-inoculated at a starting OD_{600} ~0.005 in fresh Sauton's minimal medium and grown at 37ºC for 24 hr with optical density readings recorded every 3 hr. Unlike the corresponding *Mtb* mutant (Warner et al., 2007), *Msm* Δ*metE* was able to grow with kinetics similar to the parental WT *Msm* mc^2155 (**Figure 3.1C**). This result indicated that, in contrast to *Mtb*, *Msm* Δ*metE* was capable of *de novo* cobalamin biosynthesis *in vitro* under standard (aerobic) culture conditions – an observation that was subsequently confirmed separately by mass spectrometric analyses of cobalamin production (Kipkorir et al., 2021). The phenotype of the Δ*metE* mutant therefore supported the potential utility of this mutant as experimental strain in a high-throughput approach to generating a detailed genetic map for *de novo* cobalamin biosynthesis in mycobacteria. To this end, TnSeq was applied in WT and Δ*metE* mutant strains during growth in defined minimal medium, with and without supplementation with cobalt and/or CNCbl.

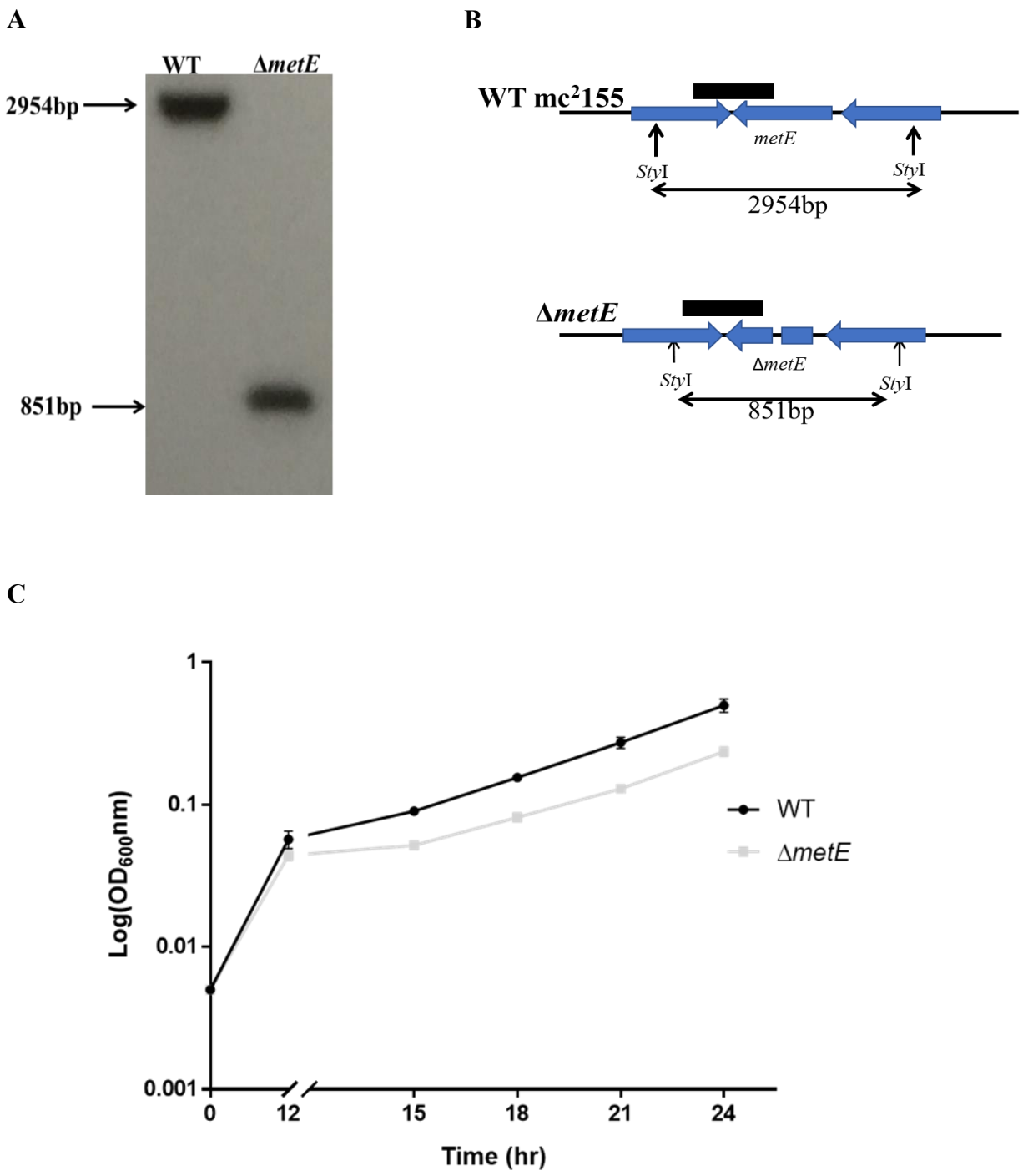


Figure 3.1: Genotypic and phenotypic characterization of *Msm* Δ*metE*. (**A**) The genotype of the *Msm* Δ*metE* mutant (kind gift of Dr. Stephanie Dawes, unpublished) was verified by Southern blot, in which *Sty*I-digested genomic DNA was probed with a 251 bp fragment complementary to downstream and intragenic *metE* sequence, yielding a 2954 bp fragment in WT *Msm* mc^{2}155 and 851 bp fragment in the Δ*metE* mutant. (**B**) Schematic representation of the WT and Δ*metE* alleles showing restriction sites and probe (black box). (**C**) Growth curves of WT *Msm* and Δ*metE* mutant generated by plotting OD_{600}nm values recorded every 3 hr over 24 hr during growth of bacilli *in vitro* in Sauton's minimal medium under aerobic conditions. Data are representative of three independent experiments. Error bars show the standard error of the mean.

In addition to CNCbl, cobalt supplementation has been shown to enable growth of *Mtb* Δ*metE* (Mashabela, Moosa *et al.*, in preparation). This suggests that cobalt might be a limiting element in standard mycobacterial growth media, a deficiency that could be exacerbated in defined media such as Sauton's. Therefore, the capacity for cobalt supplementation to enhance growth of *Msm* Δ*metE* was investigated in defined minimal medium (**Figure 3.2**). Surprisingly, no difference was observed for the Δ*metE* strain when grown in Sauton's minimal medium with/ without 10 µg/ml CNCbl or in medium supplemented with 0.65 µg/ml cobalt.

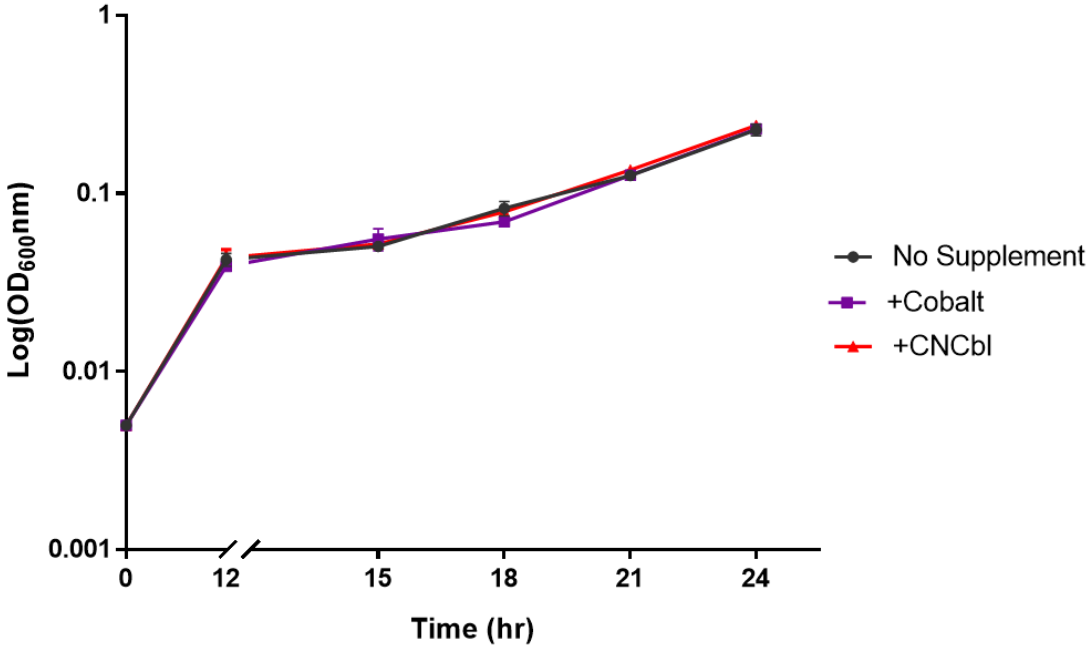


Figure 3.2: Growth of *Msm* Δ*metE* in defined minimal medium. Growth of the Δ*metE* strain in Sauton's minimal medium in the absence (No Supplement) or presence of either 10 µg/ml exogenous CNCbl (B_{12}) or 0.65 µg/ml cobalt (Cobalt). To examine the growth phenotype of the Δ*metE* knockout, cells were grown overnight until they reach log-phase. Thereafter, the cells were washed and re-inoculated at a starting OD_{600} of approximately 0.005 in fresh Sauton's minimal medium containing 10 µg/ml of CNCbl or 0.65 µg/ml cobalt grown at 37°C for 24 hr with optical density readings taken every 3 hr. Data are representative of three independent experiments. Error bars show the standard error of the mean.

3.3 Construction of Δ*metE* Tn library

Tn mutagenesis is an effective method for creating random mutations at high frequency in the genome of mycobacteria (Siegrist & Rubin, 2009; Xu et al., 2011). The Tn inserts randomly at -TA- dinucleotides (Rubin et al., 1999), disrupting the genetic sequence in that gene/locus (Hamer et al., 2001; Long et al., 2015). Transduction in fast-growing mycobacteria (such as *Msm*) has historically proven much less efficient than slow-growing bacteria (Sassetti et al., 2001). However, optimizing a modified transduction protocol enabled construction of high-density Tn libraries in *Msm* mc^2155 (Majumdar et al., 2017). So, using the transducing phage, pMycoMarT7, a Δ*metE* library comprising 400,000 CFU/ml individual Tn insertion mutants was generated (**Figure 3.3A**).

Following growth of the *metE* Tn library on nutrient-rich solid LA plates, genomic DNA was extracted. Next, the DNA was sheared, and end-repaired, sequencing adapters were added, and the Tn adjacent regions were enriched by PCR and subjected to next generation sequencing to quantify all Tn junctions. The resulting sequencing data were analysed using the TRANSIT software package (DeJesus et al., 2015). The sequence reads were aligned to a reference genome in order to determine the locations of Tn insertions. A Hidden Markov Model (HMM) incorporated into the TRANSIT platform is used to establish the exact location of each Tn insertion and the number of reads mapped to each insertion site and to determine the number of insertions/reads in a specific genomic region. This information is used to identify genes which have a significantly reduced number of insertions and reads in the library. Those genes that are underrepresented or over-represented in the mutant's library are identified as possibly essential or advantageous for growth under the specific set of conditions.

Analysis of the resulting sequencing data revealed that Tn insertions were obtained in 58,156 of the 77,755 possible -TA- sites, equating to 74.8% coverage with an average of 1,799,299 unique template counts (Tn insertion counts per -TA- site) and an average count of 30 per -TA- site; also, the identified insertions were distributed throughout the genome (**Figure 3.3B**). Thereafter, the library was cultivated in defined minimal media to enable identification of conditionally essential genes – including those required for *de novo* cobalamin biosynthesis.

A

	WT mc^{2}155	**Δ*metE***
Medium (agar)	7H10 Tw, Kan	LA medium, Tw, Kan
Library size(CFU/ml)	~240 000	~400 000
Total reads	3,716,188	1,799,299
# TA sites hit /77755 sites	59230	58156
%TA sites coverage	76.2	74.8

B

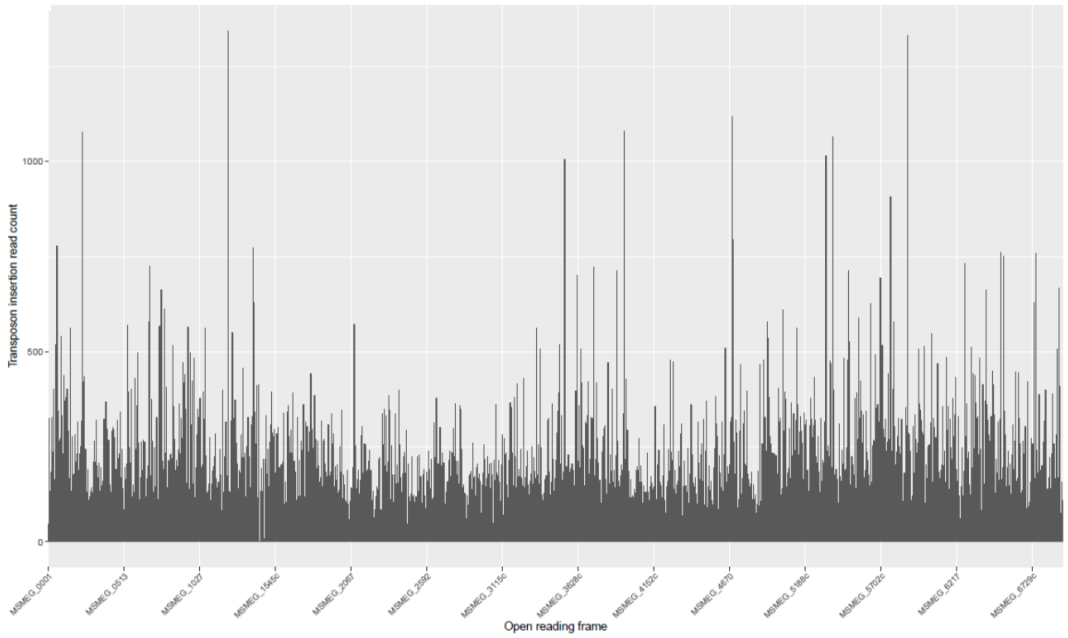


Figure 3.3: The insertion profile of the Δ*metE* Tn library following growth in rich media. (**A**) Overview of library parameters, including library size, selective conditions and -TA- sites hit for the two sequenced libraries. (**B**) Tn insertion counts across the Δ*metE* genome (LA-selected library). The height of the black bars represents the number of insertions at each individual -TA- site.

3.4 Identification of Tn insertions in *Msm* Δ*metE*

TnSeq enables elucidation of genes essential for growth under a specific set of conditions (DeJesus & Ioerger, 2013). The HMM incorporated into the TRANSIT platform (DeJesus et al., 2015) identifies essential genes and essential genomic regions based on the read count at a given site and the distribution over the surrounding sites, respectively, under a single growth condition. Of the 6,716 ORFs annotated in the *Msm* genome, 93.5% (6,277) were interrupted at least once in the *metE* Tn library (**Figure 3.4**). Another 4,8% of ORFs may have been eliminated because they were essential for growth in the LA medium. The remaining 1.7% of genes were untraceable owing to insufficient -TA- sites to allow for definitive assignment.

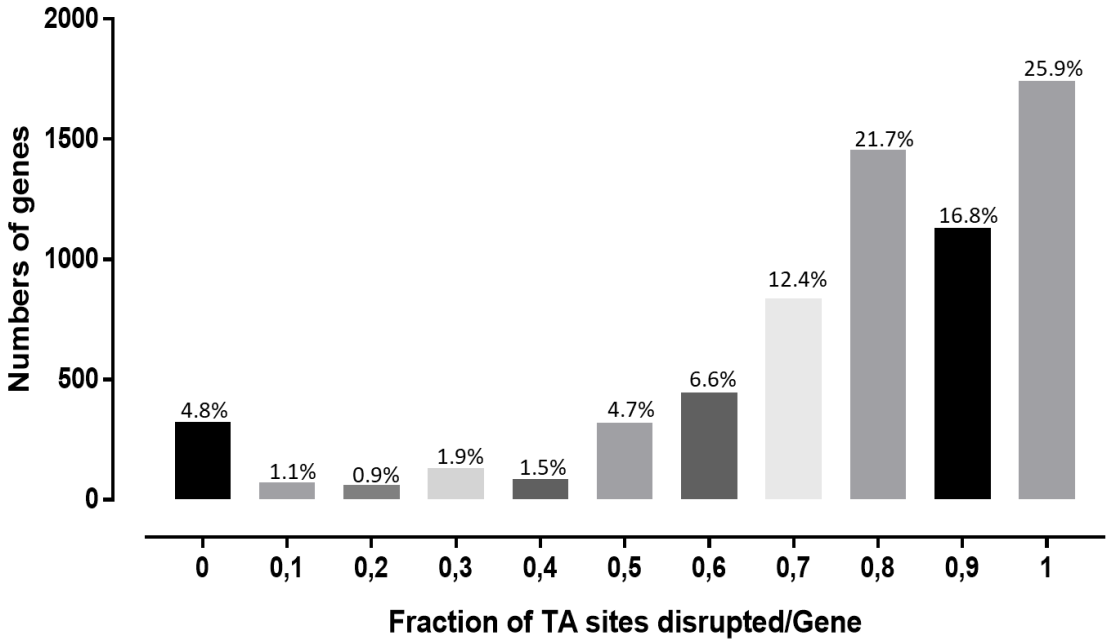


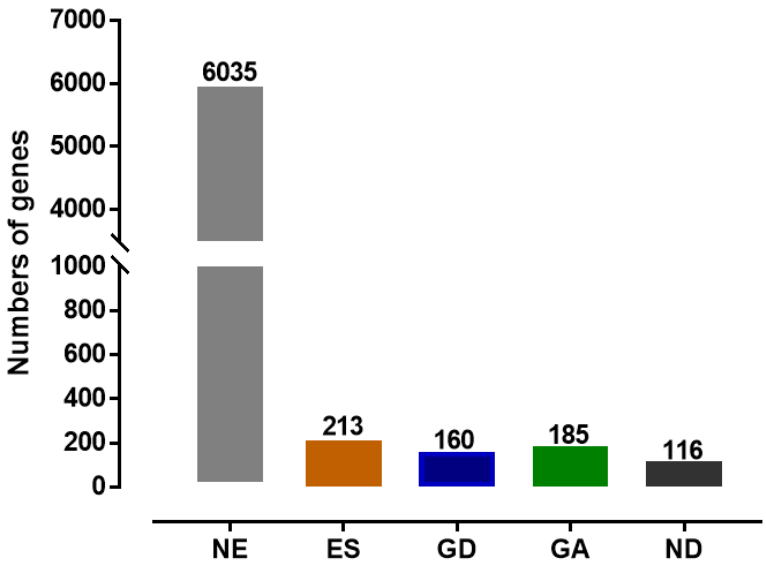
Figure 3.4: Frequency distribution of the fraction of disrupted TA sites per gene. For each gene, the fraction of disrupted -TA- sites was calculated based on the average read coverage. Displayed are the numbers of genes carrying Tn insertions in none (0) to all (1) -TA- sites.

3.5 Identification of essential genes in *Msm* Δ*metE*

Of 6716 genes, a total of 213 genes were identified as essential (ES). These included many well-known housekeeping genes which are involved in core macromolecular and cellular functions such as DNA replication, protein translation, cell growth and division, and core metabolism (**Figure 3.5B**). A further 160 genes were identified as growth defective (GD); that is, genes whose disruption was not fatal but resulted in a definite growth deficiency relative to WT. Of the remaining genes, 6035 were classified as non-essential (NE) *in vitro*, and a further 185 genes classified as growth advantage (GA), indicating that disruption of these genes improved growth kinetics relative to WT under the same conditions *in vitro*. A total of 116 *Msm* genes do not contain any -TA- sites and so were absent from the library (**Figure 3.5A**); these are listed in **Table S1**.

Tn insertions were not observed in 4.8% of genes (**Figure 3.4**), all of which were distributed throughout the *Msm* genome. These may have been missed by chance, because of sequence-specific insertion rates, or because the mutation was lethal. All genes without insertions were designated candidate essential (ES) genes. Other candidate essential genes included those with Tn insertions restricted to the termini of the ORFs (the 3' ends of some essential genes may be functionally dispensable, while some 5' ends may tolerate insertion if transcription can still proceed through the Tn) and large genes with very few insertions. In bacteria, essential genes are more evolutionarily conserved than nonessential genes (Gerdes et al., 2003). As expected, the inferred essentialities of genes involved in metabolite biosynthesis were affected by the supplemental nutrients in the LA medium, for example, histidine, aspartate, purine nucleotide biosynthesis and pantothenate biosynthesis pathways (*panB* and *panC*) were conditionally essential. However, some metabolic pathways were unaffected by the absence or presence of specific nutrient sources, perhaps indicating redundancy in gene functions. Consistent with their involvement in core cellular functions, the essential genes shared with *Mtb* included genes involved in metabolic, respiratory and cell wall-associated processes (**Figure 3.5B**) (**Table S1**).

A



B

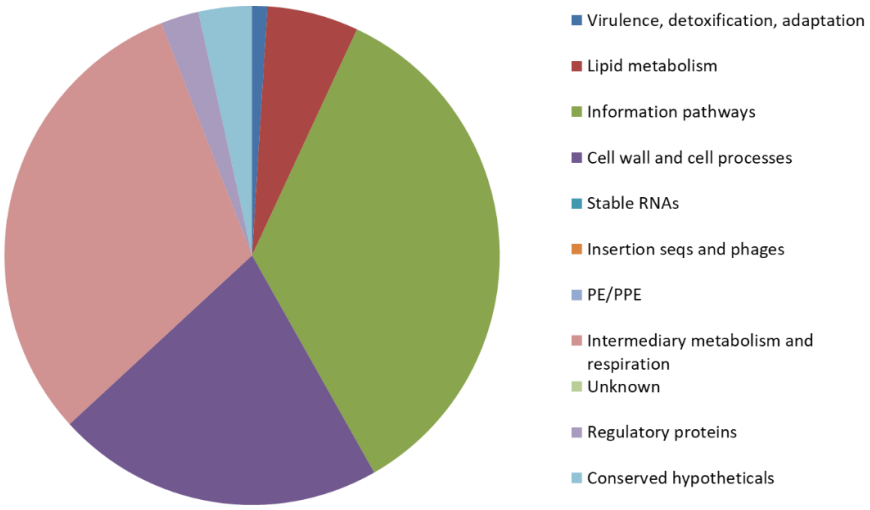


Figure 3.5: Identification of essential genes in *Msm* Δ*metE* in rich medium. (**A**) 6035 genes were identified as non-essential (NE) for growth, 213 as essential (ES), 160 as causing a growth defect (GD) when disrupted, 185 as conferring a growth advantage (GA) and 116 not defined (ND). (**B**) Functional categories (Kapopoulou et al., 2011), of predicted essential genes in *Msm* Δ*metE* during growth in rich medium.

3.6 Identification and characterization of Tn insertion within the predicted genes of the cobalamin biosynthesis pathway (rich medium-selected library)

A total of 157 -TA- sites were disrupted in genes predicted to be involved in *de novo* cobalamin biosynthesis with mean insertion density of 66% (**Table 3.1**), confirming the non-essentiality of the encoded functions in rich medium. The disrupted genes included those predicted to be involved in the biosynthesis of the tetrapyrrole precursors, ALA and Uro' III (*hemD* and *cobA*), genes involved in the aerobic corrin ring synthesis pathway (*cobI*, *cobG*, *cobIJ*, *cobM*, *cobF*, *cobK*, *cobL*, *cobH*, *cobB*, MSMEG_2615, *cobN*, *cobQ1* and *cobQ2*), genes involved in nucleotide loop assembly, synthesis of cobinamide-phosphate, and in the synthesis and attachment of DMB (*cobD*, *cobU*, *cobS*, *cobT*, *bluB* and MSMEG_4305). However, two cobalamin pathway genes, *hemD* and *cobQ2* (**Figure 3.6**), were identified as essential under these conditions, strongly suggesting an additional role(s) in another metabolic pathway(s). The putative *Msm* homologues of the predicted *Mtb* cobalamin biosynthetic genes were obtained from the Mycobrowser portal (Kapopoulou et al., 2011).

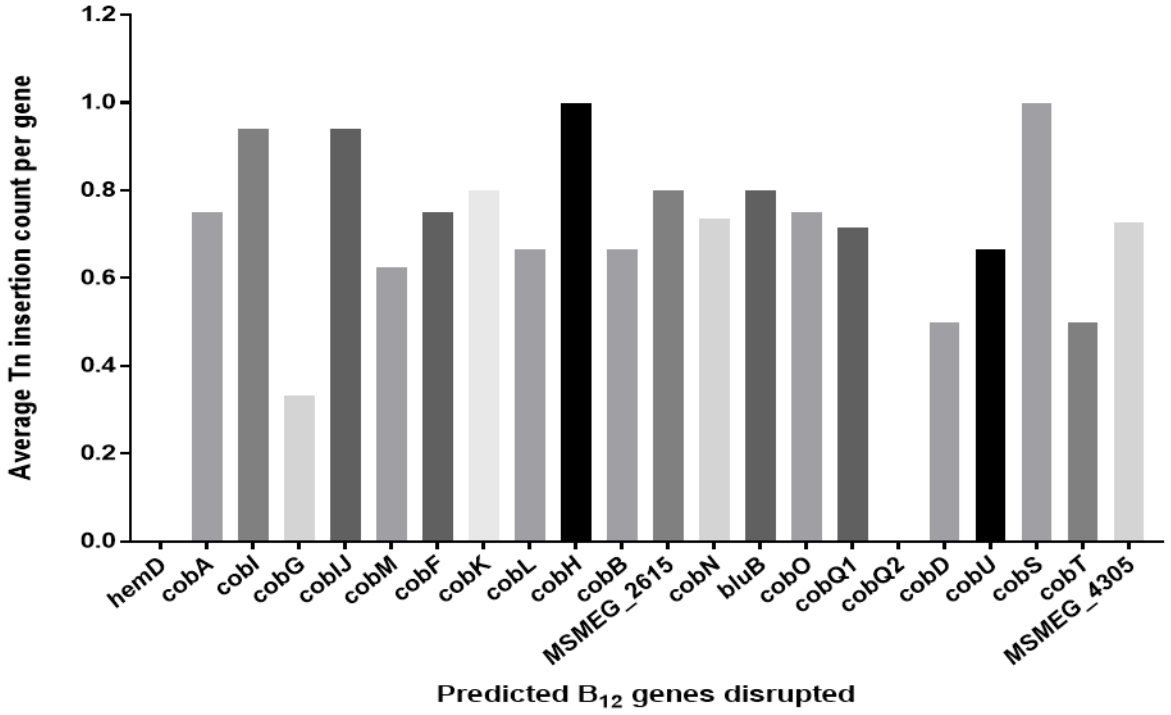


Figure 3.6: Mean insertion densities within selected genes predicted to be involved in cobalamin biosynthesis, transport, and cobalamin-related metabolism. Note that *hemD* and *cobQ2* contained no Tn insertions, suggesting essentiality under the conditions tested.

A further 29, 83, 18, and 120 insertion hits (**Table 3.1**), respectively, were identified in genes predicted to encode putative cobalamin biosynthetic genes: MSMEG_1123, *cbiM* (MSMEG_2607), MSMEG_2607, MSMEG_4934, MSMEG_6048 and *cobW* (MSMEG_6069*)*; putative cobamide transporters, *bacA* (MSMEG_3655), *bacA* (MSMEG_4380), *btuF* (MSMEG_6064), *btuC* (MSMEG_6063), *btuD* (MSMEG_6062), *btuF2* (MSMEG_4561), *btuF2* (MSMEG_4560), *btuC2* (MSMEG_4559) and *btuD2* (MSMEG_4557); putative cobalt transporters *cbiM* (MSMEG_2608), *cbiQ* (MSMEG_2609), MSMEG_2610, MSMEG_4469 and MSMEG_6724; and cobalamin-dependent enzymes *metH* (MSMEG_4185), *hemL* (MSMEG_0969), MSMEG_6318, *eutB* (MSMEG_1553), *eutC* (MSMEG_1554), MSMEG_1547, MSMEG_1546, MSMEG_1548, MSMEG_6320, MSMEG_6321, MSMEG_0496, MSMEG_0497, *mutA* (MSMEG_3158) and *mutB* (MSMEG_3159 (**Table 3.1**).

Table 3.1: Tn insertions in genes predicted to be involved in cobalamin biosynthesis, transport, and cobalamin-related metabolism during growth of Δ*metE* mutant in nutrient-rich medium

Accession no.	Gene name	Predicted product/function	Transposon insertions within the ORF/ TA sites	H37Rv homologue
		Corrin synthetic genes		
MSMEG_0954	*hemD*	uroporphyrinogen-III synthase	0/6 **(ES)**	Rv0511
MSMEG_2618	*cobA*	uroporphyrin-III C methyltransferase	6/8	Rv2847c
MSMEG_3873	*cobI*	cobalamin biosynthesis protein CobIJ	13/17	Rv2066
MSMEG_3871	*cobG*	precorrin-3B synthase	1/3	Rv2064
MSMEG_3873	*cobIJ*	cobalamin biosynthesis protein CobIJ	13/17	Rv2066
MSMEG_3877	*cobM*	precorrin-4 C11-methyltransferase	5/8	Rv2071c
MSMEG_5548	*cobF*	precorrin 6A synthase	6/8	_
MSMEG_3875	*cobK*	cobalt-precorrin-6x reductase	4/5	Rv2070c
MSMEG_3878	*cobL*	precorrin-6Y C5,15-methyltransferase (decarboxylating)	4/6	Rv2072c
MSMEG_3872	*cobH*	precorrin-8X methylmutase	4/4	Rv2065
MSMEG_2617	*cobB*	cobyrinic acid a,c-diamide synthase	6/9	Rv2848c

MSMEG_2615	_	chelatase	9/10	Rv2820c
MSMEG_3864	*cobN*	cobaltochelatase	28/38	Rv2062c
MSMEG_2616	*cobO*	cob(I)yrinic acid a,c-diamide adenosyltransferase	3/4	Rv2849c
MSMEG_2588	*cobQ*	cobyric acid synthase	5/7	Rv0255c
MSMEG_6277	*cobQ2*	cobyric acid synthase	0/6 **(ES)**	Rv3713
		α-ribazole synthetic genes		
MSMEG_6053	*bluB*	5,6-dimethylbenzimidazole synthase	4/5	Rv0306
MSMEG_4275	*cobT*	nicotinate-nucleotide-dimethylbenzimidazole phosphoribosyltransferase	2/4	Rv2207
MSMEG_4305		bifunctional RNase H/acid phosphatase	9/12	Rv2228c
		Adenosylcobalamin synthetic genes		
MSMEG_4310	*cobD*	cobalamin biosynthesis protein	4/8	Rv2236c
MSMEG_4274	*cobU*	cobinamide kinase / cobinamide phosphate guanyltransferase	2/3	Rv0254c
MSMEG_4277	*cobS*	cobalamin synthase	2/2	Rv2208
		Putative cobalamin biosynthesis protein		
MSMEG_1123		cobalamin synthesis protein	7/8	_
MSMEG_2607	*cbiM*	cobalamin biosynthesis protein CbiM	6/7	_
MSMEG_4934	_	ATP:cob(I)alamin adenosyltransferase	6/6	Rv1314c
MSMEG_6048		cobalamin synthesis protein/P47K	4/5	_
MSMEG_6069	*cobW*	CobW/P47K domain-containing protein	6/7	Rv0106
		Putative cobamide transporters		
MSMEG_3655	*bacA*	ABC transporter, permease/ATP-binding protein	17/18	Rv1819c
MSMEG_4380	*bacA2*	ABC transporter, permease/ATP-binding protein	17/20	Rv1819c
MSMEG_6064	*btuF*	lipoprotein	13/14	_
MSMEG_6063	*btuC*	Fe uptake system integral membrane protein	6/6	_
MSMEG_6062	*btuD*	Fe uptake system permease	7/9	_

MSMEG_4561	*btuF2a*	ABC Fe3+-siderophores transporter, periplasmic binding protein	8/9	_
MSMEG_4560	*btuF2b*	periplasmic binding protein	6/9	_
MSMEG_4559	*btuC2*	ABC transporter, membrane spanning protein	5/6	_
MSMEG_4557	*btuD2*	ABC transporter, ATP-binding protein	4/4	_
		Putative cobalt transporters		
MSMEG_2608	*cbiM*	cobalt transport protein CbiM	3/3	_
MSMEG_2609	*cbiQ*	cobalt ABC transporter, permease protein CbiQ	7/8	_
MSMEG_2610	_	cobalt transport protein ATP-binding subunit	2/2	_
MSMEG_4469		cobalt transport protein	3/4	Rv2325c
MSMEG_6724	_	ABC-type cobalt transport system	3/5	_
		Cobalamin-dependent enzymes		
MSMEG_4185c	*metH*	B_{12}-dependent methionine synthase	34/38	Rv2124c
MSMEG_0969	*hemL*	glutamate-1-semialdehyde aminotransferase	0/10 **(ES)**	Rv0524
MSMEG_1553	*eutB*	ethanolamine ammonia-lyase, large subunit	11/12	_
MSMEG_1554	*eutC*	ethanolamine ammonia-lyase small subunit	3/4	_
MSMEG_1547	_	glycerol dehydratase large subunit	8/11	_
MSMEG_1546	_	coenzyme B_{12}-dependent glycerol dehydrogenase small subunit	4/4	_
MSMEG_1548	_	propanediol utilization: dehydratase, medium subunit	1/3	_
MSMEG_6320	_	diol dehydrase gamma subunit	3/4	_
MSMEG_6321	_	glycerol dehydratase large subunit	11/17	_
MSMEG_0496	_	coenzyme B_{12}-dependent glycerol dehydrogenase small subunit	2/2	_
MSMEG_0497	_	glycerol dehydratase large subunit	12/19	_
MSMEG_3158	*mutA*	methylmalonyl-CoA mutase, small subunit MutA	5/11	Rv1492
MSMEG_3159	*mutB*	methylmalonyl-CoA mutase, Large subunit MutB	21/27	Rv1493

3.7 Comparison of WT *Msm* and Δ*metE* essential genes

Gene essentiality is conditional (Carey et al., 2018; Rancati et al., 2018; Rousset et al., 2021) and can be affected by genetic background and nutrient availability; therefore, it was necessary to compare predicted gene essentialities in WT *Msm* and the Δ*metE* mutant. To this end, WT *Msm* Tn libraries were selected on 7H10-OADC plates, collecting at least 240,000 colonies to ensure a saturated library. An average of 3,716,188 unique template counts (Tn insertion counts per -TA- site) was identified, covering 76.4% of the -TA- sites in the genome, with an average count of 26.2 per -TA- site. Tn insertions were distributed throughout the genome (**Figure S1**). In total, 278 *Msm* genes were classified as essential (ES) for growth, 93 as growth defective (GD), 311 growth advantage (GA), and 5931 as non-essential (NE). 83 genes were undefined (N/A), due to insufficient number of -TA- sites to robustly call their essentiality status. (**Figure 3.7A**). Genes involved in cobalamin biosynthesis and assimilation were not under-represented in this screen, which was expected given that WT *Msm* possesses a cobalamin-independent methionine synthase, MetE, which enables the production of methionine in the absence of coenzyme B_{12} (**Table 3.2**). Again, both *hemD* and *cobQ2* were predicted as essential under these conditions. In addition, another gene, *cobA*, which is involved in the biosynthesis of the tetrapyrrole precursors ALA and Uro' III, was identified as growth defective, indicating an additional role(s) in other metabolic pathway(s) (**Table 3.2**).

Table 3.2: Enzymes predicted to be involved in cobalamin biosynthesis in *Msm* cobalamin biosynthesis genes (7H10-selected library)

			Essentiality call
Accession no.	**Gene name**	**Predicted product/function**	
MSMEG_0954	*hemD*	uroporphyrinogen-III synthase	ES
MSMEG_2618	*cobA*	uroporphyrin-III C-methyltransferase	GD
MSMEG_3873	*cobI*	cobalamin biosynthesis protein cobIJ	NE
MSMEG_3871	*cobG*	precorrin-3B synthase	NE
MSMEG_3873	*cobIJ*	cobalamin biosynthesis protein cobIJ	NE
MSMEG_3877	*cobM*	precorrin-4 C11-methyltransferase	NE
MSMEG_5548	*cobF*	precorrin 6A synthase	NE
MSMEG_3875	*cobK*	cobalt-precorrin-6x reductase	NE

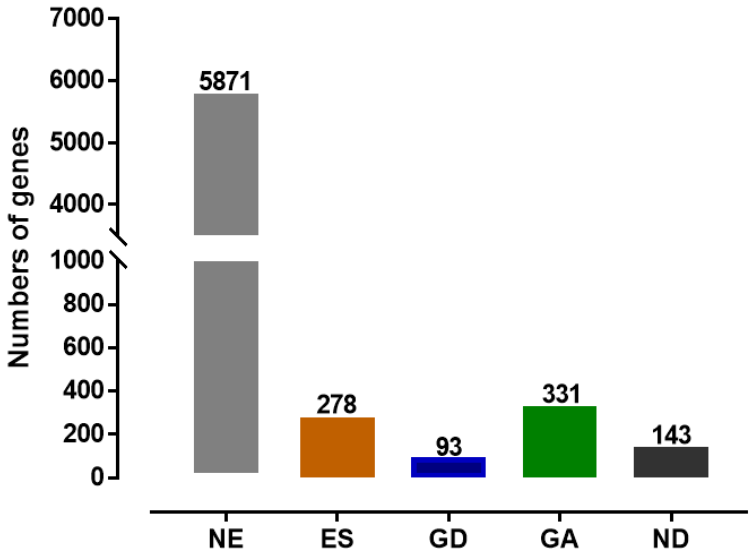
MSMEG_3878	*cobL*	precorrin-6Y C5,15-methyltransferase (decarboxylating)	NE
MSMEG_3872	*cobH*	precorrin-8X methylmutase	NE
MSMEG_2617	*cobB*	cobyrinic acid a,c-diamide synthase	NE
MSMEG_2615	-	Chelatase	NE
MSMEG_3864	*cobN*	Cobaltochelatase	NE
MSMEG_6053	*bluB*	5,6-dimethylbenzimidazole synthase	NE
MSMEG_2616	*cobO*	cob(I)yrinic acid a,c-diamide adenosyltransferase	NE
MSMEG_2588	*cobQ*	cobyric acid synthase	NE
MSMEG_6277	*cobQ2*	cobyric acid synthase	ES
MSMEG_4310	*cobD*	cobalamin biosynthesis protein	NE
MSMEG_4274	*cobU*	cobinamide kinase / cobinamide phosphate guanyltransferase	NE
MSMEG_4277	*cobS*	cobalamin synthase	NE
MSMEG_4275	*cobT*	nicotinate-nucleotide--dimethylbenzimidazole phosphoribosyltransferase	NE
MSMEG_4305	-	bifunctional RNase H/acid phosphatase	NE

To identify genes with statistically significant differences in Tn insertion count between WT and Δ*metE*, the sequencing data for the WT and Δ*metE* libraries were compared using a permutation test-based method incorporated into the TRANSIT platform (DeJesus et al., 2015). Using this so-called "Resampling" method, 23 genes were identified as differentially essential in the WT (standard media) *versus* Δ*metE* (rich media) datasets (**Figure 3.7B**). Of the 23 genes, only 7 were essential in the *metE* background but not in the parental WT. The genes identified as required for growth in Δ*metE* were MSMEG_0381, MSMEG_0382, *asnB* (MSMEG_4269), MSMEG_4716, *glgB* (MSMEG_4918), MSMEG_0389. These genes returned read counts of 2421.34, 2040.69, 37.01, 15.86, 15.86, and 0.00, respectively, *versus* 16979.6, 14800.7, 5559.9, 1876.3, 6239.3, and 445.9 in the WT *Msm* (7H10-selected library) (**Table 3.2**). The *metE* gene itself was mostly devoid of Tn insertions in Δ*metE* in this analysis. This result was expected given the absence of *metE* sequence (and, consequently, -TA- sites, in the targeted deletion mutant) but nevertheless served as a useful control for this analysis. Similarly, *metH*, encoding the cobalamin-dependent methionine synthase, was identified as essential in WT only (**Table 3.3**). This inferred conditional essentiality of *metH* was consistent with recent work

demonstrating the role of cobalamin riboswitch-mediated repression of *metE* transcription in *Msm* (Kipkorir et al., 2021), and again provided an important internal control for the resampling analysis. Of the 23 differentially essential genes, only 16 were under-represented in the WT library (7H10-selected) when compared to Δ*metE* (rich-selected library). The genes identified as necessary for growth in the WT background (7H10-selected) included MSMEG_0935, MSMEG_1937, *ilvB* (MSMEG_2372), *carB* (MSMEG_3047), *ilvE* (MSMEG_4276), MSMEG_4527, MSMEG_4646, *glyA* (MSMEG_5249), *purF* (MSMEG_5800), *purL* (MSMEG_5824), *leuA* (MSMEG_6271), MSMEG_0949, *metX* (MSMEG_1651), *argG* (MSMEG_3770) and *purH* (MSMEG_5515), all of which returned read counts that were significantly reduced (**Table 3.4**).

The lists of essential genes identified in this study (two libraries: WT and Δ*metE*) were also compared with the essential genes reported recently by Dragset et al. (2019) (**Figure 3.7C**). A total of 237 genes were identified as essential in all three libraries (**Figure 3.7C**) (**Table S1**), while 19 genes (classified as ES or GD) were essential only in the Δ*metE* mutant (**Figure 3.7C**) (**Table S2**).

