**Yohannes Derese Siyum**

# Role of Esx-3 Secretion System and Stress Response in *Mycobacteria smegmatis*

**Master's Thesis in Molecular Medicine**

**Trondheim, Norway, June 2015**

**Principal Supervisor: Magnus Steigedal**

**Co-supervisor: Marte Singsås Dragset**

**Norwegian University of Science and Technology**

**Faculty of Medicine**

**Department of Cancer Research and Molecular Medicine Center of Molecular Inflammation Research**

**Acknowledgements**

This Master's thesis was done at the Faculty of Medicine, Center of Molecular Inflammation Research (CEMIR), Department of Cancer Research and Molecular Medicine, Norwegian University of Science and Technology (NTNU), Trondheim, Norway.

This thesis would not have been completed without the contribution of several persons. My first and foremost appreciation goes to all people for their willingness to contribute to the success of my thesis work by giving their precious time and valuable information during the lab work.

I would like to acknowledge my supervisor Magnus Steigedal for his interest to mentor. He devoted his time to guide me by his constructive comments and discussions throughout the thesis work. His critical comments were wonderful, featured with a sense of humour and quick response. I would also like to extend my sincere thanks and appreciation to post doc Marte, Alexander Gido, Zekarias and PhD student Nisha Kenan for their support in the lab.

I want to express my gratitude to all the CEMIR staffs, Mycobacteria research group and my fellow Master's students. I also extend my heartfelt thanks to my beloved family Helina Seid, My father Derese Siyum, friends and colleagues for their support to reach at this stage.

Finally, I would like to appreciate and sincere thanks to the Norwegian Government for granting scholarship through the Quota Scheme Program. My sincere thanks also go to the staff member of the international office of NTNU, Anette Moen for her warm and helpful treatment.

**Abstract**

**Background:** Bacteria secrets proteins to manipulate their growth environments through various secretion systems. Type VII systems are important protein secretary systems known in mycobacteria. 5 different Type VII systems are predicted with different function which named as Esx-1 to Esx-5. Protein secretion and some function of Esx-1, Esx-3 and Esx-5 have been shown in several studies. Esx-4 and Esx-3 are the only systems conserved in every species of mycobacteria and Esx-3 has an essential role in most species. Esx-3 secretion system has an important role in mycobacterial iron uptake. Iron is an essential element important for cellular metabolism and oxygen transport. Unregulated iron could lead to harmful toxic radical formation. Mycobacteria must carefully regulate iron to survive and establish infection or stay dormant inside host. Despite the importance of Esx-3 for iron metabolism there is some evidence that Esx-3 would also be involved in other processes.

**Aim:** The aim of this study was to identify novel function of Esx-3 secretion system by evaluating gene expression of genes in central metabolic processes.

**Objectives:** To investigate transcription, we wanted to optimize high quality mycobacteria RNA extraction methods. For gene expression analysis, we used NanoString technology to identify role of mycobacterial Esx-3 system in different metabolic processes. Evaluate esx-3 mutant *M. smegmatis* stress responses in solid and liquid media and determine localization of Esx-3 proteins in *M. smegmatis* by cloning with gateway cloning system and confocal microscopy.

**Result:** 107 *M. smegmatis* genes were screened by NanoSting gene expression analysis. Most of expression in redox regulatory, iron regulatory and iron dependent repressor genes were affected in esx-3 mutant *M. smegmatis*. We revealed for the first time that under high iron media, almost all screened redox regulatory genes were down regulated in esx-3 mutant *M. smegmatis*. Expression of *ahp*C gene in esx-3 mutant *M. smegmatis* challenged with $H_2O_2$ was upregulated in both low and high iron condition whereas *kat*G was upregulated only under low iron $H_2O_2$ exposure. In the absence of $H_2O_2$, *kat*G was upregulated only in WT under high iron condition. Based on the expression data we investigated the tolerance of the esx-3 mutant towards $H_2O_2$ stressor both in solid and liquid media and under low and high iron conditions. Our experiments showed that mutant had higher tolerance to $H_2O_2$ compared to both wild type and mycobactin mutant *M. smeg*matis under all conditions tested. Part of the biological function of a protein is the correct localization. We in investigated the

localization of important Esx-3 proteins and found that $EccB_3$ protein was observed as strong fluorescent large beads on average 2 – 3 per bacteria with predominantly polar localization whereas; $EspG_3$ was seen as small green fluorescent dots scattered throughout the bacterium, with more dots per cell than in $EccB_3$.

**Conclusion:** Esx-3 could be negatively regulated by both *fur*A and *ide*R. Esx-3 could also involve in *ahp*C oxidative stress response, possibly through gene regulation. Defects in one of the repressor *fur*A or *ide*R will indirectly affects the function of Esx-3 and increase *M. smegmatis* oxidative stressor susceptibility. Further study is required to investigate how esx-3 is regulated by *fur*A and *ide*R and role of Esx-3 in oxidative regulatory genes. The experiment also should be repeated in pathogenic mycobacteria species.

## Table of Contents

List of figures

List of tables



List of supplementary tables



List of supplementary figures

# 1 Introduction

## 1.1 The challenges of Tuberculosis

Approximately, one-third of the world's population is infected by *Mycobacterium tuberculosis (M. tb)* among which 1.7 million killed annually[1-3]. Although Tuberculosis (TB) is preventable disease, it is continuing as a major health threat and is the second most cause of infectious disease death next to human immunodeficiency virus (HIV) AIDS [4].According to world health organization (WHO) declaration, TB has become as a global public health emergency [5]. Anti TB drug responds poorly in immune compromised TB patients this aggravated mortality in Multi drug resistance (MDR) TB and HIV AIDS patients [6]. In 2014, WHO designed a strategy to end tuberculosis as a major threat of health problem by 2035. Although efforts has made so far, the number of TB cases has been rising which estimated to 9 million cases and 1.5 million deaths in 2013 and half a million more prevalence by 2014 [7]. The death report of TB in 2013 was even higher than from previous years. In 2014 WHO report, the actual number of MDR-TB cases was under estimated, that was because of resource limitation in high TB burden countries such as lack of drug resistant test and treatment services. In 2013, 9% of MDR cases suspected to develop into the nearly untreatable extensively drug resistant (XDR) form [7].

## 1.2 TB treatment and emergence of drug resistance

Currently, patients diagnosed as TB are treated with a combination of 4 first line anti-TB drugs for a total period of six months [2]. MDR TB cases are treated with the second line drugs [8] which are more toxic and 100 times more expensive than first line dugs and have to be taken for about 20 months [2,8]. MDR is defined as resistance of mycobacterial isolates, at least for two most potent first line anti-TB drugs Isoniazid (INH) and Rifampicin (RIF) [8]. MDR TB patients do not respond to the standard anti-TB treatment regimen and continued to spread drug resistant TB in to the populations [4]. The main cause of treatment failure is due to incorrect use of anti-TB drugs, wrong treatment dosage or regimen and acquiring new drug resistant strain infection [2]. The emergence of anti TB drug resistant strains makes TB as a global health problem and our limited understanding of the pathobiology of *M. tb* and the interaction with the host makes it more difficult to develop efficient drugs and vaccines[9]. Some features of *M. tb* are known and promotes the drug resistance we observe. One example is the unique structure and low permeability of Mycobacterial cell wall, enzymatic inactivation and insufficient porin that inhibit permeability of antimicrobial agents[10]. In addition to the cell wall, drug resistance is also conferred by enzymatic processes like aminoacetyle transferase, β- lactamases, and ABC transporter encoded by mycbacterial genomes[11]. Developing new treatment is critically an important approach to halt the spread of TB and drug resistant strain [8]. Thus, understanding the possible mechanism of drug resistance and the biology of this pathogen will enhance the development of new anti-mycobacterial drugs[11].

### 1.3 Tuberculosis vaccine

*Mycobacteria bovis (M. bovis)* strain Bacilli Calmette Guérin (BCG) is the only available vaccine in market which developed by sub-culturing strain of *M. bovis* several times in an artificial media. In spite of effective protective vaccine in children, BCG has inconsistent protection against pulmonary infection in adults[12]. However, it has been used worldwide in the last 60 years as a main vaccine against TB. Nowadays, identification of effective TB vaccine has been becoming widely the main concern in TB research[13]. TB vaccine is categorized mainly in to live attenuated and sub unit vaccine. Attenuated vaccine is based on the hypothesis that continuous several passaging weaken the BCG by letting gene loss [14]. Experimental study on an animal model shows that complementation of lost gene back to BCG and increasing expression of specific immune-dominant gene enhances the effectiveness of BCG vaccines [15,16]. One example for this approach is over expression of the immunodominant recombinant BCG antigen, Ag85B [17]. Sub unit vaccine approach is based on increasing the efficiency of antigen uptake by immune cells. Making surplus amount of bacterial-derived antigen available to antigen presenting cells and consequently, activate cytotoxic T-cells to boost protection[16-18]. Adjutants are also required to activate cell mediated immune response in response to subunit vaccines. However, subunit vaccines cloned with viral vectors like adenovirus and vaccinia induces a strong cell mediate immune response without any adjuvant. The protection by subunit vaccines induces short term compared to the live attenuated BCG and adjutants used in tuberculosis vaccines were unsuccessful either failed to stimulate immune cells or toxic to the host[13].

### 1.4 Mycobacteria infection and host interaction

Mycobacteria are the causative agent of tuberculosis and have the capability to evade the immune system of the host[19]. Mycobacterial aerosolized droplets from infected person transmit to uninfected individual affecting primarily lung tissue then spread to extrapulmenary sites. At an early stage of TB infection, lung epithelial cells and natural killer cells are activated which in turn activates mucosal associated invariant T cells (MAIT) and macrophages respectively[20]. Besides, other inherent immune defence mechanism inhibits the spread of infection[12]. However complex host-pathogen interaction will let the continuation of disease progress. The host attempts to control the infection by coordinating innate and adaptive immune response either to kill the infected phagocytic cells or reduce replication[21,22]. Both secreted proteins and secretion compartment proteins of *M. tb* can be targeted by T-cells immune response[12].