A



B

***Msm* library**	**Number of conditionally essential genes**
Δ*metE*, rich LA medium	7
WT, standard 7H10 medium	16

C

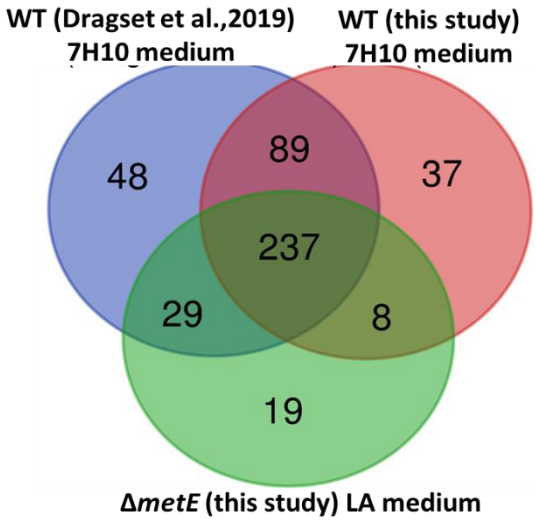


Figure 3.7: Identification of conditionally essential genes in WT *Msm* and Δ*metE* mutant Tn library. (**A**) Definition of the WT (7H10-selected) genetic requirements for growth: 5931 genes were identified as non-essential (NE), 311 as growth advantage when disrupted (GA), 278 as essential (ES), and 93 as causing growth defect when disrupted (GD). 83 genes were not defined (ND). (**B**) Results for comparative analysis between WT and Δ*metE* Tn libraries (adjusted P value of <0.05)**.** (**C**) Venn diagram comparisons of essential genes (ES + GD) inferred by TnSeq in this study and one previous study (Dragset et al., 2019).

Table 3.3: List of genes under-represented in Δ*metE* (rich-selected) when compared to WT (7H10-selected)

Accession no.	Gene name	Predicted product/function	Mtb ortholog	Adj. p-value
MSMEG_0381	*mmpl4a*	putative transport protein, MmpL4a	-	0
MSMEG_0382	*mmpL4b*	putative transport protein, MmpL4b	Rv0450c	0
MSMEG_4269	*asnB*	asparagine synthase (glutamine-hydrolyzing)	Rv2201	0
MSMEG_4916	*glgE*	alpha-amylase family protein	Rv1327c	0
MSMEG_4918	*glgB*	glycogen branching enzyme	Rv1326c	0
MSMEG_6638	*metE*	5-methyltetrahydropteroyltriglutamate--homocysteine methyltransferase	Rv1133c	0
MSMEG_0389	*gtf1*	Putative glycosyl transferase	-	0,0292

Table 3.4: List of genes under-represented in WT library (7H10-selected) when compared to Δ*metE* (rich-selected library)

Accession no.	Gene name	Predicted product/function	Mtb ortholog	Adj. p-value
MSMEG_0935	*gpm1*	2,3-bisphosphoglycerate-dependent phosphoglycerate mutase	Rv0489	0
MSMEG_1937	*moeB1*	molybdopterin biosynthesis-like protein MoeZ	Rv3206c	0
MSMEG_2372	*ilvB*	acetolactate synthase 1 catalytic subunit	Rv3003c	0
MSMEG_3047	*carB*	carbamoyl phosphate synthase large subunit	Rv1384	0
MSMEG_4185	*metH*	B12-dependent methionine synthase	Rv2124c	0
MSMEG_4276	*ilvE*	branched-chain amino acid aminotransferase	Rv2210c	0
MSMEG_4527	*sirA*	ferredoxin sulfite reductase	Rv2391	0
MSMEG_4646	-	pyruvate synthase	Rv2455c	0
MSMEG_5249	*glyA*	serine hydroxymethyltransferase		0
MSMEG_5800	*purF*	amidophosphoribosyltransferase	Rv0808	0
MSMEG_5824	*purL*	phosphoribosylformylglycinamidine synthase II	Rv0803	0
MSMEG_6271	*leuA*	2-isopropylmalate synthase	Rv3710	0
MSMEG_0949	*serB1*	HAD-superfamily protein subfamily protein IB hydrolase	Rv0505c	0,0292
MSMEG_1651	*metX*	homoserine O-acetyltransferase	Rv3341	0,0292
MSMEG_3770	*argG*	argininosuccinate synthase	Rv1658	0,0292
MSMEG_5515	*purH*	bifunctional phosphoribosylaminoimidazolecarboxamide formyltransferase/IMP cyclohydrolase	Rv0957	0,0292

3.8 TnSeq analysis of gene essentiality in Δ*metE* library

3.8.1 Identification of conditional essential genes in *Msm* Δ*metE*

To identify genes required for the biosynthesis and/or assimilation of cobalamin during growth in defined media *in vitro,* the Δ*metE* Tn library was selected on Sauton's medium with or without CNCbl and cobalt supplementation. For selection, approximately 10^6 cells were used per condition to ensure an adequately complex library was analysed consisting of all possible mutants with unique insertions (Majumdar et al., 2017). Following selection of the *metE* Tn libraries on supplemented and unsupplemented minimal media, genomic DNA was extracted, amplified by PCR, and subjected to high-throughput sequencing to quantify all Tn junctions. Sequencing these libraries yielded an average of 1.5 million unique Tn-chromosome junctions, which could be mapped to 66.5 - 76.7% of the -TA- dinucleotide sites in the chromosome in each individual experiment (**Table 3.5**), representing a significant number of clones in the mutant library. Further analysis of these data using HMM (DeJesus et al., 2015) revealed differences in genetic requirements for growth – including those likely required for *de novo* cobalamin biosynthesis.

Table 3.5: Tn library saturation for selected library replicates

Medium	Supplements	Total reads	Mapped Templates	NZ mean	#TA sites hit/77755	Coverage %
Sautons, Tw, Kan	Sautons_rep1	7,969,333	2,441,545	38.4	57,730	74.2
	Sautons_rep2	9,903,885	6,088,171	83.3	59,102	76.0
	Sautons_rep3	10,835,912	4,564,328	52.4	58,885	75.7
	+ CNCbl _rep1	7,950,793	1,865,657	15.5	51,670	66.5
	+ CNCbl _rep2	10,472,153	5,516,774	67.5	59,230	76.2
	+ CNCbl _rep3	9,614,314	5,161,043	50.0	58,548	75.3
	+Cobalt_rep1	7,881,616	2,178,145	32.1	56,981	73.3
	+Cobalt_rep2	8,942,085	5,183,288	77.7	58.880	75.7
	+Cobalt_rep3	12,190,924	7,202,275	103.3	59,218	76.2

Analysis of these data revealed that a total of 6.05% of the -TA- sites in genome were labelled as essential (ES) in defined minimal medium, whereas 5.08% and 5.82% of genes were identified as essential in CNCbl-supplemented and cobalt-supplemented Sauton's minimal media, respectively. Although not definitive, this result was encouraging since it suggested that the number of essential genes declined as a consequence of supplementation with either CNCbl or cobalt.

Again, most -TA- sites were labelled non-essential (NE) in all tested conditions (87% in unsupplemented, 88% and 82% in CNCbl-supplemented and cobalt-supplemented, respectively), with a small percentage of genes classified as growth-defect (GD) or growth-advantage (GA): 1.64% and 4.94% in unsupplemented, 2.44% and 4.26% in CNCbl-supplemented, and 2.08% and 5.13% in cobalt-supplemented Sauton's minimal medium, respectively. Essential -TA- sites averaged a very small number of insertions and read counts (0.056 and 7.90 counts in unsupplemented, 4.8 and 9.74 counts in CNCbl-supplemented, and 4.8 and 9.74 counts in cobalt-supplemented media, respectively), indicating that the HMM associated the essential state with stretches devoid of insertions, even though these locations could occasionally contain insertions with a very small number of reads. In contrast, non-essential regions exhibited a mean insertion density of ~82% and mean read counts of ~143.50 in all conditions in this dataset. Growth-defect regions had some insertions, but these were significantly fewer (~21-28% density and a ~10- fold reduction in mean read counts). Insertion densities in growth-advantage regions were almost saturated (93%) and mean read counts on average were 386, 23 counts across all conditions tested (**Table 3.6**).

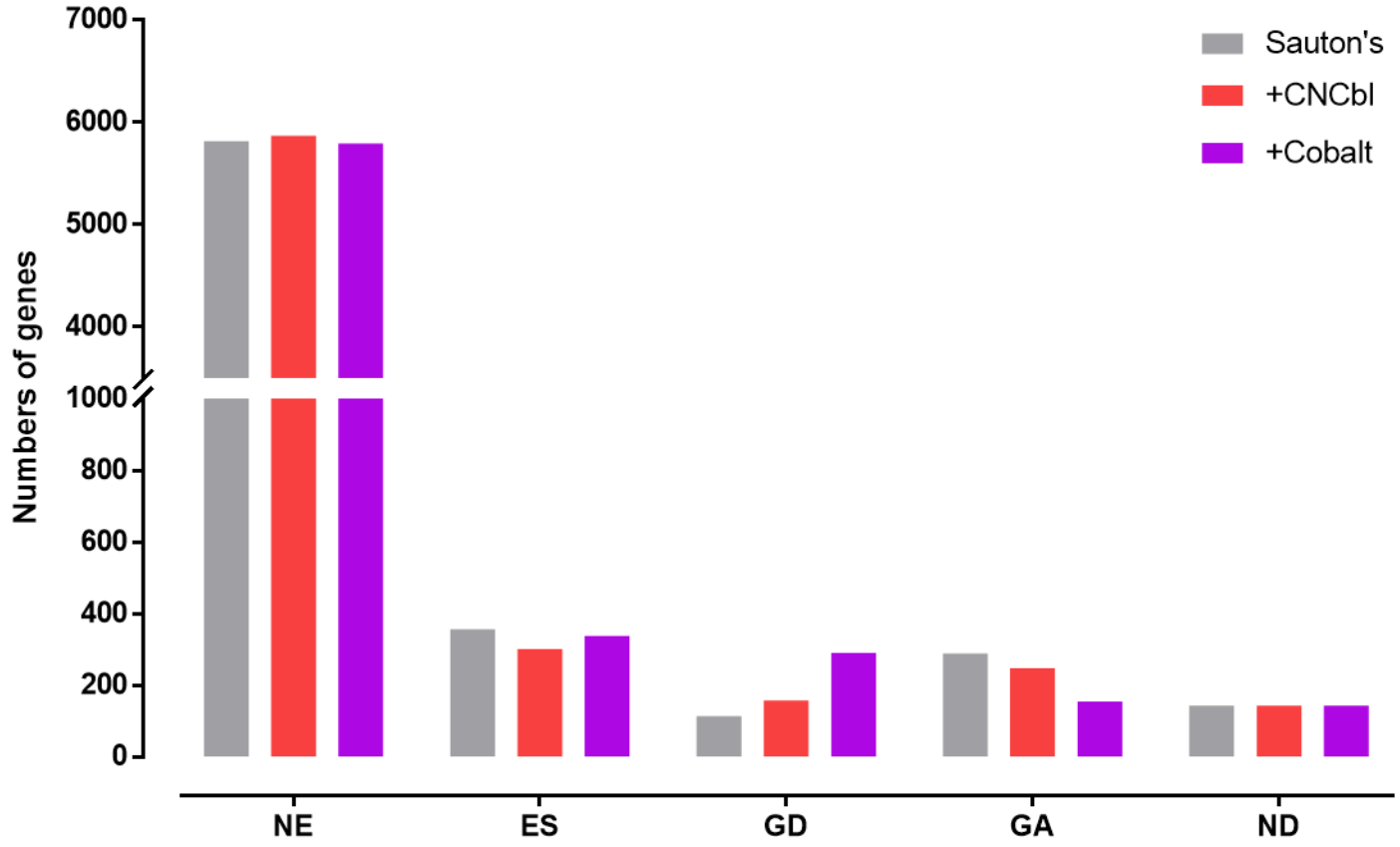
Of the predicted 6,716 genes in the *Msm* genome, 356 genes were identified as essential for growth on unsupplemented minimal medium, while 301 and 337 genes were identified as essential on CNCbl-supplemented and cobalt-supplemented minimal medium, respectively (**Figure 3.8A**). A total of 424 genes were identified as essential across all conditions tested (**Figure 3.8B**) with only 10, 13 and 24 genes (ES plus GD) uniquely required for growth in unsupplemented, CNCbl-supplemented and cobalt supplemented Sauton's minimal medium, respectively (**Figure 3.8B**) (**Table S3-8**).

Table 3.6: Distributions of essentiality calls in Sauton's medium

	State Distribution (%)	Mean insertion density	Mean read counts
	Sauton's only		
Essential	6.05	0.056	7.90
Growth defect	1.64	0.28	6.16
Non-essential	87.37	0.82	143.50
Growth advantaged	4.94	0.93	376.84
	+ CNCbl		
Essential	5.08	0.049	9.98
Growth defect	2.44	0.25	9.54
Non-essential	88.22	0.82	144.31
Growth advantaged	4.26	0.93	382.33
	+Cobalt		
Essential	5.82	0.048	9.74
Growth defect	2.08	0.21	7.30
Non-essential	86.97	0.82	144.31
Growth advantaged	5.13	0.93	399.52

Distribution of essentiality calls for the Sauton's only, CNCbl-supplemented, and cobalt-supplemented datasets obtained by the HMM method. Essential states represent those regions which are mostly devoid of insertions. Non-Essential regions contained read-counts close to the mean read-count in the dataset. Growth Defect regions and Growth Advantage regions represented, respectively, those regions which had significantly suppressed or increased read-counts.

A


Numbers of genes
7000
6000
5000
4000
1000
800
600
400
200
0
Sauton's
+CNCbl
+Cobalt
NE
ES
GD
GA
ND

B

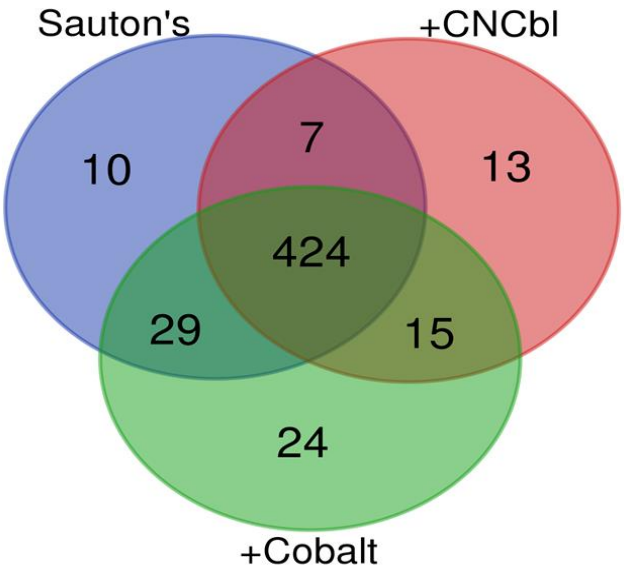


Figure 3.8: Essentiality analysis of the Δ*metE* Tn-library grown in Sauton's minimal medium with and without indicated supplements. (**A**) Genetic requirements for growth of the Δ*metE* Tn library selected under the different conditions. 5814, 5866 and 5790 genes were identified as non-essential (NE) for growth; 356, 301 and 337 as essential (ES); 114, 158 and 291 as causing growth defect (GD) when disrupted; and 289, 248 and 155 as causing growth advantage (GA) when disrupted. A total of 143 genes were not defined on Sauton's, CNCbl-supplemented and cobalt-supplemented minimal medium. (**B**) Venn diagram comparison of essential genes (ES + GD) by TnSeq.

3.9 The essentialities of cobamide biosynthetic genes in Δ*metE* (Sautons-selected library)

The identification of essential genes via TnSeq involves statistical analyses of Tn insertion location and frequency across a coding region, ultimately making a qualitative classification (essential vs non-essential) from quantitative insertion count data. In principle, the Tn insertions read counts which are observed at each -TA- site correlate with the number of replicates in which insertions are detected, meaning that the abundance of a mutant is correlated with the probability that it would be detected.

Differences in Tn mutant abundance reflect differences in the fitness cost of gene disruption in a particular genetic background or condition and indicate the relative dependency of a strain on the gene for growth and survival. In TnSeq, fitness is determined by averaging over multiple independent insertions within the same gene (Burby et al., 2017; Van Opijnen et al., 2009; van Opijnen & Levin, 2020).

Using the Resampling method (DeJesus et al., 2015), it was observed that Tn insertions in genes involved in the synthesis of the cobalamin were underrepresented in unsupplemented Sauton's minimal medium. *De novo* cobalamin biosynthesis involves approximately 30 steps (Warren et al., 2002), and the pathway can be divided into three stages. The first stage, tetrapyrrole precursor biosynthesis, contains the first five steps of the pathway, most of which are also common to the biosynthesis of heme, chlorophyll, and other tetrapyrroles. It was observed that genes involved in the biosynthesis of the tetrapyrrole precursors ALA and Uro' III (*hemA, hemL, hemB, hemD* and *cobA*) were underrepresented in all three conditions: unsupplemented Sauton's minimal medium, CNCbl-supplemented and cobalt supplemented medium (**Figure 3.9**).

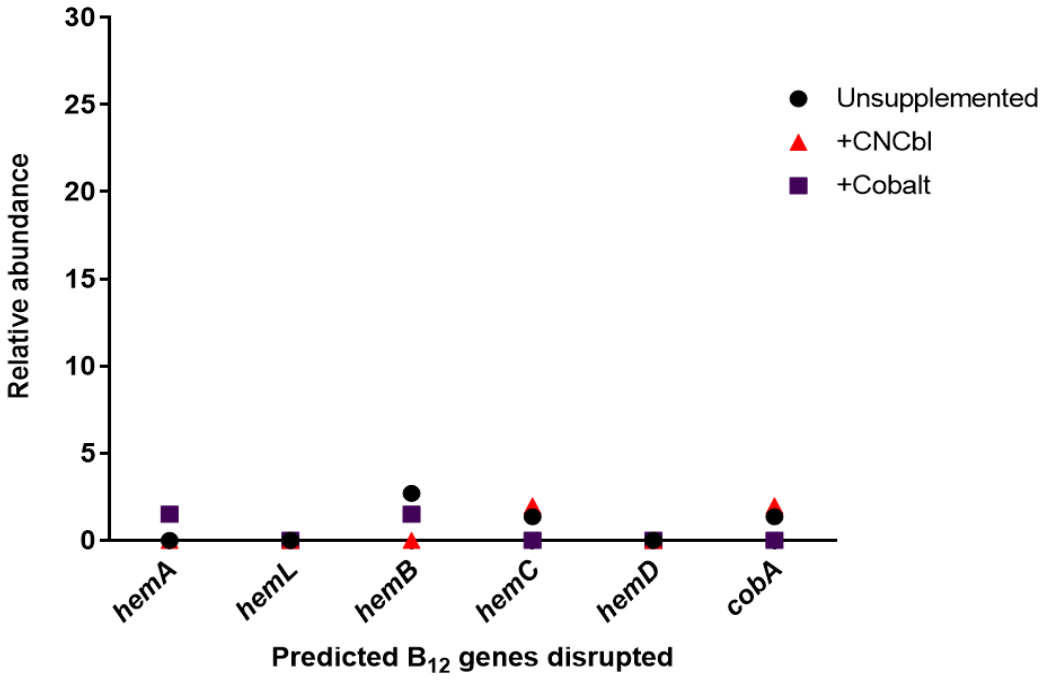


Figure 3.9: Fitness analysis of genes involved in biosynthesis of the tetrapyrrole precursors, ALA and Uro' III in *Msm*. All the genes involved in the biosynthesis of the tetrapyrrole precursors, ALA and Uro' III, were underrepresented in unsupplemented Sauton's minimal medium and both CNCbl and cobalt supplementation failed to rescue these mutants. (Relative abundance = Mean insertion densities)

The second stage in cobalamin biosynthesis consists of genes involved in corrin ring synthesis, which can be divided into anaerobic (oxygen-sensitive) and aerobic (oxygen-dependent) routes, depending on the organism. These two different pathways then converge at a late intermediate, which is further modified to form the cobamide. However, there were no hits for genes encoding the anaerobic corrin ring biosynthetic pathway, although *cobB, cobH, cobK and cobM* were found to be hypothetical orthologous replacements for *cbiA, cbiC, cibJ and cibF*, respectively, in the anaerobic pathway. Genes involved in aerobic corrin ring synthesis were underrepresented in unsupplemented Sauton's minimal medium as compared to the CNCbl-supplemented medium (**Figure 3.10**), with most genes identified as essential for growth in unsupplemented medium by Transit analysis and non-essential in the CNCbl-supplemented medium (**Table 3.7**). For example, genes involved in aerobic corrin ring synthesis pathway that were underrepresented in unsupplemented conditions included *cobI, cobG, cobIJ, cobM, cobF, cobK, cobL, cobH, cobB,* MSMEG_2615, *cobN* and *cobQ1* (**Figure 3.10**).

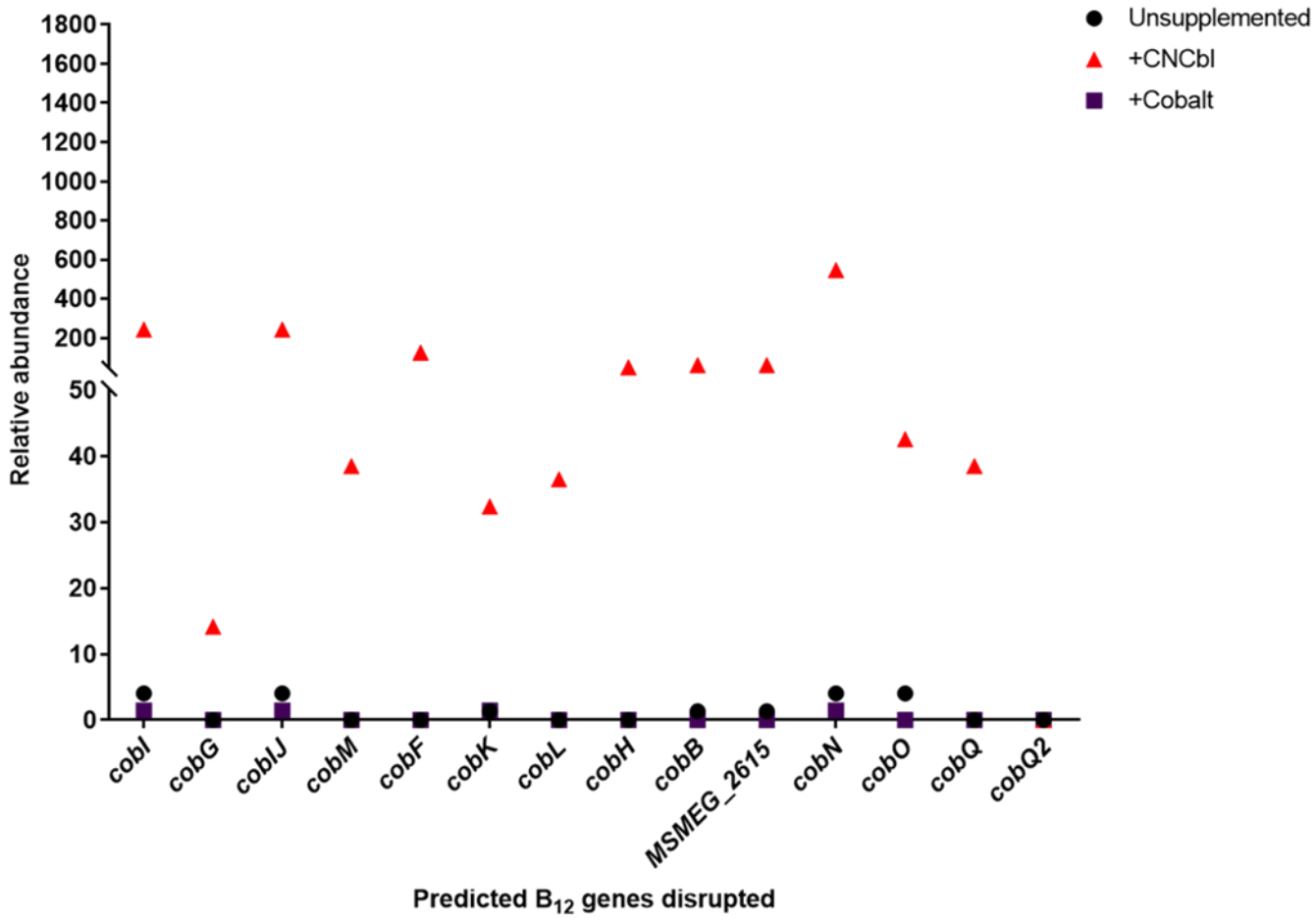


Figure 3.10: Fitness analysis of genes involved in aerobic corrin ring synthesis in *Msm*. In unsupplemented media, the cobalamin autotroph mutant were significantly suppressed in this condition. Providing vitamin B_{12} (CNCbl) in the growth medium significantly enhanced Tn mutant fitness.

The predicted CobO adenosyltranserase involved in formation of adenosylcobyrinate a,c-diamide was identified as GD in unsupplemented medium. Apart from participating in the *de novo* cobalamin biosynthesis pathway, CobO is also predicted to be involved in the salvage of cobamides in *Mtb* (Gopinath, Venclovas, et al., 2013). The last stage of the pathway involves the synthesis of cobinamide-phosphate and DMB. The genes involved in nucleotide loop assembly, synthesis of cobinamide-phosphate, and in the synthesis and attachment of DMB included *cobD, cobU, cobS,* and *cobT,* all of which were classified as essential (**Figure 3.11**). Three genes – *hemD, cobA* and *cobQ2* – were classified as essential for growth across all condition's, reinforcing their inferred requirement for other metabolic pathways.

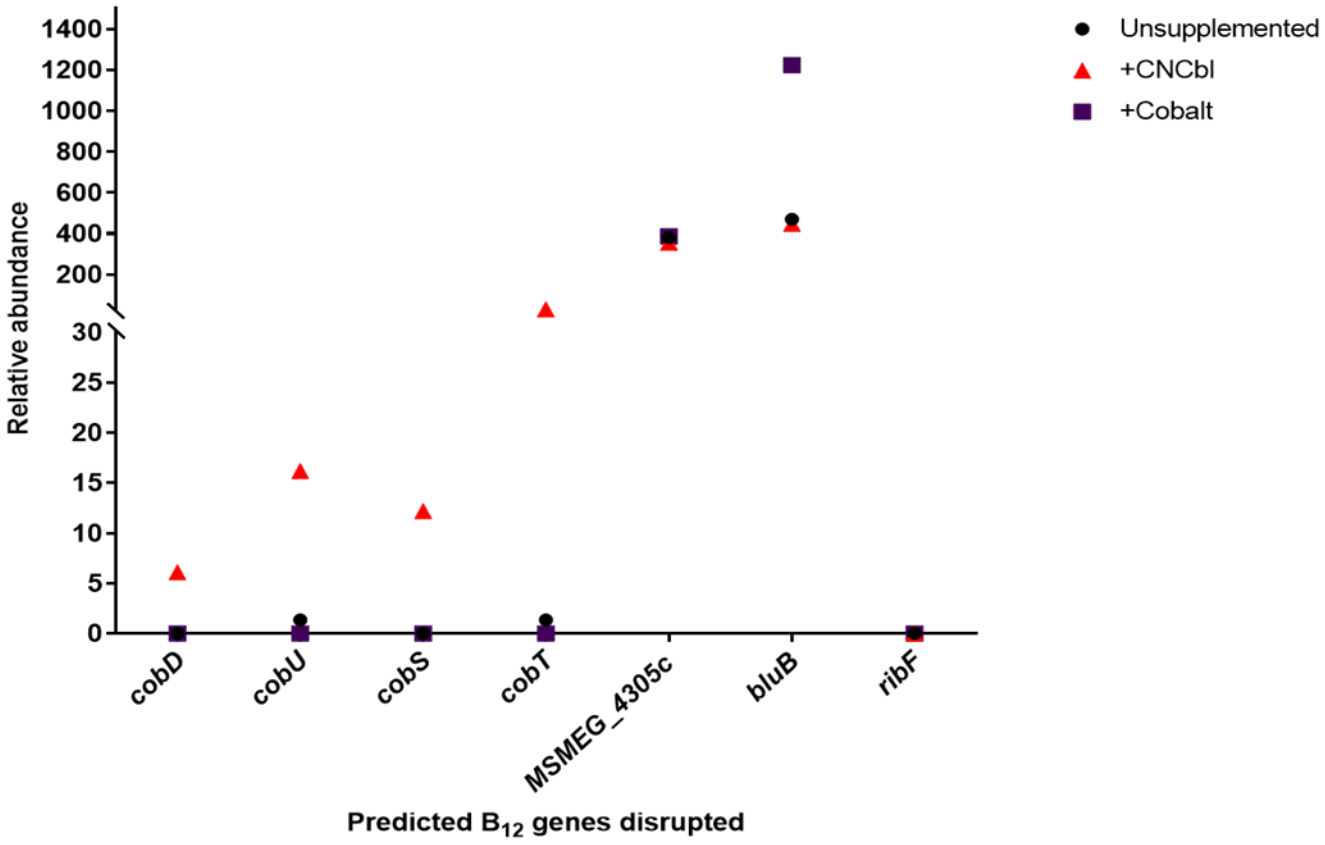


Figure 3.11: Fitness analysis of genes involved in *de novo* cobinamide-phosphate and DMB synthesis in *Msm*. Genes involved in *de novo* cobinamide-phosphate and DMB synthesis appeared enriched in CNCbl-supplemented medium as compared to unsupplemented and cobalt-supplemented media; the exceptions were *bluB* and *MSMEG_4305c* genes which were overrepresented in all conditions.

It was noteworthy that the predicted BluB (MSMEG_6053) which catalyses the formation of DMB and cob(II)yrinate a,c diamide reductase (Gopinath, Moosa, et al., 2013) was identified as non-essential for growth in both unsupplemented and CNCbl-supplemented Sauton's minimal medium (**Table 3.7**). In addition, MSMEG_4305, predicted to encode the α-ribazole phosphatase, was also classified as non-essential in both conditions. Both genes had similar read counts, potentially suggesting the presence of an alternative protein(s) that can compensate for the loss of these genes. In contrast to *Msm*, Tn disruption of Rv2228c was found to cause a growth defect in *Mtb* H37Rv *in vitro* (DeJesus et al., 2017; Griffin et al., 2011), most likely because it additionally provides critical RNase H function (Watkins & Baker, 2010).

Table 3.7: Predicted cobalamin genes identified in Tn mutagenesis screen of *Msm* Δ*metE*

			Essentiality call	
Accession no.	**Gene name**	**Annotated/predictedfunction**	**Unsupplemented**	**+CNCbl**
MSMEG_0954	*hemD*	uroporphyrinogen-III synthase	ES	ES
MSMEG_2618	*cobA*	uroporphyrin-III C-methyltransferase	ES	ES
MSMEG_3873	*cobI*	cobalamin biosynthesis protein CobIJ	ES	NE
MSMEG_3871	*cobG*	precorrin-3B synthase	ES	NE
MSMEG_3873	*cobIJ*	cobalamin biosynthesis protein CobIJ	ES	NE
MSMEG_3877	*cobM*	precorrin-4 C11-methyltransferase	ES	NE
MSMEG_5548	*cobF*	precorrin 6A synthase	ES	NE
MSMEG_3875	*cobK*	cobalt-precorrin-6x reductase	ES	NE
MSMEG_3878	*cobL*	precorrin-6Y C5,15-methyltransferase (decarboxylating)	ES	NE
MSMEG_3872	*cobH*	precorrin-8X methylmutase	ES	NE
MSMEG_2617	*cobB*	cobyrinic acid a,c-diamide synthase	ES	NE
MSMEG_2615	-	Chelatase	ES	NE
MSMEG_3864	*cobN*	Cobaltochelatase	ES	NE
MSMEG_6053	*bluB*	5,6-dimethylbenzimidazole synthase	NE	NE
MSMEG_2616	*cobO*	cob(I)yrinic acid a,c-diamide adenosyltransferase	GD	NE
MSMEG_2588	*cobQ*	cobyric acid synthase	ES	NE
MSMEG_6277	*cobQ2*	cobyric acid synthase	ES	ES
MSMEG_4310	*cobD*	cobalamin biosynthesis protein	ES	GD
MSMEG_4274	*cobU*	cobinamide kinase / cobinamide phosphate guanyltransferase	ES	NE
MSMEG_4277	*cobS*	cobalamin synthase	ES	NE
MSMEG_4275	*cobT*	nicotinate-nucleotide--dimethylbenzimidazole phosphoribosyltransferase	ES	NE
MSMEG_4305	-	bifunctional RNase H/acid phosphatase	NE	NE

3.10 Cobalt supplementation fails to rescue growth of *Msm* cobalamin Tn mutants

Previous studies in the MMRU have demonstrated that *Mtb ΔmetE*::*hyg* - a strict cobalamin auxotroph - can grow aerobically in liquid medium supplemented with cobalt (Moosa, 2013). This observation suggests that *Mtb* has retained the ability to synthesize cobalamin provided that cobalt is not limiting. In contrast, cobalt supplementation did not restore growth in any of the cobalamin Tn mutants: all genes involved in synthesis of aerobic corrin ring, nucleotide loop assembly, synthesis of cobinamide-phosphate, and the synthesis and attachment of DMB, were underrepresented in cobalt-containing medium as compared to the CNCbl-supplemented medium. These genes were classified as essential for growth, hence suggesting a fitness cost to the mutants. This suggests that the growth defect of the cobalamin auxotrophs was cobalamin specific and not a consequence of defective cobalt supply (**Figures 3.10 & 3.11**). In turn, this reinforces the conclusion that the results provide direct evidence of the functionality and involvement of these genes in cobalamin biosynthesis in *Msm*.

3.11 Inhibition of *metE* expression by endogenous cobalamin in *Msm*

Like many organisms, *Msm* encodes both cobalamin-dependent (MetH) and cobalamin-independent (MetE) methionine synthases, a feature that is conserved across diverse mycobacterial species. In *Msm* and *Mtb*, expression of *metE* is regulated by a cobalamin-responsive riboswitch located upstream of the gene (Kipkorir et al., 2021; Warner et al., 2007). By comparing the mean read counts between *metH* and *metE*, it was determined that only *metH* was essential in WT (7H10-selected library) whereas both genes were underrepresented in Δ*metE* (Sauton's minimal medium-selected) (**Figure 3.12**) (**Table 3.4**). This result was expected in Δ*metE* given the absence of, -TA- sites, in the targeted deletion mutant but nevertheless served as an important internal control for this analysis. The observed conditional essentialities were consistent with previous studies reporting the functionalities of the *metE* riboswitches in *Mtb* (Minato et al., 2019; Warner et al., 2007) and *Msm* (Kipkorir et al., 2021).

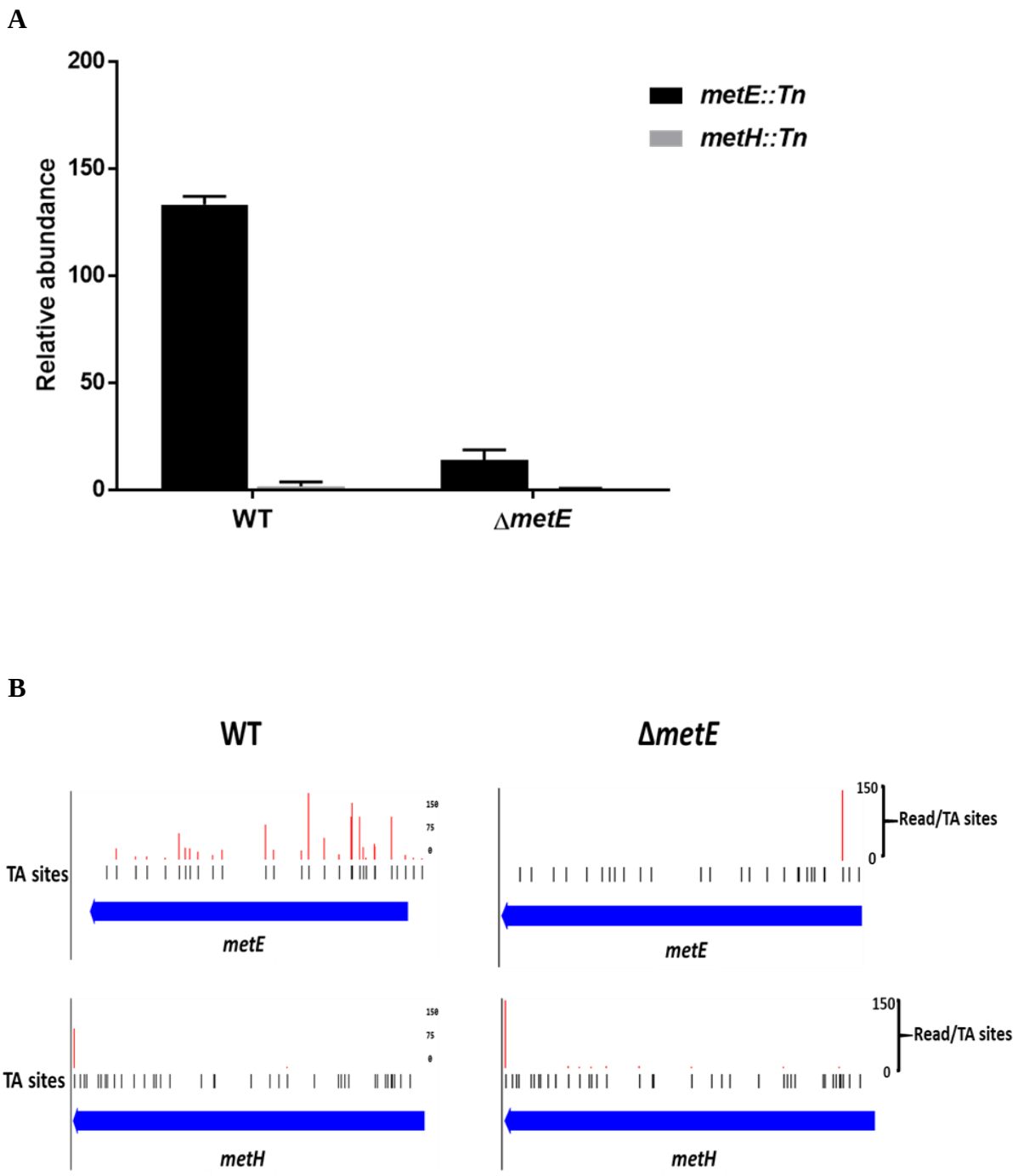


Figure 3.12: Tn insertion reads in *Msm* cobalamin-dependent (*metH*) and cobalamin-independent (*metE*) methionine synthases. (**A-B**) Quantification of Tn reads between cobalamin-dependent (*metH*) and cobalamin-independent (*metE*) methionine synthases in the WT (7H10-selected) and Δ*metE* (Sauton's-selected) Tn libraries.

3.12 Identification and characterisation of putative cobalamin biosynthesis and cobalamin-dependent enzymes in *Msm* (Sauton's minimal medium-selected)

3.12.1 Putative cobalamin biosynthetic genes in *Msm*

In contrast to *Mtb*, *Msm* appears to encode all enzymes required for *de novo* cobalamin biosynthesis and the predicted pathway is also suggestive of the aerobic type. Additionally, *Msm* encodes five other putative cobalamin biosynthesis proteins – MSMEG_1123, CbiM (MSMEG_2607), MSMEG_6048, CobW (MSMEG_6069), and MSMEG_4934 – for which enzymatic functions have not been defined. Using the Resampling method, it was observed that Tn insertion counts within these genes were the same in both experimental conditions; moreover, the first four genes were all classified as non-essential for growth (**Table 3.8**), while insertions in MSMEG_4934 conferred a growth advantage in unsupplemented and CNCbl-supplemented medium both conditions (**Figure 3.13**). This result shows that, the first four genes are dispensable for growth.

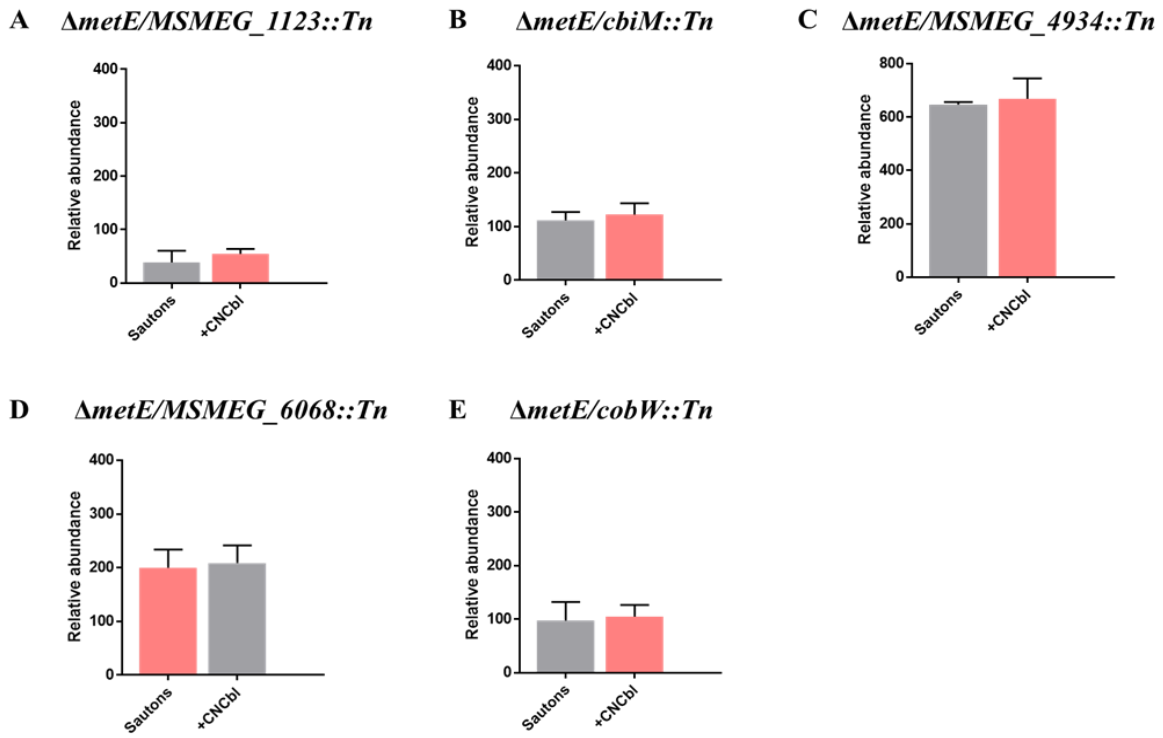


Figure 3.13: Identification of putative cobamide biosynthetic genes. Tn insertion read counts presented per gene between unsupplemented and CNCbl-supplemented medium.

3.12.2 Putative cobalt-transporters in *Msm*

Cobalt (Co) plays an important biological role as the centrally coordinated ion in cyclic tetrapyrroles known as corrin rings and in non-corrin cobalt-dependent enzymes (Cheng et al., 2011). The only known cobalt transport systems in bacteria are encoded by CbiMNQO in *Rhodobacter capsulatus* (Rodionov et al., 2006) and CbtJKL in *S. meliloti* (Cheng et al., 2011). The *Msm* genome appears to contain at least five putative cobalt transporters (CbiM, CbiQ, MSMEG_2610, MSMEG_4469c and MSMEG_6724) which exhibited average Tn insertion densities between 60-100% (**Figure 3.14A**) (**Table 3.1**). On average, these genes were associated with a high number of Tn insertions and read counts (**Figure 3.14A&B**), suggesting that cobalt acquisition in *Msm* is not limiting *in vitro*.

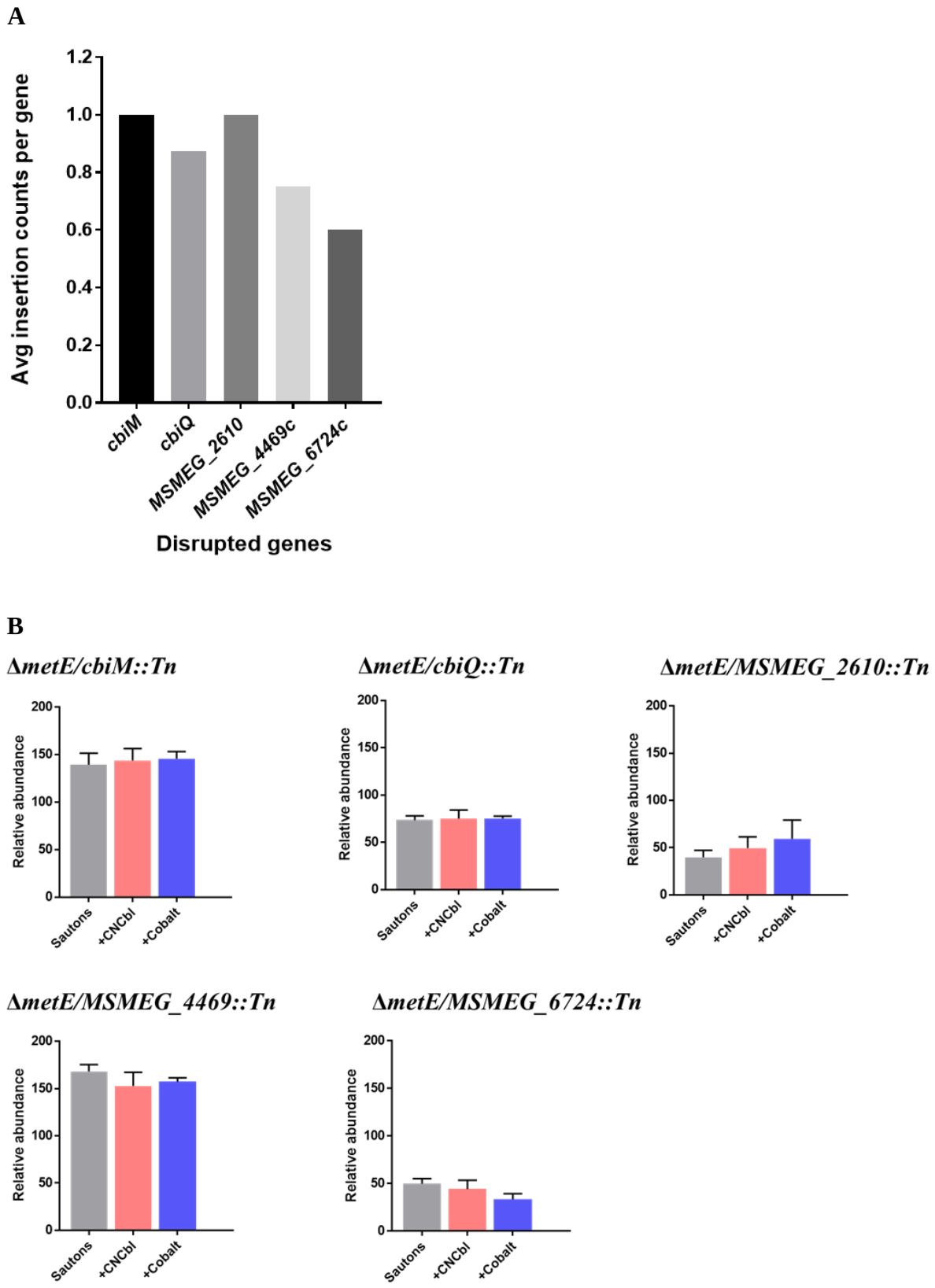


Figure 3.14: Identification of putative cobalt-transporters in *Msm*. Tn insertion read counts presented per gene between unsupplemented and CNCbl-supplemented medium. (**A**) Average insertions count per gene. (**B**) Average Tn read counts within predicted cobalt transporter genes.

3.12.3 Putative cobalamin-dependent enzymes in *Msm*

Msm contains putative homologues of cobalamin-dependent enzymes that are not encoded in the *Mtb* genome: specifically, glutamate mutase (MSMEG_0969), ethanolamine ammonium lyase made up of small and large subunits (MSMEG_1553-1554), and three glycerol dehydratases made up of large and small subunits (MSMEG_1547-1548, MSMEG_6320-6321 and MSMEG_0496-0497). Notably, Tn insertion read counts within these genes were indistinguishable in both unsupplemented and CNCbl-supplemented Sauton's minimal medium (**Figure 3.15**), indicating that none of these putative cobalamin-dependent enzymes was essential for growth (**Table 3.8**).

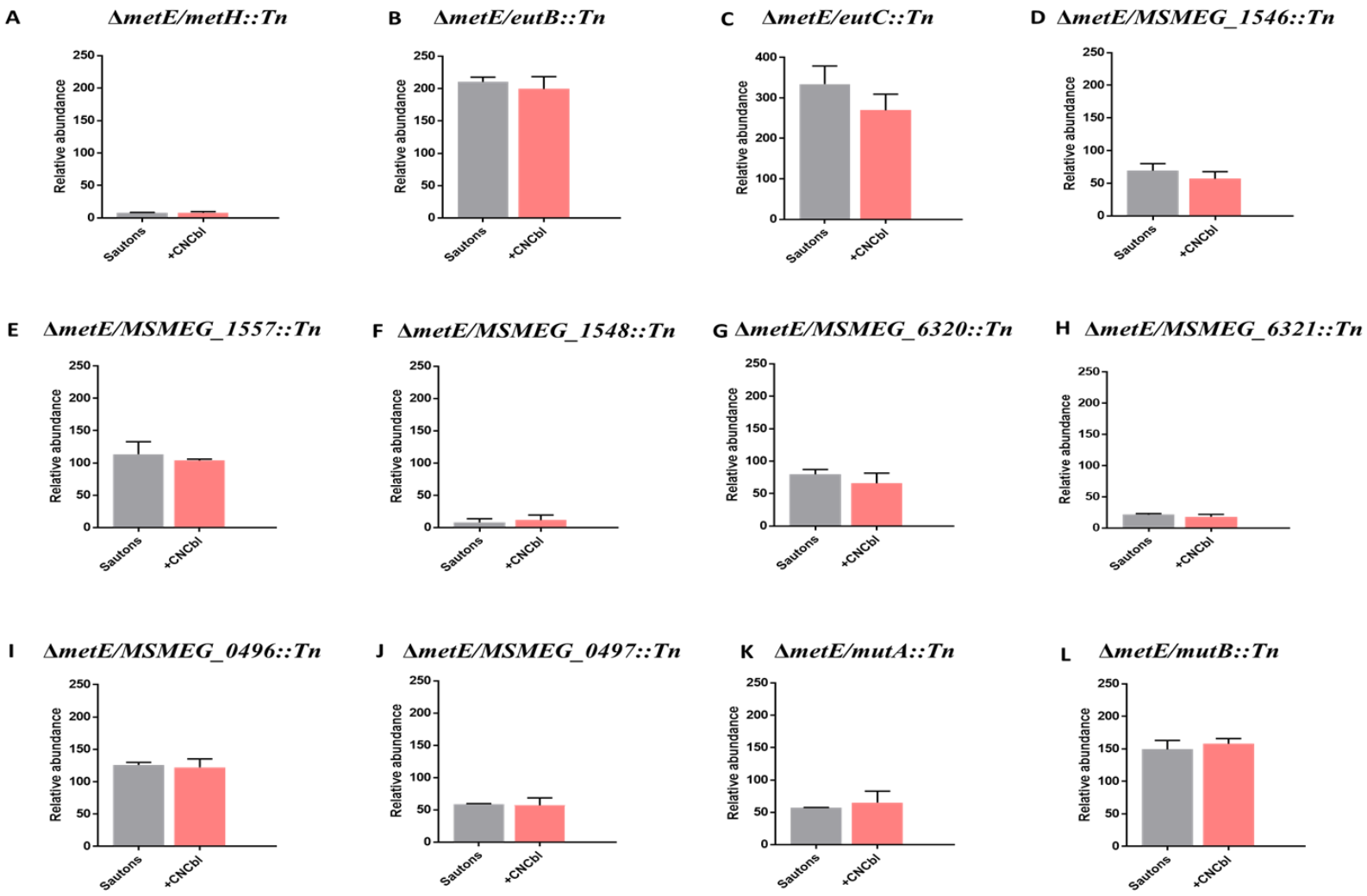


Figure 3.15: Identification of putative cobalamin-dependent enzymes in *Msm*. Tn insertion read counts presented per gene between unsupplemented and CNCbl-supplemented medium**.**

3.13 Validation of selected B_{12} pathway genes by inducible CRISPR interference

Tn screens are very powerful in providing predictions of gene function. However, many studies have demonstrated the requirement to validate Tn screen results at the individual gene level. To that end, CRISPR interference (CRISPRi) (Rock et al., 2017) was applied to a number of candidate cobalamin biosynthesis genes in a spotting assay designed to provide confirmation of inferred essentiality in the *metE* background. For this purpose, the Kan-selectable, chromosomally integrating CRISPR vector, PLJR962 (Rock et al., 2017) was used since it has been optimized for mycobacteria. This vector contains an anhydrotetracycline (ATc)-inducible deactivated CRISPR-associated protein-9 (Cas9) nuclease from *Streptococcus thermophilus* (*d*Cas9$_{Sth1}$) as well as an ATc-inducible, target-specific sgRNA handle to block transcription at the sgRNA base-pairing genomic locus (Peters et al., 2016; Qi et al., 2013) (**Figure 3.16**).

Upon induction by ATc, *d*Cas9$_{Sth1}$ is directed to the target gene by the sgRNA, where the *d*Cas9$_{Sth1}$-sgRNA complex binds a protospacer adjacent motif (PAM) on the DNA. This binding physically blocks the RNA polymerase, preventing transcription initiation or elongation (**Figure 3.16**). Transcriptional interference in a gene required for growth leads to growth inhibition and even cell death, which is inferred from the loss of colony forming ability. To assess the essentialities of putative cobalamin biosynthesis mutants, 7 sgRNA oligos were selected for each gene (**Table 2.5**). Each sgRNA had an assigned PAM score between 1 and 25 in descending order of predicted strength (Rock et al., 2017). Among the 7 sgRNAs, the highest PAM score was 11 and the lowest was 1. The sgRNAs were designed to target random positions from the transcription start site (TSS) of the selected cobalamin biosynthesis genes.

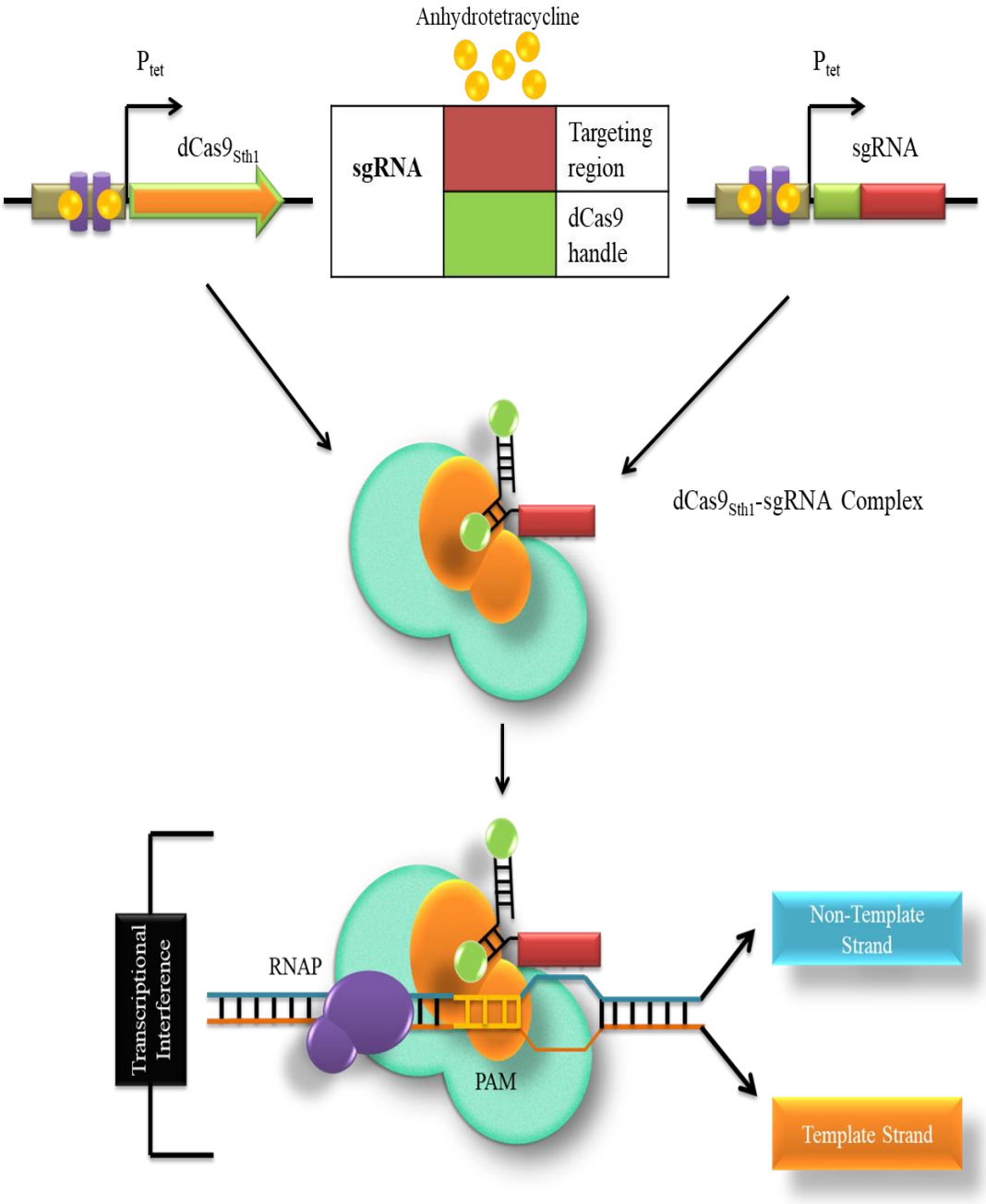


Figure 3.16: Schematic of the mycobacterial CRISPRi system (Rock et al., 2017). The PLJR962 CRISPRi vector is kanamycin selectable and contains tetracycline-inducible promoters for the catalytically-inactive Cas9 protein from *S. thermophilus* (*d*Cas9$_{Sth1}$) and a scaffold of short guide RNAs (sgRNA) specific to target DNA. Anhydrotetracycline (ATc) induces the expression of *d*Cas9$_{Sth1}$ and the sgRNA scaffold, which form a complex and bind target DNA blocking the RNA polymerase and causing transcriptional interference. Figure adapted with permission from Lucas Raphela (personal communication)

3.13.1 Transcriptional knockdown of cobalamin biosynthetic genes in *Msm* Δ*metE*

To validate the TnSeq results, 7 cobalamin biosynthetic genes were selected which are involved in different stages of the predicted pathway; one gene involved in the biosynthesis of the tetrapyrrole precursors ALA and Uro' III (*cobA*), two genes involved in aerobic corrin ring synthesis (*cobIJ* and *cobN*), and the *cobO*, *cobU*, *cobT* and *bluB* genes involved in cobinamide-phosphate synthesis and the assembly and attachment of DMB. All 7 CRISPRi knockdown mutants were constructed in the Δ*metE* mutant: *cobA*, *cobIJ cobN*, *cobU*, *cobT* (all ES according to Tn analyses), *cobO* (GD) and *bluB* (NE). The resulting CRISPRi mutants were plated on 7H10 agar containing Kan, without ATc, and a single colony was then cultured to exponential phase, serially diluted and spotted on solid 7H10 medium containing ATc and 10 µg/ml CNCbl or 1 µM L-methionine **(Figure 3.17)**. An sgRNA for *mmpL3*, a known essential gene involved in mycolic acid biosynthesis, was used as a positive control (Rock et al., 2017)).

As expected, induction of *mmpL3* knock-down (KD) on ATc-containing plates completely inhibited mycobacterial growth. Notably, CRISPRi-mediated KD of genes involved in either the biosynthesis of the tetrapyrrole precursors, ALA and Uro' III (*cobA*), aerobic corrin ring synthesis (*cobIJ* and *cobN*), or in synthesis cobinamide-phosphate and the assembly and attachment of DMB (*cobO*, *cobU/cobT*) inhibited growth of the Δ*metE* mutant in the absence of CNCbl or methionine supplementation. The sole exception was *bluB* (**Figure 3.17**). Moreover, induction of *bluB* KD in ATc-containing plates did not inhibit growth on unsupplemented medium, whereas the induction of *cobO* knockdown under similar conditions resulted in growth impairment (the same gene was identified as growth defective in unsupplemented Sauton's minimal medium) (**Figure 3.17**). The ability to rescue cobalamin KD mutants with L-methionine or CNCbl demonstrated that uptake systems for both were intact in *Msm*, as reported recently (Kipkorir et al., 2021). In summary, the CRISPRi observations provided compelling evidence of the inferred essentialities of the identified cobalamin biosynthetic genes for growth of *Msm* Δ*metE*, in turn supporting the validity of the cobalamin biosynthetic pathway constructed from the TnSeq results (**Figure S6**).

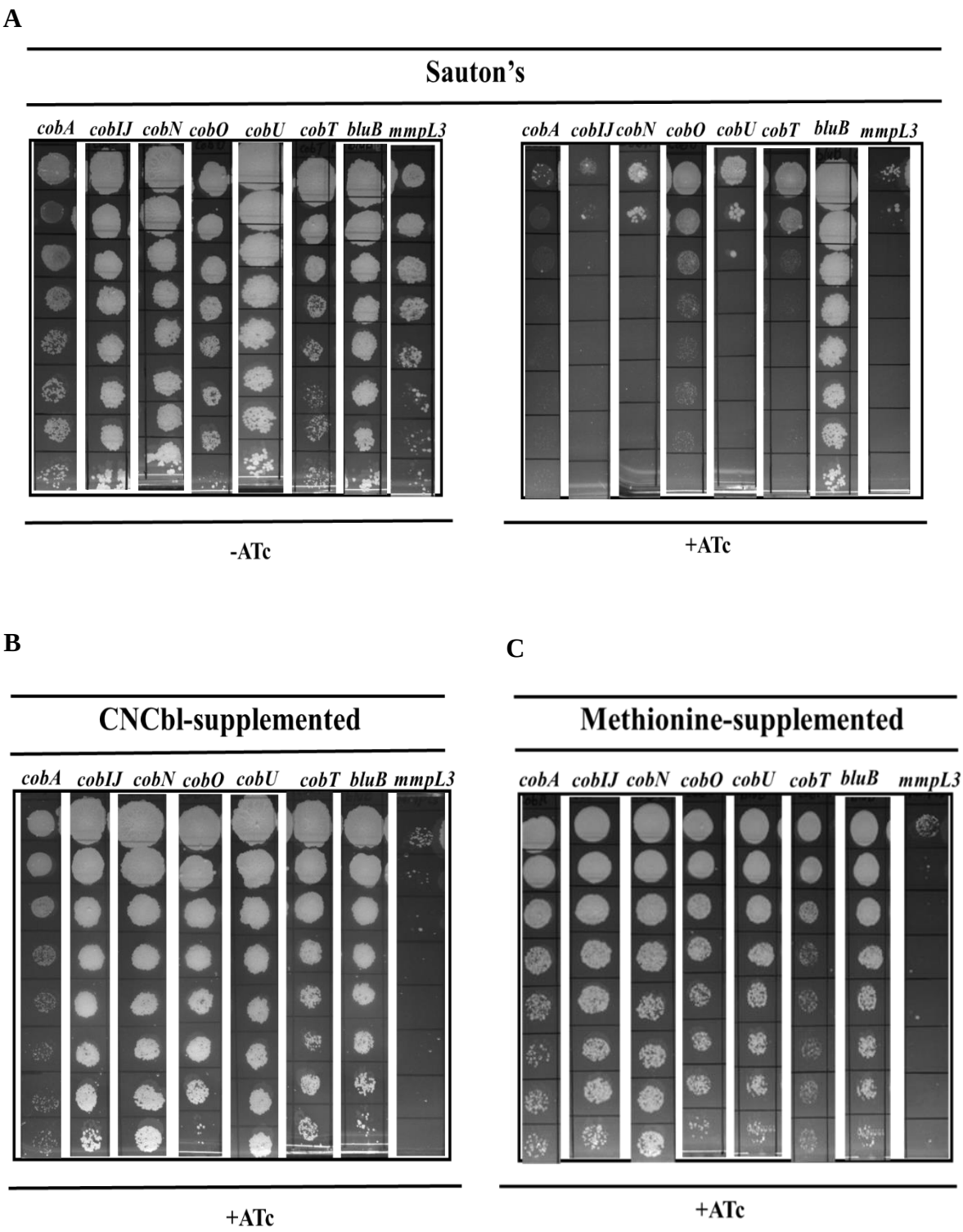


Figure 3.17: CRISPRi validation of predicted cobalamin biosynthesis genes. Spotting assays of *cobA, cobIJ, cobN, cobO, cobU, cobT* and *bluB* knockdowns in the Δ*metE* background. An exponential culture of *Msm* Δ*metE* carrying the indicated CRISPRi KD construct was serially diluted and spotted on 7H10-OADC agar containing methionine or CNCbl with (+ATc) or without (-ATc) the ATc inducer.

3.14 Putative cobalamin transporters in *Msm*

Previous studies have demonstrated that *Mtb* possesses the capacity to scavenge cobamides and can also transport and assimilate cobinamide into functional cobalamin (Gopinath, Venclovas, et al., 2013; Savvi et al., 2008; Warner et al., 2007). Our TnSeq results and CRISPRi validation data provided compelling evidence that CNCbl uptake was functional in *Msm*, even though the proteins and mechanisms involved are uncharacterized. The salvaging of exogenous cobamides starts with their transport into the cell. The only known transporters of cobamides in bacteria are the BtuB-TonB-ExbBD-BtuF-BtuCD systems found in Gram-negative bacteria (Noinaj et al., 2010; Vitreschak et al., 2003). Comparative genetic analyses revealed that homologues of the *E. coli* cobamide transport system components appear to be missing in *Mtb*. However, previous work in the MMRU implicated an atypical transporter, BacA, an ATP-binding cassette transporter protein encoded by the Rv1819c, in cobalamin uptake (Gopinath, Venclovas, et al., 2013). In contrast, to *Mtb*, *Msm* appears to encode homologues of the canonical bacterial cobamide transport system; *MSMEG_6064* (*btuF*), *MSMEG_6063* (*btuC*), and *MSMEG_6062* (*btuD*), together with their paralogs, *btuF2* (*MSMEG_4561* and *MSMEG_4560*), *btuC2* (*MSMEG_4559*) and *btuD2* (*MSMEG_4557*) found in separate operon (Novichkov et al., 2013). Furthermore, the *Msm* genome encodes two homologues of the *Mtb* BacA (*MSMEG_3655 (bacA)/MSMEG_4380 (bacA2)*). These genes appear to be regulated by cobalamin-riboswitches which are predicted upstream of *btuF* and *btuF2a.* Tn insertions in genes involved in cobalamin uptake were indistinguishable in both unsupplemented and CNCbl-supplemented Sauton's minimal medium (**Figure 3.18**), suggesting likely redundancy (**Table 3.8**).

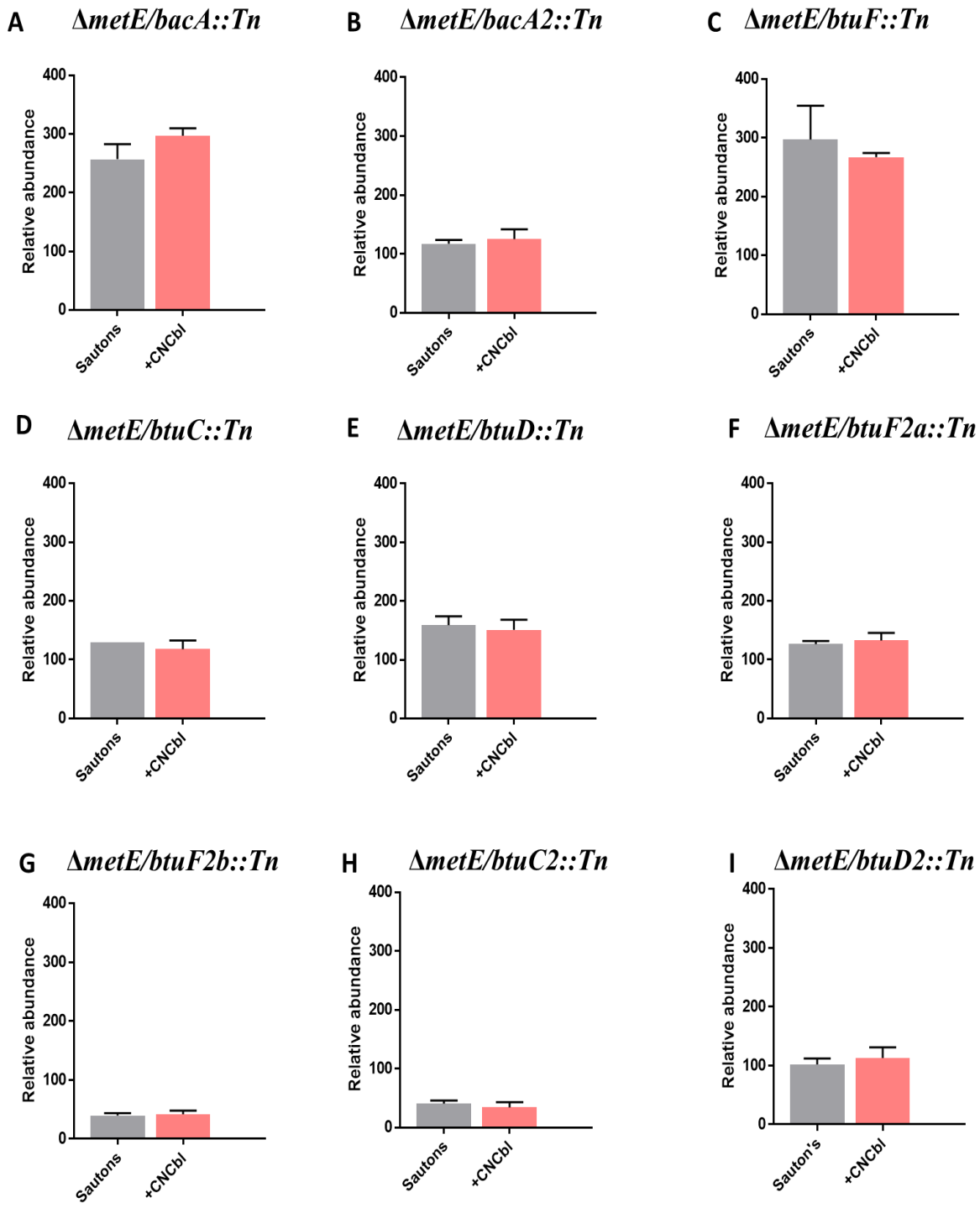


Figure 3.18: Identification of putative cobamide transporters in *Msm*. Tn insertion read counts presented per gene between unsupplemented and CNCbl-supplemented medium**.** Tn insertions in these gene were indistinguishable in both conditions. Data are representative of three independent experiments. Error bars show the standard error of the mean.

Table 3.8: Tn insertions in selected genes predicted to be involved in cobalamin biosynthesis, transport, and cobalamin-related metabolism during growth of Δ*metE* mutant in Sautons minimal medium with or without CNCbl and cobalt supplementation

Accession no.	Gene name	Predicted function	Essentiality call in Sauton's	Essentiality call in +CNCbl	Essentiality call in +Cobalt
Putative cobalamin biosynthesis protein					
MSMEG_1123		cobalamin synthesis protein	NE	NE	NE
MSMEG_2607	*cbiM*	cobalamin biosynthesis protein CbiM	NE	NE	NE
MSMEG_4934	-	ATP:cob(I)alamin adenosyltransferase	GA	GA	GA
MSMEG_6048	-	cobalamin synthesis protein/P47K	NE	NE	GA
MSMEG_6069	*cobW*	CobW/P47K domain-containing protein	NE	NE	NE
Putative cobamide transporters					
MSMEG_3655	*bacA*	ABC transporter, permease/ATP-binding protein	NE	NE	NE
MSMEG_4380	*bacA2*	ABC transporter, permease/ATP-binding protein	NE	NE	NE
MSMEG_6064	*btuF*	lipoprotein	NE	NE	NE
MSMEG_6063	*btuC*	Fe uptake system integral membrane protein	NE	NE	NE
MSMEG_6062	*btuD*	Fe uptake system permease	NE	NE	NE
MSMEG_4561	*btuF2a*	ABC Fe3+-siderophores transporter, periplasmic binding protein	NE	NE	NE
MSMEG_4560	*btuF2b*	periplasmic binding protein	NE	NE	NE
MSMEG_4559	*btuC2*	ABC transporter, membrane spanning protein	NE	NE	NE
MSMEG_4557	*btuD2*	ABC transporter, ATP-binding protein	NE	NE	NE
Putative cobalt transporters					
MSMEG_2608	*cbiM*	cobalt transport protein CbiM	NE	NE	NE
MSMEG_2609	*cbiQ*	cobalt ABC transporter, permease protein CbiQ	NE	NE	NE
MSMEG_2610	-	cobalt transport protein ATP-binding subunit	NE	NE	NE
MSMEG_4469	-	cobalt transport protein	NE	NE	NE
MSMEG_6724	-	ABC-type cobalt transport system	NE	NE	NE
Cobalamin-dependent enzymes					

MSMEG_4185	*metH*	B_{12}-dependent methionine synthase	ES	ES	ES
MSMEG_0969	*hemL*	glutamate-1-semialdehyde aminotransferase	ES	ES	ES
MSMEG_1553	*eutB*	ethanolamine ammonia-lyase, large subunit	NE	NE	NE
MSMEG_1554	*eutC*	ethanolamine ammonia-lyase small subunit	NE	NE	NE
MSMEG_1547	-	glycerol dehydratase large subunit	NE	NE	NE
MSMEG_1546	-	coenzyme B_{12}-dependent glycerol dehydrogenase small subunit	NE	NE	NE
MSMEG_1548	-	propanediol utilization: dehydratase, medium subunit	NE	NE	NE
MSMEG_6320	-	diol dehydrase gamma subunit	NE	NE	NE
MSMEG_6321	-	glycerol dehydratase large subunit	NE	NE	NE
MSMEG_0496	-	coenzyme B_{12}-dependent glycerol dehydrogenase small subunit	NE	NE	NE
MSMEG_0497	-	glycerol dehydratase large subunit	NE	NE	NE
MSMEG_3158	*mutA*	methylmalonyl-CoA mutase, small subunit MutA	NE	NE	NE
MSMEG_3159	*mutB*	methylmalonyl-CoA mutase, Large subunit MutB	NE	NE	NE

Chapter 4: Discussion

4.1 Identification of differential genetic requirements for *in vitro* growth by TnSeq

TnSeq is a powerful whole-genome genotypic screening method that has been widely used to identify essential as well as *conditionally essential* regions in a genome – both coding and non-coding – in different growth conditions and following exposure to antibiotics and various other chemical stresses (Deutschbauer et al., 2011; Gallagher et al., 2011; van Opijnen & Camilli, 2012; Yung et al., 2015). A Tn-mutagenized library is composed of mutants that contain one Tn insertion at a random location in the host cell's genome, typically disrupting and thus inactivating a gene. Highly saturated libraries can be easily generated so that, across the entire library, every gene in the genome is disrupted at multiple locations. Owing to their lethality, insertions in essential genes will be eliminated from the library during construction; this enables essential genes/genome regions to be putatively identified by the absence of insertions (Shull & Camilli, 2018; Van Opijnen et al., 2009; van Opijnen & Camilli, 2012; van Opijnen et al., 2014). Conditionally essential genes are considered an extension of the "true" essential genes in the sense that they are required for growth or survival only under the condition investigated (Kwon et al., 2016). By comparing TnSeq datasets under different experimental conditions it is thus possible to infer conditional essentialities (Chao et al., 2016; Van Opijnen et al., 2009; van Opijnen & Camilli, 2012; van Opijnen & Levin, 2020).

In this study, a combination of high-density mutagenesis and deep-sequencing was used to characterize the composition of Tn mutants derived from libraries generated in different genetic backgrounds (WT *Msm* mc^2155 and a derivative Δ*metE* mutant) and exposed to different growth media *in vitro* (±CNCbl, $\pm Co^{2+}$). A total of 372 genes (classified as ES or GD) were identified as critical for growth of *Msm* on 7H10 agar, whereas 292 genes (classified as ES or GD) were identified as essential for growth of the Δ*metE* mutant on rich (LB) agar. By comparing the list of essential genes generated in this study with the corresponding essential gene list reported recently by Dragset et al. (2019), 237 genes were commonly identified as essential for survival of *Msm in vitro*, representing 3.5% of total coding sequences. This number is significantly lower than the 15% essential identified in *Mtb* (DeJesus et al., 2017; Griffin et al., 2011), a deviation that seems likely to reflect the specialisation of *Mtb* for human colonization/infection – whereas the environmental mycobacterium possesses a larger genome encoding alternatives for otherwise essential genes (Weerdenburg et al., 2015).

TnSeq involves statistical analyses of Tn insertion location and frequency across coding and non-coding regions, ultimately providing a quantifiable assessment of Tn library composition from quantitative insertion count. However, performing discrete comparisons of essential genes between strains or conditions overlooks quantitative differences in Tn count between genes with the same qualitative classification (Carey et al., 2018). For example, a gene able to tolerate Tn insertions in two strains could be classified as non-essential in both strains, even if the relative abundance of these Tn mutants in their respective library pools differs significantly. Such differences in Tn mutant abundance reflect differences in the fitness cost of gene disruption in a particular genetic background and indicate the relative dependency of a strain on the gene for growth and survival (Sassetti et al., 2003; Van Opijnen et al., 2009; Xu et al., 2017). To overcome this problem, a permutation test-based method incorporated into the TRANSIT platform (DeJesus et al., 2015) was used to define genes with such differential requirements for *in vitro* growth. Employing this methodology, 23 genes were identified which were differentially essential in the WT (7H10-standard medium) *versus* Δ*metE* (LA-rich medium) datasets. These included genes involved in fundamental cellular processes, including methionine biosynthesis, purine and pyrimidine nucleotide biosynthesis pathways and one-carbon metabolism (**Figure 4.1A**).

Among the genes identified with differential essentiality assignments were cobalamin-dependent (MetH) and cobalamin-independent (MetE) methionine synthases, and the bifunctional PurH enzyme which catalyses the final two steps in *de novo* purine biosynthesis, AICAR transformylase and IMP cyclohydrolase. The *glyA* gene, encoding a serine hydroxymethyltransferase required for the reversible interconversion of serine to glycine utilizing tetrahydrofolate (THF) as the one-carbon carrier, was also found to be differentially essential in this study. PurH and GlyA play key roles in maintaining the cellular pool of folate in bacteria. These enzymes are involved in recycling 5,10-methylenetetrahydrofolate (5,10-CH_2-THF) back to free THF which can be further converted to other folate forms (Nijhout et al., 2004). 5,10-CH_2-THF can be terminally reduced to form 5-CH_3-THF, a methyl donor used for conversion of homocysteine to methionine in the final step of methionine biosynthesises by cobalamin-dependent (MetH) or cobalamin-independent (MetE) methionine synthases (Blanco et al., 1998). Folate and methionine metabolism intersect at the methylation of homocysteine. Besides the conversion of homocysteine to generates methionine, MetH and MetE recycles 5-CH_3-THF back to THF. Recently, Guzzo *et al*. showed that the disruption of MetH activity or cobalamin biosynthesis fusion protein CobIJ, led to the accumulation of 5-CH_3-THF in *Msm*,

forming the methyl folate trap, thus interrupting the normal flow of the one-carbon metabolic reactions (Guzzo et al., 2016).

This reactions serves as the major source of one-carbon groups required for the biosynthesis of purines, thymidylate, methionine, and other important biomolecules; the inferred essentiality in WT *Msm* but not Δ*metE* therefore implies alterations in one-carbon utilization as a consequence of disrupted homeostasis in the methionine synthase-deficient mutant (Guzzo et al., 2016). Performing discrete and pairwise comparisons of the Tn insertion profile (Chao et al., 2013; Pritchard et al., 2014) of the WT and Δ*metE* mutant strains revealed both essential and enriched genes in each genetic background. Comparing the essentiality assignments determined in this study with recent work by Dragset et al. (2019), it was noted that the majority of essential genes were consistent across both studies. However, significant differences in these predictions were also observed, since 45 genes were identified as conditionally essential in this study, while 77 genes were essential in Dragset et al. (2019) study , and many of these differences might be due to technical and analytical refinements, for example variability in mutant's abundance during library constructions (variations in Tn insertions at each particular sites), as well as, stochastic variations among replicates and biases due to sample preparation protocol and sequencing technology (**Figure 4.1B**). Like Dragset, the current study identified genes whose disruption in both WT and *metE* seemed beneficial to the cell.

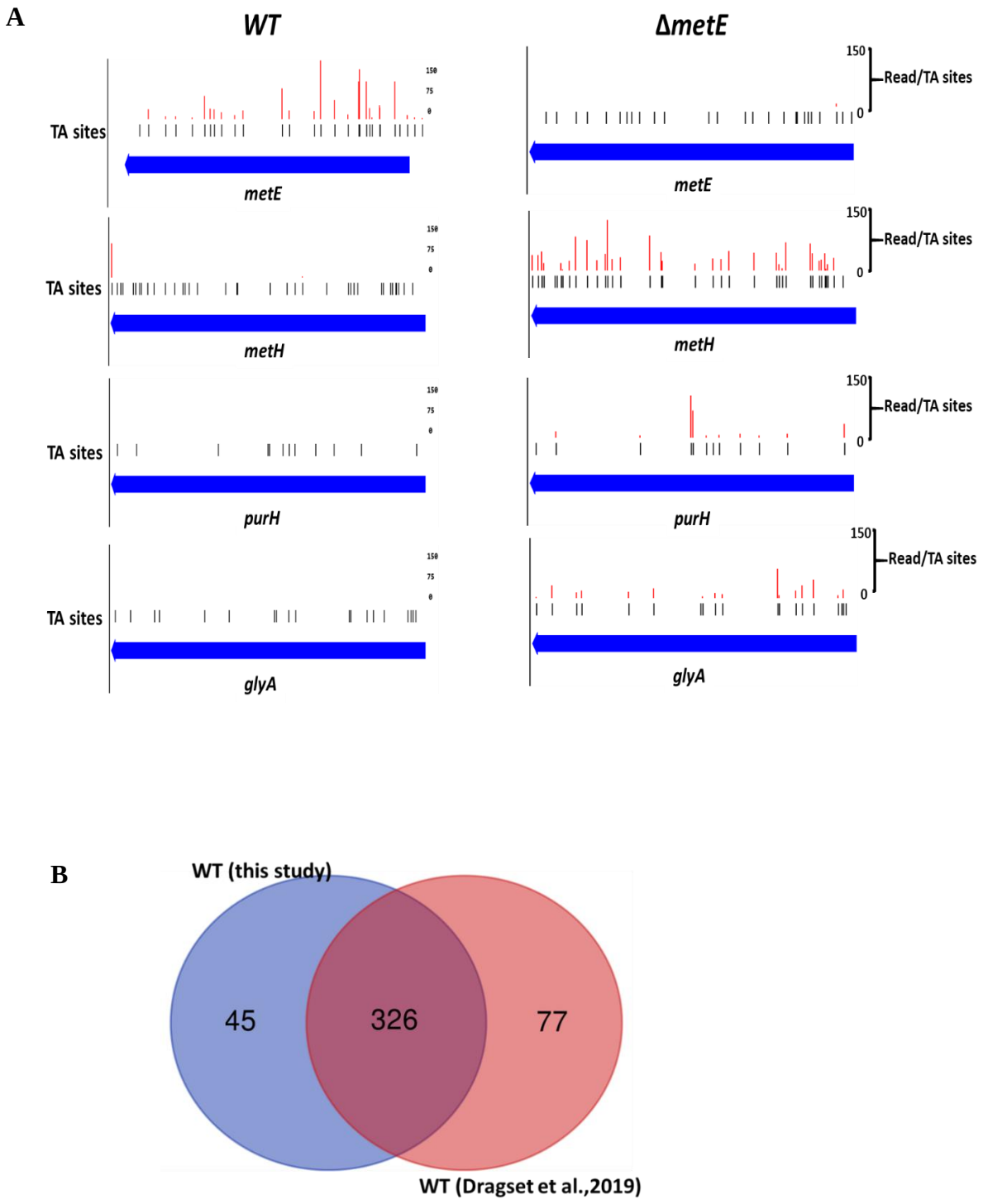


Figure 4.1: Tn insertions determined from high-throughput sequencing. (**A**) Differences in Tn insertion counts and distributions were observed in the *metE* gene (insertions in WT only) and *metH*, *purH* and *glyA* (which were devoid of any insertions in WT). (**B**) Venn diagram comparing essential genes (ES + GD) identified by TnSeq in this study and in recent work by Dragset et al. (2019).