Inhalation of pathogenic *M.tb* is the main route of infection where alveolar macrophages engulf upon encounter the bacilli in lung tissue [23]. Receptor mediated internalization of the bacilli by phagocytic cells is initiated by complement receptor, (CR), Mannose receptors (MR) and dendritic cell specific intercellular adhesion molecule [24]. Mycobacteria activates both classical and alternative complement pathway to be up taken by both opsonisation dependent and independent mechanism[25]. MR predominantly expressed on alveolar macrophages which recognize the bacterial surface mannose. Mannose inhibits the phagolysosomal fusion and this enhances anti-inflammatory response and replication of the bacteria. The interaction between mycobacterial surface to phagocytic cell receptors induces the production of anti-inflammatory IL-10 which down regulate multiple signaling pathways and pathogen receptors [24]. Outer mycobacterial surface molecules like trehalose dimycolate also activate the innate immune response[26]. Cytotoxic T-helper immune cells (Th1) are the main immune response against many intracellular pathogens including mycobacteria[27]. Although it was well characterized that, the interaction of CD4 T cells and activated macrophages plays crucial role in the control of TB, it has not been well understood how the interaction of immune cells protect against mycobacteria [12]. However, considerable amount of mycobacterial secretion system proteins were identified as immunodominant antigens that activated T cells[28].

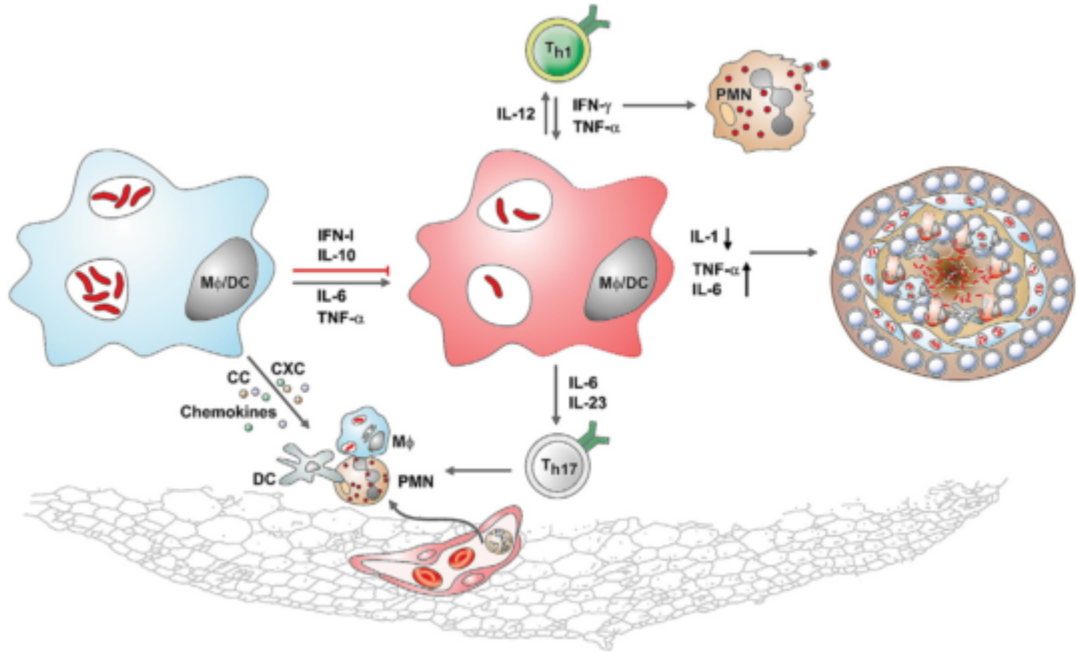



Figure 1.1 Mycobacteria infection and host immune regulatory response**.** Infected phagocytic cells release regulatory cytokine IFN-1and IL-10 and the pro-inflammatory cytokines IL-6 and TNF-alpha which kills the bacilli by over activation of macrophages. Some infected macrophages/DCs produce chemokines that increases the accumulation of *M. tb* permissive cells locally. The IL-12 cytokine from infected cells activates Th1 arm of the T-cells which in turn produce IFN-gamma usually participated in activation of the neutrophil apoptotic and the potent pro-inflammatory booster TNF-alpha. The IFN-gamma from Th-1 is not only promotes anti-inflammatory through apoptosis, but also aggravates the pro-inflammatory signal. The activated DCs/macrophages also secrets IL-6 and IL-23 for the activation of Th-17 arm which recruits further phagocytic and inflammatory cells to the surrounding tissues. The balance between cytoprotective IL-1 cytokine and pro-inflammatory TNF-alpha/IL-6 culminates in granuloma formation. If this two antagonistic cytokines are unbalanced, caseation and cavitation will be formed which enhances dissemination of the bacilli [23].

A study by Sweeney and colleagues mouse model experiment had shown that, the involvement of esx-3 locus in evading the innate immune response was promising in mycobacterial target. Although *M. smegmatis* was considered as saprophytic, it has the potential to induce infection and death at high dosage in mice. Compensation of *M. smegmatis* esx-3 mutant with *M. tb* esx-3 gene locus creates attenuated strains and susceptible to innate and adaptive immune response which makes esx-3 as promising vaccine candidate in future[27]. Unlike *M. smegmatis*, *M. tb* is unable to survive in the absence of esx-3 and thus impossible to construct esx-3 mutant *M. tb*[29]. But, replacing esx-3 locus from *M. tb* to *M. smegmatis* esx-3 mutant induced a similar immune response as that of wild type *M. smegmatis* strain with high and immediate IL-12, IFN-γ cytokine response indicating that esx-3 is conserved at species level. *ecc*$A_3$ mutant *M. smegmatis* had also shown similar effects as the whole esx-3 mutant strains with regards to virulence attenuation and host immune response. *M. tb* and *M. smegmatis* have a homologues esx-3 gene locus with about 44% sequence similarity and 85% at protein level. All the esx-3 characteristics studied above could help to design possible potential TB vaccine[27]. To effective control of the disease, identifying new molecular approach besides modifying the previous treatments and understanding host pathogen interaction are mandatory [19]

## 1.5 Mycobacterial cell wall structure

The genus Mycobacteria contains more than 150 species[30]. Mycobacteria are acid fast bacilli which on the bases of growth rate classified in to slow growing and fast growing group. The well-known human and animal pathogen species *Mycobacteria tuberculosis, Mycobacteria bovis, Mycobacteria africanum, Mycobacterium canettii, Mycobacteria microti, Mycobacteria laprae* and *Mycobacteria pinnipedii* are collectively named as Mycobacteria tuberculosis complex (MTBC) which are grouped under slow growing while nonpathogenic *Mycobacteria smegmatis* grouped under fast growing [31]. The slow growing pathogenic *M. tb* has doubling time of about 24 hours and take up to 1 month to observe visible colony in solid media whereas the fast growing saprophytic *M. smegmatis* has doubling time of 3-4 hours with much lesser colony forming period. *M. smegmatis* genome is highly similar and 1.7 times larger than *M. tb* genome [32]. The cell wall structure of mycobacterium is similar to gram positive bacteria but has an additional outer membrane made of lipid and polysaccharide[11]. Both *M. smegmatis* and *M. tb* have similar complex cell wall structure. *M. smegmatis* can also be used as drug susceptibility test and pathogenicity model despite nonpathogenic, it can cause infection at high dose in animal model[27]. All the above mentioned property of the lab strain *M. smegmatis* $MC^2$ 155 made preferable and important model to study the biology of pathogenic *M. tb* and from which results can be extrapolated [33].

Mycobacteria are recently categorized phylogeneticaly under order corynebacteriales which are high GC DNA gram positive bacteria and further classified by the presence of covalently linked mycolic acids complex cell wall structures [34]. The cell envelop has a unique hydrophobic outer membrane incorporated with variable size mycolic acid (depending on species)[35,36]. In general, the cell wall structure of mycobacteria is a double lipid membrane with the outer lipid membrane composed of mycolic acid covalently linked to arabinogalactan-peptidoglycan polymer which in turn externally layered with capsule[37]. Briefly, the core cell wall component are mycolic acids comprising long hydroxyl fatty acids, alpha alkyl chain, C30 – C90 and covalently attached to peptidoglycan-arabinogalactan complex [38,39]. The mycolic acids together with lipids form outer membrane layer. In addition, there is a complex of proteins, glycans and glycolipid which forms the outer thick layer capsule[39,40]. Despite structural complexity of cell wall, proteins and virulence factors are yet secreted out of the bacteria [41]. The structure and thickness of outer membrane of *Mycobacteria* is similar with that of gram negative bacteria but different from gram positive bacteria[38] .The presence of outer membrane in Mycobacteria differentiates from gram positive bacteria which prevent nutrient permeability [42,43]. Moreover, This waxy thick and complex cell wall makes the bacteria impermeable to hydrophilic substances and contributes for drug resistance [44]. The bacteria also protected against phagocytic cell effectors and notorious environmental substances [42].