As expected, it was found that the representation of cobalamin Tn mutants was similar in both libraries, WT and Δ*metE*. Statistical analysis of Δ*metE Himar1*-Tn mutants generated in LA medium identified a collection of cobalamin mutants that displayed normal growth in this condition, confirming the non-essentiality of the encoded functions in rich medium. Of the 58,156 insertions into 77,755 possible -TA- sites, 157 insertions occurred in genes predicted to be involved in *de novo* cobalamin biosynthesis, while 250 insertions were identified in genes predicted to be involved in cobamide metabolism or related pathways. The disrupted genes included those predicted to be involved in the biosynthesis of the tetrapyrrole precursors, genes involved in the aerobic corrin ring synthesis pathway, genes involved in nucleotide loop assembly, synthesis of cobinamide-phosphate, and in the synthesis and attachment of DMB. Other genes whose functions are related to cobamide metabolism or related pathways included those involved in cobamide and cobalt scavenging, cobalamin-dependent reactions and one-carbon metabolism. Similar results were observed in the analysis of the WT Tn library, in which cobalamin biosynthesis mutants displayed a normal growth implying the non-essentiality of the corresponding genes in this background. Since the WT genome encodes the cobalamin-independent MetE, which enables the biosynthesis of methionine in the absence of cobalamin, this finding was not surprising. Moreover, the identification of many genes whose functions are related to cobamide metabolism indicated that the screen was successful.

With the exception of genes involved in one-carbon metabolism, the Tn-mediated inactivation of genes involved in cobalamin metabolism was not detrimental to bacterial growth in either background, WT or Δ*metE*. Furthermore, of the 23 identified nonessential genes involved in the cobamide biosynthetic pathway in *Msm*, only two genes – *hemD* and *cobQ2* – were identified as critical for growth in both backgrounds, strongly suggesting an additional role(s) in other metabolic pathways. Even though these two genes contain a large number of -TA- sites, significant gaps were observed in Tn insertion coverage (**Figure 3.6**). Identifying essential genes (and regions) by searching for statistically significant gaps in Tn insertion coverage, instead of simply looking for those that are completely devoid of insertions, is crucial. This is because other likely essential genes could harbour a small number of insertions and could be falsely classified as nonessential (DeJesus et al., 2017; Griffin et al., 2011). *De novo* cobalamin biosynthetic pathway can be divided into three stages. The first stage, tetrapyrrole precursor biosynthesis, contains the first five steps of the pathway, most of which are also common to the biosynthesis of heme, chlorophyll, and other tetrapyrroles. We observed that genes involved in the biosynthesis of heme (*hemA, hemC, hemD, hemB, hemL,*

hemE, *hemG* and *hemH* were identified as essential for growth in both backgrounds, (WT and Δ*metE*) (**Table S1**) and disruption of this genes in *Mtb* was found to be detrimental to bacterial survival *in vitro* (**Table S1**) (DeJesus et al., 2017). Heme is an essential cofactor and serves as source of nutritional iron for many bacterial pathogens, including *Mtb* (Choby & Skaar, 2016).

4.2. Analysis of conditional gene fitness of *Msm* Δ*metE* Tn mutants during growth in minimal defined medium.

TnSeq uses the bacterial cell as a sensor to report directly on the genes and pathways required for microbial growth. It simultaneous allows the assessment of the relative abundance of pool of insertion mutants in a population and identification of genes whose disruptions is disastrous for cell growth under selective conditions (Carey et al., 2018; Samant et al., 2008). In this study, it was notable that the abundance of each mutant strain changed as result of the impact of the underlying gene mutation.

To identify phenotypes and gain insight into gene functions, *Msm* Δ*metE* Tn mutants were subjected to competitive growth assays in defined media. In these experiments, replicate portions of this library (approximately 10^6 cells were used per condition) were grown for 24 generations in Sauton's medium with or without CNCbl or cobalt supplementation. As expected, the relative abundances of most mutants were significantly different between supplemented and unsupplemented minimal media, including those required for *de novo* cobalamin biosynthesis. Only 10, 13 and 24 genes (ES plus GD) were identified as uniquely required for growth in unsupplemented, CNCbl-supplemented and cobalt-supplemented Sauton's minimal medium, respectively. Furthermore, because gene essentiality can be affected by nutrient availability, it was necessary to compare gene essentialities in rich medium (LA) and minimal medium (Sautons).

As expected, a much larger number of mutants was underrepresented in the Sauton's-grown pool. So, 191 genes were identified as conditionally essential in this minimal medium compared to those found in rich medium (**Figure 4.2**). Unexpectedly, 14 genes were conditionally essential in rich medium compared to those found in minimal medium. Of the 14, only 10 were identified in the WT Tn library from this study and Dragset et al. (2019). Example genes included two putative acyl-CoA dehydrogenases (*MSMEG_5621c* and *MSMEG_5622c*). Acyl-CoA dehydrogenase (ACD) introduces unsaturation into fatty acids in

lipid metabolism pathways (Chen et al., 2020). Other genes include three hypothetical proteins (MSMEG_0271c, MSMEG_0274c and MSMEG_1241), and the signal-transduction histidine kinase (*senX3*) (**TableS1**).

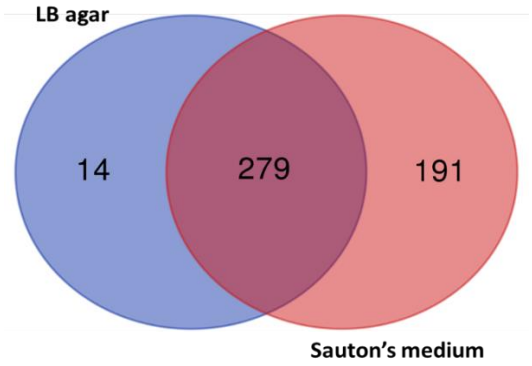


Figure 4.2: Venn diagram comparisons of essential genes (ES + GD) identified in LA-selected Tn library and Sauton's minimal medium-selected Tn library.

The principal goal of this study was to provide direct evidence for the involvement of predicted genes in the *de novo* cobalamin biosynthesis in *Msm* (**Figure S6**). To this end, selection was performed in triplicate using highly saturated Tn libraries, thereby avoiding bottleneck effects. To validate the reproducibility of TnSeq, biological replicates were performed, enabling comparison of two or more independently established and selected-on libraries (Van Opijnen et al., 2009). The datasets included three replicates grown on Sauton's minimal medium, and three replicates grown on the same medium supplemented with CNCbl or cobalt. The quality of TnSeq experiments can be assessed in multiple ways: firstly, by looking at the number of reads containing the Tn and the number of reads mapping to the genome; secondly, from the number of insertion sites recovered and mean read count; and thirdly, from the correlation between the numbers of reads recovered for each gene in replicated experiments (DeJesus et al., 2015). This ensures that, if the loss of particular mutants was purely due to selection, one might expect a high correlation across libraries and experimental conditions, as these losses should presumably be reproducible under the same experimental conditions.

Summary statistics of the sequencing runs for this study are presented in Table 5. Total read counts yield varied from ~7 to ~12.2 million reads. Sequencing these libraries yielded an average of 1.5 million unique Tn-chromosome junctions (Mapped Templates), which could be mapped to 66.5–76.7% of the -TA- dinucleotide sites in the chromosome in each individual experiment; this was deemed to provide adequate coverage for the experimental assays. The

mean read counts of the replicates were in the range 38.4 to 103.3 per TA site. Reproducibility between biological replicates was always high with a Pearson correlation coefficient between 0.76–0.96 (**Figure 4.3**). These data support the inference that TnSeq is highly reproducible both within and between independent experiments (Van Opijnen et al., 2009; van Opijnen & Levin, 2020).

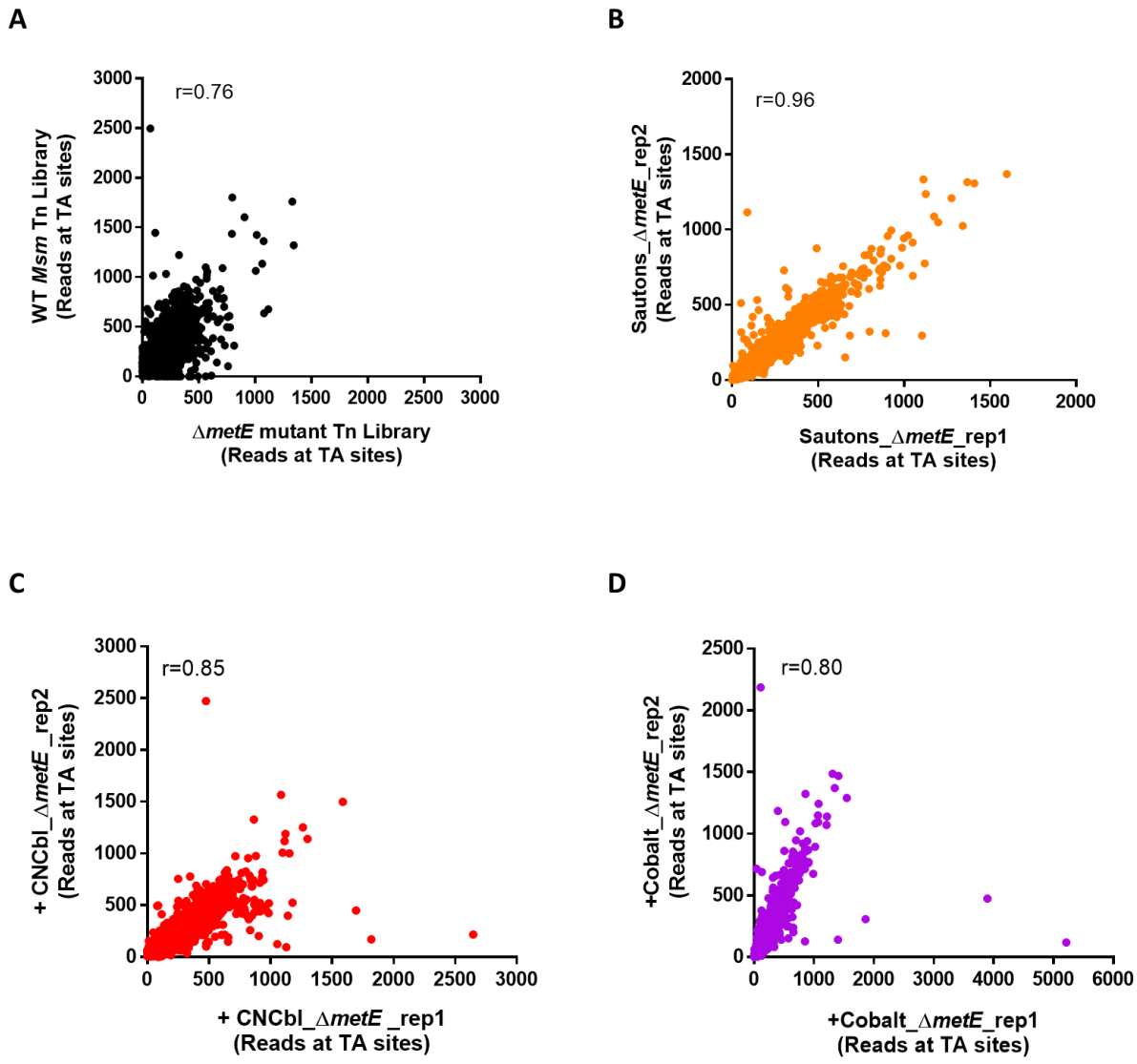


Figure 4.3: Validation of TnSeq fitness measurements. (**A**) Technical replicates comparing two independently generated Tn libraries, (**B-D**) selected upon and sequenced libraries. Pearson correlation coefficients between biological replicates ranged between 0.76 – 0.96.

4.3 Cobalamin biosynthetic enzymes are functional in *Msm*

Cobamides are essential nutrients, required by humans, bacteria, and other organisms as enzyme cofactors. The *de novo* biosynthesis of cobamides occurs only in bacteria and archaea (Doxey et al., 2015; Zhang et al., 2009), even though the co-factors are used by all domains of life (Martens et al., 2002; Nielsen et al., 2012; Rodionov et al., 2003). Cobamides are best known for their role as cofactors of enzymes that mediate methyl transfer reactions, and they regulate protein expression by binding to ligand-specific riboswitches to prevent protein translation (Rodionov et al., 2003). Genomic analyses indicate that *Msm* possesses the complete pathway for *de novo* cobalamin biosynthesis, comprising nearly 30 enzyme-catalysed steps (Kipkorir et al., 2021). In contrast, key absences in gene complement suggest that *Mtb* encodes a near-complete cobamide biosynthetic pathway (Gopinath, Moosa, et al., 2013; Rodionov et al., 2003; Young et al., 2015). For example, *Mtb* lacks a homologue of the precorrin-6A synthase encoded by CobF in *P. denitrificans* (Debussche et al., 1993; Min et al., 1993), the loss of which is considered a defining molecular event in the evolution of MTBC strains (Boritsch et al., 2014; Brosch et al., 2002; Ngabonziza et al., 2020; Supply & Brosch, 2017).

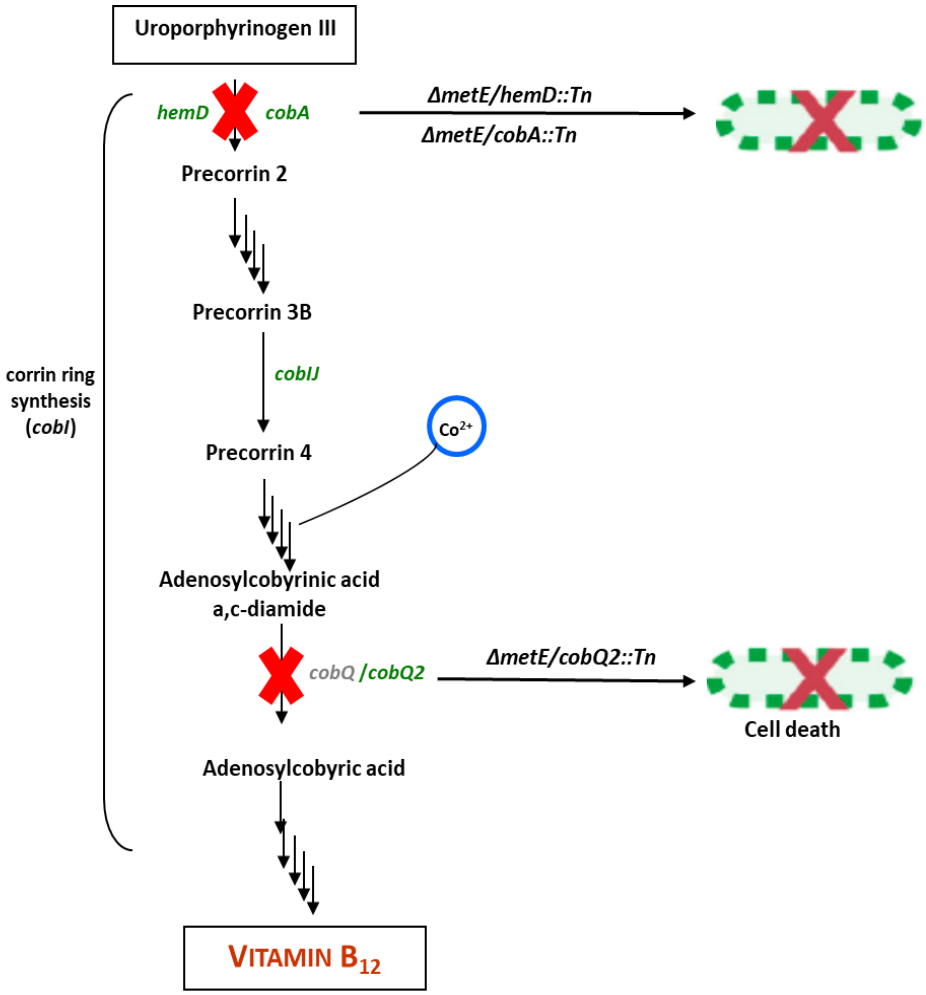
In this study, TnSeq was applied to obtain direct evidence for the functionality of genes in the *de novo* cobalamin biosynthesis and the essentiality of cobalamin-dependent regulation of methionine biosynthesis in *Msm*. To identify differences in the genetic requirements for *de novo* cobalamin biosynthesis between conditions, a permutation test-based method was used to identify genes with statistically significant differences in Tn insertion count. A large fraction of cobalamin biosynthetic genes in *Msm* were among those significantly under-represented in the *metE* mutant during growth in minimal media, confirming the essentiality of cobalamin biosynthesis under these conditions (**Figure 3.10**-**3.12**). Notably, definitive elucidation of cobalamin biosynthetic and assimilatory genes required the analysis of libraries exposed to CNCbl-supplemented minimal media for extended durations, probably to ensure exhaustion of the organism's capacity for co-factor storage and recycling. In addition, the results from our study add further evidence in support of the functionality of the cobalamin riboswitch upstream of *metE* (**Figure 4.1A**).

Although our TnSeq analysis was highly reproducible, the reliability of the data was further validated by testing seven cobalamin knockdown strains in the Δ*metE* background: *cobA, cobIJ cobN, cobU, cobT* (all ES), *cobO* (GD) and *bluB* (NE). CRISPRi-mediated transcriptional

silencing of *cobA, cobIJ cobN, cobU, cobT* abrogated growth of the Δ*metE* mutant in unsupplemented 7H10 medium, whereas induction of *cobO* knockdown resulted in growth impairment, and the induction of *bluB* knockdown had no suppressive effect. These data confirmed the inferred essentiality of cobalamin biosynthetic genes for the growth of *Msm* Δ*metE*, in turn validating the utility of TnSeq to conduct genotypic profiling to identify genes required for fitness under any condition of interest.

The observation that CNCbl supplementation could rescue the growth defect of all cobalamin biosynthesis mutants – except for *hemD, cobA* and *cobQ2* Tn mutants (**Figure 4.4**) – supported the functionality of cobamide transporters in *Msm*. The first step in the cobamide biosynthetic pathway, cyclization of the linear tetrapyrrole, hydroxymethylbilane, to the macrocyclic uroporphyrinogen III, is catalysed by the *hemD*-encoded uroporphyrinogen-III synthase. CobA serves as a rate-limiting enzyme that converts uroporphyrinogen II to precorrin-2, which is eventually incorporated with DMB to form cobalamin (Piwowarek et al., 2018). Interestingly, the *cobA* gene was initially annotated as *cysG* in some bacteria, including *Propionibacterium spp*, due to high similarity to C-terminal portion of the *E. coli* CysG, which is responsible for cabaltochelatase activity. (Escalante-Semerena et al., 1990; Li et al., 2020; Sattler et al., 1995). CobQ2 has been identified as a putative GatD amidase involved in cell wall biosynthesis (Pavelka Jr et al., 2014) and disruption of this gene (*cobQ2*; Rv3713) in *Mtb* was found to be detrimental to bacterial survival *in vitro* (DeJesus et al., 2017; Griffin et al., 2011; Sassetti et al., 2003). Notably, while *cobA* Tn mutants fail to grow in CNCbl-supplemented medium, it was observed that the *cobA* CRISPRi knockdown was partially rescued in CNCbl-supplemented medium. Therefore, the functionality of this gene in cobamide biosynthesis seems unlikely and further work is required to determine its role in *Mtb*.

A


Uroporphyrinogen III
hemD
cobA
ΔmetE/hemD::Tn
ΔmetE/cobA::Tn
Precorrin 2
Precorrin 3B
cobIJ
corrin ring synthesis (cobI)
Co2+
Precorrin 4
Adenosylcobyrinic acid a,c-diamide
cobQ /cobQ2
ΔmetE/cobQ2::Tn
Cell death
Adenosylcobyric acid
VITAMIN B12

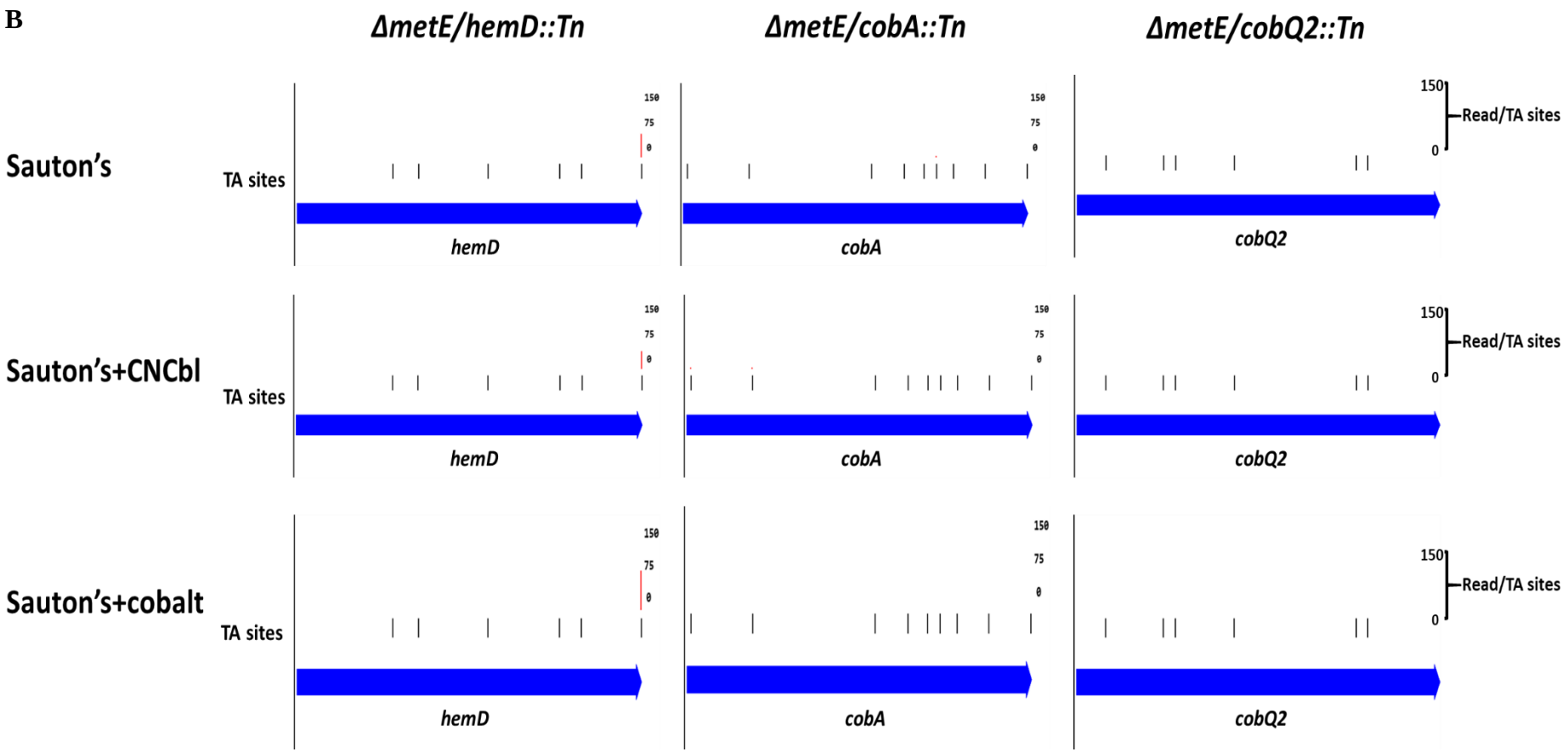


Figure 4.4: Comparison of Tn insertion frequencies among selected cobalamin biosynthesis mutants under different selective condition.

The *Msm* genome encodes homologues of classic BtuFCD-type cobamide transporters (Noinaj et al., 2010) as well as the atypical transport protein BacA (Gopinath, Venclovas, et al., 2013). Consistent with the presence of multiple candidate cobamide transporters in *Msm*, disruption of *bacA* did not cause any growth impairment in this study. Further experiments are necessary to verify the functions of these putative cobamide transport proteins in cobamide salvage. In addition, *Msm* genome appears to encode at least five putative cobalt transport systems, three of which are homologues of well-defined cobalt transporters including CbiMQO, CbtAB and CorA (Nelson & Kennedy, 1971; Rodionov et al., 2003; Roth et al., 1993). Tn insertions were identified in the predicted cobalt transporters; however, following selection in unsupplemented and cobalt supplemented media, none of these genes were required for growth (**Table 3.8**), supporting their functional redundancy.

4.4 BluB and MSMEG_4305 are dispensable for growth in *Msm*

Although most of the cobalamin genes were found to be essential in unsupplemented minimal medium, two genes predicted to be involved in α-ribazole synthesis, *bluB* and *MSMEG_4305*, were found not to be required for growth in all tested conditions. Both genes had higher levels of Tn insertions in the Δ*metE* Tn library selected on unsupplemented and CNCbl or cobalt supplemented minimal medium (**Figure 4.5**). As expected, *cobT* (involved in the attachment of DMB to cobinamide-phosphate) was identified as essential for growth in supplemented minimal medium and not in CNCbl-supplemented minimal medium (**Figure 4.5A**). Therefore, because the inactivation of the genes was not detrimental to bacterial growth in all tested conditions, it was concluded that these genes were dispensable for growth in *Msm*. These results raised the possibility that *Msm* might possess an alternative route for the synthesis of the lower ligand or that *Msm* might possess the capacity to synthesize (and utilize) pseudocobamides. Previous work in *Salmonella* has established that the synthesis of pseudocobalamin requires the same set of enzymes (CobT, CobU, CobS and CobC) that are involved in the synthesis of DMB-containing cobalamin (Anderson et al., 2008). Notably, BluB was previously predicted to be a cob(II)yrinic acid a,c-diamide reductase ("cobalt reductase") in the cobamide biosynthetic pathway. However, bioinformatic analyses identified Rv0306 and MSMEG_6053 as the putative homologue of *S. meliloti* BluB in *Mtb* and *Msm*, respectively (Campbell et al., 2006; Gray & Escalante-Semerena, 2007; Rodionov et al., 2003; Taga et al., 2007). MSMEG_4305 is a two-domain protein containing a signature motif of phosphoglycerate mutase family. The N-terminal domain of MSMEG_4305 encodes an RNase

H type I. The C-terminal domain is a presumed CobC, predicted to be involved in the aerobic synthesis of cobalamin (Czubat et al., 2020). MSMEG_4305 and Rv2228c are the putative homologues of *S. enterica*, a-ribazole-5-phosphatase (CobC) in *Msm* and *Mtb*, respectively. In contrast to *Msm*, Rv2228c was predicted to be essential in *Mtb* H37Rv *in vitro* (DeJesus et al., 2017; Griffin et al., 2011).

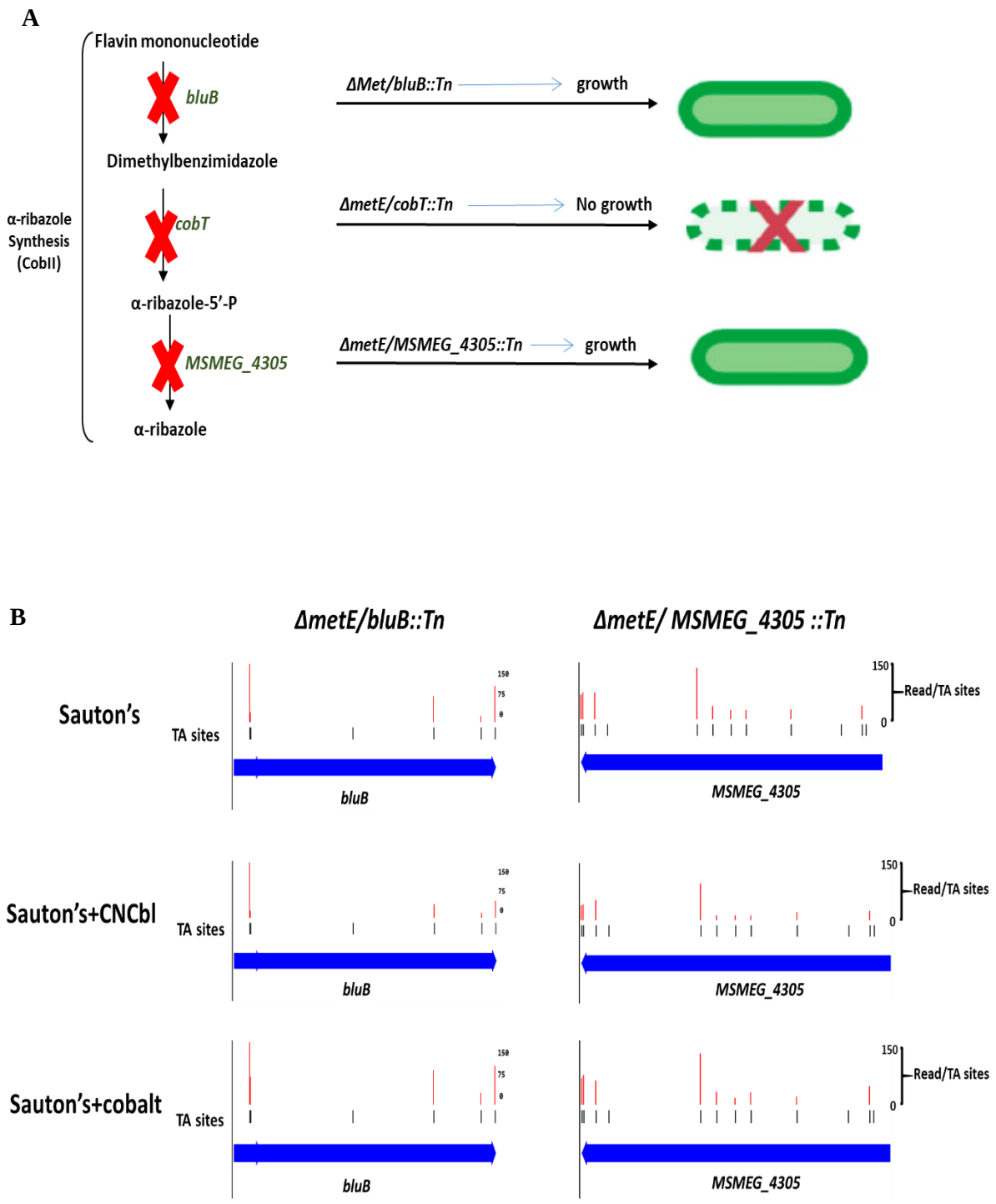


Figure 4.5: Comparison of Tn insertion frequencies among genes predicted to function in α-ribazole biosynthesis.

5 Conclusion

Cobamide biosynthetic genes and cobamide-dependent enzymes are widespread among mycobacteria. However, most mycobacteria encode partial biosynthesis pathways which are associated with corrionoid uptake and assimilations. In nature, cobamides are synthesized *de novo* using one of two alternative pathways: the aerobic or anaerobic pathway. While the near-complete cobalamin biosynthetic pathway appears predominantly of the aerobic type, there are some features that are characteristic of anaerobic biosynthesis – perhaps implying that mycobacteria might be able to partially use the anaerobic pathway (Gopinath, Moosa, et al., 2013), a possibility that requires further investigation.

Cobamides fulfil central roles in cellular metabolism as cofactors, enzyme subnstrates, and/or gene regulators in different essential biochemical processes including the biosynthesis of deoxribonucleotides and methionine and the metabolism of propionyl-CoA, a by-product of β-oxidation of cholesterol and odd-chain fatty acids (Young et al., 2015). This study focused on elucidating the predicted cobalamin biosynthetic genes, including putative cobamide and cobalt transporters, in the non-pathogenic *Msm*. Although in silico predictions had yielded candidate genes for most predicted biosynthetic steps, the combination of high-density mutagenesis and deep-sequencing applied here enabled a functional genomics study of the genetic requirements for cobamide biosynthesis during growth of *Msm* in different conditions. Based on the data presented here, both *de novo* cobalamin biosynthesis and cobamide uptake/assimilation pathways are functional in *Msm*. This is an important result as it suggests the potential to apply an analogous approach in *Mtb*. Elucidating the genetic requirements for cobalamin biosynthesis, uptake and assimilation in mycobacteria for optimal growth under specific conditions can inform our basic understanding of mycobacterial physiology and pathogenicity, identifying potential vulnerabilities for novel anti-TB therapeutics.

The functionality of aerobic cobalamin pathway in *Msm* reinforces the potential importance of cobamide in mycobacterial physiology. The data further demonstrate that biosynthesis of methionine via cobalamin-independent enzyme, MetE, is under the control of a cobalamin riboswitch; this was confirmed by the observation that *metH*, encoding the cobalamin-dependent methionine synthase, was required for growth in media lacking methionine. The inferred conditional essentiality was consistent with recent work demonstrating the role of cobalamin riboswitch-mediated repression in *metH* essentiality in *Msm* (Kipkorir et al., 2021).

Hypothetically, the cobalamin riboswitch-mediated repression of *metE* and the conditional essentiality of *metH* indicate that *Msm* synthesizes AdoCbl and MeCbl, both having DMB as the lower ligand. However, additional work is required to determine whether *Msm* possesses the ability to synthesize pseudocobalamin or whether pseudocobalamin might be a relevant cofactor in other mycobacteria.

Other areas for future study include the apparent redundancy in cobamide transporters in the environmental mycobacterium. As noted recently (Kipkorir et al., 2021), elucidating the contributions of these putative cobamide transporters might provide useful insight into the genetic factors which have shaped the adaptation of mycobacteria to environments (including intracellular) that might be characterized by very different cobamide availabilities. The inferred essentiality of *cobQ2* also requires further investigation, with its predicted role as GatD-type amidase (Pavelka Jr et al., 2014) offering a more plausible rationale for the phenotype observed in this study, and in morphological analyses of CRISPRi knockdown mutants (de Wet et al., 2020).

The role(s) of cobamides in mycobacterial pathogenicity remain uncertain. However, by providing insight into the genes involved in the pathway for *de novo* cobalamin biosynthesis in *Msm*, the data contained in this thesis represent a valuable addition to our current understanding of these pathways in pathogenic and non-pathogenic mycobacteria, and should provide a useful foil against which equivalent investigations in *Mtb* might be considered.

6 Appendices: Culture media

All media are made up to a final volume of 1L with deionised water, and sterilised by autoclaving at 121^{o}C for 15min, unless otherwise stated.

Luria-Bertani Broth (LB)

5g yeast, 10g tryptone, 10g sodium chloride

Luria-Bertani Agar (LA)

5g yeast, 10g tryptone, 10g sodium chloride, 15g agar

2TY

5g sodium chloride, 10g yeast extract, 16g tryptone

Middlebrook 7H9

4.7g DifcoTM Middlebrook 7H9 broth, 2ml glycerol, 100ml OADC supplement and 0.05% Tween 80 added after autoclaving

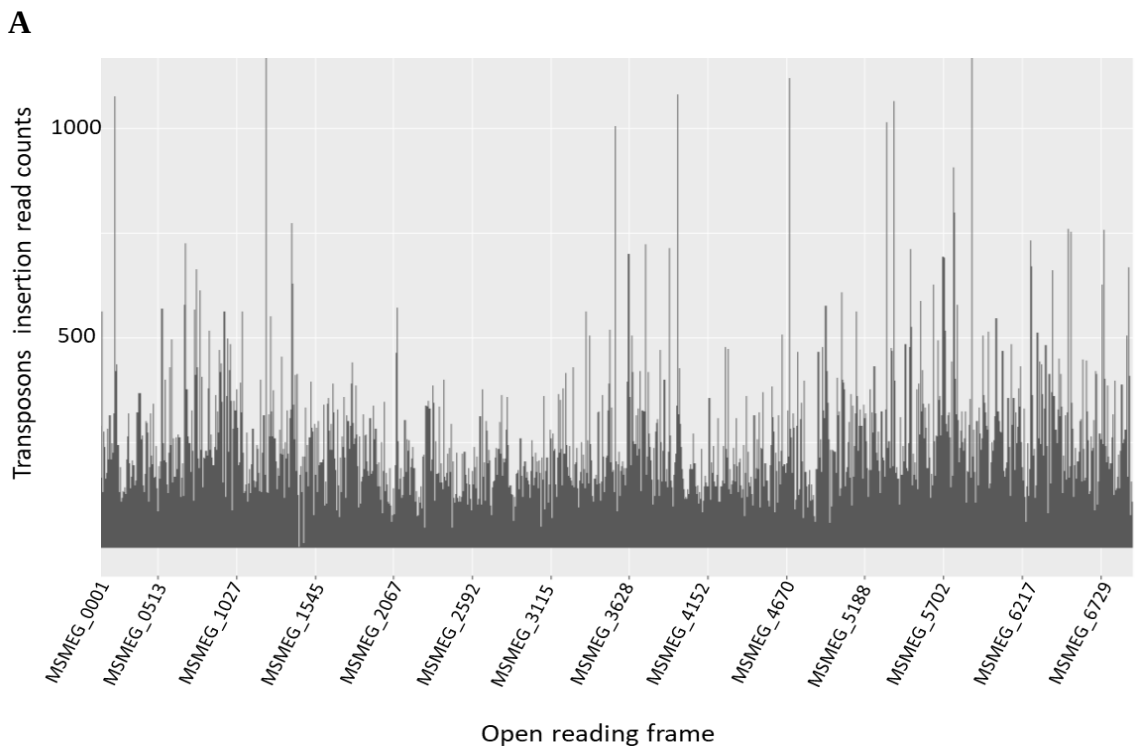
Middlebrook 7H10

19g DifcoTM Middlebrook 7H10 agar, 5ml glycerol, 100ml OADC supplement and 0.05% Tween 80 added after autoclaving

Sauton's minimal medium (pH 7.2)

4g asparagine, 0.5g magnesium sulphate, 2g citric acid, 0.5g potassium dihydrogen orthophosphate, 0.05g ammonium ferric citrate, 60ml glycerol. Sterilised by filtration.

7 Supplementary information



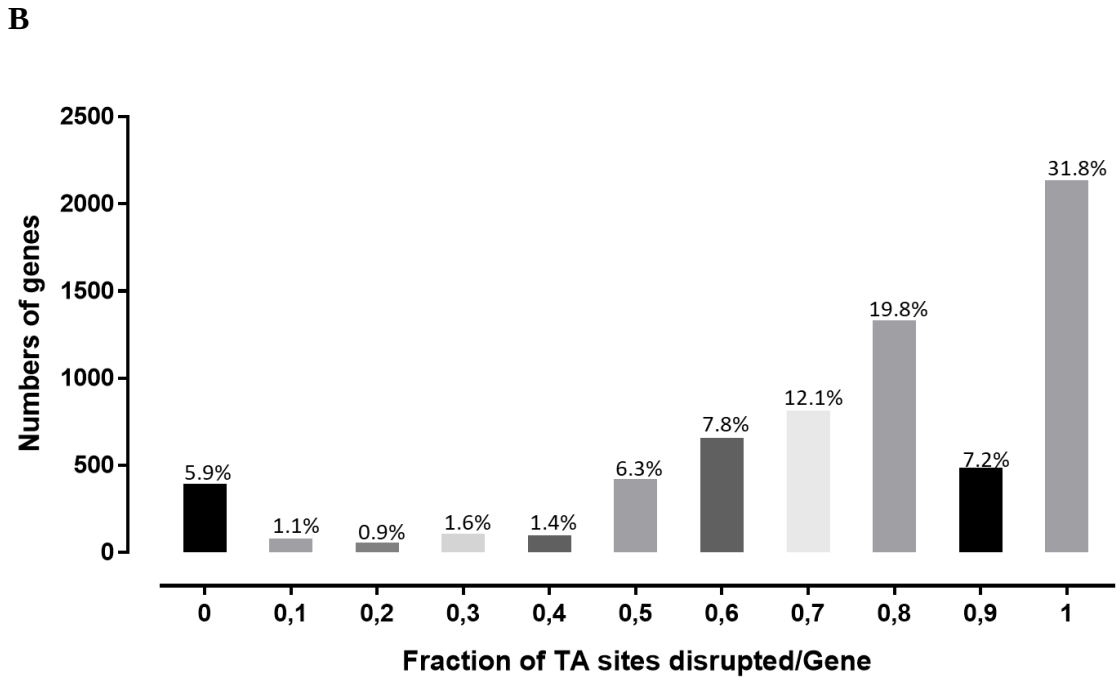


Figure S1: The insertion profile of the WT *Msm* Tn library. (**A**) Tn insertion counts across the WT *Msm* genome (7H10-selected library). The height of the black bars represents the number of insertions at each individual -TA- site. (**B)** For each gene, the fraction of disrupted -TA- sites was calculated based on the average read coverage. Displayed are the numbers of genes carrying Tn insertions in none (0) to all (1) -TA- sites.

Table S1: Tn sequencing analysis of *Msm* genes required for growth *in vitro*. A list of genes of *Msm* classified as essential and growth defect, the essentiality analysis was obtained by applying the HMM Method incorporated into the software TRANSIT

Accession no.	Gene name	Predicted product/function	Essentiality call in Δ*metE*	Essentiality call WT (this study)	Essentiality call WT (Dragset et al. (2019))	Mtb ortholog	Essentiality call in Mtb DeJesus et al. (2017)
MSMEG_0001	*dnaN*	DNA polymerase III subunit beta	ES	ES	ES	Rv0002	ES
MSMEG_0005	*gyrB*	DNA gyrase subunit B	GD	GD	GD	Rv0005	ES
MSMEG_0006	*gyrA*	DNA gyrase subunit A	ES	ES	ES	Rv0006	ES
MSMEG_0028	*pknB*	serine-threonine protein kinase	ES	GD	GD	Rv0014c	ES
MSMEG_0030	*pknA*	serine/threonine protein kinase PknA	ES	GD	ES	Rv0015c	ES
MSMEG_0250	*mmpL3*	membrane protein, MmpL family protein	ES	ES	ES	Rv0206c	ESD
MSMEG_0311	-	glycosyltransferase	ES	ES	ES	Rv0225	ES
MSMEG_0315	-	transmembrane protein	ES	GD	ES	Rv0226c	ES
MSMEG_0317	-	hypothetical protein MSMEG_0317	ES	ES	ES	Rv0227c	ES
MSMEG_0319	-	acyltransferase	ES	ES	ES	Rv0228	ES
MSMEG_0359	*aftD*	hypothetical protein MSMEG_0359	ES	ES	ES	Rv0236c	ES
MSMEG_0384	*rfbA*	glucose-1-phosphate thymidylyltransferase	ES	ES	ES	Rv0334	ES
MSMEG_0690	-	iron-sulfur cluster-binding protein	GD	GD	GD	Rv0338c	ES
MSMEG_0879	-	hypothetical protein MSMEG_0879	ES	ES	ES	-	?
MSMEG_0880	*groEL*	chaperonin GroEL	ES	ES	ES	Rv0440	ES
MSMEG_0903	*lpdA*	dihydrolipoamide dehydrogenase	ES	ES	ES	Rv0462	GD
MSMEG_0937	*regX3*	DNA-binding response regulator RegX3	ES	ES	ES	Rv0491	NE
MSMEG_0952	*hemA*	glutamyl-tRNA reductase	ES	ES	ES	Rv0509	ES
MSMEG_0953	*hemC*	porphobilinogen deaminase	ES	ES	ES	Rv0510	ES
MSMEG_0954	*hemD*	uroporphyrinogen-III synthase	ES	ES	ES	Rv0511	GD
MSMEG_0956	*hemB*	delta-aminolevulinic acid dehydratase	GD	ES	ES	Rv0512	ES
MSMEG_0969	*hemL*	glutamate-1-semialdehyde aminotransferase	GD	ES	GD	Rv0524	ES

MSMEG_0970	-	phosphoglycerate mutase family protein	GD	ES	GD	Rv0525	ES
MSMEG_0988c	*menA*	1,4-dihydroxy-2-naphthoate octaprenyltransferase	GD	ES	ES	Rv0534c	ES
MSMEG_1109	*menD*	2-succinyl-6-hydroxy-2,4-cyclohexadiene-1-carboxylic acid synthase/2-oxoglutarate decarboxylase	GD	ES	ES	Rv0555	ES
MSMEG_1133	*grcC1*	bifunctional short chain isoprenyl diphosphate synthase	ES	ES	ES	Rv0562	ES
MSMEG_1344	*secE*	preprotein translocase subunit SecE	ES	ES	GD	Rv0638	ES
MSMEG_1345	*nusG*	transcription antitermination protein NusG	ES	ES	GD	Rv0639	ES
MSMEG_1367	*rpoB*	DNA-directed RNA polymerase subunit beta	ES	ES	ES	Rv0667	ES
MSMEG_1368	*rpoC*	DNA-directed RNA polymerase subunit beta'	ES	ES	ES	Rv0668	ES
MSMEG_1398	*rpsL*	30S ribosomal protein S12	ES	ES	ES	Rv0682	ES
MSMEG_1399	*rpsG*	30S ribosomal protein S7	ES	ES	ES	Rv0683	ES
MSMEG_1401	*tuf*	elongation factor Tu	ES	GD	GD	Rv0685	ES
MSMEG_1436	*rplC*	50S ribosomal protein L3	ES	ES	ES	Rv0701	ES
MSMEG_1437	*rplD*	50S ribosomal protein L4	ES	ES	ES	Rv0702	ES
MSMEG_1438	*rplW*	50S ribosomal protein L23	ES	ES	ES	Rv0703	ES
MSMEG_1439	*rplB*	50S ribosomal protein L2	ES	ES	ES	Rv0704	ES
MSMEG_1441	*rplV*	50S ribosomal protein L22	ES	ES	ES	Rv0706	ES
MSMEG_1442	*rpsC*	30S ribosomal protein S3	ES	ES	ES	Rv0707	ES
MSMEG_1443	*rplP*	50S ribosomal protein L16	ES	ES	ES	Rv0708	ES
MSMEG_1444	*rpmC*	50S ribosomal protein L29	ES	ES	ES	Rv0709	GD
MSMEG_1445	*rpsQ*	30S ribosomal protein S17	ES	ES	ES	Rv0710	ES
MSMEG_1465	*rplN*	ribosomal protein L14	ES	GD	ES	Rv0714	ES
MSMEG_1466	*rplX*	50S ribosomal protein L24	ES	GD	ES	Rv0715	ES
MSMEG_1467	*rplE*	50S ribosomal protein L5	ES	GD	ES	Rv0716	ES
MSMEG_1468	*rpsN*	ribosomal protein S14p/S29e	ES	GD	ES	Rv0717	Uncertain

MSMEG_1469	*rpsH*	30S ribosomal protein S8	ES	GD	ES	Rv0718	ES
MSMEG_1470	*rplF*	50S ribosomal protein L6	ES	ES	ES	Rv0719	ES
MSMEG_1471	*rplR*	50S ribosomal protein L18	ES	ES	ES	Rv0720	ES
MSMEG_1472	*rpsE*	30S ribosomal protein S5	ES	ES	ES	Rv0721	ES
MSMEG_1473	*rpmD*	ribosomal protein L30	ES	ES	ES	Rv0722	Uncertain
MSMEG_1483	*secY*	preprotein translocase subunit SecY	ES	ES	ES	Rv0732	ES
MSMEG_1510	*rmlC*	dTDP-4-dehydrorhamnose 3,5-epimerase	ES	GD	ES	Rv3465	ES
MSMEG_1512	*rfbB*	dTDP-glucose 4,6-dehydratase	ES	GD	ES	Rv3464	ES
MSMEG_1521	*rpsM*	30S ribosomal protein S13	GD	ES	GD	Rv3460c	ES
MSMEG_1522	*rpsK*	30S ribosomal protein S11	GD	ES	GD	Rv3459c	Uncertain
MSMEG_1523	*rpsD*	30S ribosomal protein S4	GD	ES	GD	Rv3458c	ES
MSMEG_1524	*rpoA*	DNA-directed RNA polymerase subunit alpha	ES	ES	ES	Rv3457c	ES
MSMEG_1556	*rplM*	50S ribosomal protein L13	ES	ES	ES	Rv3443c	ES
MSMEG_1557	*rpsI*	ribosomal protein S9	ES	ES	ES	Rv3442c	ES
MSMEG_1568	*glmS*	D-fructose-6-phosphate amidotransferase	ES	ES	ES	Rv3436c	ES
MSMEG_1602	*guaB*	inositol-5-monophosphate dehydrogenase	ES	ES	ES	Rv3411c	ES
MSMEG_1657	*trpS*	tryptophanyl-tRNA synthetase	ES	ES	ES	Rv3336c	ES
MSMEG_1669	*sdhB*	succinate dehydrogenase iron-sulfur subunit	ES	ES	ES	Rv3319	NE
MSMEG_1670	*sdhA*	succinate dehydrogenase flavoprotein subunit	ES	ES	ES	Rv3318	NE
MSMEG_1671	*sdhD*	succinate dehydrogenase hydrophobic membrane anchor protein SdhD	ES	ES	ES	Rv3317	NE
MSMEG_1672	*sdhC*	succinate dehydrogenase, cytochrome b556 subunit	ES	ES	ES	Rv3316	NE
MSMEG_1813	*accD5*	propionyl-CoA carboxylase beta chain	ES	GD	ES	Rv3280	ES
MSMEG_1825	*rfbD*	dTDP-4-dehydrorhamnose reductase	GD	GD	GD	Rv3266c	ES
MSMEG_1826	*wbbL1*	dTDP-RhA:a-D-GlcNAc-diphosphoryl polyprenol, a-3-L-rhamnosyl transferase	ES	ES	ES	Rv3265c	ES
MSMEG_1834	*manB*	phosphomannomutase/phosphoglucomutase	ES	ES	ES	Rv3257c	ES

MSMEG_1836	*manA*	mannose-6-phosphate isomerase, class I	ES	ES	GD	Rv3255c	ES
MSMEG_1868	-	hypothetical protein MSMEG_1868	ES	ES	ES	-	?
MSMEG_1873	*tmk*	thymidylate kinase	ES	ES	ES	Rv3247c	ES
MSMEG_1881	*secA*	preprotein translocase subunit SecA	ES	ES	ES	Rv3240c	ESD
MSMEG_1952	*uvrD2*	ATP-dependent DNA helicase	ES	ES	ES	Rv3198c	ESD
MSMEG_2086	*prfB*	peptide chain release factor 2	GD	ES	GD	Rv3105c	ES
MSMEG_2362	*ligA*	NAD-dependent DNA ligase LigA	ES	ES	ES	Rv3014c	ES
MSMEG_2365	*gatA*	aspartyl/glutamyl-tRNA amidotransferase subunit A	ES	ES	ES	Rv3011c	ES
MSMEG_2367	*gatB*	aspartyl/glutamyl-tRNA amidotransferase subunit B	ES	ES	ES	Rv3009c	ES
MSMEG_2383	*gltX*	glutamyl-tRNA synthetase	ES	ES	ES	Rv2992c	ES
MSMEG_2395	*ddl*	D-alanyl-alanine synthetase A	ES	ES	ES	Rv2981c	ES
MSMEG_2410	-	putative serine-threonine protein kinase	ES	ES	ES	Rv2969c	ES
MSMEG_2411	-	integral membrane protein	ES	ES	ES	Rv2968c	ES
MSMEG_2441	*lepB*	signal peptidase I	ES	ES	GD	Rv2903c	ES
MSMEG_2519	*rpsB*	30S ribosomal protein S2	ES	ES	ES	Rv2890c	ES
MSMEG_2520	*tsf*	elongation factor Ts	ES	ES	ES	Rv2889c	ES
MSMEG_2540	*pyrH*	uridylate kinase	ES	GD	ES	Rv2883c	ES
MSMEG_2541	*frr*	ribosome recycling factor	ES	GD	ES	Rv2882c	ES
MSMEG_2543	*cdsA*	phosphatidate cytidylyltransferase	ES	GD	ES	Rv2881c	ES
MSMEG_2578	*dxr*	1-deoxy-D-xylulose 5-phosphate reductoisomerase	ES	ES	ES	Rv2870c	ES
MSMEG_2613	*mqo*	malate:quinone oxidoreductase	GD	GD	ES	Rv2852c	NE
MSMEG_2621	*proS*	prolyl-tRNA synthetase	ES	ES	ES	Rv2845c	ES
MSMEG_2627	-	hypothetical protein MSMEG_2627	ES	ES	ES	-	?
MSMEG_2628	*infB*	translation initiation factor IF-2	ES	ES	ES	Rv2839c	ES
MSMEG_2653	*ribF*	bifunctional riboflavin kinase/FMN adenylyltransferase	ES	ES	ES	Rv2786c	ES

MSMEG_2656	*gpsI*	polynucleotide phosphorylase/polyadenylase	ES	ES	ES	Rv2783c	ES
MSMEG_2670	*thyA*	thymidylate synthase	ES	ES	GD	Rv2764c	GD
MSMEG_2684	*dapA*	dihydrodipicolinate synthase	ES	ES	GD	Rv2753c	ES
MSMEG_2690	*ftsK*	DNA translocase FtsK	ES	ES	ES	Rv2748c	ES
MSMEG_2758	*sigA*	RNA polymerase sigma factor	ES	ES	ES	Rv2703	ES
MSMEG_2776	*dxs*	1-deoxy-D-xylulose-5-phosphate synthase	ES	ES	ES	Rv2682c	ES
MSMEG_2780	*hemE*	uroporphyrinogen decarboxylase	GD	ES	ES	Rv2678c	ES
MSMEG_2781	*hemG*	protoporphyrinogen oxidase	GD	ES	ES	Rv2677c	ES
MSMEG_2782	-	hypothetical protein MSMEG_2782	GD	GD	ES	Rv2676c	ES
MSMEG_2785	*aftC*	putative integral membrane protein	ES	ES	ES	Rv2673	ES
MSMEG_2931	*thrS*	threonyl-tRNA synthetase	ES	ES	ES	Rv2614c	ES
MSMEG_2935	*pimA*	phosphatidylinositol alpha-mannosyltransferase	ES	ES	ES	Rv2610c	ES
MSMEG_2961	*secD*	preprotein translocase subunit SecD	GD	ES	ES	Rv2587c	ES
MSMEG_2976	*hisS*	histidyl-tRNA synthetase	ES	ES	ES	Rv2580c	ES
MSMEG_3003	*aspS*	aspartyl-tRNA synthetase	ES	ES	ES	Rv2572c	ES
MSMEG_3025	*alaS*	alanyl-tRNA synthetase	ES	ES	ES	Rv2555c	ES
MSMEG_3030	*aroC*	chorismate synthase	GD	ES	ES	Rv2540c	ES
MSMEG_3031	*aroK*	shikimate kinase	GD	ES	ES	Rv2539c	ES
MSMEG_3033	*aroB*	3-dehydroquinate synthase	GD	ES	ES	Rv2538c	ES
MSMEG_3054	*coaBC*	bifunctional phosphopantothenoylcysteine decarboxylase/phosphopantothenate synthase	ES	ES	GD	Rv1391	ES
MSMEG_3055	*metK*	S-adenosylmethionine synthetase	ES	ES	ES	Rv1392	ES
MSMEG_3061	*priA*	primosome assembly protein PriA	ES	ES	ES	Rv1402	ES
MSMEG_3072	*ribA2*	bifunctional 3,4-dihydroxy-2-butanone 4-phosphate synthase/GTP cyclohydrolase II protein	ES	ES	ES	Rv1415	ES
MSMEG_3081	-	hypothetical protein MSMEG_3081	ES	ES	ES	Rv1423	ES

MSMEG_3120	-	hypothetical protein MSMEG_3120	ES	ES	ES	Rv1459c	ESD
MSMEG_3122	*sufB*	FeS assembly protein SufB	ES	ES	ES	Rv1461	ES
MSMEG_3123	*sufD*	FeS assembly protein SufD	GD	ES	ES	Rv1462	ES
MSMEG_3124	*sufC*	FeS assembly ATPase SufC	GD	ES	ES	Rv1463	ES
MSMEG_3143	*acnA*	aconitate hydratase 1	ES	ES	ES	Rv1475c	ES
MSMEG_3150	*fabG*	3-oxoacyl-(acyl-carrier-protein) reductase	ES	ES	ES	Rv1483	ES
MSMEG_3151	*inhA*	enoyl-(acyl carrier protein) reductase	ES	ES	ES	Rv1484	ES
MSMEG_3152	*hemH*	ferrochelatase	ES	ES	ES	Rv1485	ES
MSMEG_3169	*ileS*	isoleucyl-tRNA synthetase	ES	ES	ES	Rv1536	ES
MSMEG_3178	*dnaE*	DNA polymerase III subunit alpha	ES	ES	ES	Rv1547	ES
MSMEG_3213	-	hypothetical protein MSMEG_3213	ES	ES	ES	Rv3263	NE
MSMEG_3217	*trpE*	anthranilate synthase component I	GD	ES	ES	Rv1609	ES
MSMEG_3219	*trpC*	indole-3-glycerol-phosphate synthase	GD	ES	GD	Rv1611	ES
MSMEG_3220	*trpB*	tryptophan synthase subunit beta	GD	ES	ES	Rv1612	ES
MSMEG_3221	*trpA*	tryptophan synthase subunit alpha	GD	ES	ES	Rv1613	ES
MSMEG_3621	*ndh*	NADH dehydrogenase	ES	ES	ES	Rv1854c	GD
MSMEG_3746	*pyrG*	CTP synthetase	ES	ES	ES	Rv1699	ES
MSMEG_3758	*tyrS*	tyrosyl-tRNA synthetase	ES	ES	ES	Rv1689	ES
MSMEG_3777	*pheT*	phenylalanyl-tRNA synthetase subunit beta	ES	ES	ES	Rv1650	ES
MSMEG_3778	*pheS*	phenylalanyl-tRNA synthetase subunit alpha	ES	ES	ES	Rv1649	ES
MSMEG_3833	*rpsA*	30S ribosomal protein S1	GD	ES	ES	Rv1630	ES
MSMEG_3859	-	glycosyl transferase, group 2 family protein	GD	ES	ES	-	?
MSMEG_4219	-	hypothetical protein MSMEG_4219	ES	ES	ES	Rv2147c	ES
MSMEG_4225	*ftsQ*	putative cell division protein FtsQ	ES	ES	ES	Rv2151c	ES
MSMEG_4226	*murC*	UDP-N-acetylmuramate--L-alanine ligase	ES	ES	ES	Rv2152c	ES
MSMEG_4227	*murG*	N-acetylglucosaminyl transferase	ES	ES	ES	Rv2153c	ES
MSMEG_4228	*ftsW*	cell division protein FtsW	ES	ES	ES	Rv2154c	ESD

MSMEG_4229	*murD*	UDP-N-acetylmuramoyl-L-alanyl-D-glutamate synthetase	ES	ES	ES	Rv2155c	ES
MSMEG_4230	*mraY*	phospho-N-acetylmuramoyl-pentapeptide-transferase	ES	ES	ES	Rv2156c	ES
MSMEG_4231	*murF*	UDP-N-acetylmuramoyl-tripeptide--D-alanyl-D-alanine ligase	ES	ES	ES	Rv2157c	ES
MSMEG_4232	*murE*	UDP-N-acetylmuramoylalanyl-D-glutamate--2,6-diaminopimelate ligase	ES	ES	ES	Rv2158c	ES
MSMEG_4233	*pbpB*	penicillin binding protein transpeptidase domain-containing protein	ES	ES	ES	Rv2163c	ESD
MSMEG_4234	-	hypothetical protein MSMEG_4234	ES	ES	ES	Rv2164c	GD
MSMEG_4253	-	glycosyl transferase, group 1 family protein	ES	GD	ES	Rv2188c	ES
MSMEG_4260	*ctaE*	cytochrome c oxidase subunit 3	GD	GD	GD	Rv2193	GD
MSMEG_4261	*qcrC*	ubiquinol-cytochrome c reductase cytochrome c subunit	GD	GD	GD	Rv2194	GD
MSMEG_4262	*qcrA*	ubiquinol-cytochrome c reductase iron-sulfur subunit	GD	GD	GD	Rv2195	GD
MSMEG_4263	*qcrB*	ubiquinol-cytochrome c reductase cytochrome b subunit	GD	GD	GD	Rv2196	GD
MSMEG_4283	*sucB*	dihydrolipoamide acetyltransferase	ES	ES	ES	Rv2215	NE
MSMEG_4286	*lipA*	lipoyl synthase	ES	GD	ES	Rv2218	ES
MSMEG_4290	*glnA*	glutamine synthetase, type I	ES	ES	ES	Rv2220	ES
MSMEG_4294	*glnA*	glutamine synthetase, type I	ES	ES	ES	Rv2222c	NE
MSMEG_4326	*acpP*	acyl carrier protein	ES	ES	ES	Rv2244	ES
MSMEG_4327	*kasA*	3-oxoacyl-(acyl carrier protein) synthase II	ES	ES	ES	Rv2245	ES
MSMEG_4482	*dnaG*	DNA primase	ES	ES	GD	Rv2343c	ES
MSMEG_4485	*glyS*	glycyl-tRNA synthetase	ES	GD	ES	Rv2357c	ES
MSMEG_4490	*uppS*	undecaprenyl diphosphate synthase	ES	GD	ES	Rv2361c	ES
MSMEG_4617	*nadE*	NAD synthetase	ES	ES	ES	Rv2438c	ES
MSMEG_4623	*obgE*	GTPase ObgE	ES	ES	GD	Rv2440c	ES
MSMEG_4626	*rne*	ribonuclease, Rne/Rng family protein	ES	ES	ES	Rv2444c	ESD

MSMEG_4630	*valS*	valyl-tRNA synthetase	ES	ES	ES	Rv2448c	ES
MSMEG_4671	*clpX*	ATP-dependent protease ATP-binding subunit	ES	ES	ES	Rv2457c	ES
MSMEG_4672	*clpP*	ATP-dependent Clp protease proteolytic subunit	ES	ES	ES	Rv2460c	ES
MSMEG_4673	*clpP*	ATP-dependent Clp protease proteolytic subunit	ES	ES	ES	Rv2461c	ES
MSMEG_4700	-	putative ABC transporter ATP-binding protein	ES	GD	ES	Rv2477c	ES
MSMEG_4757	*fas*	fatty acid synthase	ES	ES	ES	Rv2524c	ES
MSMEG_4947	-	glycosyl trasferase	ES	ES	ES	Rv1302	ES
MSMEG_4950	*prfA*	peptide chain release factor 1	ES	ES	ES	Rv1299	ES
MSMEG_4954	*rho*	transcription termination factor Rho	ES	ES	ES	Rv1297	ES
MSMEG_4955	*thrB*	homoserine kinase	GD	ES	ES	Rv1296	ES
MSMEG_4956	*thrC*	threonine synthase	GD	ES	ES	Rv1295	ES
MSMEG_4957	*thrA*	homoserine dehydrogenase	GD	ES	ES	Rv1294	ES
MSMEG_4959	*argS*	arginyl-tRNA synthetase	ES	ES	ES	Rv1292	ES
MSMEG_5049	*kgd*	alpha-ketoglutarate decarboxylase	ES	GD	GD	Rv1248c	NE
MSMEG_5103	*dapE*	dipeptidase	ES	ES	ES	Rv1202	ES
MSMEG_5252	*coaA*	pantothenate kinase	ES	ES	ES	Rv1092c	ES
MSMEG_5413	-	exopolyphosphatase	GD	GD	GD	Rv1026	NE
MSMEG_5414	-	septum formation initiator subfamily protein, putative	GD	GD	GD	Rv1025	ES
MSMEG_5415	*eno*	phosphopyruvate hydratase	GD	ES	ES	Rv1023	GD
MSMEG_5426	*glmU*	UDP-N-acetylglucosamine pyrophosphorylase	GD	ES	ES	Rv1018c	ES
MSMEG_5435	-	acyl-CoA synthetase	ES	ES	GD	Rv1013	NE
MSMEG_5441	*metG*	methionyl-tRNA synthetase	ES	ES	ES	Rv1007c	ES
MSMEG_5773	*desA1*	fatty acid desaturase	ES	ES	ES	Rv0824c	ES
MSMEG_5847	*purB*	adenylosuccinate lyase	ES	ES	ES	Rv0777	ES

MSMEG_6074	*cysS*	cysteinyl-tRNA synthetase	ES	ES	ES	Rv3580c	ES
MSMEG_6075	*ispF*	2-C-methyl-D-erythritol 2,4-cyclodiphosphate synthase	ES	ES	ES	Rv3581c	ES
MSMEG_6076	*ispD*	2-C-methyl-D-erythritol 4-phosphate cytidylyltransferase	ES	ES	ES	Rv3582c	ES
MSMEG_6077	-	transcriptional regulator, CarD family protein	ES	ES	ES	Rv3583c	GD
MSMEG_6081	-	hypothetical protein MSMEG_6081	GD	ES	GD	Rv3587c	ES
MSMEG_6091	-	negative regulator of genetic competence ClpC/mecB	ES	GD	ES	Rv3596c	ES
MSMEG_6094	*lysS*	lysyl-tRNA synthetase	ES	ES	ES	Rv3598c	ES
MSMEG_6101	*folK*	2-amino-4-hydroxy-6-hydroxymethyldihydropteridine pyrophosphokinase	ES	ES	ES	Rv3606c	GD
MSMEG_6102	*folB*	dihydroneopterin aldolase	ES	ES	ES	Rv3607c	GD
MSMEG_6103	*folP*	dihydropteroate synthase	ES	ES	ES	Rv3608c	ESD
MSMEG_6104	*folE*	GTP cyclohydrolase I	ES	ES	ES	Rv3609c	ES
MSMEG_6157	*topA*	DNA topoisomerase I	ES	ES	ES	Rv3646c	ES
MSMEG_6184	-	hydrolase, alpha/beta fold family protein	ES	GD	ES	Rv3670	NE
MSMEG_6256	*asd*	aspartate-semialdehyde dehydrogenase	ES	ES	ES	Rv3708c	GD
MSMEG_6257	*ask*	aspartate kinase	ES	ES	ES	Rv3709c	ESD
MSMEG_6276	-	mur ligase family protein	ES	ES	ES	Rv3712	ES
MSMEG_6277	*cobQ2*	cobyric acid synthase	ES	ES	ES	Rv3713	ES
MSMEG_6285	*dnaZX*	DNA polymerase III subunits gamma and tau	ES	ES	ES	Rv3721c	ES
MSMEG_6286	-	aspartate transaminase	GD	ES	ES	Rv3722c	ES
MSMEG_6366	*rfbE*	O-antigen export system, ATP-binding protein	ES	ES	ES	Rv3781	ES
MSMEG_6367	-	glycosyl transferase, group 2 family protein	ES	ES	ES	Rv3782	ES
MSMEG_6369	*rfbD*	O-antigen export system, permease protein	ES	ES	ES	Rv3783	ES
MSMEG_6382	-	oxidoreductase, FAD-binding	ES	ES	ES	Rv3790	ES

MSMEG_6386	-	membrane protein	ES	ES	ES	Rv3792	ES
MSMEG_6389	*embB*	arabinosyltransferase A	GD	GD	ES	Rv3795	ES
MSMEG_6391	*accD4*	propionyl-CoA carboxylase beta chain	ES	ES	ES	Rv3799c	ES
MSMEG_6392	*pks13*	polyketide synthase	ES	ES	ES	Rv3800c	ES
MSMEG_6393	-	acyl-CoA synthetase	ES	GD	ES	Rv3801c	ES
MSMEG_6400	-	transmembrane protein	GD	GD	ES	Rv3805c	ES
MSMEG_6401	*ubiA*	phosphoribose diphosphate:decaprenyl-phosphate phosphoribosyltransferase	ES	ES	ES	Rv3806c	ES
MSMEG_6403	*glfT2*	bifunctional udp-galactofuranosyl transferase *glfT*	ES	ES	ES	Rv3808c	ES
MSMEG_6404	*glf*	UDP-galactopyranose mutase	ES	ES	ES	Rv3809c	ES
MSMEG_6413	*serS*	seryl-tRNA synthetase	ES	ES	ES	Rv3834c	ES
MSMEG_6892	*dnaB*	replicative DNA helicase	ES	ES	ES	Rv0058	ES
MSMEG_6900	*ponA1*	penicillin-binding protein 1	ES	ES	ES	Rv0050	NE
MSMEG_6904	*ino1*	myo-inositol-1-phosphate synthase	GD	GD	ES	Rv0046c	ES
MSMEG_6917	*leuS*	leucyl-tRNA synthetase	ES	ES	ES	Rv0041	ES
MSMEG_6926	-	tRNA adenylyltransferase	GD	ES	ES	Rv3907c	ES
MSMEG_6928	-	hypothetical protein MSMEG_6928	ES	ES	ES	Rv3909	ES
MSMEG_6929	-	integral membrane protein MviN, putative	ES	ES	ES	Rv3910	ESD
MSMEG_6935	-	N-acetylmuramoyl-L-alanine amidase	ES	ES	ES	Rv3915	ES
MSMEG_6942	-	putative inner membrane protein translocase component YidC	GD	ES	ES	Rv3921c	ESD
MSMEG_6947	*dnaA*	chromosomal replication initiation protein	ES	ES	ES	Rv0001	ES

Table S2: List of essential genes only in Δ*metE* (rich-selected library) when compared to WT libraries (7H10-selected)

Accession no.	Gene name	Predicted product/function	Essentiality call
MSMEG_0936	*senX3*	sensor histidine kinase SenX3	ES
MSMEG_0274	-	hypothetical protein MSMEG_0274	GD
MSMEG_1241	-	hypothetical protein MSMEG_1241	ES
MSMEG_2734	*miaA*	tRNA delta(2)-isopentenylpyrophosphate transferase	GD
MSMEG_2733	-	hypothetical protein MSMEG_2733	GD
MSMEG_2735	*dapF*	diaminopimelate epimerase	GD
MSMEG_3080	-	hypothetical protein MSMEG_3080	ES
MSMEG_3646	-	transcriptional regulator, MerR family protein	ES
MSMEG_3647	*garA*	forkhead-associated protein	ES
MSMEG_3708	-	catalase	ES
MSMEG_4311	-	hypothetical protein MSMEG_4311	ES
MSMEG_4541	-	ABC transporter, ATP-binding protein	ES
MSMEG_4877	-	CaiB/BaiF family protein	GD
MSMEG_4878	-	2Fe-2S iron-sulfur cluster binding domain-containing protein	GD
MSMEG_4916	*glgE*	alpha-amylase family protein	GD
MSMEG_4918	*glgB*	glycogen branching enzyme	GD
MSMEG_5621	-	putative acyl-CoA dehydrogenase	ES
MSMEG_5622	-	putative acyl-CoA dehydrogenase	ES
MSMEG_6638	*metE*	5-methyltetrahydropteroyltriglutamate--homocysteine methyltransferase	GD

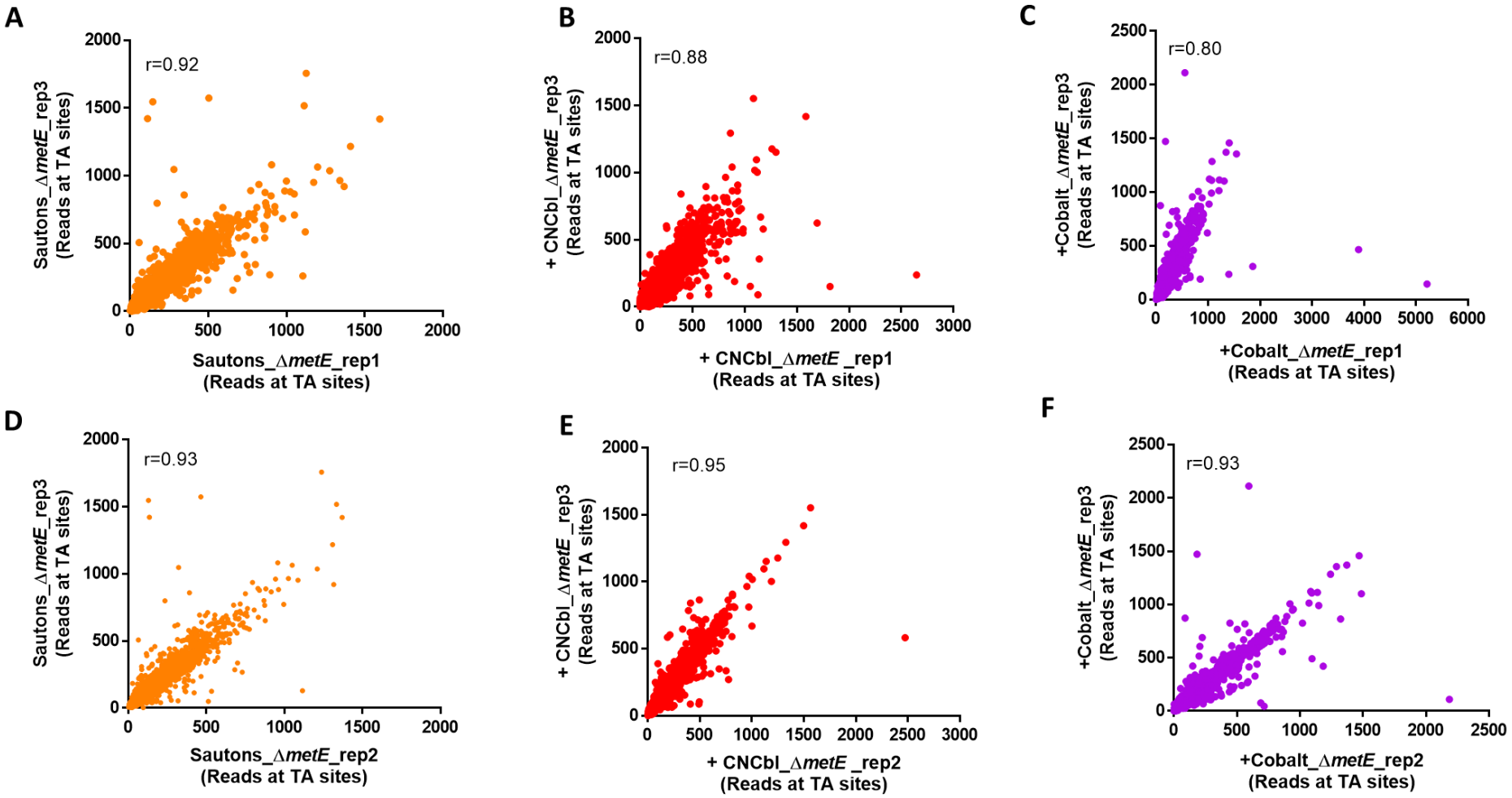


Figure S2: Validation of TnSeq fitness measurements. (**A-F**) A biological replicate comparing three independently, selected upon (Sauton's in yellow, +CNCbl in red and +cobalt in purple) and sequenced libraries. Pearson correlation coefficients between biological replicates ranged between 0.76 – 0.96.

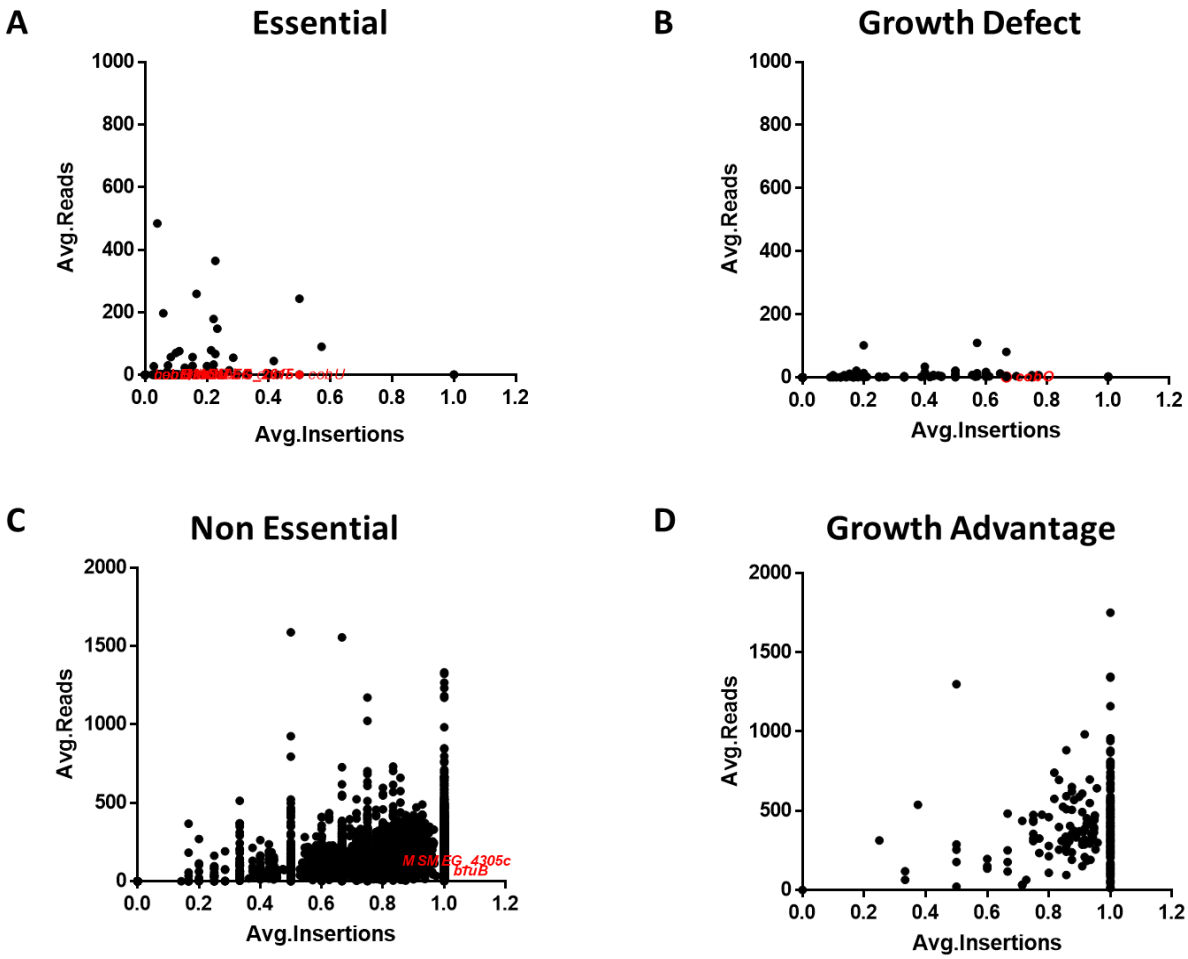


Figure S3: Identification of conditionally essential genes in unsupplemented conditions. (**A-D**) Scatter plot of mean non-zero read counts and mean insertion frequency for the regions identified by the HMM.

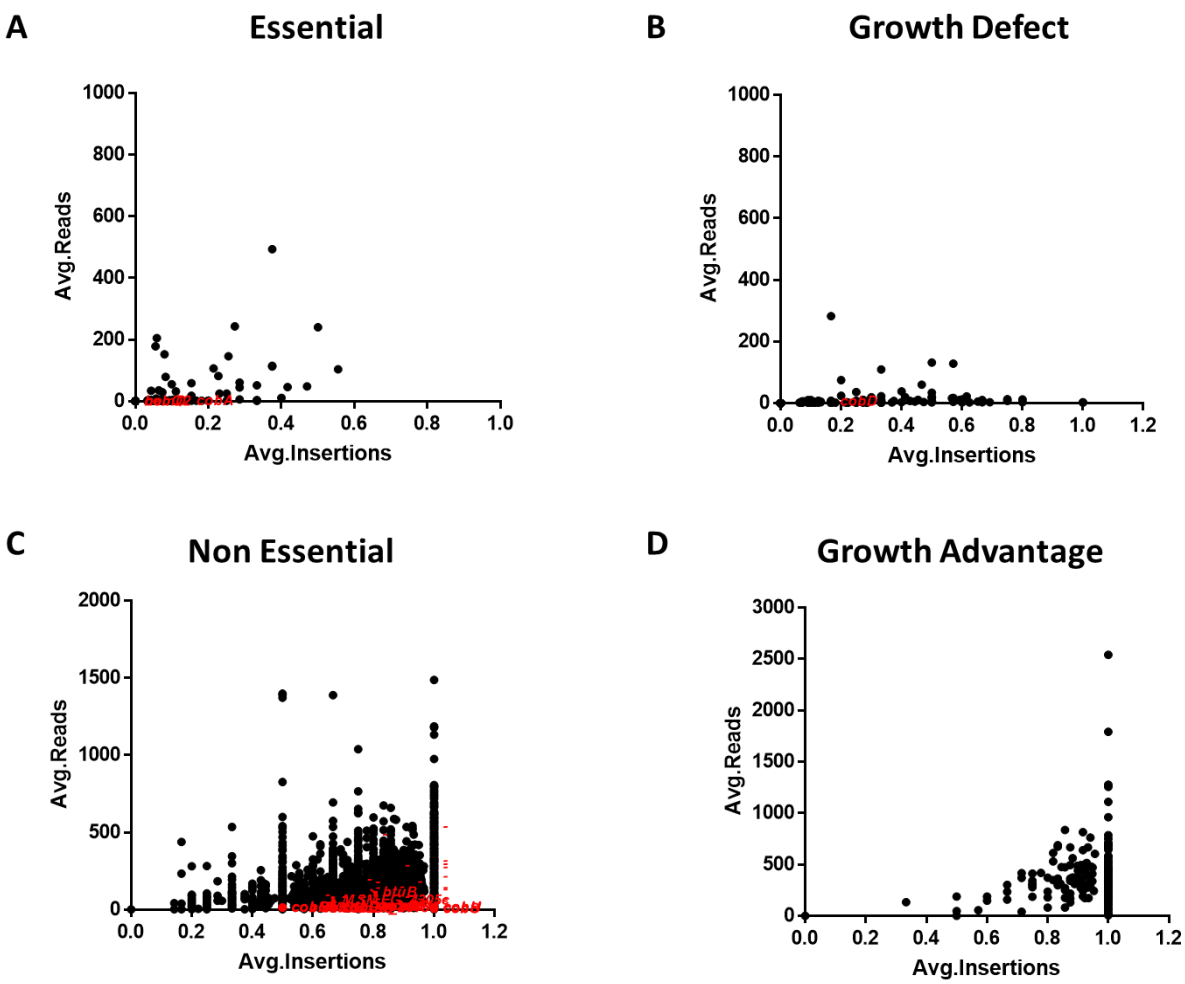


Figure S4: Identification of conditionally essential genes in CNCbl supplemented conditions. (**A-B**) Scatter plot of mean non-zero read counts and mean insertion frequency for the regions identified by the HMM.