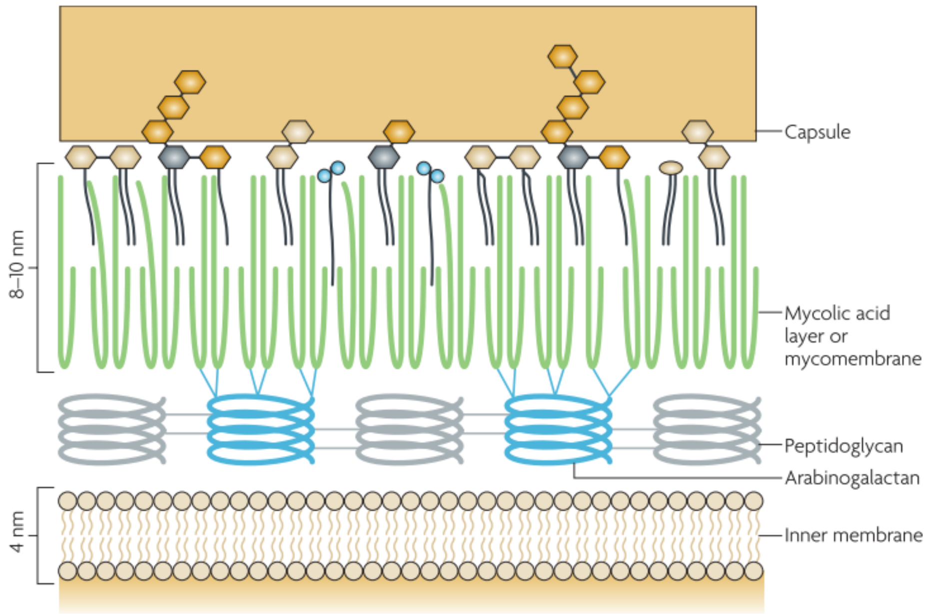



**Figure 1.2 Mycobacterial complex cell envelope structure and components.**The cytoplamic membrane is the inner most membrane followed by the thicker outer membrane composed of lipids interspersed with mycolic acids which is covalently linked with arabinogalactan-peptidoglycan complex. The outer membrane is layered with a composition of Proteins, glycan and glycolipids complex called capsule [45].

### 1.6 Mycobacteria iron acquisition and regulation mechanism

Although iron is the second most abundant element in the earth's crust, it is not readily available element for living organism[43]. Iron is essential for survival of nearly all living cells which function as a cofactor in DNA synthesis, amino acid synthesis, detoxification and respiration[46,47]. *M. tb* uses iron as a cofactor for more than 40 different enzymes and crucial for virulence factors[9]. Iron is also not freely found in sufficient soluble form in the host; instead it binds with host protein like feritin and transferrin[5,48-50]. Pathogenic mycobacterium have adapted with controlled iron uptake system to overcome the shortness and surplus iron. During infection, microorganisms use different mechanism to compensate this iron shortcoming[46,47]. Excess iron can cause irreversible oxidative damage and risk for survival of Mycobacteria, unless tightly regulated in multiple mechanisms[46,51]. For iron uptake mycobacteria use small molecular weight compounds called siderophores. One of them, mycobaction comes in two forms; the hydrophilic secreted carboxymycobactin and cell associated lipophylic mycobactin. Both have been shown to function as iron chelators and even has been shown to be important as *M. tb* virulence factors[9,46,51-53]. Non-pathogenic saprophytic mycobacteria, like *Mycobacterium smegmatis*, produce a second siderophore called exochellin in addition to the mycobactins. Mycobactin and Exochellins are two structurally related iron scavenging molecules that are synthesized by non-ribosomal peptides polyketides synthase mechanism[11].

*mbt*-1 and *mbt*-2 are the two important mycobacterial gene clusters involved in mycobactein biosynthesis. The *mbt*-1 cluster designated as *mbt*A – *mbt*J are responsible for mycobactin core scaffolding synthesis while the *mbt*-2 cluster which designated as *mbt*K – *mbt*N are involved in assembly of mycobactin side chains. Similarly, the saprophytic *M. smegmatis* genes *fxb*A, *fxb*B and *fxb*C are responsible for exochelin biosynthesis[54,55].

Carboxymycobactin is secreted and is presumed to have a role in capturing iron from the host proteins. The ferric ($Fe^{3+}$) bound mycobactin transported in to the cytosol binding specifically with receptors probably by reduction and facilitated by the help of ATP binding cassette (ABC) transporter like *irt*AB[56]. Although many bacteria acquire iron from the surrounding through a siderophore dependent pathway, many also rely on the use of siderophore independent mechanisms like diffusion through porin[43]. In the absence or limited siderophores, bacteria can maintain the required amount of iron which implies the presence of alternative pathway. In *M. smegmatis*, pore is formed on the outer membrane by Msp porins which are the most abundant protein in *M. smegmatis* and important for import-export of small hydrophilic molecules. Defect in the Msp affects the growth of *M. smegmatis.* The production of high siderophores in Msp porin deficient *M. smegmatis* under low iron media, clearly show that porin is involved in influx of iron. Under high iron media, Msp porins are involved in iron uptake independent of the major siderophore exochellin in *M. smegmatis* whereas under low iron condition, acquisition is mainly affinity dependent transport mechanism which is iron bound siderohores. Under high iron, expression of siderophore encoding genes are down regulated and affinity independent iron transport mechanism like

Msp porin contributes to fulfil the required iron uptake. *ide*R is an iron dependent repressor gene which repress transcription of about 92 mycobacteria genes involved in iron acquisition. The activity of *ide*R is increased in response to decreasing the number of porins [43] which shows porin is responsible for iron uptake. In a study conducted on *M. smegmatis* shown that, the availability of surplus iron makes the *ide*R represses the synthesis of siderohores. *ide*R mutant *M. smegmatis* was also shown susceptible to $H_2O_2$ stress and the expression of *kat*G was down regulated [57].

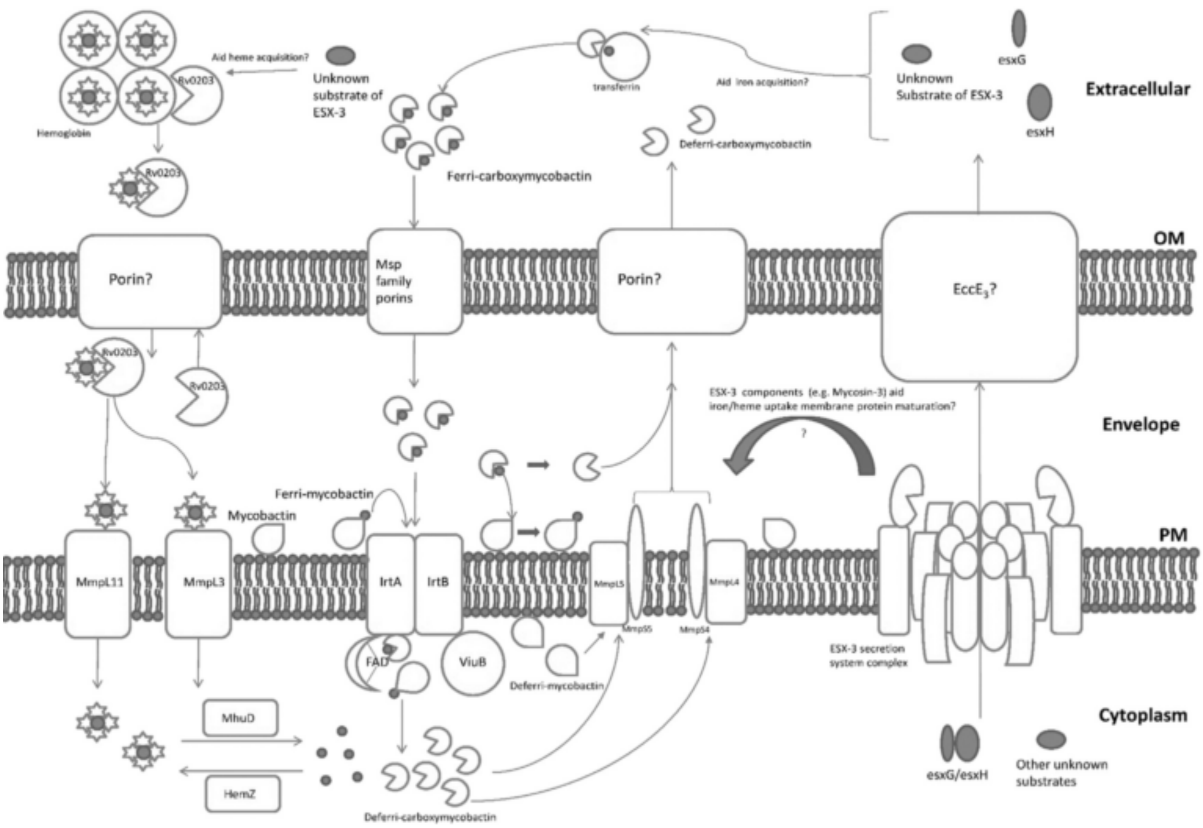



Figure 1.3. Model of mycobacterial iron acquisition systems and Esx-3 involvement. Iron sources are from host proteins like transferrin. Insoluble iron from transferin captured by soluble mycobactin and transported in to the periplasmic membrane through Msp family porin. Then the relay continued either through irtAB membrane proteins or insoluble mycobactin transportation process. The enzyme FAD encoded by irtA, reduce iron bound carboxymycobactin and release biological usable iron form (Fe++) ferrous ion. The soluble carboxymacobactin exported back to the surrounding through the MmpL5/S5 and MmpL4/S4 membrane complex to import further iron if necessary. Presumed exported proteins encoded from rv0203 released outside of the bacteria and bind with heme to initiate iron acquisition through MmpL11and MmpL3 heme pathway. The source of heme is not only from the host but also synthesized by mycobacteria ferrochelatase, HemZ. In the process of iron uptake, Esx-3 might be involved either by making secretary/transport channels or releasing important substrates that facilitate the transportation process or both [58]

## 1.7 Bacterial secretion system

One third of bacterial proteins are secreted out of the bacteria or incorporated as part of membrane [59]. Secreted substances are usually enzymes, proteins, lipoproteins, toxins and other appendage surface proteins[59,60]. In gram negative bacteria, 8 different secretion systems are known which named as Type I – VI secretion system in addition to the two signal dependent general secretion system (Sec) or Twin-arginine translocation (Tat) system. Type I, III, IV and VI secretion system directly translocate proteins across the double membrane without need of posttranslational modification. Type II and V secretion systems are chaperon dependent which translocate proteins only across the outer membrane after signal dependent proteins are translocated to periplasmic area by Sec or Tat system[61]. In addition to gram negative secretion system, type VII secretion system is possessed by gram positive Actinobacteria including mycobacteria group and in distantly related firmicutes group [62]

Many bacterial proteins with N- terminal signal sequence are translocated to the inner membrane through Sec and Tat system although there is a limited study in mycobacteria. Tat supports the translocation of folded and oligomer proteins and important for optimum growth. In the absence of the system growth of *M. tb* was much more suppressed compared to *M. smegmatis*[63,64]. Similarly, Sec translocate N-terminal signal dependent unfolded proteins [65]. Mycobacterium cell wall incorporates SecA1 and SecA2 with different localization which encoded from the two corresponding (secA1 and secA2) genes. It was reported that SecA1 is important for survival of bacteria while SecA2 which has partial protein similarity with secA1 is for virulence[63,64]. In addition, to accomplish full secretion of substance across the complex double membrane, mycobacteria use its signal independent specialized type 7 secretion system (T7SS)[62,64].

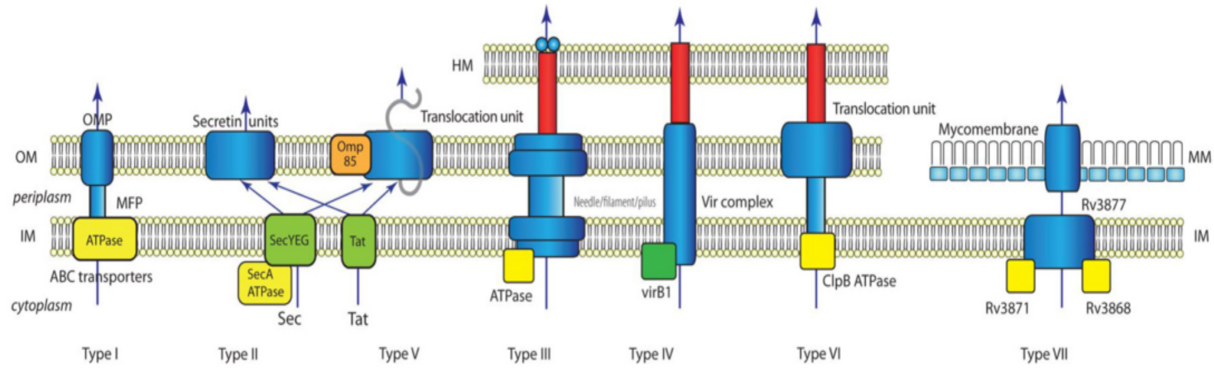



Figure 1.4. Structural model of known bacteria secretion systems. Type III and V receive folded proteins from Sec/Tat and translocate through outer membrane (OM). Type III- VI has a potential to secrete out proteins directly through double membrane and inject in to host membrane (HM) as most Gram negative bacteria do. Type VII secretion system secrete substances through both inner membrane (IM) and outer mycomembrane complex although outer membrane component is yet to be found. Membrane fusion proteins (MFP) is secretion apparatus that bridges the two membrane components as it has shown in the model for Type I, III,IV and VI [65].

### 1.7.1 **Type VII secretion system (T7SS)**

T7SS or Esx secretion system is the main secretary mechanism of proteins in mycobacteria[35]. There are 5 different gene clusters of T7SS named as *esx-1* to *esx-5* with varied size and number of genes at species level. The 5 gene clusters identified were possibly evolved by gene duplication event from the ancestor *esx-4* followed by *esx-1*, *esx-3*, and *esx-2* and at last most recently *esx-5*[35,66]. The presence of gene components in at least 4 *esx* gene clusters named by suffix esx conserved component (ecc), while all the other components are named esx-specific protein (esp), except the subtilisin-like protease family MycP, PE/PPE proteins and the small couple secreted proteins[66]. The *esx* gene cluster arrangement and composition looks similar although have a different functional role. Most mycobacterial proteins are secreted by means of protein encoded from this clusters of genes[67].

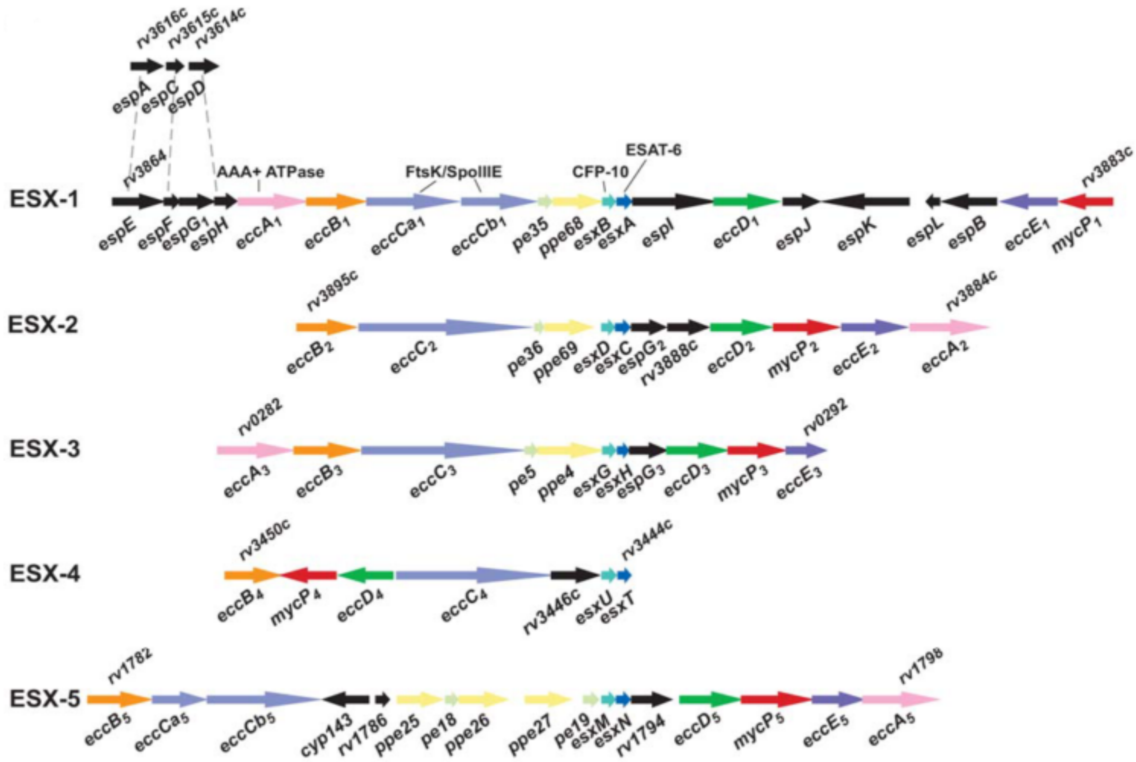



Figure 1.5. Type 7 secretion system gene clusters and organization. Model shows proposed organization of the 5 esx gene clusters in mycobacteria. eccC locus in esx-1 and esx-5 is splitted in to eccCa and eccCb gene 66.

T7SS substrate Esx-Esx proteins form a heterodimer protein complex each of them belong to WxG100 protein families formed by two conserved amino acids tryptophan (W) and glycine (G) separated by any single amino acids until an average of 100 residues. This heterodimer protein folds to secrete out of the bacteria. Similarly, Esx proteins also pair with non *esx* encoded protein in a similar fashion as that of Esx-Esx heterodimer which supports the possible occurrence of recent gene duplication event. Another important substance secreted through Esx apparatus is PE/PPE protein family which forms heterodimer in similar way of other Esx proteins and mainly encoded from esx gene clusters[35].

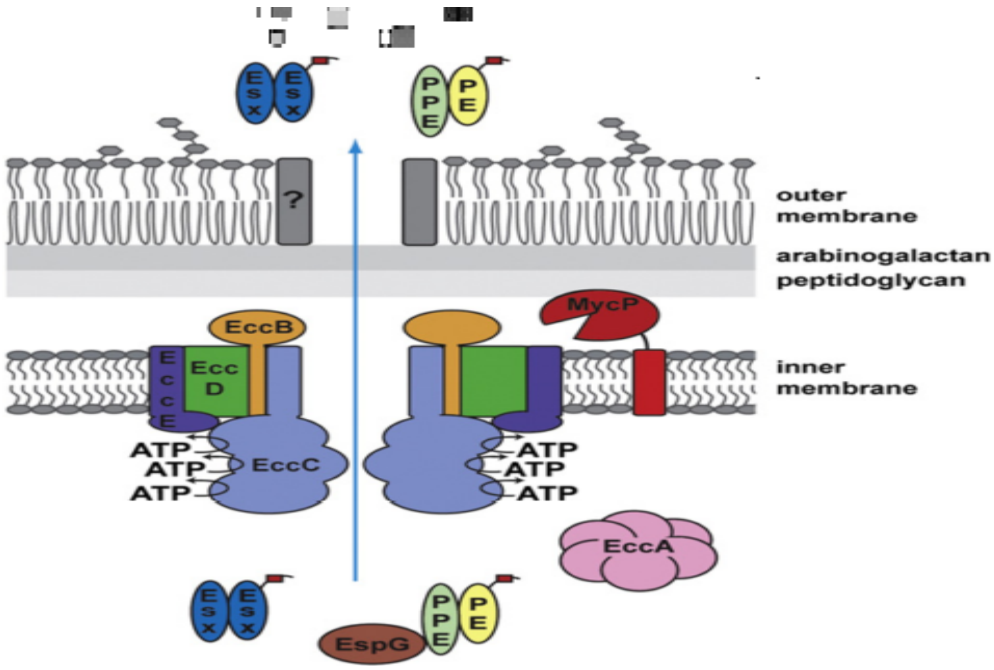



Figure 1.6. A model depicting components of mycobacterial type VII secretion system. esx encoded proteins including PE/PPE families are secreted through type VII secretion system by forming a dimer. EccB, EccC, EccD and EccE are possibly among the putative membrane channels important for secretion of substances like PE/PPE and iron bound molecules. EccC has 3 ATP binding sites that could produce energy for active translocation of substances out of the bacteria by interacting with the other putative membrane protein, EccB. Another long turn EccD protein with 11 transmembrane region predictions presumed to form channel on inner membrane. EspG could be putative chaperon protein important for secretion of substances. Although esx encoded proteins are assumed to form the inner membrane channel, the outer membrane protein channels also not yet understood35.