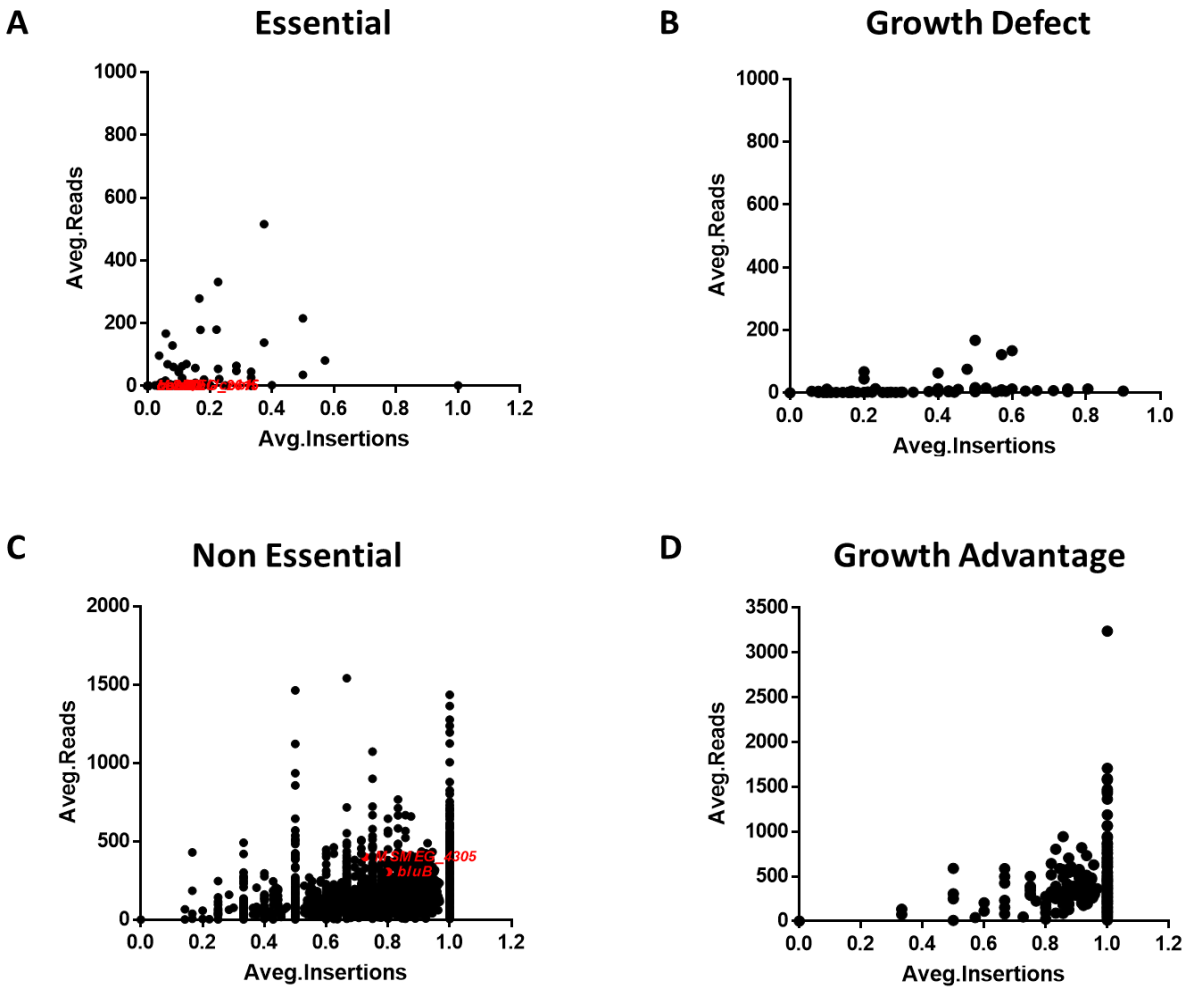


Figure S5: Identification of conditionally essential genes in cobalt supplemented conditions. (A-B) Scatter plot of mean non-zero read counts and mean insertion frequency for the regions identified by the HMM.

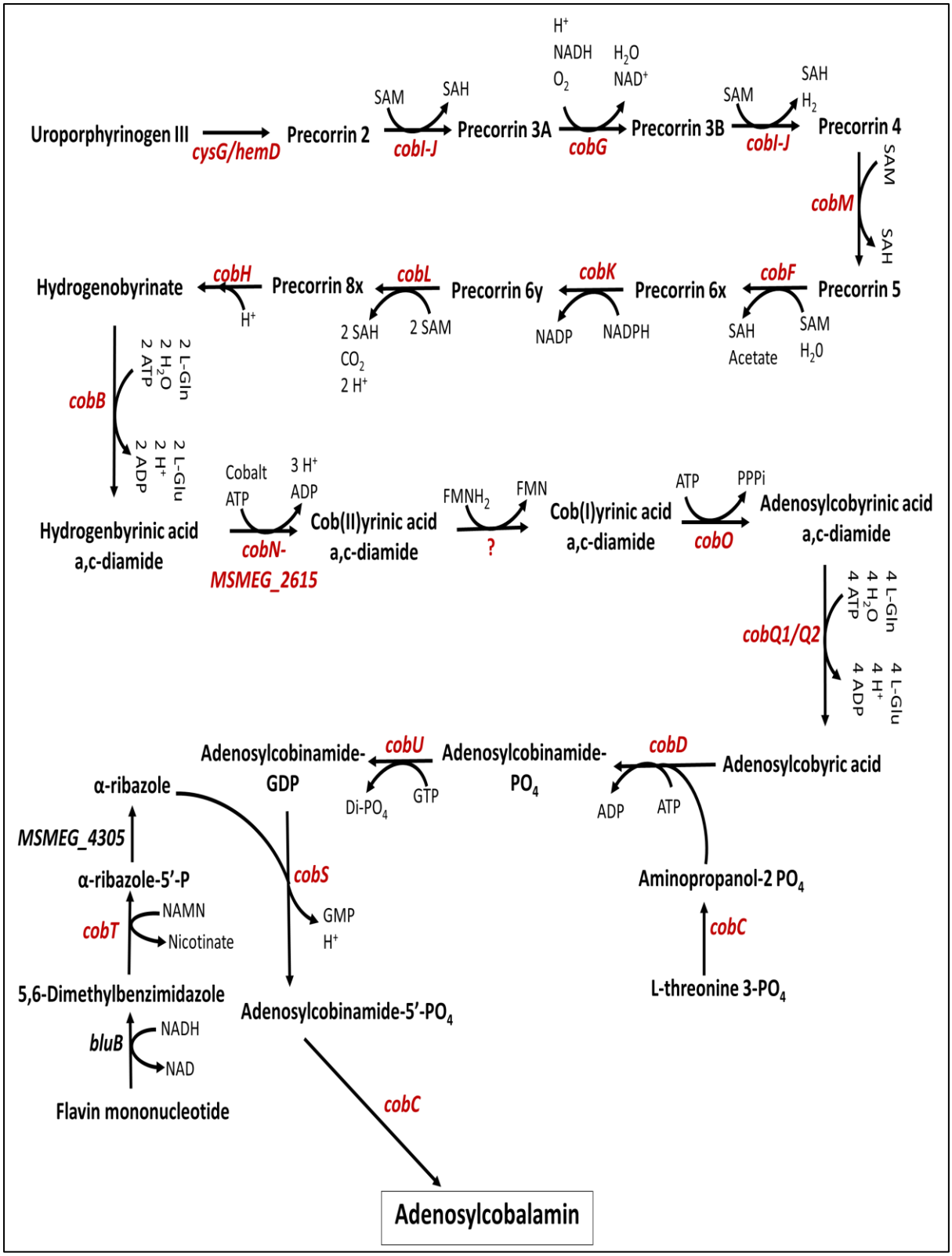


Figure S6: Predicted cobalamin biosynthetic pathway in *Msm*. Essential genes are show in red and non-essential genes are show in black. Essentiality classifications are from Tn insertion data for Δ*metE* library grown in unsupplemented Sauton's minimal medium.

Table S3: The 10 genes identified as essential only in Sauton's

Accession no.	Gene name	Predicted product/function	TAs	ES	GD	NE	GA	NZsites	NZmean	Essentiality call
MSMEG_0129	-	cyclase/dehydrase family protein	4	4	0	0	0	0	0	ES
MSMEG_1435	*rpsJ*	30S ribosomal protein S10	4	0	4	0	0	0	0	GD
MSMEG_2373	*ilvH*	acetolactate synthase 3 regulatory subunit	2	2	0	0	0	0	0	ES
MSMEG_2388	*leuD*	isopropylmalate isomerase small subunit	8	7	0	1	0	0,25	2	ES
MSMEG_2578	*dxr*	1-deoxy-D-xylulose 5-phosphate reductoisomerase	9	9	0	0	0	0	0	ES
MSMEG_4293	*glnE*	glutamate-ammonia-ligase adenylyltransferase	25	0	25	0	0	0,6	8,87	GD
MSMEG_4571	*rpsT*	30S ribosomal protein S20	3	0	3	0	0	1	3,33	GD
MSMEG_4572	-	hypothetical protein MSMEG_4572	5	0	5	0	0	0,2	1	GD
MSMEG_4581	*nadD*	nicotinate (nicotinamide) nucleotide adenylyltransferase	7	7	0	0	0	0	0	ES
MSMEG_4751	-	hypothetical protein MSMEG_4751	3	0	2	1	0	0,6667	81	GD

Table S4: The 13 genes identified as essential only in Sauton's +CNCbl

Accession no.	Gene name	Predicted product/function	TAs	ES	GD	NE	GA	NZsites	NZmean	Essentiality call
MSMEG_1647	*folD*	tetrahydrofolate dehydrogenase/cyclohydrolase FolD	2	0	2	0	0	0	0	GD
MSMEG_1646	-	ribosomal RNA adenine dimethylase family protein	9	0	9	0	0	0,3333	10,67	GD
MSMEG_2380	-	sugar transporter family protein	6	0	4	2	0	0,5	131,67	GD
MSMEG_2936	-	hydrolase, nudix family protein	9	4	0	5	0	0,5556	103,8	ES
MSMEG_4512	*mbtD*	polyketide synthetase mbtd	17	7	0	10	0	0,4706	47,88	ES

MSMEG_4876	-	acyl-CoA synthase	13	0	7	6	0	0,6154	22,12	GD
MSMEG_5268	-	hypothetical protein MSMEG_5268	16	5	0	11	0	0,375	115,17	ES
MSMEG_5269	-	hypothetical protein MSMEG_5269	12	6	0	6	0	0,25	25	ES
MSMEG_5618	-	acyl-CoA dehydrogenase	15	0	11	4	0	0,4667	60,14	GD
MSMEG_6153	-	DNA polymerase III subunit delta'	3	0	3	0	0	0	0	GD
MSMEG_6567	-	iron-dependent peroxidase	6	0	6	0	0	0,6667	4	GD
MSMEG_6569	-	hypothetical protein MSMEG_6569	3	0	3	0	0	0,6667	9	GD
MSMEG_6626	-	short-chain dehydrogenase/reductase SDR	7	5	0	2	0	0,2857	44,5	ES

Table S5: The 7 genes common to Sauton's and +CNCbl

Sauton's										
Accession no.	**Gene name**	**Predicted product/function**	**TAs**	**ES**	**GD**	**NE**	**GA**	**NZsites**	**NZmean**	**Essentiality call**
MSMEG_0935	*gpm1*	2,3-bisphosphoglycerate-dependent phosphoglycerate mutase	11	0	10	1	0	0,4545	4,6	GD
MSMEG_1550	*pduH*	PduH protein	2	0	2	0	0	0	0	GD
MSMEG_4217	-	DivIVA protein	6	6	0	0	0	0	0	ES
MSMEG_4955	*thrB*	homoserine kinase	6	0	6	0	0	0	0	GD
MSMEG_5684	*serC*	phosphoserine aminotransferase	7	7	0	0	0	0	0	ES
MSMEG_6101	*folK*	2-amino-4-hydroxy-6-hydroxymethyldihydropteridine pyrophosphokinase	2	2	0	0	0	0	0	ES
MSMEG_6402	-	PAP2 superfamily protein	2	1	1	0	0	0,5	22	GD
+CNCbl										

Accession no.	Gene name	Predicted product/function	TAs	ES	GD	NE	GA	NZsites	NZmean	Essentiality call
MSMEG_0935	*gpm1*	2,3-bisphosphoglycerate-dependent phosphoglycerate mutase	11	0	10	1	0	0,4545	11	GD
MSMEG_1550	*pduH*	PduH protein	2	0	2	0	0	0	0	GD
MSMEG_4217	-	DivIVA protein	6	6	0	0	0	0	0	ES
MSMEG_4955	*thrB*	homoserine kinase	6	0	6	0	0	0	0	GD
MSMEG_5684	*serC*	phosphoserine aminotransferase	7	7	0	0	0	0	0	ES
MSMEG_6101	*folK*	2-amino-4-hydroxy-6-hydroxymethyldihydropteridine pyrophosphokinase	2	2	0	0	0	0	0	ES
MSMEG_6402	-	PAP2 superfamily protein	2	1	1	0	0	0,5	34	GD

Table S6: The 24 genes identified as essential in +Cobalt

Accession no.	Gene name	Predicted product/function	TAs	ES	GD	NE	GA	NZsites	NZmean	Essentiality call
MSMEG_0003	*recF*	recombination protein F	11	0	11	0	0	0,4545	11,4	GD
MSMEG_0004	-	hypothetical protein MSMEG_0004	2	0	2	0	0	0,5	17	GD
MSMEG_0792	*thiS*	sulfur carrier protein ThiS	2	0	2	0	0	0	0	GD
MSMEG_0793	*thiG*	thiazole synthase	5	0	5	0	0	0,4	2,5	GD
MSMEG_0955	-	hypothetical protein MSMEG_0955	5	0	3	0	2	0,6	134,67	GD
MSMEG_0957	-	hypothetical protein MSMEG_0957	4	0	3	0	1	0,5	168	GD
MSMEG_1474	*rplO*	50S ribosomal protein L15	3	3	0	0	0	0,3333	2	ES
MSMEG_1959	-	hypothetical protein MSMEG_1959	52	0	52	0	0	0,7115	7,49	GD
MSMEG_2771	*trkA*	TrkA protein	7	0	7	0	0	0,2857	2	GD
MSMEG_3026	-	Holliday junction resolvase-like protein	3	0	3	0	0	0	0	GD
MSMEG_3027	-	hypothetical protein MSMEG_3027	10	0	10	0	0	0,9	5,89	GD

MSMEG_3028	*aroE*	shikimate 5-dehydrogenase	6	0	6	0	0	0	0	GD
MSMEG_3118	-	ABC transporter efflux protein, DrrB family protein	1	1	0	0	0	0	0	ES
MSMEG_3119	-	ABC transporter, ATP-binding subunit	7	6	0	1	0	0,1429	2	ES
MSMEG_3188	*bioA*	adenosylmethionine--8-amino-7-oxononanoate transaminase	8	0	8	0	0	0,25	2	GD
MSMEG_4258	*trpD*	anthranilate phosphoribosyltransferase	2	0	2	0	0	0	0	GD
MSMEG_4323	*aceE*	pyruvate dehydrogenase subunit E1	41	0	22	19	0	0,8049	12,91	GD
MSMEG_4599	-	hypothetical protein MSMEG_4599	23	0	13	10	0	0,4783	75,82	GD
MSMEG_5239	*glpX*	fructose 1,6-bisphosphatase II	6	0	6	0	0	0,5	9,33	GD
MSMEG_5240	*fumC*	fumarate hydratase	3	0	3	0	0	0	0	GD
MSMEG_5694	-	hypothetical protein MSMEG_5694	7	0	7	0	0	0,5714	5,75	GD
MSMEG_5832	*purS*	phosphoribosylformylglycinamidine synthase subunit PurS	1	1	0	0	0	0	0	ES
MSMEG_6073	-	RNA methyltransferase, TrmH family protein, group 3	3	3	0	0	0	0	0	ES
MSMEG_6938	*parB*	ParB-like partition proteins	6	0	6	0	0	0,3333	2	GD

Table S7: The 15 genes common to +Cobalt and +CNCbl

+Cobalt										
Accession no.	**Gene name**	**Predicted product/function**	**TAs**	**ES**	**GD**	**NE**	**GA**	**NZsites**	**NZmean**	**Essentiality call**
MSMEG_0068	*eccD1*	ESX conserved component *eccD1*. Probable transmembrane protein	8	4	0	3	1	0,375	516	ES
MSMEG_0272	*pduA*	propanediol utilization protein PduA	1	0	1	0	0	0	0	GD
MSMEG_0273	*eutN*	ethanolamine utilization protein EutN	1	0	1	0	0	0	0	GD

MSMEG_0274	-	hypothetical protein MSMEG_0274	2	0	2	0	0	0	0	GD
MSMEG_1896	-	acetyl-CoA acetyltransferase	5	0	4	1	0	0,4	63,5	GD
MSMEG_1925	-	isochorismate synthase DhbC	5	5	0	0	0	0	0	ES
MSMEG_3449	-	DNA-binding protein	6	3	0	3	0	0,3333	45,5	ES
MSMEG_4309	*ptpA*	low molecular weight protein-tyrosine-phosphatase	4	0	4	0	0	0,75	8,33	GD
MSMEG_4313	-	glyoxalase/bleomycin resistance protein/dioxygenase	4	0	3	1	0	0,5	13,5	GD
MSMEG_4878	-	2Fe-2S iron-sulfur cluster binding domain-containing protein	8	7	0	1	0	0,125	70	ES
MSMEG_5621	-	putative acyl-CoA dehydrogenase	4	4	0	0	0	0	0	ES
MSMEG_5622	-	putative acyl-CoA dehydrogenase	8	5	0	3	0	0,375	138	ES
MSMEG_5831	*purQ*	phosphoribosylformylglycinamidine synthase subunit I	5	5	0	0	0	0	0	ES
MSMEG_6520	-	orotate phosphoribosyltransferase	7	0	7	0	0	0,1429	2	GD
MSMEG_6933	*trxB*	thioredoxin-disulfide reductase	7	7	0	0	0	0	0	ES
CNCbl										
Accession no.	**Gene name**	**Predicted product/function**	**TAs**	**ES**	**GD**	**NE**	**GA**	**NZsites**	**NZmean**	**Essentiality call**
MSMEG_0068	*eccD1*	ESX conserved component *eccD1*. Probable transmembrane protein	8	4	0	3	1	0,375	493,33	ES
MSMEG_0272	*pduA*	propanediol utilization protein PduA	1	0	1	0	0	0	0	GD
MSMEG_0273	*eutN*	ethanolamine utilization protein EutN	1	0	1	0	0	0	0	GD
MSMEG_0274	-	hypothetical protein MSMEG_0274	2	0	2	0	0	0	0	GD
MSMEG_1896	-	acetyl-CoA acetyltransferase	5	0	4	1	0	0,4	38,5	GD
MSMEG_1925	-	isochorismate synthase DhbC	5	5	0	0	0	0	0	ES
MSMEG_3449	-	DNA-binding protein	6	3	0	3	0	0,3333	51,5	ES
MSMEG_4309	*ptpA*	low molecular weight protein-tyrosine-phosphatase	4	0	4	0	0	0,75	12,67	GD

MSMEG_4313	-	glyoxalase/bleomycin resistance protein/dioxygenase	4	0	4	0	0	0,5	6	GD
MSMEG_4878	-	2Fe-2S iron-sulfur cluster binding domain-containing protein	8	0	7	1	0	0,25	36,5	GD
MSMEG_5621	-	putative acyl-CoA dehydrogenase	4	4	0	0	0	0	0	ES
MSMEG_5622	-	putative acyl-CoA dehydrogenase	8	5	0	3	0	0,375	112,67	ES
MSMEG_5831	*purQ*	phosphoribosylformylglycinamidine synthase subunit I	5	5	0	0	0	0	0	ES
MSMEG_6520	-	orotate phosphoribosyltransferase	7	7	0	0	0	0	0	ES
MSMEG_6933	*trxB*	thioredoxin-disulfide reductase	7	7	0	0	0	0	0	ES

Table S8: The 29 genes common to Sauton's and +Cobalt

Sauton's										
Accession no.	**Gene name**	**Predicted product/function**	**TAs**	**ES**	**GD**	**NE**	**GA**	**NZsites**	**NZmean**	**Essentiality call**
MSMEG_0042	-	TetR-family protein transcriptional regulator	6	0	6	0	0	0,5	8	GD
MSMEG_0107	-	cell envelope-related function transcriptional attenuator	18	0	18	0	0	0,6111	4	GD
MSMEG_0710	*grpE*	co-chaperone GrpE	5	0	4	1	0	0,2	102	GD
MSMEG_1063	-	hypothetical protein MSMEG_1063	2	2	0	0	0	0	0	ES
MSMEG_1519	*infA*	translation initiation factor IF-1	3	0	3	0	0	0	0	GD
MSMEG_1945	-	ion channel membrane protein	9	0	9	0	0	0,4444	6,5	GD
MSMEG_2588	*cobQ*	cobyric acid synthase	7	7	0	0	0	0	0	ES
MSMEG_2615	-	chelatase	8	8	0	0	0	0,125	1	ES
MSMEG_2617	*cobB*	cobyrinic acid a,c-diamide synthase	8	8	0	0	0	0,125	1	ES

MSMEG_3150	*fabG*	3-oxoacyl-(acyl-carrier-protein) reductase	4	4	0	0	0	0	0	ES
MSMEG_3637	-	CBS domain-containing protein	5	0	5	0	0	0,2	3	GD
MSMEG_3638	-	CBS domain-containing protein	6	0	6	0	0	0,6667	4,75	GD
MSMEG_3864	*cobN*	cobaltochelatase	34	34	0	0	0	0,0882	1	ES
MSMEG_3871	*cobG*	precorrin-3B synthase	2	0	2	0	0	0	0	ES
MSMEG_3872	*cobH*	precorrin-8X methylmutase	3	0	3	0	0	0	0	ES
MSMEG_3873	*cobIJ*	cobalamin biosynthesis protein cobIJ	14	0	14	0	0	0,1429	2	ES
MSMEG_3875	*cobK*	cobalt-precorrin-6x reductase	5	5	0	0	0	0,2	1	ES
MSMEG_3877	*cobM*	precorrin-4 C11-methyltransferase	6	6	0	0	0	0	0	ES
MSMEG_3878	*cobL*	precorrin-6Y C5,15-methyltransferase (decarboxylating)	6	6	0	0	0	0	0	ES
MSMEG_4274	*cobU*	cobinamide kinase / cobinamide phosphate guanyltransferase	2	2	0	0	0	0,5	1	ES
MSMEG_4275	*cobT*	nicotinate-nucleotide--dimethylbenzimidazole phosphoribosyltransferase	3	3	0	0	0	0,3333	1	ES
MSMEG_4277	*cobS*	cobalamin synthase	2	2	0	0	0	0	0	ES
MSMEG_5069	*tatB*	sec-independent translocase	1	1	0	0	0	0	0	ES
MSMEG_5070	-	trypsin	5	5	0	0	0	0	0	ES
MSMEG_5122	*fdxC*	ferredoxin	5	5	0	0	0	0	0	ES
MSMEG_5548	*cobF*	precorrin 6A synthase	6	6	0	0	0	0	0	ES
MSMEG_5796	-	glycine cleavage T-protein (aminomethyl transferase)	6	6	0	0	0	0	0	ES
MSMEG_6154	-	adenylate cyclase, family protein 3	10	0	10	0	0	0,7	4,43	GD
MSMEG_6893	-	hypothetical protein MSMEG_6893	7	3	0	4	0	0,5714	90,25	ES
Cobalt										
Accession no.	**Gene name**	**Predicted product/function**	**TAs**	**ES**	**GD**	**NE**	**GA**	**NZsites**	**NZmean**	**Essentiality call**
MSMEG_0042	-	TetR-family protein transcriptional regulator	6	0	6	0	0	0,5	3,67	GD

MSMEG_0107	-	cell envelope-related function transcriptional attenuator	18	0	18	0	0	0,5556	3,5	GD
MSMEG_0710	*grpE*	co-chaperone GrpE	5	0	4	1	0	0,2	68	GD
MSMEG_1063	-	hypothetical protein MSMEG_1063	2	0	2	0	0	0	0	GD
MSMEG_1519	*infA*	translation initiation factor IF-1	3	0	3	0	0	0	0	GD
MSMEG_1945	-	ion channel membrane protein	9	0	9	0	0	0,4444	2,5	GD
MSMEG_2588	*cobQ*	cobyric acid synthase	7	7	0	0	0	0	0	ES
MSMEG_2615	-	chelatase	8	8	0	0	0	0	0	ES
MSMEG_2617	*cobB*	cobyrinic acid a,c-diamide synthase	8	8	0	0	0	0	0	ES
MSMEG_3150	*fabG*	3-oxoacyl-(acyl-carrier-protein) reductase	4	4	0	0	0	0,25	2	ES
MSMEG_3637	-	CBS domain-containing protein	5	0	5	0	0	0,2	2	GD
MSMEG_3638	-	CBS domain-containing protein	6	0	6	0	0	0,5	2	GD
MSMEG_3864	*cobN*	cobaltochelatase	34	34	0	0	0	0,0294	2	ES
MSMEG_3871	*cobG*	precorrin-3B synthase	2	2	0	0	0	0	0	ES
MSMEG_3872	*cobH*	precorrin-8X methylmutase	3	3	0	0	0	0	0	ES
MSMEG_3873	*cobIJ*	cobalamin biosynthesis protein cobIJ	14	14	0	0	0	0,0714	2	ES
MSMEG_3875	*cobK*	cobalt-precorrin-6x reductase	5	5	0	0	0	0,2	2	ES
MSMEG_3877	*cobM*	precorrin-4 C11-methyltransferase	6	6	0	0	0	0	0	ES
MSMEG_3878	*cobL*	precorrin-6Y C5,15-methyltransferase (decarboxylating)	6	6	0	0	0	0	0	ES
MSMEG_4274	*cobU*	cobinamide kinase / cobinamide phosphate guanyltransferase	2	2	0	0	0	0	0	ES
MSMEG_4275	*cobT*	nicotinate-nucleotide--dimethylbenzimidazole phosphoribosyltransferase	3	3	0	0	0	0	0	ES
MSMEG_4277	*cobS*	cobalamin synthase	2	2	0	0	0	0	0	ES
MSMEG_5069	*tatB*	sec-independent translocase	1	1	0	0	0	0	0	ES
MSMEG_5070	-	trypsin	5	5	0	0	0	0	0	ES
MSMEG_5122	*fdxC*	ferredoxin	5	5	0	0	0	0	0	ES
MSMEG_5548	*cobF*	precorrin 6A synthase	6	6	0	0	0	0	0	ES

MSMEG_5796	-	glycine cleavage T-protein (aminomethyl transferase)	6	6	0	0	0	0	0	ES
MSMEG_6154	-	adenylate cyclase, family protein 3	10	0	10	0	0	0,5	3,8	GD
MSMEG_6893	-	hypothetical protein MSMEG_6893	7	3	0	4	0	0,5714	81,25	ES

8 References

Abubakar, I., Zignol, M., Falzon, D., Raviglione, M., Ditiu, L., Masham, S., Adetifa, I., Ford, N., Cox, H., & Lawn, S. D. (2013). Drug-resistant tuberculosis: time for visionary political leadership. *The Lancet infectious diseases*, *13*(6), 529-539.

Akerley, B. J., Rubin, E. J., Camilli, A., Lampe, D. J., Robertson, H. M., & Mekalanos, J. J. (1998). Systematic identification of essential genes by in vitro mariner mutagenesis. *Proceedings of the National Academy of Sciences*, *95*(15), 8927-8932.

Akerley, B. J., Rubin, E. J., Novick, V. L., Amaya, K., Judson, N., & Mekalanos, J. J. (2002). A genome-scale analysis for identification of genes required for growth or survival of Haemophilus influenzae. *Proceedings of the National Academy of Sciences*, *99*(2), 966-971.

Allen, B. W. (1998). Mycobacteria. In T. Parish & N. G. Stoker (Eds.), *Mycobacteria Protocols* (pp. 15-30). Humana Press.

Allen, R. H., Stabler, S. P., Savage, D. G., & Lindenbaum, J. (1993). Metabolic abnormalities in cobalamin (vitamin B12) and folate deficiency. *The FASEB journal*, *7*(14), 1344-1353.

Anderson, P. J., Lango, J., Carkeet, C., Britten, A., Kräutler, B., Hammock, B. D., & Roth, J. R. (2008). One pathway can incorporate either adenine or dimethylbenzimidazole as an α-axial ligand of B12 cofactors in Salmonella enterica. *Journal of bacteriology*, *190*(4), 1160-1171.

Baba, T., Ara, T., Hasegawa, M., Takai, Y., Okumura, Y., Baba, M., Datsenko, K. A., Tomita, M., Wanner, B. L., & Mori, H. (2006). Construction of Escherichia coli K-12 in-frame, single-gene knockout mutants: the Keio collection. *Molecular systems biology*, *2*(1), 2006.0008.

Banerjee, R., & Ragsdale, S. W. (2003). The many faces of vitamin B12: catalysis by cobalamin-dependent Enzymes1. *Annual review of biochemistry*, *72*.

Banerjee, R. V., & Matthews, R. G. (1990). Cobalamin-dependent methionine synthase. *The FASEB journal*, *4*(5), 1450-1459.

Bao, Z., Qi, X., Hong, S., Xu, K., He, F., Zhang, M., Chen, J., Chao, D., Zhao, W., & Li, D. (2017). Structure and mechanism of a group-I cobalt energy coupling factor transporter. *Cell research*, *27*(5), 675-687.

Barker, H., Weissbach, H., & Smyth, R. (1958). A coenzyme containing pseudovitamin B12. *Proceedings of the National Academy of Sciences of the United States of America*, *44*(11), 1093.

Barquist, L., Boinett, C. J., & Cain, A. K. (2013). Approaches to querying bacterial genomes with transposon-insertion sequencing. *RNA biology*, *10*(7), 1161-1169.

Barquist, L., Mayho, M., Cummins, C., Cain, A. K., Boinett, C. J., Page, A. J., Langridge, G. C., Quail, M. A., Keane, J. A., & Parkhill, J. (2016). The TraDIS toolkit: sequencing and analysis for dense transposon mutant libraries. *Bioinformatics*, *32*(7), 1109-1111.

Barry, C. E. (2001). Mycobacterium smegmatis: an absurd model for tuberculosis? Response from Barry, III. *Trends in microbiology*, *9*(10), 473-474.

Behar, S., Martin, C., Booty, M., Nishimura, T., Zhao, X., Gan, H., Divangahi, M., & Remold, H. (2011). Apoptosis is an innate defense function of macrophages against Mycobacterium tuberculosis. *Mucosal immunology*, *4*(3), 279-287.

Bermingham, A., & Derrick, J. P. (2002). The folic acid biosynthesis pathway in bacteria: evaluation of potential for antibacterial drug discovery. *Bioessays*, *24*(7), 637-648.

Berry, M. P., Graham, C. M., McNab, F. W., Xu, Z., Bloch, S. A., Oni, T., Wilkinson, K. A., Banchereau, R., Skinner, J., & Wilkinson, R. J. (2010). An interferon-inducible neutrophil-driven blood transcriptional signature in human tuberculosis. *Nature*, *466*(7309), 973-977.

Biémont, C. (2010). A brief history of the status of transposable elements: from junk DNA to major players in evolution. *Genetics*, *186*(4), 1085-1093.

Blanche, F., Cameron, B., Crouzet, J., Debussche, L., Thibaut, D., Vuilhorgne, M., Leeper, F. J., & Battersby, A. R. (1995). Vitamin B12: how the problem of its biosynthesis was solved. *Angewandte Chemie International Edition in English*, *34*(4), 383-411.

Blanco, J., Coque, J. J. R., & Martin, J. F. (1998). The folate branch of the methionine biosynthesis pathway in Streptomyces lividans: Disruption of the 5, 10-methylenetetrahydrofolate reductase gene leads to methionine auxotrophy. *Journal of bacteriology*, *180*(6), 1586-1591.

Blankley, S., Graham, C. M., Turner, J., Berry, M. P., Bloom, C. I., Xu, Z., Pascual, V., Banchereau, J., Chaussabel, D., & Breen, R. (2016). The transcriptional signature of active tuberculosis reflects symptom status in extra-pulmonary and pulmonary tuberculosis. *PloS one*, *11*(10), e0162220.

Bloom, B. R., Atun, R., Cohen, T., Dye, C., Fraser, H., Gomez, G. B., Knight, G., Murray, M., Nardell, E., Rubin, E., Salomon, J., Vassall, A., Volchenkov, G., White, R., Wilson, D., & Yadav, P. (2017). Tuberculosis. In K. K. Holmes, S. Bertozzi, B. R. Bloom, & P. Jha (Eds.), *Major Infectious Diseases* (pp. 233-313).

Bloss, E., Kukša, L., Holtz, T., Riekstina, V., Skripčonoka, V., Kammerer, S., & Leimane, V. (2010). Adverse events related to multidrug-resistant tuberculosis treatment, Latvia, 2000–2004. *The International journal of tuberculosis and lung disease*, *14*(3), 275-281.

Blount, K. F., & Breaker, R. R. (2006). Riboswitches as antibacterial drug targets. *Nature biotechnology*, *24*(12), 1558-1564.

Boissier, F., Bardou, F., Guillet, V., Uttenweiler-Joseph, S., Daffé, M., Quémard, A., & Mourey, L. (2006). Further insight into S-adenosylmethionine-dependent methyltransferases: structural characterization of Hma, an enzyme essential for the biosynthesis of oxygenated mycolic acids in Mycobacterium tuberculosis. *Journal of Biological Chemistry*, *281*(7), 4434-4445.

Boom, W. H., Schaible, U. E., & Achkar, J. M. (2021). The knowns and unknowns of latent Mycobacterium tuberculosis infection. *Journal of Clinical Investigation*, *131*(3), e136222.

Boritsch, E. C., Frigui, W., Cascioferro, A., Malaga, W., Etienne, G., Laval, F., Pawlik, A., Le Chevalier, F., Orgeur, M., & Ma, L. (2016). pks5-recombination-mediated surface remodelling in Mycobacterium tuberculosis emergence. *Nature microbiology*, *1*(2), 1-11.

Boritsch, E. C., Supply, P., Honoré, N., Seeman, T., Stinear, T. P., & Brosch, R. (2014). A glimpse into the past and predictions for the future: the molecular evolution of the tuberculosis agent. *Molecular microbiology*, *93*(5), 835-852.

Borovok, I., Gorovitz, B., Schreiber, R., Aharonowitz, Y., & Cohen, G. (2006). Coenzyme B12 controls transcription of the Streptomyces class Ia ribonucleotide reductase nrdABS operon via a riboswitch mechanism. *Journal of bacteriology*, *188*(7), 2512-2520.

Borths, E. L., Poolman, B., Hvorup, R. N., Locher, K. P., & Rees, D. C. (2005). In vitro functional characterization of BtuCD-F, the Escherichia coli ABC transporter for vitamin B12 uptake. *Biochemistry*, *44*(49), 16301-16309.

Bridwell-Rabb, J., & Drennan, C. L. (2017). Vitamin B12 in the spotlight again. *Current opinion in chemical biology*, *37*, 63-70.

Brosch, R., Gordon, S. V., Marmiesse, M., Brodin, P., Buchrieser, C., Eiglmeier, K., Garnier, T., Gutierrez, C., Hewinson, G., & Kremer, K. (2002). A new evolutionary scenario for the Mycobacterium tuberculosis complex. *Proceedings of the National Academy of Sciences*, *99*(6), 3684-3689.

Bruchfeld, J., Correia-Neves, M., & Källenius, G. (2015). Tuberculosis and HIV coinfection. *Cold Spring Harbor perspectives in medicine*, *5*(7), a017871.

Burby, P. E., Nye, T. M., Schroeder, J. W., & Simmons, L. A. (2017). Implementation and data analysis of Tn-seq, whole-genome resequencing, and single-molecule real-time sequencing for bacterial genetics. *Journal of bacteriology*, *199*(1).

Cadena, A. M., Fortune, S. M., & Flynn, J. L. (2017). Heterogeneity in tuberculosis. *Nature Reviews Immunology*, *17*(11), 691-702.

Cain, A. K., Barquist, L., Goodman, A. L., Paulsen, I. T., Parkhill, J., & van Opijnen, T. (2020). A decade of advances in transposon-insertion sequencing. *Nature Reviews Genetics*, *21*(9), 526-540.

Calmette, A., Boquet, A., & Negre, L. (1921). Contribution à l'étude du bacille tuberculeux bilié. *Ann Inst Pasteur*, *9*, 561-570.

Campbell, G. R., Taga, M. E., Mistry, K., Lloret, J., Anderson, P. J., Roth, J. R., & Walker, G. C. (2006). Sinorhizobium meliloti bluB is necessary for production of 5, 6-dimethylbenzimidazole, the lower ligand of B12. *Proceedings of the National Academy of Sciences*, *103*(12), 4634-4639.

Capyk, J. K., Kalscheuer, R., Stewart, G. R., Liu, J., Kwon, H., Zhao, R., Okamoto, S., Jacobs Jr, W. R., Eltis, L. D., & Mohn, W. W. (2009). Mycobacterial cytochrome p450 125 (cyp125) catalyzes the terminal hydroxylation of c27 steroids. *Journal of Biological Chemistry*, *284*(51), 35534-35542.

Carey, A. F., Rock, J. M., Krieger, I. V., Chase, M. R., Fernandez-Suarez, M., Gagneux, S., Sacchettini, J. C., Ioerger, T. R., & Fortune, S. M. (2018). TnSeq of Mycobacterium tuberculosis clinical isolates reveals strain-specific antibiotic liabilities. *PLoS pathogens*, *14*(3), e1006939.

Casabon, I., Crowe, A. M., Liu, J., & Eltis, L. D. (2013). FadD 3 is an acyl-CoA synthetase that initiates catabolism of cholesterol rings C and D in actinobacteria. *Molecular microbiology*, *87*(2), 269-283.

Casabon, I., Swain, K., Crowe, A. M., Eltis, L. D., & Mohn, W. W. (2014). Actinobacterial acyl coenzyme A synthetases involved in steroid side-chain catabolism. *Journal of bacteriology*, *196*(3), 579-587.

Chao, M. C., Abel, S., Davis, B. M., & Waldor, M. K. (2016). The design and analysis of transposon insertion sequencing experiments. *Nature Reviews Microbiology*, *14*(2), 119.

Chao, M. C., Pritchard, J. R., Zhang, Y. J., Rubin, E. J., Livny, J., Davis, B. M., & Waldor, M. K. (2013). High-resolution definition of the Vibrio cholerae essential gene set with hidden Markov model–based analyses of transposon-insertion sequencing data. *Nucleic acids research*, *41*(19), 9033-9048.

Chaoui, I., Sabouni, R., Kourout, M., Jordaan, A. M., Lahlou, O., Elouad, R., Akrim, M., Victor, T. C., & Elmzibri, M. (2009). Analysis of isoniazid, streptomycin and ethambutol resistance in Mycobacterium tuberculosis isolates from Morocco. *The Journal of Infection in Developing Countries*, *3*(04), 278-284.

Cheng, J., Poduska, B., Morton, R. A., & Finan, T. M. (2011). An ABC-type cobalt transport system is essential for growth of Sinorhizobium melilotiat trace metal concentrations. *Journal of bacteriology*, *193*(17), 4405-4416.

Chetty, S., Ramesh, M., Singh-Pillay, A., & Soliman, M. E. (2017). Recent advancements in the development of anti-tuberculosis drugs. *Bioorganic & medicinal chemistry letters*, *27*(3), 370-386.

Chevalier, F. L., Cascioferro, A., Majlessi, L., Herrmann, J. L., & Brosch, R. (2014). Mycobacterium tuberculosis evolutionary pathogenesis and its putative impact on drug development. *Future microbiology*, *9*(8), 969-985.

Chiang, S. L., & Rubin, E. J. (2002). Construction of a mariner-based transposon for epitope-tagging and genomic targeting. *Gene*, *296*(1-2), 179-185.

Choby, J. E., & Skaar, E. P. (2016). Heme synthesis and acquisition in bacterial pathogens. *Journal of molecular biology*, *428*(17), 3408-3428.

Choudhary, P. K., Duret, A., Rohrbach-Brandt, E., Holliger, C., Sigel, R. K., & Maillard, J. (2013). Diversity of cobalamin riboswitches in the corrinoid-producing organohalide respirer Desulfitobacterium hafniense. *Journal of bacteriology*, *195*(22), 5186-5195.

Claes, W. A., Pühler, A., & Kalinowski, J. (2002). Identification of two prpDBC gene clusters in Corynebacterium glutamicum and their involvement in propionate degradation via the 2-methylcitrate cycle. *Journal of bacteriology*, *184*(10), 2728-2739.

Clark-Curtiss, J. E., & Haydel, S. E. (2003). Molecular genetics of Mycobacterium tuberculosis pathogenesis. *Annual Reviews in Microbiology*, *57*(1), 517-549.

Cole, S., Brosch, R., Parkhill, J., Garnier, T., Churcher, C., Harris, D., Gordon, S., Eiglmeier, K., Gas, S., & Barry, C. r. (1998). Deciphering the biology of Mycobacterium tuberculosis from the complete genome sequence. *Nature*, *396*(6707), 190-190.

Cole, S., Eiglmeier, K., Parkhill, J., James, K., Thomson, N., Wheeler, P., Honore, N., Garnier, T., Churcher, C., & Harris, D. (2001). Massive gene decay in the leprosy bacillus. *Nature*, *409*(6823), 1007-1011.

Comas, I., Coscolla, M., Luo, T., Borrell, S., Holt, K. E., Kato-Maeda, M., Parkhill, J., Malla, B., Berg, S., & Thwaites, G. (2013). Out-of-Africa migration and Neolithic

coexpansion of Mycobacterium tuberculosis with modern humans. *Nature genetics*, *45*(10), 1176-1182.

Conrad, W. H., Osman, M. M., Shanahan, J. K., Chu, F., Takaki, K. K., Cameron, J., Hopkinson-Woolley, D., Brosch, R., & Ramakrishnan, L. (2017). Mycobacterial ESX-1 secretion system mediates host cell lysis through bacterium contact-dependent gross membrane disruptions. *Proceedings of the National Academy of Sciences*, *114*(6), 1371-1376.

Conradie, F., Diacon, A. H., Ngubane, N., Howell, P., Everitt, D., Crook, A. M., Mendel, C. M., Egizi, E., Moreira, J., & Timm, J. (2020). Treatment of highly drug-resistant pulmonary tuberculosis. *New England Journal of Medicine*, *382*(10), 893-902.

Corbett, E. L., Watt, C. J., Walker, N., Maher, D., Williams, B. G., Raviglione, M. C., & Dye, C. (2003). The growing burden of tuberculosis: global trends and interactions with the HIV epidemic. *Archives of internal medicine*, *163*(9), 1009-1021.