Conserved proteins encoded from *ecc*B, *ecc*C, *ecc*D, *ecc*E and *myc*P probably form membrane secretion channel. *esx-1* and *esx-5* gene clusters encode the EccC protein belonging to FtsK/SpoIIIE ATPase family from a split gene *ecc*Ca and *ecc*Cb. Similarly, EccD also predicted as transmebrane channel forming proteins which crosses the membrane 12 times. MycP is grouped under Subtilisin-like serine protease family and predicted as a membrane protein with its catalytic active site protruding outside. *ecc*A and *ecc*E genes are not the component of the ancestral esx-4 cluster indicating that they probably incorporated recently in the system [64]. In mycobacteria protein secretion and some function of Esx-1, Esx-3 and Esx-5 has been shown in several studies. Esx-1, Esx-3 and Esx-5 have also been shown to be virulence factors. For the remaining two type-VII systems, Esx-2 and Esx-4, there is no evidence of any functional role yet. Nevertheless, Esx-4 and Esx-3 are the only systems present in every species of mycobacteria and Esx-3 has an essential role in most species indicating that also Esx-4 may have an important function[35].

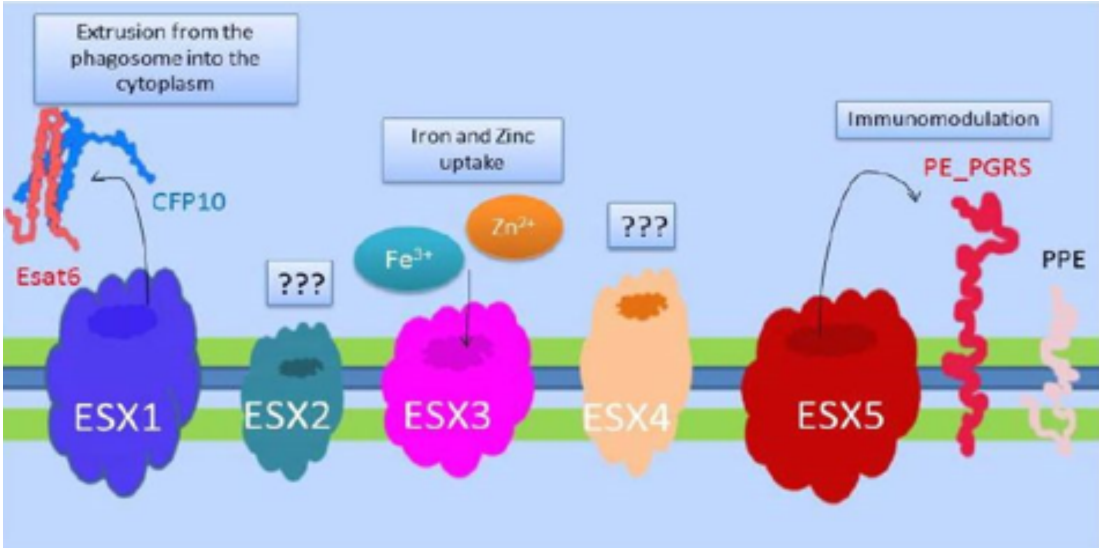



Figure 1.7. Function model of Type VII Secretion system. Mycobacteria T7SS proteins are encoded by Esx-1 to Esx-5 gene clusters. Proteins encoded from Esx-1, Esx-3 and Esx-5 has been revealed as important in the secretion of substances and virulence factors whereas the Esx-2 and Esx-4 encoded proteins have not yet found any functional role. Esx-1 proteins inhibit phagosome maturation by rupturing the phagosome membrane and let the bacteria to escape in to the cytosol before phagolysosome formation. Iron and zinc homeostasis is dependent on Esx-3 and Mycobacteria tuberculosis with Esx-3 deficiency is unable to grow in artificial media. The virulent proteins PE/PPE are secreted by Esx-5 which is predominantly found in slow growing mycobacteria Esx-5 is presumed to have an immmunomodulation function 68.

#### 1.7.1.1 **Esx-1 secretion system**

During earlier T7SS experiment, it was shown that Esx-1 secretion system is responsible for the secretion of small proteins important for virulence like ESAT-6 (EsxA) and CFP-10 (EsxB) [69]. Role of Esx-1 in *M. tb* and *M. smegmatis* may be different. In *M. tb*, its function is as secretion system for virulence factors while in *M. smegmatis* a role in conjugation has also been proposed[70]. It has also been shown that deletion of the Esx-1system is an important part of the attenuation of the vaccine strain of *M. bovis* which lost its virulence after several sub-cultures to become BCG vaccine[35,71]. In some mycobacterial species, partial loss of gene encoding Esx-1 substrate compensated by secretion of non esx encoded toxin. For example, loss of Esx-1 substrate in *Mycobacterium ulcerans* compensated by secreting cytotoxic toxin mycolactone which is potent virulence factor whereas, glycopeptidolipids (GPL) is synthesized from the only source pathogenic *Mycobacterium avium* in compensation of partial loss of esx-1[35]. A study in *M. tb* and *Mycobacterium marinum* showed that macrophage lysosome secretion and proinflamatory IL-1beta and IL18 cytokine discharge was dependent on Esx-1 system[36]. It was suggested that phagocytic cells autophagy machinery system could altered by Esx-1 system[72]. Esx-1 is also required for translocation of pathogenic mycobacteria in to cytosol before phago-lysosomal fusion which contributes for its virulence factors[73].

#### 1.7.1.2 **Esx-3 secretion system**

Esx-3 has been identified in all mycobacterial species. Mycobacterial Esx-3 secretion system has an important role in iron limited growth condition and involved in mycobacterial iron uptake both *in vitro* and *in vivo*[29,74]. Previously, *in vitro* study revealed that under sufficient iron the growth of *M. smegmatis* was not affected by absence of Esx-3[29,75]. However, growth was inhibited in mutant *M. smegmatis* where both siderophore systems where knocked out when grown in low iron. Interestingly, the same phenotype was observed when only one siderophore, exochellin, and esx-3 was knocked out. Mutants with mutations both in mycobaction synthesis and Esx-3 were able to grow under low iron conditions. This indicated that, Esx-3 secretion system is required for efficient utilization of iron bound with mycobactin [29]. Deleting one of *esx-3* components of a gene, impaired utilization of iron bound mycobactin similar to that of the entire esx-3 mutant[29]. Esx-3 secretion system also involved in the secretion of its product like EsxH and EsxG [75] which are homologous to EsxA/EsxB of Esx-1 and form similar structural and conformational fold [35]. Unlike *M. smegmatis*, *M. tb* is unable to survive in the absence of Esx-3 and thus it is not possible to construct esx-3 mutants of *M. tb* in a simple manner and conducting experiments in the pathogen. This could be because, mycobactin is the only iron scavenging siderophore in *M. tb* which is utilized through the Esx-3 secretion system where as saprophytic *M. smegmatis* acquire iron through both siderophores ( mycobactin and exochellin) predominantly by Esx-3 independent exochellin. *M. tb* and *M. smegmatis* have homologues *esx*-3 gene locus with about 44% sequence similarity and 85% at protein level[27].

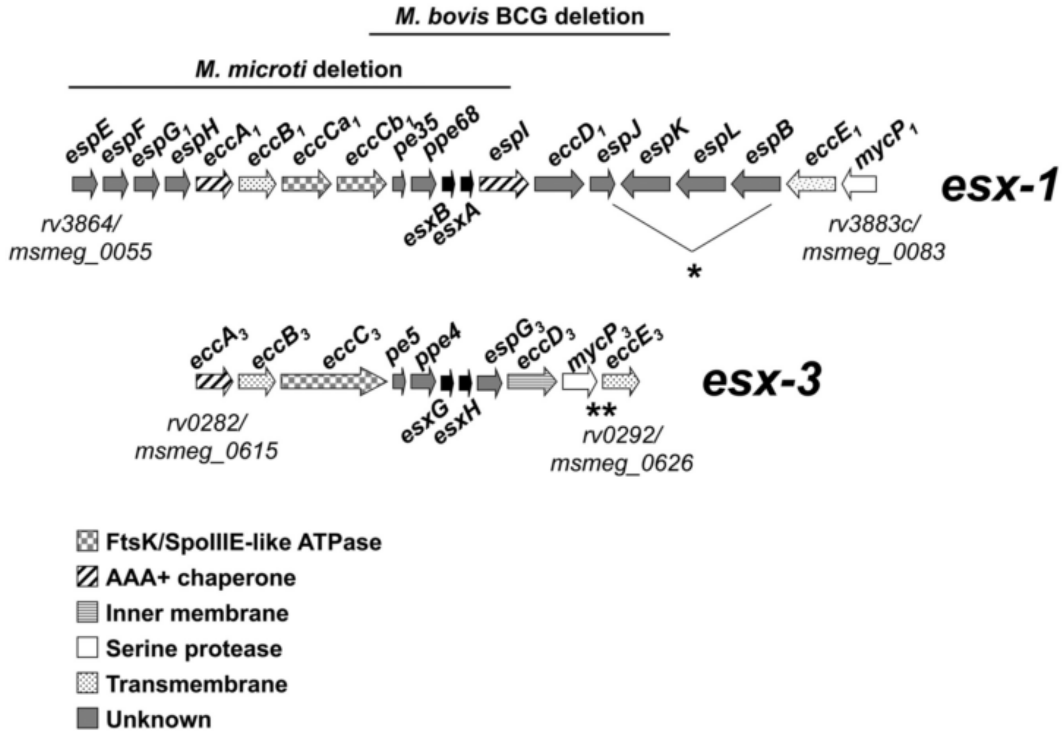



Figure 1.8. Arrangement and conservation of *M. semgmatis* esx-1 and esx-3 gene clusters. Putative encoded protein function/ location was described in legend. Esx-1 is one of well-studied T7SS and part of gene cluster has deleted in *M. bovis* and *M. microti* species. *M. smegmatis* esx-1 *esp*J,K,L and B gene arrangement was slightly different from *M. tb.*. In esx-3 gene cluster, *M.smeg*_0625 was missing while *M.smeg*_0624 and *M.smeg*_0626 are named as mycP3 and *ecc*E3 respectively [75].

### 1.7.1.3 **Esx-5 secretion system**

Esx-5 is identified as major PE/PPE protein complex secretory system and important in uptake of carbon from hydrophobic source[62]. The presence of Esx-5 in slow growing mycobacteria indicated as characteristics of most pathogenic mycobacteria species, and its absence in fast growing (saprophytic) seems to be the differentiation mechanism between this groups[35].

In a study by Di Luka *et al.,* showed that both membrane protein $EccB_5$ and $EccC_5$ are essential for *M. tb* growth *in vivo* as well as *in vitro* and also revealed that both proteins suggested being involved in translocation of substance across the membrane[76]. A study revealed that the majority of PE/PPE complexes are encoded by *esx*-5 gene clusters and highly involved in secretion of these proteins. Generally, 10% of mycobacterial genomic DNA encodes PE/PPE proteins which predicted to be secreted out by T7SS mainly through Esx-5[35].

## 1.8 Mycobacteria Stress response

Iron is an essential element important for cellular metabolism and oxygen transport and thus readily available to be used by the organism. Despite oxidation and reduction is an important phenomena in function of iron, it can produces free radicals and other reactive oxygen species which can damage the organism itself through different mechanism. The reaction between iron and hydrogen peroxide causes the initiation of Fenton reaction that can end up with biological damage. Fenton reaction is a reaction of iron with hydrogen peroxide that produces a higher oxidation state of iron ($Fe^{+4}$) and hydroxyl radical. Unregulated iron provoke the formation of radicals which can causes autoxidation, lipid peroxidation and interact with hydrogen peroxide to produce a potent reactive species [77]. To protect from such damage the organism produces several enzymes like alky hydroxide reductase C (AhpC) [78] superoxide dismutase (SOD) and catalase (KatG) besides to the regulation of iron chelators [77].

Pathogenic mycobacteria uses iron as a cofactor for oxidative detoxifying enzymes and many other metabolic processes. *M. tb* has two main family of metal dependent regulon which are Feric uptake regulator (Fur) protein family encoded from two related genes *fur*A and *fur*B wheras DtxR (diphtheria toxin repressor) protein family encoded from *ide*R and *sir*R.The regulatory activity of *ide*R is dependent on iron which in turn bind to DNA to which it regulates. Microarray study revealed that the transcriptions of several genes are regulated by *ide*R. To mention some, genes involved in siderophore synthesis, iron storage, transporters proteins, PE/PPE protein family, transcription regulators and lipid metabolism. Under high iron environment, *bfr*A is also positively regulated by *ide*R [79]. DNA sequence of *fur*A in *M. smegmatis* has shown homologous with *M. tb* [80]. Mutation of *kat*G in *M. tb* is one of the common challenges in first line TB INH drug resistant [81]

Under oxidative stress, *M. smegmatis* induced expression of gene analogous to the main gram negative bacteria oxidative stress regulator gene *oxy*R which act as sensor of transcription activator and ROS by inducing expression of enzyme encoding genes involved in oxidative stress. *M. tb* is adapted to resist toxic oxidative radicals and persist in phagocytic cells and in granulomatous caseous lesion by mechanism different from the one mentioned in gram negative defence mechanism *oxy*R. This indicates that pathogenic mycobacteria has a unique defense mechanism against oxidative stressor which able to survive in potent oxidative environment in an alveolar macrophage. Despite the presence of intact *oxy*R in *M. tb*, numerous mutations could make the gene inactive in oxidative response[82].

Mycobacterial *ahp*C gene encodes alkyl hydro peroxide reductase C which involved in the activity of peroxinitrite reductase and peroxidase[83]. AhpC contain two cysteine residues responsible in catalytic activity. The first one is peroxidatic cysteine residue which reduces peroxides or peroxinirite in to cysteine sulfenic acid. The second cysteine residue is sulfhydryl group also called resolving cysteine which attacks the first catalytic product sulfenic acid in to disulfide form. The end product from the second cysteine catalytic activity disulfide bond is reduced and recycled back by another enzyme AhpD peroxiredoxin reductase encoded by ahpD gene located downstream of *ahp*C gene. AhpC also can be reduced by NADH dependent thioredoxin reductase and thioredoxin C (TrxC) [84]. Mycobacteria ahpC is peroxidoxin NADH dependent peroxidase and peroxynitrite protein that reduced by AhpD [85]. *ahp*C and *ahp*D form as operon [86]. On upstream of *ahp*CD operon, most mycobcateria possess *oxy*R homologue LysR family regulators which is also responsible in oxidative stress response by inducing the expression of *kat*G and *ahp*C. Despite it is inactivated in some *M.tb* complex species due to several mutations such as in *M. tb, M. bovis, M. africanum, M. microti* [87]. *oxy*R is totally absent in *M. smegmatis* while present both in *M. bovis* and *M. lepreae*. This shows the oxidative stress response in pathogenic mycobacteria looks quite different from the saprophytic *M. smegmatis* [51]

In the absence *oxy*R, mycobacteria *ahp*C gene expression is increased in response to different oxidative stress inducers such as hydrogen peroxide, diamide, organic hydroperoxides (Cumene hydropeoxide), tert-butyl hydroperoxide [88,89]. A study, in *M. tb* revealed *ahp*C mutant exposed for susceptibility to organic hydroperoxides. It was also suggested under oxygen limited growth condition such as in static growth condition, *ahp*C expression was activated in pathogenic *M. tb*. Microarray expression study revealed that mutation of *M. tb* crp (RV3676) gene which encode cAMP receptor protein and mutation of two component system led the downregualtion of *ahp*C [90] whereas upregulated in *whi*B transcription factors mutants [78]. It was also demonstrated that upregulation of *oxy*S gene induce the down regulation of *ahpC* in *M. tb*. In previous study, mutation of *M. smegmatis sig*F didn't affect the expression of *ahp*C [78]. In *M. smegmatis* exposure to oxidative stressor $H_2O_2$ increases the level of intracellular cAMP which in turn activated expression of *ahp*C. cAMP is activated *ahp*C via binding with Crp (cAMP receptor protein ). *ahp*C is negatively regulated by *fur*A and under oxidative stress condition the activation of Crp is activated due to increasing the level of cAMP [78].

In a study, level of *sod*A and *kat*G encoded proteins was reduced in *ide*R mutant *M .smegmatis* strains and susceptible to the anti-tuberculosis drug INH. These data confirms that IdeR is responsible in the regulation of oxidative stress response and has a protective function against ROS and INH in *M. smegmatis*. Reactive oxygen species (ROS) is one of the well-studied innate immune response induced by phagocytic cells and demolish the function of proteins, lipids and DNA. Oxidative exposure response in mycobacteria was not well characterized as of the gram negative bacteria. Gram negative bacteria respond to hydrogen peroxide stress response through *oxy*R regulation. *M. smegmatis* oxidative response against $H_2O_2$ exposure was comparable to *oxy*R of gram negative despite; no any similar gene has been found yet. *ahpC* gene mutation in *M. smegmatis* exposed for highly susceptibility to INH which induces reactive oxygen intermediates. Besides to *oxy*R, it was also presumed that *kat*G is positively regulated by *fur* as the mutation of both genes down regulate the expression of katG [51].

Mycobacteria survive and evade the killing mechanism of the host cells by inhibiting phagosome maturation, adapting with toxic substance of the host and inhibit apoptotic cell death[91,92]. Up on ingestion of bacteria, phagocytic cells produces toxic reactive nitrogen species (RNS), reactive oxygen species (ROS) which are efficient in killing microorganism. AhpC, SodC, TpX and KatG are among the notable *M. tb* complex enzymes which detoxify directly the toxic effects from ROS and RNS in addition to other antioxidant proteins like mycothiol [91]. Mycothiol production helps as cytosolic redox buffering. Different study shows that the absence of mycothiol exposed *M. tb* for the susceptibility of antibiotics [92]. Organic molecules like DNA, protein and carbohydrates are also affected by reacting with ROS and RNS intermediates. However, most intracellular microorganisms adapted to evade the toxic effect of this reaction with different strategy. Mycobacteria have a unique defense mechanism compared with other intracellular bacteria. For example, OxyR is one of the defense mechanism for most intracellular bacteria which is not functional in all *M. tb* complex [91] while it is totally absent from the saprophytic *M. smegmatis* genome [78]