Costa, F. G. (2020). *Coenzyme Biosynthesis and Delivery in the Ethanolamine Utilization Metabolosome* University of Georgia].

Creswell, J., Raviglione, M., Ottmani, S., Migliori, G. B., Uplekar, M., Blanc, L., Sotgiu, G., & Lönnroth, K. (2011). Tuberculosis and noncommunicable diseases: neglected links and missed opportunities. *European Respiratory Journal*, *37*(5), 1269-1282.

Crofts, T. S., Seth, E. C., Hazra, A. B., & Taga, M. E. (2013). Cobamide structure depends on both lower ligand availability and CobT substrate specificity. *Chemistry & biology*, *20*(10), 1265-1274.

Crowe, A. M., Casabon, I., Brown, K. L., Liu, J., Lian, J., Rogalski, J. C., Hurst, T. E., Snieckus, V., Foster, L. J., & Eltis, L. D. (2017). Catabolism of the last two steroid rings in Mycobacterium tuberculosis and other bacteria. *MBio*, *8*(2).

Czubat, B., Minias, A., Brzostek, A., Żaczek, A., Struś, K., Zakrzewska-Czerwińska, J., & Dziadek, J. (2020). Functional Disassociation Between the Protein Domains of MSMEG_4305 of Mycolicibacterium smegmatis (Mycobacterium smegmatis) in vivo. *Frontiers in microbiology*, *11*, 2008.

D'Souza, G., Shitut, S., Preussger, D., Yousif, G., Waschina, S., & Kost, C. (2018). Ecology and evolution of metabolic cross-feeding interactions in bacteria. *Natural product reports*, *35*(5), 455-488.

Danilchanka, O., Sun, J., Pavlenok, M., Maueröder, C., Speer, A., Siroy, A., Marrero, J., Trujillo, C., Mayhew, D. L., & Doornbos, K. S. (2014). An outer membrane channel protein of Mycobacterium tuberculosis with exotoxin activity. *Proceedings of the National Academy of Sciences*, *111*(18), 6750-6755.

Dassanayake, R. S., Cabelli, D. E., & Brasch, N. E. (2013). Pulse radiolysis studies on the reaction of the reduced vitamin B12 complex Cob (II) alamin with superoxide. *ChemBioChem*, *14*(9), 1081-1083.

Davidson, A. L., Dassa, E., Orelle, C., & Chen, J. (2008). Structure, function, and evolution of bacterial ATP-binding cassette systems. *Microbiology and molecular biology reviews*, *72*(2), 317-364.

Davis, B. D. (1952). Utilization of pseudovitamin B12 by mutants of Escherichia coli. *Journal of bacteriology*, *64*(3), 432.

Dawes, S. S., & Mizrahi, V. (2001). DNA metabolism in Mycobacterium leprae. *Leprosy review*, *72*(4), 408-414.

Dawes, S. S., Warner, D. F., Tsenova, L., Timm, J., McKinney, J. D., Kaplan, G., Rubin, H., & Mizrahi, V. (2003). Ribonucleotide reduction in Mycobacterium tuberculosis: function and expression of genes encoding class Ib and class II ribonucleotide reductases. *Infection and immunity*, *71*(11), 6124-6131.

de Carvalho, L. P. S., Fischer, S. M., Marrero, J., Nathan, C., Ehrt, S., & Rhee, K. Y. (2010). Metabolomics of Mycobacterium tuberculosis reveals compartmentalized co-catabolism of carbon substrates. *Chemistry & biology*, *17*(10), 1122-1131.

de Martino, M., Galli, L., & Chiappini, E. (2014). Reflections on the immunology of tuberculosis: will we ever unravel the skein? *BMC infectious diseases*, *14*(1), 1-6.

de Martino, M., Lodi, L., Galli, L., & Chiappini, E. (2019). Immune response to Mycobacterium tuberculosis: a narrative review. *Frontiers in pediatrics*, *7*, 350.

de wet, T. J., Gobe, I., Mhlanga, M. M., & Warner, D. F. (2018). CRISPRi-Seq for the identification and characterisation of essential mycobacterial genes and transcriptional units. *BioRxiv*, 358275.

de Wet, T. J., Winkler, K. R., Mhlanga, M., Mizrahi, V., & Warner, D. F. (2020). Arrayed CRISPRi and quantitative imaging describe the morphotypic landscape of essential mycobacterial genes. *Elife*, *9*, e60083.

Debussche, L., Couder, M., Thibaut, D., Cameron, B., Crouzet, J., & Blanche, F. (1992). Assay, purification, and characterization of cobaltochelatase, a unique complex enzyme catalyzing cobalt insertion in hydrogenobyrinic acid a, c-diamide during coenzyme B12 biosynthesis in Pseudomonas denitrificans. *Journal of bacteriology*, *174*(22), 7445-7451.

Debussche, L., Thibaut, D., Cameron, B., Crouzet, J., & Blanche, F. (1993). Biosynthesis of the corrin macrocycle of coenzyme B12 in Pseudomonas denitrificans. *Journal of bacteriology*, *175*(22), 7430-7440.

Degnan, P. H., Barry, N. A., Mok, K. C., Taga, M. E., & Goodman, A. L. (2014). Human gut microbes use multiple transporters to distinguish vitamin B12 analogs and compete in the gut. *Cell host & microbe*, *15*(1), 47-57.

DeJesus, M. A., Ambadipudi, C., Baker, R., Sassetti, C., & Ioerger, T. R. (2015). TRANSIT-a software tool for Himar1 TnSeq analysis. *PLoS Comput Biol*, *11*(10), e1004401.

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. (2017). Comprehensive essentiality analysis of the Mycobacterium tuberculosis genome via saturating transposon mutagenesis. *MBio*, *8*(1).

DeJesus, M. A., & Ioerger, T. R. (2013). A Hidden Markov Model for identifying essential and growth-defect regions in bacterial genomes from transposon insertion sequencing data. *BMC bioinformatics*, *14*(1), 1-12.

Deutschbauer, A., Price, M. N., Wetmore, K. M., Shao, W., Baumohl, J. K., Xu, Z., Nguyen, M., Tamse, R., Davis, R. W., & Arkin, A. P. (2011). Evidence-based annotation of gene function in Shewanella oneidensis MR-1 using genome-wide fitness profiling across 121 conditions. *PLoS Genet*, *7*(11), e1002385.

DeVeaux, L. C., & Kadner, R. J. (1985). Transport of vitamin B12 in Escherichia coli: cloning of the btuCD region. *Journal of bacteriology*, *162*(3), 888-896.

Dheda, K., Gumbo, T., Gandhi, N. R., Murray, M., Theron, G., Udwadia, Z., Migliori, G., & Warren, R. (2014). Global control of tuberculosis: from extensively drug-resistant to untreatable tuberculosis. *The lancet Respiratory medicine*, *2*(4), 321-338.

Dheda, K., Gumbo, T., Maartens, G., Dooley, K. E., McNerney, R., Murray, M., Furin, J., Nardell, E. A., London, L., & Lessem, E. (2017). The epidemiology, pathogenesis, transmission, diagnosis, and management of multidrug-resistant, extensively drug-resistant, and incurable tuberculosis. *The lancet Respiratory medicine*, *5*(4), 291-360.

Domenech, P., Kobayashi, H., LeVier, K., Walker, G. C., & Barry, C. E. (2009). BacA, an ABC transporter involved in maintenance of chronic murine infections with Mycobacterium tuberculosis. *Journal of bacteriology*, *191*(2), 477-485.

Dowling, D. P., Miles, Z. D., Köhrer, C., Maiocco, S. J., Elliott, S. J., Bandarian, V., & Drennan, C. L. (2016). Molecular basis of cobalamin-dependent RNA modification. *Nucleic acids research*, *44*(20), 9965-9976.

Doxey, A. C., Kurtz, D. A., Lynch, M. D., Sauder, L. A., & Neufeld, J. D. (2015). Aquatic metagenomes implicate Thaumarchaeota in global cobalamin production. *The ISME journal*, *9*(2), 461-471.

Dragset, M. S., Ioerger, T. R., Zhang, Y. J., Mærk, M., Ginbot, Z., Sacchettini, J. C., Flo, T. H., Rubin, E. J., & Steigedal, M. (2019). Genome-wide phenotypic profiling identifies and categorizes genes required for mycobacterial low iron fitness. *Scientific reports*, *9*(1), 1-11.

Drain, P. K., Bajema, K. L., Dowdy, D., Dheda, K., Naidoo, K., Schumacher, S. G., Ma, S., Meermeier, E., Lewinsohn, D. M., & Sherman, D. R. (2018). Incipient and subclinical tuberculosis: a clinical review of early stages and progression of infection. *Clinical microbiology reviews*, *31*(4).

Drennan, C. L., Huang, S., Drummond, J. T., Matthews, R. G., & Lidwig, M. (1994). How a protein binds B12: A 3.0 A X-ray structure of B12-binding domains of methionine synthase. *Science*, *266*(5191), 1669-1674.

Duarte, R., Lönnroth, K., Carvalho, C., Lima, F., Carvalho, A., Muñoz-Torrico, M., & Centis, R. (2018). Tuberculosis, social determinants and co-morbidities (including HIV). *Pulmonology*, *24*(2), 115-119.

Dubnau, E., Chan, J., Mohan, V., & Smith, I. (2005). Responses of Mycobacterium tuberculosis to growth in the mouse lung. *Infection and immunity*, *73*(6), 3754-3757.

Ducati, R. G., Ruffino-Netto, A., Basso, L. A., & Santos, D. S. (2006). The resumption of consumption: a review on tuberculosis. *Memórias do Instituto Oswaldo Cruz*, *101*(7), 697-714.

Ehrt, S., Schnappinger, D., & Rhee, K. Y. (2018). Metabolic principles of persistence and pathogenicity in Mycobacterium tuberculosis. *Nature Reviews Microbiology*, *16*(8), 496-507.

Eitinger, T., Rodionov, D. A., Grote, M., & Schneider, E. (2011). Canonical and ECF-type ATP-binding cassette importers in prokaryotes: diversity in modular organization and cellular functions. *FEMS microbiology reviews*, *35*(1), 3-67.

Eitinger, T., Suhr, J., Moore, L., & Smith, J. A. C. (2005). Secondary transporters for nickel and cobalt ions: theme and variations. *Biometals, 18*(4), 399-405.

Eoh, H., & Rhee, K. Y. (2014). Methylcitrate cycle defines the bactericidal essentiality of isocitrate lyase for survival of Mycobacterium tuberculosis on fatty acids. *Proceedings of the National Academy of Sciences, 111*(13), 4976-4981.

Ernst, J. D. (1998). Macrophage receptors for Mycobacterium tuberculosis. *Infection and immunity, 66*(4), 1277-1281.

Escalante-Semerena, J., Suh, S., & Roth, J. (1990). cobA function is required for both de novo cobalamin biosynthesis and assimilation of exogenous corrinoids in Salmonella typhimurium. *Journal of bacteriology, 172*(1), 273-280.

Escalante-Semerena, J. C. (2007). Conversion of cobinamide into adenosylcobamide in bacteria and archaea. *Journal of bacteriology, 189*(13), 4555-4560.

Esmail, A., Sabur, N. F., Okpechi, I., & Dheda, K. (2018). Management of drug-resistant tuberculosis in special sub-populations including those with HIV co-infection, pregnancy, diabetes, organ-specific dysfunction, and in the critically ill. *Journal of thoracic disease, 10*(5), 3102.

Espinal, M. A., & Dye, C. (2005). Can DOTS control multidrug-resistant tuberculosis? *The Lancet, 365*(9466), 1206-1209.

Eum, S.-Y., Kong, J.-H., Hong, M.-S., Lee, Y.-J., Kim, J.-H., Hwang, S.-H., Cho, S.-N., Via, L. E., & Barry III, C. E. (2010). Neutrophils are the predominant infected phagocytic cells in the airways of patients with active pulmonary TB. *Chest, 137*(1), 122-128.

Evans, J. C., & Mizrahi, V. (2018). Priming the tuberculosis drug pipeline: new antimycobacterial targets and agents. *Current opinion in microbiology, 45*, 39-46.

Falzon, D., Gandhi, N., Migliori, G. B., Sotgiu, G., Cox, H. S., Holtz, T. H., Hollm-Delgado, M.-G., Keshavjee, S., DeRiemer, K., & Centis, R. (2013). Resistance to fluoroquinolones and second-line injectable drugs: impact on multidrug-resistant TB outcomes. *European Respiratory Journal, 42*(1), 156-168.

Fang, H., Kang, J., & Zhang, D. (2017). Microbial production of vitamin B 12: a review and future perspectives. *Microbial cell factories, 16*(1), 1-14.

Farhat, M. R., Jacobson, K. R., Franke, M. F., Kaur, D., Sloutsky, A., Mitnick, C. D., & Murray, M. (2016). Gyrase mutations are associated with variable levels of fluoroquinolone resistance in Mycobacterium tuberculosis. *Journal of clinical microbiology, 54*(3), 727-733.

Fenhalls, G., Stevens, L., Moses, L., Bezuidenhout, J., Betts, J. C., van Helden, P., Lukey, P. T., & Duncan, K. (2002). In situ detection of Mycobacterium tuberculosis transcripts in human lung granulomas reveals differential gene expression in necrotic lesions. *Infection and immunity, 70*(11), 6330-6338.

Feschotte, C., & Pritham, E. J. (2007). DNA transposons and the evolution of eukaryotic genomes. *Annu. Rev. Genet., 41*, 331-368.

Forrellad, M. A., Klepp, L. I., Gioffré, A., Sabio y Garcia, J., Morbidoni, H. R., Santangelo, M. D. L. P., Cataldi, A. A., & Bigi, F. (2013). Virulence factors of the Mycobacterium tuberculosis complex. *Virulence, 4*(1), 3-66.

Frank, D. J., Madrona, Y., & de Montellano, P. R. O. (2014). Cholesterol ester oxidation by mycobacterial cytochrome P450. *Journal of Biological Chemistry*, *289*(44), 30417-30425.

Friedmann, H. C., & Fyfe, J. A. (1969). Pseudovitamin B12 Biosynthesis: Enzymatic formation of a new adenylic acid, 7-α-D-ribofuranosyladenine 5′-phosphate. *Journal of Biological Chemistry*, *244*(7), 1667-1671.

Froese, D. S., Fowler, B., & Baumgartner, M. R. (2019). Vitamin B12, folate, and the methionine remethylation cycle—biochemistry, pathways, and regulation. *Journal of inherited metabolic disease*, *42*(4), 673-685.

Gallagher, L. A., Shendure, J., & Manoil, C. (2011). Genome-scale identification of resistance functions in Pseudomonas aeruginosa using Tn-seq. *MBio*, *2*(1).

Gandhi, N. R., Moll, A., Sturm, A. W., Pawinski, R., Govender, T., Lalloo, U., Zeller, K., Andrews, J., & Friedland, G. (2006). Extensively drug-resistant tuberculosis as a cause of death in patients co-infected with tuberculosis and HIV in a rural area of South Africa. *The Lancet*, *368*(9547), 1575-1580.

Gao, L.-Y., Groger, R., Cox, J. S., Beverley, S. M., Lawson, E. H., & Brown, E. J. (2003). Transposon mutagenesis of Mycobacterium marinum identifies a locus linking pigmentation and intracellular survival. *Infection and immunity*, *71*(2), 922-929.

Gawronski, J. D., Wong, S. M., Giannoukos, G., Ward, D. V., & 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), 16422-16427.

Gengenbacher, M., & Kaufmann, S. H. (2012). Mycobacterium tuberculosis: success through dormancy. *FEMS microbiology reviews*, *36*(3), 514-532.

Gerdes, S., Scholle, M., Campbell, J., Balazsi, G., Ravasz, E., Daugherty, M., Somera, A., Kyrpides, N., Anderson, I., & Gelfand, M. (2003). Experimental determination and system level analysis of essential genes in Escherichia coli MG1655. *Journal of bacteriology*, *185*(19), 5673-5684.

Gevers, D., Vandepoele, K., Simillion, C., & Van de Peer, Y. (2004). Gene duplication and biased functional retention of paralogs in bacterial genomes. *Trends in microbiology*, *12*(4), 148-154.

Gonzalez, J. C., Banerjee, R. V., Huang, S., Sumner, J. S., & Matthews, R. G. (1992). Comparison of cobalamin-independent and cobalamin-dependent methionine synthases from Escherichia coli: two solutions to the same chemical problem. *Biochemistry*, *31*(26), 6045-6056.

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

Gopinath, K., Moosa, A., Mizrahi, V., & Warner, D. F. (2013). Vitamin B12 metabolism in Mycobacterium tuberculosis. *Future microbiology*, *8*(11), 1405-1418.

Gopinath, K., Venclovas, Č., Ioerger, T. R., Sacchettini, J. C., McKinney, J. D., Mizrahi, V., & Warner, D. F. (2013). A vitamin B12 transporter in Mycobacterium tuberculosis. *Open biology*, *3*(2), 120175.

Gopinath, K., Warner, D. F., & Mizrahi, V. (2015). Targeted gene knockout and essentiality testing by homologous recombination. In *Mycobacteria Protocols* (pp. 131-149). Springer.

Gordon, S. V., Bottai, D., Simeone, R., Stinear, T. P., & Brosch, R. (2009). Pathogenicity in the tubercle bacillus: molecular and evolutionary determinants. *Bioessays*, *31*(4), 378-388.

Goryshin, I. Y., & Reznikoff, W. S. (1998). Tn5 in vitro transposition. *Journal of Biological Chemistry*, *273*(13), 7367-7374.

Gould, T. A., Van De Langemheen, H., Muñoz-Elías, E. J., McKinney, J. D., & Sacchettini, J. C. (2006). Dual role of isocitrate lyase 1 in the glyoxylate and methylcitrate cycles in Mycobacterium tuberculosis. *Molecular microbiology*, *61*(4), 940-947.

Gray, M. J., & Escalante-Semerena, J. C. (2007). Single-enzyme conversion of FMNH2 to 5, 6-dimethylbenzimidazole, the lower ligand of B12. *Proceedings of the National Academy of Sciences*, *104*(8), 2921-2926.

Green, J. M., & Matthews, R. G. (2007). Folate Biosynthesis, Reduction, and Polyglutamylation and the Interconversion of Folate Derivatives. *EcoSal Plus*, *2*(2).

Griffin, J. E., Gawronski, J. D., DeJesus, M. A., Ioerger, T. R., Akerley, B. J., & Sassetti, C. M. (2011). High-resolution phenotypic profiling defines genes essential for mycobacterial growth and cholesterol catabolism. *PLoS Pathog*, *7*(9), e1002251.

Griffin, J. E., Pandey, A. K., Gilmore, S. A., Mizrahi, V., Mckinney, J. D., Bertozzi, C. R., & Sassetti, C. M. (2012). Cholesterol catabolism by Mycobacterium tuberculosis requires transcriptional and metabolic adaptations. *Chemistry & biology*, *19*(2), 218-227.

Grosset, J. (2003). Mycobacterium tuberculosis in the extracellular compartment: an underestimated adversary. *Antimicrobial agents and chemotherapy*, *47*(3), 833-836.

Gruber, K., Puffer, B., & Kräutler, B. (2011). Vitamin B 12-derivatives—enzyme cofactors and ligands of proteins and nucleic acids. *Chemical Society Reviews*, *40*(8), 4346-4363.

Guzzo, M. B., Nguyen, H. T., Pham, T. H., Wyszczelska-Rokiel, M., Jakubowski, H., Wolff, K. A., Ogwang, S., Timpona, J. L., Gogula, S., & Jacobs, M. R. (2016). Methylfolate trap promotes bacterial thymineless death by sulfa drugs. *PLoS pathogens*, *12*(10), e1005949.

Haller, A., Souliere, M. F., & Micura, R. (2011). The dynamic nature of RNA as key to understanding riboswitch mechanisms. *Accounts of chemical research*, *44*(12), 1339-1348.

Halpern, J. (1985). Mechanisms of coenzyme B12-dependent rearrangements. *Science*, *227*(4689), 869-875.

Hamer, L., DeZwaan, T. M., Montenegro-Chamorro, M. V., Frank, S. A., & Hamer, J. E. (2001). Recent advances in large-scale transposon mutagenesis. *Current opinion in chemical biology*, *5*(1), 67-73.

Harding, C. V., & Boom, W. H. (2010). Regulation of antigen presentation by Mycobacterium tuberculosis: a role for Toll-like receptors. *Nature Reviews Microbiology*, *8*(4), 296-307.

Hazra, A. B., Han, A. W., Mehta, A. P., Mok, K. C., Osadchiy, V., Begley, T. P., & Taga, M. E. (2015). Anaerobic biosynthesis of the lower ligand of vitamin B12. *Proceedings of the National Academy of Sciences*, *112*(34), 10792-10797.

Held, K., Ramage, E., Jacobs, M., Gallagher, L., & Manoil, C. (2012). Sequence-verified two-allele transposon mutant library for Pseudomonas aeruginosa PAO1. *Journal of bacteriology*, *194*(23), 6387-6389.

Helliwell, K. E., Lawrence, A. D., Holzer, A., Kudahl, U. J., Sasso, S., Kräutler, B., Scanlan, D. J., Warren, M. J., & Smith, A. G. (2016). Cyanobacteria and eukaryotic algae use different chemical variants of vitamin B12. *Current Biology*, *26*(8), 999-1008.

Hendrixson, D. R., Akerley, B. J., & DiRita, V. J. (2001). Transposon mutagenesis of Campylobacter jejuni identifies a bipartite energy taxis system required for motility. *Molecular microbiology*, *40*(1), 214-224.

Hestvik, A. L. K., Hmama, Z., & Av-Gay, Y. (2005). Mycobacterial manipulation of the host cell. *FEMS microbiology reviews*, *29*(5), 1041-1050.

Hickman, S. P., Chan, J., & Salgame, P. (2002). Mycobacterium tuberculosis induces differential cytokine production from dendritic cells and macrophages with divergent effects on naive T cell polarization. *The Journal of Immunology*, *168*(9), 4636-4642.

Hmama, Z., Peña-Díaz, S., Joseph, S., & Av-Gay, Y. (2015). Immunoevasion and immunosuppression of the macrophage by M ycobacterium tuberculosis. *Immunological reviews*, *264*(1), 220-232.

Ho, N. A. T., Dawes, S. S., Crowe, A. M., Casabon, I., Gao, C., Kendall, S. L., Baker, E. N., Eltis, L. D., & Lott, J. S. (2016). The structure of the transcriptional repressor KstR in complex with CoA thioester cholesterol metabolites sheds light on the regulation of cholesterol catabolism in Mycobacterium tuberculosis. *Journal of Biological Chemistry*, *291*(14), 7256-7266.

Hobbs, E. C., Astarita, J. L., & Storz, G. (2010). Small RNAs and small proteins involved in resistance to cell envelope stress and acid shock in Escherichia coli: analysis of a bar-coded mutant collection. *Journal of bacteriology*, *192*(1), 59-67.

Hodgkin, D. C., Kamper, J., Mackay, M., Pickworth, J., Trueblood, K. N., & White, J. G. (1956). Structure of vitamin B12. *Nature*, *178*(4524), 64-66.

Hoffmann, B., Oberhuber, M., Stupperich, E., Bothe, H., Buckel, W., Konrat, R., & Kräutler, B. (2000). Native corrinoids from Clostridium cochlearium are adeninylcobamides: spectroscopic analysis and identification of pseudovitamin B12 and factor A. *Journal of bacteriology*, *182*(17), 4773-4782.

Högbom, M., Stenmark, P., Voevodskaya, N., McClarty, G., Gräslund, A., & Nordlund, P. (2004). The radical site in chlamydial ribonucleotide reductase defines a new R2 subclass. *Science*, *305*(5681), 245-248.

Horswill, A. R., & Escalante-Semerena, J. C. (1999). Salmonella typhimurium LT2 catabolizes propionate via the 2-methylcitric acid cycle. *Journal of bacteriology*, *181*(18), 5615-5623.

Ignatov, D., & Johansson, J. (2017). RNA-mediated signal perception in pathogenic bacteria. *Wiley Interdisciplinary Reviews: RNA*, *8*(6), e1429.

Jacobs, M. A., Alwood, A., Thaipisuttikul, I., Spencer, D., Haugen, E., Ernst, S., Will, O., Kaul, R., Raymond, C., & Levy, R. (2003). Comprehensive transposon mutant library of Pseudomonas aeruginosa. *Proceedings of the National Academy of Sciences*, *100*(24), 14339-14344.

James, B., Williams, A., & Marsh, P. (2000). The physiology and pathogenicity of Mycobacterium tuberculosis grown under controlled conditions in a defined medium. *Journal of applied microbiology*, *88*(4), 669-677.

Jangam, D., Feschotte, C., & Betrán, E. (2017). Transposable element domestication as an adaptation to evolutionary conflicts. *Trends in Genetics*, *33*(11), 817-831.

Jayachandran, R., Sundaramurthy, V., Combaluzier, B., Mueller, P., Korf, H., Huygen, K., Miyazaki, T., Albrecht, I., Massner, J., & Pieters, J. (2007). Survival of mycobacteria in macrophages is mediated by coronin 1-dependent activation of calcineurin. *Cell*, *130*(1), 37-50.

Johnson Jr, J. E., Reyes, F. E., Polaski, J. T., & Batey, R. T. (2012). B 12 cofactors directly stabilize an mRNA regulatory switch. *Nature*, *492*(7427), 133-137.

Johnson, R., Streicher, E. M., Louw, G. E., Warren, R. M., Van Helden, P. D., & Victor, T. (2007). Drug resistance in M. tuberculosis. *Understanding the mechanisms of drug resistance in enhancing rapid molecular detection of drug resistance in Mycobacterium tuberculosis*, *7*.

Jordan, A., & Reichard, P. (1998). Ribonucleotide reductases. *Annual review of biochemistry*, *67*(1), 71-98.

Kang, P. B., Azad, A. K., Torrelles, J. B., Kaufman, T. M., Beharka, A., Tibesar, E., DesJardin, L. E., & Schlesinger, L. S. (2005). The human macrophage mannose receptor directs Mycobacterium tuberculosis lipoarabinomannan-mediated phagosome biogenesis. *The Journal of experimental medicine*, *202*(7), 987-999.

Kang, Z., Zhang, J., Zhou, J., Qi, Q., Du, G., & Chen, J. (2012). Recent advances in microbial production of δ-aminolevulinic acid and vitamin B12. *Biotechnology advances*, *30*(6), 1533-1542.

Kapopoulou, A., Lew, J. M., & Cole, S. T. (2011). The MycoBrowser portal: a comprehensive and manually annotated resource for mycobacterial genomes. *Tuberculosis*, *91*(1), 8-13.

Kaufmann, S. H., Lange, C., Rao, M., Balaji, K. N., Lotze, M., Schito, M., Zumla, A. I., & Maeurer, M. (2014). Progress in tuberculosis vaccine development and host-directed therapies—a state of the art review. *The lancet Respiratory medicine*, *2*(4), 301-320.

Keck, B., Munder, M., & Renz, P. (1998). Biosynthesis of cobalamin in Salmonella typhimurium: transformation of riboflavin into the 5, 6-dimethylbenzimidazole moiety. *Archives of microbiology*, *171*(1), 66-68.

Keck, B., & Renz, P. (2000). Salmonella typhimurium forms adenylcobamide and 2-methyladenylcobamide, but no detectable cobalamin during strictly anaerobic growth. *Archives of microbiology*, *173*(1), 76-77.

Kidwell, M. G., & Lisch, D. R. (2001). Perspective: transposable elements, parasitic DNA, and genome evolution. *Evolution*, *55*(1), 1-24.

Kipkorir, T., Mashabela, G. T., Timothy, J., Koch, A., Wiesner, L., Mizrahi, V., & Warner, D. F. (2021). De Novo Cobalamin Biosynthesis, Transport, and Assimilation and Cobalamin-Mediated Regulation of Methionine Biosynthesis in Mycobacterium smegmatis. *Journal of bacteriology*, *203*(7).

Knuth, K., Niesalla, H., Hueck, C. J., & Fuchs, T. M. (2004). Large-scale identification of essential Salmonella genes by trapping lethal insertions. *Molecular microbiology*, *51*(6), 1729-1744.

Kobayashi, K., Ehrlich, S. D., Albertini, A., Amati, G., Andersen, K., Arnaud, M., Asai, K., Ashikaga, S., Aymerich, S., & Bessieres, P. (2003). Essential Bacillus subtilis genes. *Proceedings of the National Academy of Sciences*, *100*(8), 4678-4683.

Kobayashi, M., & Shimizu, S. (1999). Cobalt proteins. *European journal of biochemistry*, *261*(1), 1-9.

Kochi, A., Vareldzis, B., & Styblo, K. (1993). Multidrug-resistant tuberculosis and its control. *Research in microbiology*, *144*(2), 104-110.

Korkhov, V. M., Mireku, S. A., & Locher, K. P. (2012). Structure of AMP-PNP-bound vitamin B 12 transporter BtuCD–F. *Nature*, *490*(7420), 367-372.

Koutmos, M., Datta, S., Pattridge, K. A., Smith, J. L., & Matthews, R. G. (2009). Insights into the reactivation of cobalamin-dependent methionine synthase. *Proceedings of the National Academy of Sciences*, *106*(44), 18527-18532.

Kräutler, B. (2005). Vitamin B12: chemistry and biochemistry. *Biochemical Society Transactions*, *33*(4), 806-810.

Krautler, B., Moll, J., & Thauer, R. K. (1987). The corrinoid from Methanobacterium thermoautotrophicum (Marburg strain) Spectroscopic structure analysis and identification as Coβ-cyano-5′-hydroxybenzimidazolyl-cobamide (factor III). *European journal of biochemistry*, *162*(2), 275-278.

Krithika, R., Marathe, U., Saxena, P., Ansari, M. Z., Mohanty, D., & Gokhale, R. S. (2006). A genetic locus required for iron acquisition in Mycobacterium tuberculosis. *Proceedings of the National Academy of Sciences*, *103*(7), 2069-2074.

Kubodera, T., Watanabe, M., Yoshiuchi, K., Yamashita, N., Nishimura, A., Nakai, S., Gomi, K., & Hanamoto, H. (2003). Thiamine-regulated gene expression of Aspergillus oryzae thiA requires splicing of the intron containing a riboswitch-like domain in the 5′-UTR. *FEBS letters*, *555*(3), 516-520.

Kulasekara, H. D. (2014). Transposon mutagenesis. In *Pseudomonas Methods and Protocols* (pp. 501-519). Springer.

Kwon, Y. M., Ricke, S. C., & Mandal, R. K. (2016). Transposon sequencing: methods and expanding applications. *Applied microbiology and biotechnology*, *100*(1), 31-43.

Lamm, L., Heckmann, G., & Renz, P. (1982). Biosynthesis of Vitamin B12 in Anaerobic Bacteria: Mode of Incorporation of Glycine into the 5, 6-Dimethylbenzimidazole Moiety in Eubactevium limosum. *European journal of biochemistry*, *122*(3), 569-571.

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. (2009). Simultaneous assay of every Salmonella Typhi gene using one million transposon mutants. *Genome research*, *19*(12), 2308-2316.

Larsen, M. (2000). Appendix 1. Molecular Genetics of Mycobacteria. GF Hatfull and WR Jacobs, Jr. In: Washington, DC, ASM Press.

Laurenzi, M., Ginsberg, A., & Spigelman, M. (2007). Challenges associated with current and future TB treatment. *Infectious Disorders-Drug Targets (Formerly Current Drug Targets-Infectious Disorders)*, *7*(2), 105-119.

Lawrence, A. D., Deery, E., McLean, K. J., Munro, A. W., Pickersgill, R. W., Rigby, S. E., & Warren, M. J. (2008). Identification, characterization, and structure/function analysis of a corrin reductase involved in adenosylcobalamin biosynthesis. *Journal of Biological Chemistry*, *283*(16), 10813-10821.

Lawrence, J. G., & Roth, J. R. (1995). The cobalamin (coenzyme B12) biosynthetic genes of Escherichia coli. *Journal of bacteriology*, *177*(22), 6371-6380.

Lee, W., VanderVen, B. C., Fahey, R. J., & Russell, D. G. (2013). Intracellular Mycobacterium tuberculosis exploits host-derived fatty acids to limit metabolic stress. *Journal of Biological Chemistry*, *288*(10), 6788-6800.

Lenhert, P. G., & Hodgkin, D. C. (1961). Structure of the 5, 6-dimethylbenzimidazolylcobamide coenzyme. *Nature*, *192*(4806), 937-938.

Lerner, T. R., Borel, S., & Gutierrez, M. G. (2015). The innate immune response in human tuberculosis. *Cellular microbiology*, *17*(9), 1277-1285.

Lezius, A., & Barker, H. (1965). Corrinoid compounds of Methanobacillus omelianskii. I. Fractionation of the corrinoid compounds and identification of factor III and factor III coenzyme. *Biochemistry*, *4*(3), 510-518.

Li, J., Ge, Y., Zadeh, M., Curtiss, R., & Mohamadzadeh, M. (2020). Regulating vitamin B12 biosynthesis via the cbiMCbl riboswitch in Propionibacterium strain UF1. *Proceedings of the National Academy of Sciences*, *117*(1), 602-609.

Liberati, N. T., Urbach, J. M., Miyata, S., Lee, D. G., Drenkard, E., Wu, G., Villanueva, J., Wei, T., & Ausubel, F. M. (2006). An ordered, nonredundant library of Pseudomonas aeruginosa strain PA14 transposon insertion mutants. *Proceedings of the National Academy of Sciences*, *103*(8), 2833-2838.

Lin, P. L., & Flynn, J. L. (2018). The end of the binary era: revisiting the spectrum of tuberculosis. *The Journal of Immunology*, *201*(9), 2541-2548.

Liu, C. H., Liu, H., & Ge, B. (2017). Innate immunity in tuberculosis: host defense vs pathogen evasion. *Cellular & molecular immunology*, *14*(12), 963-975.

Liu, Z.-M., Tucker, A. M., Driskell, L. O., & Wood, D. O. (2007). Mariner-based transposon mutagenesis of Rickettsia prowazekii. *Applied and environmental microbiology*, *73*(20), 6644-6649.

Locher, K. P. (2004). Structure and mechanism of ABC transporters. *Current opinion in structural biology*, *14*(4), 426-431.

Loh, E., Dussurget, O., Gripenland, J., Vaitkevicius, K., Tiensuu, T., Mandin, P., Repoila, F., Buchrieser, C., Cossart, P., & Johansson, J. (2009). A trans-acting riboswitch controls expression of the virulence regulator PrfA in Listeria monocytogenes. *Cell*, *139*(4), 770-779.

Long, J. E., DeJesus, M., Ward, D., Baker, R. E., Ioerger, T., & Sassetti, C. M. (2015). Identifying essential genes in Mycobacterium tuberculosis by global phenotypic profiling. In *Gene Essentiality* (pp. 79-95). Springer.

Lönnroth, K., Castro, K. G., Chakaya, J. M., Chauhan, L. S., Floyd, K., Glaziou, P., & Raviglione, M. C. (2010). Tuberculosis control and elimination 2010–50: cure, care, and social development. *The Lancet*, *375*(9728), 1814-1829.

Lovewell, R. R., Sassetti, C. M., & VanderVen, B. C. (2016). Chewing the fat: lipid metabolism and homeostasis during M. tuberculosis infection. *Current opinion in microbiology*, *29*, 30-36.

Lu, R., Schmitz, W., & Sampson, N. S. (2015). α-methyl acyl CoA racemase provides Mycobacterium tuberculosis catabolic access to cholesterol esters. *Biochemistry*, *54*(37), 5669-5672.

Ludwig, M. L., Drennan, C. L., & Matthews, R. G. (1996). The reactivity of B12 cofactors: the proteins make a difference. *Structure*, *4*(5), 505-512.

Lugo-Villarino, G., Vérollet, C., Maridonneau-Parini, I., & Neyrolles, O. (2011). Macrophage polarization: convergence point targeted by mycobacterium tuberculosis and HIV. *Frontiers in immunology*, *2*, 43.

Lukácsi, S., Mácsik-Valent, B., Nagy-Baló, Z., Kovács, K. G., Kliment, K., Bajtay, Z., & Erdei, A. (2020). Utilization of complement receptors in immune cell–microbe interaction. *FEBS letters*, *594*(16), 2695-2713.

Lundrigan, M. D., Köster, W., & Kadner, R. J. (1991). Transcribed sequences of the Escherichia coli btuB gene control its expression and regulation by vitamin B12. *Proceedings of the National Academy of Sciences*, *88*(4), 1479-1483.

Maguire, M. E. (2006). Magnesium transporters: properties, regulation and structure. *Front Biosci*, *11*, 3149-3163.

Majumdar, G., Mbau, R., Singh, V., Warner, D. F., Dragset, M. S., & Mukherjee, R. (2017). Genome-wide transposon mutagenesis in Mycobacterium tuberculosis and Mycobacterium smegmatis. In *In Vitro Mutagenesis* (pp. 321-335). Springer.

Malik, Z. A., Denning, G. M., & Kusner, D. J. (2000). Inhibition of Ca2+ signaling by Mycobacterium tuberculosis is associated with reduced phagosome–lysosome fusion and increased survival within human macrophages. *The Journal of experimental medicine*, *191*(2), 287-302.

Mandal, M., & Breaker, R. R. (2004). Gene regulation by riboswitches. *Nature reviews Molecular cell biology*, *5*(6), 451-463.

Marahatta, S. (2010). Multi-drug resistant tuberculosis burden and risk factors: an update. *Kathmandu University medical journal*, *8*(1), 116-125.

Marrero, J., Rhee, K. Y., Schnappinger, D., Pethe, K., & Ehrt, S. (2010). Gluconeogenic carbon flow of tricarboxylic acid cycle intermediates is critical for Mycobacterium tuberculosis to establish and maintain infection. *Proceedings of the National Academy of Sciences*, *107*(21), 9819-9824.

Marsh, E. (1999). Coenzyme B12 (cobalamin)-dependent enzymes. *Essays in biochemistry*, *34*, 139-154.

Martens, J.-H., Barg, H., Warren, M. a., & Jahn, D. (2002). Microbial production of vitamin B 12. *Applied microbiology and biotechnology*, *58*(3), 275-285.

Mashabela, G. T., Wet, T. J. D., & Warner, D. F. (2019). Mycobacterium tuberculosis metabolism. *Gram-Positive Pathogens*, 1107-1128.

Mattes, T., Deery, E., Warren, M., & Escalante-Semerena, J. (2017). Cobalamin Biosynthesis and Insertion In: Encyclopedia of Inorganic and Bioinorganic Chemistry. Scott RA. In. Chichester, UK: John Wiley & Sons, Ltd.

Matthews, R. G., & Goulding, C. W. (1997). Enzyme-catalyzed methyl transfers to thiols: the role of zinc. *Current opinion in chemical biology*, *1*(3), 332-339.

Matthews, R. G., Smith, A. E., Zhou, Z. S., Taurog, R. E., Bandarian, V., Evans, J. C., & Ludwig, M. (2003). Cobalamin-dependent and cobalamin-independent methionine synthases: Are there two solutions to the same chemical problem? *Helvetica chimica acta*, *86*(12), 3939-3954.

McKinney, J. D., Zu Bentrup, K. H., Muñoz-Elías, E. J., Miczak, A., Chen, B., Chan, W.-T., Swenson, D., Sacchettini, J. C., Jacobs, W. R., & Russell, D. G. (2000). Persistence of Mycobacterium tuberculosis in macrophages and mice requires the glyoxylate shunt enzyme isocitrate lyase. *Nature*, *406*(6797), 735-738.

Mellin, J., Koutero, M., Dar, D., Nahori, M.-A., Sorek, R., & Cossart, P. (2014). Sequestration of a two-component response regulator by a riboswitch-regulated noncoding RNA. *Science*, *345*(6199), 940-943.

Mihret, A. (2012). The role of dendritic cells in Mycobacterium tuberculosis infection. *Virulence*, *3*(7), 654-659.

Miller, B. H., Fratti, R. A., Poschet, J. F., Timmins, G. S., Master, S. S., Burgos, M., Marletta, M. A., & Deretic, V. (2004). Mycobacteria inhibit nitric oxide synthase recruitment to phagosomes during macrophage infection. *Infection and immunity*, *72*(5), 2872-2878.

Min, C., Atshaves, B. P., Roessner, C. A., Stolowich, N. J., Spencer, J. B., & Scott, A. I. (1993). Isolation, structure, and genetically engineered synthesis of precorrin-5, the pentamethylated intermediate of vitamin B12 biosynthesis. *Journal of the American Chemical Society*, *115*(22), 10380-10381.

Minato, Y., Gohl, D. M., Thiede, J. M., Chacón, J. M., Harcombe, W. R., Maruyama, F., & Baughn, A. D. (2019). Genomewide assessment of Mycobacterium tuberculosis conditionally essential metabolic pathways. *Msystems*, *4*(4).

Minato, Y., Thiede, J. M., Kordus, S. L., McKlveen, E. J., Turman, B. J., & Baughn, A. D. (2015). Mycobacterium tuberculosis folate metabolism and the mechanistic basis for para-aminosalicylic acid susceptibility and resistance. *Antimicrobial agents and chemotherapy*, *59*(9), 5097-5106.

Minias, A., Minias, P., Czubat, B., & Dziadek, J. (2018). Purifying selective pressure suggests the functionality of a vitamin B12 biosynthesis pathway in a global population of Mycobacterium tuberculosis. *Genome biology and evolution*, *10*(9), 2326-2337.

Mironov, A. S., Gusarov, I., Rafikov, R., Lopez, L. E., Shatalin, K., Kreneva, R. A., Perumov, D. A., & Nudler, E. (2002). Sensing small molecules by nascent RNA: a mechanism to control transcription in bacteria. *Cell*, *111*(5), 747-756.