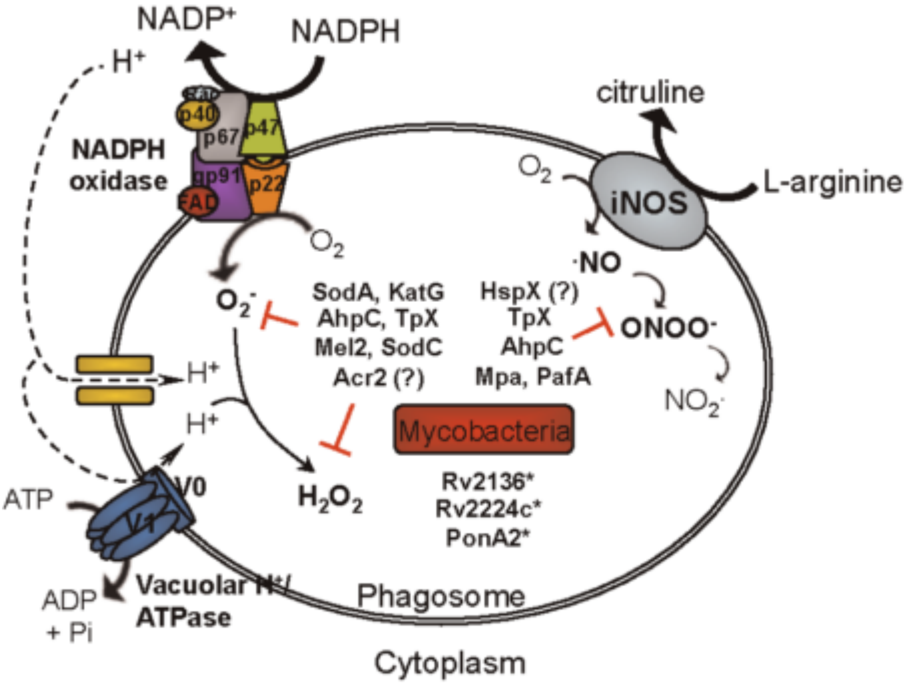



Figure 1.9. Mycobacteria Oxidative and Nitrosative stress response inside phagocytic cell phagosome. Mycobacteria redox regulatory genes encode proteins that detoxify reactive oxygen and nitrogen species intermediate. Some of the gene directly hydrolyze $H_2O_2$ [31].

The transcription factors *oxy*R and SoRs regulate bacterial genes involved on the oxidative stress response. Where *oxy*R and *sox*RS are activates these genes in response to peroxide and superoxide respectively[93]. Besides to microbial enzymes, uric acid, beta carotene, alpha tocopherol, ascorbic acid and glutathione are some other antioxidant substances. Although non mutant gene was found on *M. smegmatis*, *M. bovis* and *M. leprae*, the functional role is replaced by other stress response like *ahpC*. Mycobacteria could have unique defense mechanism against oxidant which makes to survive inside the host cells[94].

# 2 Aim of the study

*M. tb* is a very successful pathogen, both because of unique defence mechanisms and ability to escape from the host immune system. The emergence of MDR and XDR strains compromised treatment efficacy of anti-tuberculosis. Several years has been lapsed since the last effective anti-tuberculosis Rifampicin discovered. Thus, finding a novel susceptible target, effective drug and vaccine discovery is an important task in the control of devastating TB disease [9]. Detailed understanding the function of mycobacteria secretion system proteins could facilitate the root of new drug target against mycobacterium infection. Novel anti TB drugs targeted to the system protein could also significantly shorten the duration of standard drug treatment as well as limit potential side effects that are normally associated with the current available anti TB treatment. Therefore, identification function of essential mycobacterial proteins that maintains the survival of bacteria *in vivo* and *in vitro* will find the possible solution against TB disease. Previous study showed that esx-3 is involved in iron acquisition[74,75] and suggested as it could have also other unidentified role. Here, our aim is to identify novel function of esx-3 secretion system.

This study has the following objectives:

- Optimize mycobacteria RNA extraction method to extract high quality RNA.
- Identify role of *esx-3* in metabolic pathways other than iron uptake by NanoString and quantitative real time PCR gene expression analysis
- Investigate role of *Mycobacterial smegmatis* esx-3 in oxidative stress response under low and high iron condition.
- Identify localization of EspG3, EccB3 and EccC3 proteins in *Mycobacteria smegmatis.*

# 3 Materials and methods

## 3.1 Bacteria strains and growth condition

### 3.1.1 Mycobacteria smegmatis (M. smegmatis)

Non-pathogenic wild type (WT) 155 $mc^2$, Δ*esp*$G_3$, and Δ*mbt*D strains[75] were used for all proceeding mycobacteria experiments. Bacteria strains were incubated at 37 $^0$C in liquid or solid media. As a starting culture, mycobacterial glycerol stock strains kept at -80 $^0$C were thawed and diluted 1:20 with 7H9 media (Section 2.1.1) and incubated in a shaking incubator until the culture was reached stationary phase. Sub culturing was done in either 7H9 or Sauton's (Section 2.1.2 and 2.1.3) with or without iron from an initial starter culture for subsequent experiments. All liquid media culture was incubated in a constantly shaking incubator until the desired optical density $OD_{600}$ was reached. Antibiotic concentration used for bacterial selection was 50 µg/ml hygromycin, 25 µg/ml kanamycin and 100 µg/ml ampicillin as required.

Bacterial suspensions were also plated on solid plate media using a sterile glass rod or spotted, depending on the experimental requirement. The plates were incubated at 37 $^0$C incubator. Single colonies were picked up and grown in liquid media for immediate use or stored as glycerol stocks at -80 $^0$C if required.

### 3.1.2 **Escherichia coli (E. coli)**

Chemically competent DH5α *E. coli* strain (Randi Vik, NTNU) was used for Gateway cloning purposes and plasmid transformations. Plasmids containing antibiotic resistant genes were selected and the strains were incubated at 37 $^0$C in Luria Broth (LB) or Luria Agar (LA) media in constantly shaking or non-shaking incubators respectively. Antibiotic concentration used for bacterial selection was 150 µg/ml hygromycin, 50 µg/ml kanamycin or 200 µg/ml ampicillin as required.

## 3.2 Growth chamber

BIOSCREEN C Automated Growth curve machine was used to continuously follow the growth of bacteria in liquid media. A 10 x10 wells honeycomb microplate was used to pipette 200 µl bacterial suspensions in to each well. In order to minimize error each culture suspension was run in triplicates. Optical density was measured in 2 hours intervals at constant temperature 37$^0$C incubation. The machine was built in horizontally constantly shaking incubator and maintained 37$^0$C. The growth period was programmed for 3-5 days and data analysis was done using graphPad Prisim.

## 3.3 Media

### 3.3.1 **Liquid Media**

#### 3.3.1.1 **Middlebrook 7H9 broth enriched with Albumin Dextrose Catalase (ADC)**

7H9 media was prepared according to the manufacturer instructions (BD Difco). For 1liter media, 4.7 grams of 7H9 powder was diluted with 900 ml water containing 2 ml glycerol. After 15 minutes autoclaving at 121 $^0$C, sterile Albumin Dextrose Catalase (ADC) enrichment (BD) and Tween 80 (sigma) were aseptically added to a final concentration of 10% and 0.05% respectively. The complete media was kept in the fridge until use.

#### 3.3.1.2 **Sauton's growth Media (Low iron)**

0.5 g potassium di-hydrogen phosphate, 2.2g citric acid monohydrate (Merck), and 4 g L-aspargine (sigma) salts were dissolved with the help of magnetic stirrer in distilled water containing 60 ml 85% glycerol (Merck) until the volume reached 1 liter. The PH was adjusted to 7.2 by titrating with 1M sodium hydroxide and the solution was autoclaved at 121 $^0$C for 15 minutes. After cooling, sterile 1 g $MgSO_4.7H_2O$ and 20% Tween 80 (Sigma) were added to make 0.05% final concentration of Tween 80. This media was used for the growth of *M. smegmatis* without iron supplement.

#### 3.3.1.3 **Sauton's growth media with iron**

Filtered sterile 15 mM $FeCl_3$ solution was added to the Sauton's media, described in section 2.1.2 to make a final concentration of 150 µM. This media was used for *M. smegmatis* growth with iron supplement.

#### 3.3.1.4 **Luria Broth (LB) media**

5 g Yeast extract (Oxoid), 10 g Tryptone (Oxoid) and 5 g NaCl (Merck) were dissolved in 1 liter purified water. Magnetic stirrer was used to dissolve the solute completely and autoclaved at 121 $^0$C for 15 minutes. This LB media was used for cultivating *E. coli.* Selective antibiotics were added when needed.

#### 3.3.1.5 **10% Glycerol**

10% Glycerol was prepared from 85% glycerol (Merck) stock and autoclaved at 121 $^0$C for 15 minutes. This 10% glycerol was used to wash electro-competent *M. smegmatis* pellet before electro-proration to remove extra salts from the cell suspension and enhance electro transformation efficiency.

### 3.3.2 **Solid Media**

#### 3.3.2.1 **Middlebrook 7H10 agar media**

Based on the manufacturer's instructions, 19 gram of 7H10 (BD Difco**)** powder was dissolved in 900 ml purified water containing 5ml (85%) glycerol. The mixture was completely dissolved by magnetic stirrer and autoclaved at 121$^0$C for 15 minutes. 100 ml ADC and if required antibiotics were added after cooling down to 45 $^0$C. Then dispensed to the Petri dishes to solidify and kept in plastic bag in the fridge until used. 7H10 agar media was used for seeding successful transformant *M. smegmatis,* isolating single colonies and spotting bacterial suspension.

#### 3.3.2.2 **Luria Agar (LA) media**

5 g Yeast extract (Oxoid), 10 g Tryptone (Oxoid), 5 g NaCl (MERCK) and 15 g agar (Oxoid) were dissolved in 1 liter purified water using magnetic stirrer and autoclaved at 121 $^0$C for 15 minutes. When the temperature cooled down at around 45$^0$C, desired antibiotics were added as required and poured in sterile plate. Solidified and cooled agar plates were kept upside down at 4 $^0$C in a plastic bag until used for bacterial growth. Concentration of NaCl in media

was kept below 5 g/ml if hygromycin was used as it is salt sensitive and less stable in high salt concentration. LA media was used for seeding *E. coli* strains after transforming drug resistant plasmids in to bacteria.