Moore, E., Mander, A., Ames, D., Carne, R., Sanders, K., & Watters, D. (2012). Cognitive impairment and vitamin B12: a review. *International psychogeriatrics*, *24*(4), 541.

Moore, S. J., Lawrence, A. D., Biedendieck, R., Deery, E., Frank, S., Howard, M. J., Rigby, S. E., & Warren, M. J. (2013). Elucidation of the anaerobic pathway for the corrin component of cobalamin (vitamin B12). *Proceedings of the National Academy of Sciences*, *110*(37), 14906-14911.

Moosa, A. (2013). *Molecular mechanisms of transport and metabolism of vitamin B12 in mycobacteria*

Mortaz, E., Adcock, I. M., Tabarsi, P., Masjedi, M. R., Mansouri, D., Velayati, A. A., Casanova, J.-L., & Barnes, P. J. (2015). Interaction of pattern recognition receptors with Mycobacterium tuberculosis. *Journal of clinical immunology*, *35*(1), 1-10.

Mukhopadhyay, S., Nair, S., & Ghosh, S. (2012). Pathogenesis in tuberculosis: transcriptomic approaches to unraveling virulence mechanisms and finding new drug targets. *FEMS microbiology reviews*, *36*(2), 463-485.

Muñoz-López, M., & García-Pérez, J. L. (2010). DNA transposons: nature and applications in genomics. *Current genomics*, *11*(2), 115-128.

Muñoz-Elías, E. J., & McKinney, J. D. (2006). Carbon metabolism of intracellular bacteria. *Cellular microbiology*, *8*(1), 10-22.

Muñoz-Elías, E. J., Upton, A. M., Cherian, J., & McKinney, J. D. (2006). Role of the methylcitrate cycle in Mycobacterium tuberculosis metabolism, intracellular growth, and virulence. *Molecular microbiology*, *60*(5), 1109-1122.

Murray, G. L., Morel, V., Cerqueira, G. M., Croda, J., Srikram, A., Henry, R., Ko, A. I., Dellagostin, O. A., Bulach, D. M., & Sermswan, R. W. (2009). Genome-wide transposon mutagenesis in pathogenic Leptospira species. *Infection and immunity*, *77*(2), 810-816.

Murry, J. P., Sassetti, C. M., Lane, J. M., Xie, Z., & Rubin, E. J. (2008). Transposon site hybridization in Mycobacterium tuberculosis. In *Microbial Gene Essentiality: Protocols and Bioinformatics* (pp. 45-59). Springer.

Nahvi, A., Barrick, J. E., & Breaker, R. R. (2004). Coenzyme B12 riboswitches are widespread genetic control elements in prokaryotes. *Nucleic acids research*, *32*(1), 143-150.

Nahvi, A., Sudarsan, N., Ebert, M. S., Zou, X., Brown, K. L., & Breaker, R. R. (2002). Genetic control by a metabolite binding mRNA. *Chemistry & biology*, *9*(9), 1043-1049.

Nelson, D. L., & Kennedy, E. P. (1971). Magnesium transport in Escherichia coli: inhibition by cobaltous ion. *Journal of Biological Chemistry*, *246*(9), 3042-3049.

Ngabonziza, J. C. S., Loiseau, C., Marceau, M., Jouet, A., Menardo, F., Tzfadia, O., Antoine, R., Niyigena, E. B., Mulders, W., & Fissette, K. (2020). A sister lineage of the Mycobacterium tuberculosis complex discovered in the African Great Lakes region. *Nature communications*, *11*(1), 1-11.

Niegowski, D., & Eshaghi, S. (2007). The CorA family: structure and function revisited. *Cellular and molecular life sciences*, *64*(19), 2564-2574.

Nielsen, M. J., Rasmussen, M. R., Andersen, C. B., Nexø, E., & Moestrup, S. K. (2012). Vitamin B 12 transport from food to the body's cells—a sophisticated, multistep pathway. *Nature reviews Gastroenterology & hepatology*, *9*(6), 345-354.

Nies, D. H. (2003). Efflux-mediated heavy metal resistance in prokaryotes. *FEMS microbiology reviews*, *27*(2-3), 313-339.

Nijhout, H. F., Reed, M. C., Budu, P., & Ulrich, C. M. (2004). A mathematical model of the folate cycle: new insights into folate homeostasis. *Journal of Biological Chemistry*, *279*(53), 55008-55016.

Noinaj, N., Guillier, M., Barnard, T. J., & Buchanan, S. K. (2010). TonB-dependent transporters: regulation, structure, and function. *Annual review of microbiology*, *64*, 43-60.

Nordlund, P., & Reichard, P. (2006). Ribonucleotide reductases. *Annu. Rev. Biochem.*, *75*, 681-706.

Nou, X., & Kadner, R. J. (2000). Adenosylcobalamin inhibits ribosome binding to btuB RNA. *Proceedings of the National Academy of Sciences*, *97*(13), 7190-7195.

Novichkov, P. S., Kazakov, A. E., Ravcheev, D. A., Leyn, S. A., Kovaleva, G. Y., Sutormin, R. A., Kazanov, M. D., Riehl, W., Arkin, A. P., & Dubchak, I. (2013). RegPrecise 3.0–a resource for genome-scale exploration of transcriptional regulation in bacteria. *BMC genomics*, *14*(1), 1-12.

Oh-Hama, T., Stolowich, N., & Scott, A. (1988). 5-Aminolevulinic acid formation from glutamate via the C5 pathway in Clostridium thermoaceticum. *FEBS letters*, *228*(1), 89-93.

Okamoto, S., & Eltis, L. D. (2011). The biological occurrence and trafficking of cobalt. *Metallomics*, *3*(10), 963-970.

Osman, D., Cooke, A., Young, T. R., Deery, E., Robinson, N. J., & Warren, M. J. (2020). The requirement for cobalt in vitamin B12: A paradigm for protein metalation. *Biochimica et Biophysica Acta (BBA)-Molecular Cell Research*, 118896.

Palmer, A. M., Kamynina, E., Field, M. S., & Stover, P. J. (2017). Folate rescues vitamin B12 depletion-induced inhibition of nuclear thymidylate biosynthesis and genome instability. *Proceedings of the National Academy of Sciences*, *114*(20), E4095-E4102.

Parish, T., & Stoker, N. G. (2000). Use of a flexible cassette method to generate a double unmarked Mycobacterium tuberculosis tlyA plcABC mutant by gene replacement. *Microbiology*, *146*(8), 1969-1975.

Parveen, N., & Cornell, K. A. (2011). Methylthioadenosine/S-adenosylhomocysteine nucleosidase, a critical enzyme for bacterial metabolism. *Molecular microbiology*, *79*(1), 7-20.

Passemar, C., Arbués, A., Malaga, W., Mercier, I., Moreau, F., Lepourry, L., Neyrolles, O., Guilhot, C., & Astarie-Dequeker, C. (2014). Multiple deletions in the polyketide synthase gene repertoire of M ycobacterium tuberculosis reveal functional overlap of cell envelope lipids in host–pathogen interactions. *Cellular microbiology*, *16*(2), 195-213.

Pavelka Jr, M. S., Mahapatra, S., & Crick, D. C. (2014). Genetics of peptidoglycan biosynthesis. *Molecular Genetics of Mycobacteria*, 511-533.

Payne, K. A., Quezada, C. P., Fisher, K., Dunstan, M. S., Collins, F. A., Sjuts, H., Levy, C., Hay, S., Rigby, S. E., & Leys, D. (2015). Reductive dehalogenase structure suggests a mechanism for B12-dependent dehalogenation. *Nature*, *517*(7535), 513-516.

Pejchal, R., & Ludwig, M. L. (2004). Cobalamin-independent methionine synthase (MetE): a face-to-face double barrel that evolved by gene duplication. *PLoS Biol*, *3*(2), e31.

Pepper, D. J., Meintjes, G. A., McIlleron, H., & Wilkinson, R. J. (2007). Combined therapy for tuberculosis and HIV-1: the challenge for drug discovery. *Drug discovery today*, *12*(21-22), 980-989.

Perdrizet, G. A., Artsimovitch, I., Furman, R., Sosnick, T. R., & Pan, T. (2012). Transcriptional pausing coordinates folding of the aptamer domain and the expression platform of a riboswitch. *Proceedings of the National Academy of Sciences*, *109*(9), 3323-3328.

Pérez, A. A., Liu, Z., Rodionov, D. A., Li, Z., & Bryant, D. A. (2016). Complementation of cobalamin auxotrophy in Synechococcus sp. strain PCC 7002 and validation of a putative cobalamin riboswitch in vivo. *Journal of bacteriology*, *198*(19), 2743-2752.

Pérez, A. A., Rodionov, D. A., & Bryant, D. A. (2016). Identification and regulation of genes for cobalamin transport in the cyanobacterium Synechococcus sp. strain PCC 7002. *Journal of bacteriology*, *198*(19), 2753-2761.

Peselis, A., & Serganov, A. (2012). Structural insights into ligand binding and gene expression control by an adenosylcobalamin riboswitch. *Nature structural & molecular biology*, *19*(11), 1182.

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. (2016). A comprehensive, CRISPR-based functional analysis of essential genes in bacteria. *Cell*, *165*(6), 1493-1506.

Piwowarek, K., Lipińska, E., Hać-Szymańczuk, E., Kieliszek, M., & Ścibisz, I. (2018). Propionibacterium spp.—source of propionic acid, vitamin B12, and other metabolites important for the industry. *Applied microbiology and biotechnology*, *102*(2), 515-538.

Pol, A., van der Drift, C., & Vogels, G. D. (1982). Corrinoids from Methanosarcina barkeri: structure of the α-ligand. *Biochemical and biophysical research communications*, *108*(2), 731-737.

Polaski, J. T., Webster, S. M., Johnson Jr, J. E., & Batey, R. T. (2017). Cobalamin riboswitches exhibit a broad range of ability to discriminate between methylcobalamin and adenosylcobalamin. *Journal of Biological Chemistry*, *292*(28), 11650-11658.

Pritchard, J. R., Chao, M. C., Abel, S., Davis, B. M., Baranowski, C., Zhang, Y. J., Rubin, E. J., & Waldor, M. K. (2014). ARTIST: high-resolution genome-wide assessment of fitness using transposon-insertion sequencing. *PLoS Genet*, *10*(11), e1004782.

Qi, L. S., Larson, M. H., Gilbert, L. A., Doudna, J. A., Weissman, J. S., Arkin, A. P., & Lim, W. A. (2013). Repurposing CRISPR as an RNA-guided platform for sequence-specific control of gene expression. *Cell*, *152*(5), 1173-1183.

Queval, C. J., Brosch, R., & Simeone, R. (2017). The macrophage: a disputed fortress in the battle against Mycobacterium tuberculosis. *Frontiers in microbiology*, *8*, 2284.

Rachman, H., Strong, M., Ulrichs, T., Grode, L., Schuchhardt, J., Mollenkopf, H., Kosmiadi, G. A., Eisenberg, D., & Kaufmann, S. H. (2006). Unique transcriptome signature of Mycobacterium tuberculosis in pulmonary tuberculosis. *Infection and immunity*, *74*(2), 1233-1242.

Ramakrishnan, L. (2012). Revisiting the role of the granuloma in tuberculosis. *Nature Reviews Immunology*, *12*(5), 352-366.

Rancati, G., Moffat, J., Typas, A., & Pavelka, N. (2018). Emerging and evolving concepts in gene essentiality. *Nature Reviews Genetics*, *19*(1), 34.

Randaccio, L., Geremia, S., Demitri, N., & Wuerges, J. (2010). Vitamin B12: unique metalorganic compounds and the most complex vitamins. *Molecules*, *15*(5), 3228-3259.

Raux, E., Lanois, A., Warren, M. J., Rambach, A., & Thermes, C. (1998). Cobalamin (vitamin B12) biosynthesis: identification and characterization of a Bacillus megaterium cobI operon. *Biochemical Journal*, *335*(1), 159-166.

Raux, E., Schubert, H., & Warren, M. (2000). Biosynthesis of cobalamin (vitamin B12): a bacterial conundrum. *Cellular and Molecular Life Sciences CMLS*, *57*(13), 1880-1893.

Raux, E., Schubert, H. L., Roper, J. M., Wilson, K. S., & Warren, M. J. (1999). Vitamin B12: insights into biosynthesis's mount improbable. *Bioorganic chemistry*, *27*(2), 100-118.

Ravnum, S., & Andersson, D. (1997). Vitamin B12 repression of the btuB gene in Salmonella typhimurium is mediated via a translational control which requires leader and coding sequences. *Molecular microbiology*, *23*(1), 35-42.

Rees, D. C., Johnson, E., & Lewinson, O. (2009). ABC transporters: the power to change. *Nature reviews Molecular cell biology*, *10*(3), 218-227.

Rempel, S., Gati, C., Nijland, M., Thangaratnarajah, C., Karyolaimos, A., de Gier, J. W., Guskov, A., & Slotboom, D. (2020). A mycobacterial ABC transporter mediates the uptake of hydrophilic compounds. *Nature*, *580*(7803), 409-412.

Renz, P. (1970). Riboflavin as precursor in the biosynthesis of the 5, 6-dimethylbenzimidazole-moiety of vitamin B12. *FEBS letters*, *6*(3), 187-189.

Renz, P. (1999). Biosynthesis of the 5, 6-dimethylbenzimidazole moiety of cobalamin and of the other bases found in natural corrinoids. *Chemistry and Biochemistry of B12.*, 557-575.

Reyrat, J.-M., & Kahn, D. (2001). Mycobacterium smegmatis: an absurd model for tuberculosis? *Trends in microbiology*, *9*(10), 472-474.

Reznikoff, W. S. (1993). The TN5 transposon. *Annual review of microbiology*, *47*(1), 945-963.

Reznikoff, W. S., & Winterberg, K. M. (2008). Transposon-based strategies for the identification of essential bacterial genes. *Microbial Gene Essentiality: Protocols and Bioinformatics*, 13-26.

Richter-Dahlfors, A., Ravnum, S., & Andersson, D. (1994). Vitamin B12 repression of the cob operon in Salmonella typhimurium: translational control of the cbiA gene. *Molecular microbiology*, *13*(3), 541-553.

Rickes, E. L., Brink, N. G., Koniuszy, F. R., Wood, T. R., & Folkers, K. (1948). Crystalline vitamin B12. *Science (Washington)*, *107*, 396-397.

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. (2017). Programmable transcriptional repression in mycobacteria using an orthogonal CRISPR interference platform. *Nature microbiology*, *2*(4), 1-9.

Rodionov, D. A., Hebbeln, P., Eudes, A., Ter Beek, J., Rodionova, I. A., Erkens, G. B., Slotboom, D. J., Gelfand, M. S., Osterman, A. L., & Hanson, A. D. (2009). A novel

class of modular transporters for vitamins in prokaryotes. *Journal of bacteriology*, *191*(1), 42-51.

Rodionov, D. A., Hebbeln, P., Gelfand, M. S., & Eitinger, T. (2006). Comparative and functional genomic analysis of prokaryotic nickel and cobalt uptake transporters: evidence for a novel group of ATP-binding cassette transporters. *Journal of bacteriology*, *188*(1), 317-327.

Rodionov, D. A., Vitreschak, A. G., Mironov, A. A., & Gelfand, M. S. (2003). Comparative genomics of the vitamin B12 metabolism and regulation in prokaryotes. *Journal of Biological Chemistry*, *278*(42), 41148-41159.

Roman, R., Iluc, E., Mustea, A., Neacsu, A., & Asandului, V. (2001). Optimization of Medium Components in Vitamin B~ 1~ 2 Biosynthesis. *ROMANIAN BIOTECHNOLOGICAL LETTERS*, *6*, 343-350.

Roth, A., & Breaker, R. R. (2009). The structural and functional diversity of metabolite-binding riboswitches. *Annual review of biochemistry*, *78*, 305-334.

Roth, J. R., Lawrence, J., & Bobik, T. (1996). Cobalamin (coenzyme B12): synthesis and biological significance. *Annual review of microbiology*, *50*(1), 137-181.

Roth, J. R., Lawrence, J. G., Rubenfield, M., Kieffer-Higgins, S., & Church, G. M. (1993). Characterization of the cobalamin (vitamin B12) biosynthetic genes of Salmonella typhimurium. *Journal of bacteriology*, *175*(11), 3303-3316.

Rousset, F., Cabezas-Caballero, J., Piastra-Facon, F., Fernández-Rodríguez, J., Clermont, O., Denamur, E., Rocha, E. P., & Bikard, D. (2021). The impact of genetic diversity on gene essentiality within the Escherichia coli species. *Nature microbiology*, *6*(3), 301-312.

Rubin, E. J., Akerley, B. J., Novik, V. N., Lampe, D. J., Husson, R. N., & 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), 1645-1650.

Russell, D. G. (2007). Who puts the tubercle in tuberculosis? *Nature Reviews Microbiology*, *5*(1), 39-47.

Russell, D. G. (2011). Mycobacterium tuberculosis and the intimate discourse of a chronic infection. *Immunological reviews*, *240*(1), 252-268.

Russell, D. G., Cardona, P.-J., Kim, M.-J., Allain, S., & Altare, F. (2009). Foamy macrophages and the progression of the human tuberculosis granuloma. *Nature immunology*, *10*(9), 943-948.

Russell, D. G., VanderVen, B. C., Lee, W., Abramovitch, R. B., Kim, M.-j., Homolka, S., Niemann, S., & Rohde, K. H. (2010). Mycobacterium tuberculosis wears what it eats. *Cell host & microbe*, *8*(1), 68-76.

Russell, D. W., & Sambrook, J. (2001). *Molecular cloning: a laboratory manual* (Vol. 1). Cold Spring Harbor Laboratory Cold Spring Harbor, NY.

Ruvkun, G. B., Sundaresan, V., & Ausubel, F. M. (1982). Directed transposon Tn5 mutagenesis and complementation analysis of Rhizobium meliloti symbiotic nitrogen fixation genes. *Cell*, *29*(2), 551-559.

Sakamoto, K. (2012). The pathology of Mycobacterium tuberculosis infection. *Veterinary pathology*, *49*(3), 423-439.

Samant, S., Lee, H., Ghassemi, M., Chen, J., Cook, J. L., Mankin, A. S., & Neyfakh, A. A. (2008). Nucleotide biosynthesis is critical for growth of bacteria in human blood. *PLoS Pathog*, *4*(2), e37.

Sambrook, J., Fritsch, E. F., & Maniatis, T. (1989). *Molecular cloning: a laboratory manual*. Cold spring harbor laboratory press.

Sandgren, A., Strong, M., Muthukrishnan, P., Weiner, B. K., Church, G. M., & Murray, M. B. (2009). Tuberculosis drug resistance mutation database. *PLoS Med*, *6*(2), e1000002.

Santillán, M., & Mackey, M. C. (2005). Dynamic behaviour of the B12 riboswitch. *Physical biology*, *2*(1), 29.

Santos, F., Vera, J. L., Lamosa, P., de Valdez, G. F., de Vos, W. M., Santos, H., Sesma, F., & Hugenholtz, J. (2007). Pseudovitamin is the corrinoid produced by Lactobacillus reuteri CRL1098 under anaerobic conditions. *FEBS letters*, *581*(25), 4865-4870.

Santos, J. A., Rempel, S., Mous, S. T., Pereira, C. T., Ter Beek, J., de Gier, J.-W., Guskov, A., & Slotboom, D. J. (2018). Functional and structural characterization of an ECF-type ABC transporter for vitamin B12. *Elife*, *7*, e35828.

Sassetti, C. M., Boyd, D. H., & Rubin, E. J. (2001). Comprehensive identification of conditionally essential genes in mycobacteria. *Proceedings of the National Academy of Sciences*, *98*(22), 12712-12717.

Sassetti, C. M., Boyd, D. H., & Rubin, E. J. (2003). Genes required for mycobacterial growth defined by high density mutagenesis. *Molecular microbiology*, *48*(1), 77-84.

Satoh, T., & Akira, S. (2017). Toll-like receptor signaling and its inducible proteins. *Myeloid Cells in Health and Disease: A Synthesis*, 447-453.

Sattler, I., Roessner, C. A., Stolowich, N. J., Hardin, S. H., Harris-Haller, L. W., Yokubaitis, N. T., Murooka, Y., Hashimoto, Y., & Scott, A. I. (1995). Cloning, sequencing, and expression of the uroporphyrinogen III methyltransferase cobA gene of Propionibacterium freudenreichii (shermanii). *Journal of bacteriology*, *177*(6), 1564-1569.

Savvi, S., Warner, D. F., Kana, B. D., McKinney, J. D., Mizrahi, V., & Dawes, S. S. (2008). Functional characterization of a vitamin B12-dependent methylmalonyl pathway in Mycobacterium tuberculosis: implications for propionate metabolism during growth on fatty acids. *Journal of bacteriology*, *190*(11), 3886-3895.

Schnappinger, D., Ehrt, S., Voskuil, M. I., Liu, Y., Mangan, J. A., Monahan, I. M., Dolganov, G., Efron, B., Butcher, P. D., & Nathan, C. (2003). Transcriptional adaptation of Mycobacterium tuberculosis within macrophages: insights into the phagosomal environment. *The Journal of experimental medicine*, *198*(5), 693-704.

Seth, E. C., & Taga, M. E. (2014). Nutrient cross-feeding in the microbial world. *Frontiers in microbiology*, *5*, 350.

Shelton, A. N., Seth, E. C., Mok, K. C., Han, A. W., Jackson, S. N., Haft, D. R., & Taga, M. E. (2019). Uneven distribution of cobamide biosynthesis and dependence in bacteria predicted by comparative genomics. *The ISME journal*, *13*(3), 789-804.

Shull, L. M., & Camilli, A. (2018). Transposon Sequencing of Vibrio cholerae in the Infant Rabbit Model of Cholera. In *Vibrio Cholerae* (pp. 103-116). Springer.

Shultis, D. D., Purdy, M. D., Banchs, C. N., & Wiener, M. C. (2006). Outer membrane active transport: structure of the BtuB: TonB complex. *Science*, *312*(5778), 1396-1399.

Siegrist, M. S., & Rubin, E. J. (2009). Phage transposon mutagenesis. In *Mycobacteria Protocols* (pp. 311-323). Springer.

Sly, L. M., Hingley-Wilson, S. M., Reiner, N. E., & McMaster, W. R. (2003). Survival of Mycobacterium tuberculosis in host macrophages involves resistance to apoptosis dependent upon induction of antiapoptotic Bcl-2 family member Mcl-1. *The Journal of Immunology*, *170*(1), 430-437.

Smith, A. D., Warren, M. J., & Refsum, H. (2018). Vitamin B12. *Advances in food and nutrition research*, *83*, 215-279.

Smith, I. (2003). Mycobacterium tuberculosis pathogenesis and molecular determinants of virulence. *Clinical microbiology reviews*, *16*(3), 463-496.

Snapper, S., Melton, R., Mustafa, S., Kieser, T., & Jr, W. J. (1990). Isolation and characterization of efficient plasmid transformation mutants of Mycobacterium smegmatis. *Molecular microbiology*, *4*(11), 1911-1919.

Sokolovskaya, O. M., Mok, K. C., Park, J. D., Tran, J. L., Quanstrom, K. A., & Taga, M. E. (2019). Cofactor selectivity in methylmalonyl coenzyme A mutase, a model cobamide-dependent enzyme. *MBio*, *10*(5).

Sokolovskaya, O. M., Shelton, A. N., & Taga, M. E. (2020). Sharing vitamins: Cobamides unveil microbial interactions. *Science*, *369*(6499).

Soukup, J. K., & Soukup, G. A. (2004). Riboswitches exert genetic control through metabolite-induced conformational change. *Current opinion in structural biology*, *14*(3), 344-349.

Stadtman, T. C. (1960). Synthesis of adenine-B12 coenzyme by Clostridium sticklandii: relationship to one-carbon metabolism. *Journal of bacteriology*, *79*(6), 904.

Stamford, N. (1994). Genetics and enzymology of the B12 pathway. Ciba Found. Symp,

Stupperich, E., Eisinger, H.-J., & Schurr, S. (1990). Corrinoids in anaerobic bacteria. *FEMS microbiology reviews*, *7*(3-4), 355-359.

Stupperich, E., Eisinger, H.J., & Kräutler, B. (1989). Identification of phenolyl cobamide from the homoacetogenic bacterium Sporomusa ovata. *European journal of biochemistry*, *186*(3), 657-661.

Stupperich, E., & Kräutler, B. (1988). Pseudo vitamin B 12 or 5-hydroxybenzimidazolyl-cobamide are the corrinoids found in methanogenic bacteria. *Archives of microbiology*, *149*(3), 268-271.

Sudarsan, N., Wickiser, J. K., Nakamura, S., Ebert, M. S., & Breaker, R. R. (2003). An mRNA structure in bacteria that controls gene expression by binding lysine. *Genes & development*, *17*(21), 2688-2697.

Sun, J., Singh, V., Lau, A., Stokes, R. W., Obregón-Henao, A., Orme, I. M., Wong, D., Av-Gay, Y., & Hmama, Z. (2013). Mycobacterium tuberculosis nucleoside diphosphate kinase inactivates small GTPases leading to evasion of innate immunity. *PLoS Pathog*, *9*(7), e1003499.

Sun, J., Siroy, A., Lokareddy, R. K., Speer, A., Doornbos, K. S., Cingolani, G., & Niederweis, M. (2015). The tuberculosis necrotizing toxin kills macrophages by hydrolyzing NAD. *Nature structural & molecular biology*, *22*(9), 672-678.

Supply, P., & Brosch, R. (2017). The biology and epidemiology of Mycobacterium canettii. *Strain Variation in the Mycobacterium tuberculosis Complex: Its Role in Biology, Epidemiology and Control*, 27-41.

Supply, P., Marceau, M., Mangenot, S., Roche, D., Rouanet, C., Khanna, V., Majlessi, L., Criscuolo, A., Tap, J., & Pawlik, A. (2013). Genomic analysis of smooth tubercle bacilli provides insights into ancestry and pathoadaptation of Mycobacterium tuberculosis. *Nature genetics*, *45*(2), 172-179.

Taga, M. E., Larsen, N. A., Howard-Jones, A. R., Walsh, C. T., & Walker, G. C. (2007). BluB cannibalizes flavin to form the lower ligand of vitamin B 12. *Nature*, *446*(7134), 449-453.

Tailleux, L., Waddell, S. J., Pelizzola, M., Mortellaro, A., Withers, M., Tanne, A., Castagnoli, P. R., Gicquel, B., Stoker, N. G., & Butcher, P. D. (2008). Probing host pathogen cross-talk by transcriptional profiling of both Mycobacterium tuberculosis and infected human dendritic cells and macrophages. *PloS one*, *3*(1), e1403.

Tak, U., Vlach, J., Garza-Garcia, A., William, D., Danilchanka, O., de Carvalho, L. P. S., Saad, J. S., & Niederweis, M. (2019). The tuberculosis necrotizing toxin is an NAD+ and NADP+ glycohydrolase with distinct enzymatic properties. *Journal of Biological Chemistry*, *294*(9), 3024-3036.

Tanioka, Y., Yabuta, Y., Yamaji, R., Shigeoka, S., Nakano, Y., Watanabe, F., & Inui, H. (2009). Occurrence of pseudovitamin B12 and its possible function as the cofactor of cobalamin-dependent methionine synthase in a cyanobacterium Synechocystis sp. PCC6803. *Journal of nutritional science and vitaminology*, *55*(6), 518-521.

Teng, O., Ang, C. K. E., & Guan, X. L. (2017). Macrophage–bacteria interactions—a lipid-centric relationship. *Frontiers in immunology*, *8*, 1836.

Textor, S., Wendisch, V. F., De Graaf, A. A., Müller, U., Linder, M. I., Linder, D., & Buckel, W. (1997). Propionate oxidation in Escherichia coli: evidence for operation of a methylcitrate cycle in bacteria. *Archives of microbiology*, *168*(5), 428-436.

Thomas, C., Aller, S. G., Beis, K., Carpenter, E. P., & Martinoia, E. (2020). Structural and functional diversity calls for a new classification of ABC transporters. *FEBS letters*, *594*(23), 3767-3775.

Thomas, S. T., & Sampson, N. S. (2013). Mycobacterium tuberculosis utilizes a unique heterotetrameric structure for dehydrogenation of the cholesterol side chain. *Biochemistry*, *52*(17), 2895-2904.

Thomas, S. T., VanderVen, B. C., Sherman, D. R., Russell, D. G., & Sampson, N. S. (2011). Pathway profiling in Mycobacterium tuberculosis: elucidation of cholesterol-derived catabolite and enzymes that catalyze its metabolism. *Journal of Biological Chemistry*, *286*(51), 43668-43678.

Timm, J., Post, F. A., Bekker, L.-G., Walther, G. B., Wainwright, H. C., Manganelli, R., Chan, W.-T., Tsenova, L., Gold, B., & Smith, I. (2003). Differential expression of iron-, carbon-, and oxygen-responsive mycobacterial genes in the lungs of chronically infected mice and tuberculosis patients. *Proceedings of the National Academy of Sciences*, *100*(24), 14321-14326.

Trivedi, O. A., Arora, P., Sridharan, V., Tickoo, R., Mohanty, D., & Gokhale, R. S. (2004). Enzymic activation and transfer of fatty acids as acyl-adenylates in mycobacteria. *Nature*, *428*(6981), 441-445.

Tufariello, J. M., Chan, J., & Flynn, J. L. (2003). Latent tuberculosis: mechanisms of host and bacillus that contribute to persistent infection. *The Lancet infectious diseases*, *3*(9), 578-590.

Upton, A. M., & McKinney, J. D. (2007). Role of the methylcitrate cycle in propionate metabolism and detoxification in Mycobacterium smegmatis. *Microbiology*, *153*(12), 3973-3982.

van Crevel, R., Ottenhoff, T. H., & van der Meer, J. W. (2003). Innate immunity to Mycobacterium tuberculosis. *Tropical Diseases*, 241-247.

Van der Geize, R., Yam, K., Heuser, T., Wilbrink, M. H., Hara, H., Anderton, M. C., Sim, E., Dijkhuizen, L., Davies, J. E., & Mohn, W. W. (2007). A gene cluster encoding cholesterol catabolism in a soil actinomycete provides insight into Mycobacterium tuberculosis survival in macrophages. *Proceedings of the National Academy of Sciences*, *104*(6), 1947-1952.

van Helden, P. D., Victor, T. C., Warren, R. M., & van Helden, E. G. (2001). Isolation of DNA from Mycobacterium tubercolosis. In *Mycobacterium tuberculosis protocols* (pp. 19-30). Springer.

Van Opijnen, T., Bodi, K. L., & Camilli, A. (2009). Tn-seq: high-throughput parallel sequencing for fitness and genetic interaction studies in microorganisms. *Nature methods*, *6*(10), 767-772.

van Opijnen, T., & Camilli, A. (2012). A fine scale phenotype–genotype virulence map of a bacterial pathogen. *Genome research*, *22*(12), 2541-2551.

Van Opijnen, T., & Camilli, A. (2013). Transposon insertion sequencing: a new tool for systems-level analysis of microorganisms. *Nature Reviews Microbiology*, *11*(7), 435-442.

van Opijnen, T., Lazinski, D. W., & Camilli, A. (2014). Genome-wide fitness and genetic interactions determined by Tn-seq, a high-throughput massively parallel sequencing method for microorganisms. *Current protocols in molecular biology*, *106*(1), 7.16. 11-17.16. 24.

van Opijnen, T., & Levin, H. L. (2020). Transposon insertion sequencing, a global measure of gene function. *Annual review of genetics*, *54*, 337-365.

Van Scoy, R. E., & Wilkowske, C. J. (1999). Antimycobacterial therapy. Mayo Clinic Proceedings,

VanderVen, B. C., Huang, L., Rohde, K. H., & Russell, D. G. (2016). The minimal unit of infection: Mycobacterium tuberculosis in the macrophage. *Microbiology spectrum*, *4*(6), 4.6. 32.

Vergne, I., Chua, J., Lee, H.-H., Lucas, M., Belisle, J., & Deretic, V. (2005). Mechanism of phagolysosome biogenesis block by viable Mycobacterium tuberculosis. *Proceedings of the National Academy of Sciences*, *102*(11), 4033-4038.

Vey, J. L., & Drennan, C. L. (2011). Structural insights into radical generation by the radical SAM superfamily. *Chemical Reviews*, *111*(4), 2487-2506.

Villeneuve, C., Gilleron, M., Maridonneau-Parini, I., Daffé, M., Astarie-Dequeker, C., & Etienne, G. (2005). Mycobacteria use their surface-exposed glycolipids to infect human macrophages through a receptor-dependent process. *Journal of lipid research*, *46*(3), 475-483.

Vitreschak, A. G., Rodionov, D. A., Mironov, A. A., & Gelfand, M. S. (2003). Regulation of the vitamin B12 metabolism and transport in bacteria by a conserved RNA structural element. *Rna*, *9*(9), 1084-1097.

Vrieling, F., Kostidis, S., Spaink, H. P., Haks, M. C., Mayboroda, O. A., Ottenhoff, T. H., & Joosten, S. A. (2020). Analyzing the impact of Mycobacterium tuberculosis infection on primary human macrophages by combined exploratory and targeted metabolomics. *Scientific reports*, *10*(1), 1-13.

Vromman, F., & Subtil, A. (2014). Exploitation of host lipids by bacteria. *Current opinion in microbiology*, *17*, 38-45.

Warner, D. F. (2015). Mycobacterium tuberculosis metabolism. *Cold Spring Harbor perspectives in medicine*, *5*(4), a021121.

Warner, D. F., & Mizrahi, V. (2006). Tuberculosis chemotherapy: the influence of bacillary stress and damage response pathways on drug efficacy. *Clinical microbiology reviews*, *19*(3), 558-570.

Warner, D. F., Savvi, S., Mizrahi, V., & Dawes, S. S. (2007). A riboswitch regulates expression of the coenzyme B12-independent methionine synthase in Mycobacterium tuberculosis: implications for differential methionine synthase function in strains H37Rv and CDC1551. *Journal of bacteriology*, *189*(9), 3655-3659.

Warren, M. J., Raux, E., Schubert, H. L., & Escalante-Semerena, J. C. (2002). The biosynthesis of adenosylcobalamin (vitamin B12). *Natural product reports*, *19*(4), 390-412.

Watkins, H. A., & Baker, E. N. (2010). Structural and functional characterization of an RNase HI domain from the bifunctional protein Rv2228c from Mycobacterium tuberculosis. *Journal of bacteriology*, *192*(11), 2878-2886.

Way, J., Davis, M., Morisato, D., Roberts, D., & Kleckner, N. (1984). New Tn10 derivatives for transposon mutagenesis and for construction of lacZ operon fusions by transposition. *Gene*, *32*(3), 369-379.

Weerdenburg, E. M., Abdallah, A. M., Rangkuti, F., Abd El Ghany, M., Otto, T. D., Adroub, S. A., Molenaar, D., Ummels, R., Ter Veen, K., & van Stempvoort, G. (2015). Genome-wide transposon mutagenesis indicates that Mycobacterium marinum customizes its virulence mechanisms for survival and replication in different hosts. *Infection and immunity*, *83*(5), 1778-1788.

Wheeler, P. (1987). Biosynthesis and scavenging of purines by pathogenic mycobacteria including Mycobacterium leprae. *Microbiology*, *133*(11), 2999-3011.

Wheeler, P. R., & Blanchard, J. S. (2005). General metabolism and biochemical pathways of tubercle bacilli. In *Tuberculosis and the Tubercle Bacillus* (pp. 309-340). American Society of Microbiology.

Wheeler, P. R., & Ratledge, C. (1994). Metabolism of Mycobacterium tuberculosis. *Tuberculosis: Pathogenesis, Protection, and Control*, 353-385.

WHO. (2017). Guidelines for treatment of drug-susceptible tuberculosis and patient care.World Health Organization: Geneva, Switzerland.

WHO. (2020). *Global tuberculosis report* 2020. Geneva.

Wiener, M. C. (2005). TonB-dependent outer membrane transport: going for Baroque? *Current opinion in structural biology*, *15*(4), 394-400.

Wilburn, K. M., Fieweger, R. A., & VanderVen, B. C. (2018). Cholesterol and fatty acids grease the wheels of Mycobacterium tuberculosis pathogenesis. *Pathogens and disease*, *76*(2), fty021.

Williams, K. J., Boshoff, H. I., Krishnan, N., Gonzales, J., Schnappinger, D., & Robertson, B. D. (2011). The Mycobacterium tuberculosis β-oxidation genes echA5 and fadB3 are dispensable for growth in vitro and in vivo. *Tuberculosis*, *91*(6), 549-555.

Winkler, W. C., & Breaker, R. R. (2005). Regulation of bacterial gene expression by riboswitches. *Annu. Rev. Microbiol.*, *59*, 487-517.

Winkler, W. C., Nahvi, A., Roth, A., Collins, J. A., & Breaker, R. R. (2004). Control of gene expression by a natural metabolite-responsive ribozyme. *Nature*, *428*(6980), 281-286.

Winkler, W. C., Nahvi, A., Sudarsan, N., Barrick, J. E., & Breaker, R. R. (2003). An mRNA structure that controls gene expression by binding S-adenosylmethionine. *Nature structural & molecular biology*, *10*(9), 701-707.

Wipperman, M. F., Sampson, N. S., & Thomas, S. T. (2014). Pathogen roid rage: cholesterol utilization by Mycobacterium tuberculosis. *Critical reviews in biochemistry and molecular biology*, *49*(4), 269-293.

Witney, A. A., Cosgrove, C. A., Arnold, A., Hinds, J., Stoker, N. G., & Butcher, P. D. (2016). Clinical use of whole genome sequencing for Mycobacterium tuberculosis. *BMC medicine*, *14*(1), 1-7.

Xu, T., Bharucha, N., & Kumar, A. (2011). Genome-wide transposon mutagenesis in Saccharomyces cerevisiae and Candida albicans. In *Strain Engineering* (pp. 207-224). Springer.

Xu, Z., Wang, Y., Chater, K. F., Ou, H.-Y., Xu, H. H., Deng, Z., & Tao, M. (2017). Large-scale transposition mutagenesis of Streptomyces coelicolor identifies hundreds of genes influencing antibiotic biosynthesis. *Applied and environmental microbiology*, *83*(6).

Yan, J., Şimşir, B., Farmer, A. T., Bi, M., Yang, Y., Campagna, S. R., & Löffler, F. E. (2016). The corrinoid cofactor of reductive dehalogenases affects dechlorination rates and extents in organohalide-respiring Dehalococcoides mccartyi. *The ISME journal*, *10*(5), 1092-1101.

Yang, M., Guja, K. E., Thomas, S. T., Garcia-Diaz, M., & Sampson, N. S. (2014). A distinct MaoC-like enoyl-CoA hydratase architecture mediates cholesterol catabolism in Mycobacterium tuberculosis. *ACS chemical biology*, *9*(11), 2632-2645.

Yang, M., Lu, R., Guja, K. E., Wipperman, M. F., St. Clair, J. R., Bonds, A. C., Garcia-Diaz, M., & Sampson, N. S. (2015). Unraveling cholesterol catabolism in Mycobacterium tuberculosis: ChsE4-ChsE5 α2β2 acyl-CoA dehydrogenase initiates β-oxidation of 3-oxo-cholest-4-en-26-oyl CoA. *ACS infectious diseases*, *1*(2), 110-125.

Yang, X., Nesbitt, N. M., Dubnau, E., Smith, I., & Sampson, N. S. (2009). Cholesterol metabolism increases the metabolic pool of propionate in Mycobacterium tuberculosis. *Biochemistry*, *48*(18), 3819-3821.

Yin, L., & Bauer, C. E. (2013). Controlling the delicate balance of tetrapyrrole biosynthesis. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *368*(1622), 20120262.

Young, D. B., Comas, I., & de Carvalho, L. P. (2015). Phylogenetic analysis of vitamin B12-related metabolism in Mycobacterium tuberculosis. *Frontiers in molecular biosciences*, *2*, 6.

Yung, M. C., Park, D. M., Overton, K. W., Blow, M. J., Hoover, C. A., Smit, J., Murray, S. R., Ricci, D. P., Christen, B., & Bowman, G. R. (2015). Transposon mutagenesis paired with deep sequencing of Caulobacter crescentus under uranium stress reveals genes essential for detoxification and stress tolerance. *Journal of bacteriology*, *197*(19), 3160-3172.

Zelder, O., Beatrix, B., Kroll, F., & Buckel, W. (1995). Coordination of a histidine residue of the protein-component S to the cobalt atom in coenzyme B12-dependent glutamate mutase from Clostridium cochlearium. *FEBS letters*, *369*(2-3), 252-254.

Zhang, J. (2003). Evolution by gene duplication: an update. *Trends in ecology & evolution*, *18*(6), 292-298.

Zhang, Y., Rodionov, D. A., Gelfand, M. S., & Gladyshev, V. N. (2009). Comparative genomic analyses of nickel, cobalt and vitamin B12 utilization. *BMC genomics*, *10*(1), 1-26.

Zheng, J., Rubin, E. J., Bifani, P., Mathys, V., Lim, V., Au, M., Jang, J., Nam, J., Dick, T., & Walker, J. R. (2013). para-Aminosalicylic acid is a prodrug targeting dihydrofolate reductase in Mycobacterium tuberculosis. *Journal of Biological Chemistry*, *288*(32), 23447-23456.

Zheng, X., & Av-Gay, Y. (2016). New era of TB drug discovery and its impact on disease management. *Current Treatment Options in Infectious Diseases*, *8*(4), 299-310.

zu Bentrup, K. H., & Russell, D. G. (2001). Mycobacterial persistence: adaptation to a changing environment. *Trends in microbiology*, *9*(12), 597-605.

Zuniga, J., Torres-García, D., Santos-Mendoza, T., Rodriguez-Reyna, T. S., Granados, J., & Yunis, E. J. (2012). Cellular and humoral mechanisms involved in the control of tuberculosis. *Clinical and Developmental Immunology*, *2012*.