## 3.4 RNA and DNA related work

### 3.4.1 ***M. smegmatis* RNA extraction**

Mycobacterial RNA was extracted by combining mechanical, chemical and enzymatic purification methods. TRIZOL (Invitrogen) and bead beating methods were combined with Qiagen column DNase digest and TURBO DNase digest protocol. 600µl TRIZOL was added on bacterial pellet taken from 10ml mid-log phase growth with $OD_{600}$ between 0.6 – 1.2. Then bacterial suspension was transferred to 2 ml screw-caped tubes containing 0.1mm glass beads of ~ 300µl volume. Using bead beating machine (FastPrep®-24 Instrument) at ~5000 RPM, the tubes were beaten for 2 minutes with 5 minutes pause between each minutes of beating. After 5 minutes room temperature incubation, 150µl chloroform was added to the lysate followed by 15 seconds vigorous shaking and then centrifuged at 12000 RCF for 15 minutes at 4 $^0$C temperature. The addition of chloroform separates the mixture in to 3 phases where, proteins remains at the bottom in organic phase, DNA at the interface and RNA suspended in uppermost aqueous phase. The aqueous phase containing RNA was carefully transferred in to 1.5 ml RNase free Eppendorf tubes and 300 µl cold isopropanol (Teknisk) was added and kept at -20 $^0$C for 15 minutes. The mixture was again spun down at 12000 RCF for 15 minutes at 4 $^0$C and the supernatant was discarded. RNA pellet was washed by 700 µl of 70% RNAse free ethanol and left for about 5 minutes to air dry. Then double DNase digest protocol was followed to maximize the purity of RNA. First the RNasy mini spin column DNase digest protocol (Qiagen kits) and then Turbo DNase digest protocol (TURBO kits) were followed. Quality of the purified RNA was checked by NanoDrop 1000 spectrophotometer and kept at -80 $^0$C for the experiments planned.

### 3.4.2 **cDNA synthesis**

High- capacity cDNA reverse transcription kits was used for synthesizing cDNA from RNA. The reaction components were as follows and equal volume of RNA samples was mixed with master mix.

| Components of additives | Volume per reaction mix |
|---|---|
| 10X RT buffer | 2 µl |
| 100mM 25X dNTPs mix | 0.8 µl |
| 10X RT random primers | 2 µl |
| Reverse transcriptase | 1 µl |
| Nuclease free water | 4.2 µl |
| Master mix total | 10 µl |
| RNA samples | 10 µl |
| Total volume | 20 µl |

The cDNA was normalized to working concentration of 2 ng/µl or 5 ng/µl by diluting with nuclease-free water depending on the required amount of template. The thermo cycler program for cDNA synthesis according to Applied Biosystems High-Capacity cDNA reverse transcription kit, shown below, was followed.

| | Step 1 | Step 2 | Step 3 | Step 4 |
|---|---|---|---|---|
| Temperature | 25 $^{0}$C | 37 $^{0}$C | 85 $^{0}$C | 4 $^{0}$C |
| Time | 10 minutes | 120 minutes | 5 minutes | forever |

### 3.4.3 **NanoString mRNA expression (nCounter NanoString technology)**

NanoString nCounter is a novel digital technology for measuring the quantity of nucleic acids. The ability of this method to directly detect and measure absolute quantities of RNA , digital analysis and direct sequence examination makes it advantageous over other PCR based experiments [95]. NanoString detects target molecules by labelling with molecular barcode and color coded paired probes. The two probes are reporter probe and Capture probes. The signal of reporter probe is carried on its 5’end whereas capture probe is labelled with biotin on its 3’ end. The reporter probe has a total of seven color codes and each position can be one of four colors. This allows specific target hybridization among large diverse tags in a single reaction mixture.

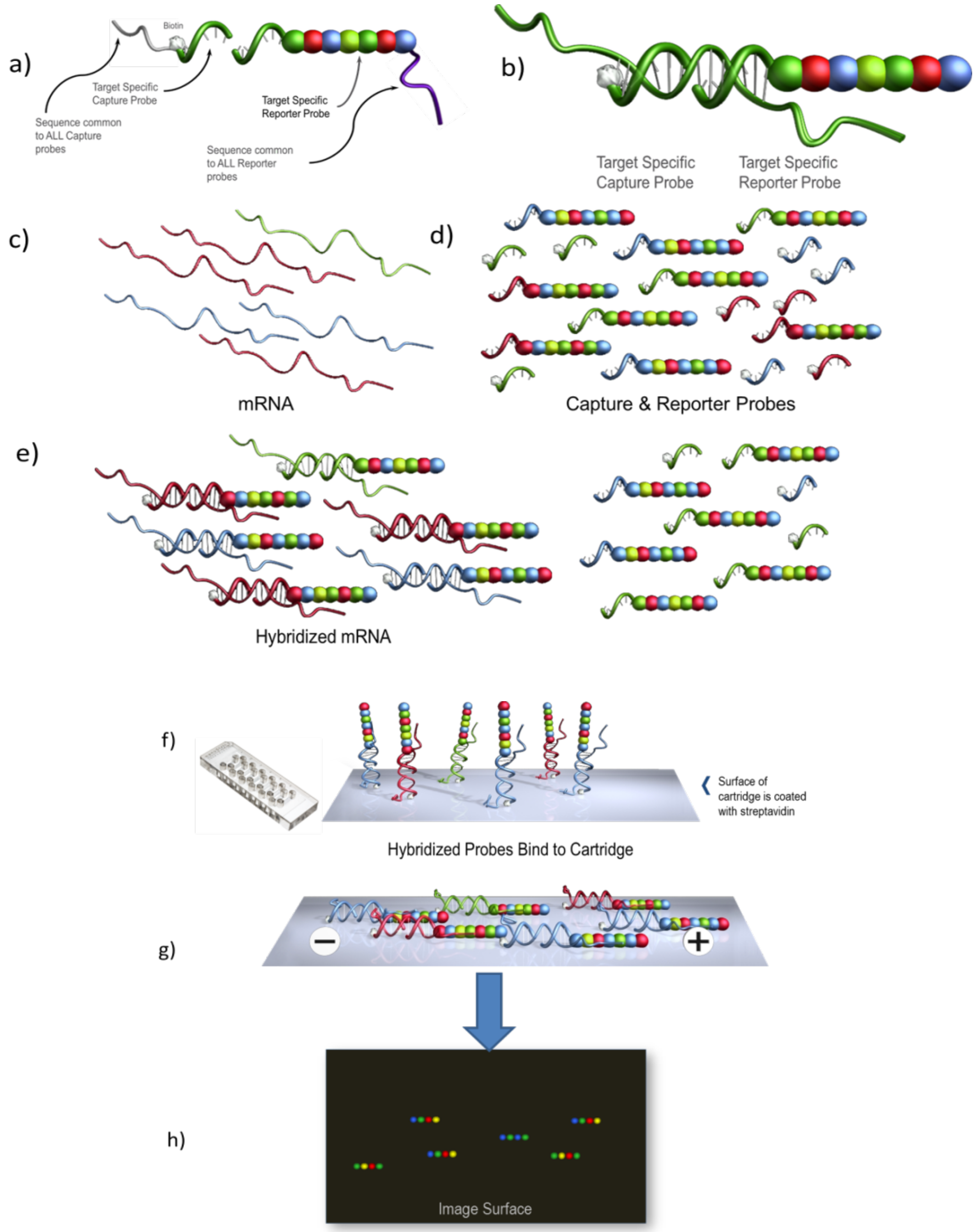



Figure 3.1. Illustration of components and steps in NanoString nCounter assay. (a) Structure of capture probe labeled with biotin and common sequence to all capture probes (left) and color coded target specific reporter probe with common sequence to all reporter probes (right). (b) Target specific capture probes and reporter probes specifically hybridized with unique mRNA sequence. (c) Different mRNA from biological sample. (d) Variety of reporter probes and capture probes that was incorporated in hybridization mix. Each unique CodeSet specifically hybridized with its corresponding specific target sequence. (e) Specific mRNA sequence from the sample hybridized with both target specific sequence of capture and reporter probes with one to one ratio at $65^{0}$C without any addition of enzyme. Any excess and non-hybridized probes were removed by robotic prep station washing steps. (f) Washed mRNA probe complex transferred in to 12 different wells of cartridge (left). The surface of cartage labeled with streptavidin has high affinity with biotin labeled on hybridized complex and stand vertically. (g) In the final prep station steps, the hybridized complex immobilized by electric current align from anode to cathode direction which prepared ready for imaging and counting. (h) After all prep station steps (a-g) the cartridge was transferred to separate

nCounter Digital analyzer for imaging and data collection. One unique color coded reporter probes interpreted as a single mRNA (adapted from NanoString Technologies nCounter® analysis system).

Based on manufacturer's instructions, 100 ng purified RNA samples were hybridized with specific biotinylated capture probe and color coded molecular barcode overnight (20 hours) at 65 $^0$C programmed thermo cycler. The hybridized mix and reagents required for the reaction were fed in to fully automated NanoString robotic prep station. Unbound reporter probes were washed out and cartridge bound reporter were immobilized with an electric current in a prep station robotic machine. During the reaction, Sterptavidin labeled over cartridge was bound with the biotin labeled on hybridized RNA reporter probe mixture. The prep station process took about 3 hours. Then the cartridge was taken carefully to digital analyzer which captured thousands of images of reporter probes bound with RNA. One unique color coded reporter probe hybridized with a specific target mRNA sequence and the digital analyzer count the number of reporter image that corresponds to the amount of reporter specific RNA within about 5 hours. The data was normalized with *M. smegmatis* housekeeping gene (*sig*A) in addition to positive and negative controls before analysis and interpretation.

### 3.4.4 **Quantitative real time PCR (qPCR)**

Quantitative or real-time PCR is a method that determines the amount of DNA based on monitoring the intensity of fluorescent double strand DNA intercalating dye[96]. In this experiment SYBR green fluorescent dye was used which intercalated with minor grove of any double strand DNA in the sample. As specific primers anneal to target cDNA template, polymerase precede the amplification and the intensity of fluorescent SYBR green dye increase corresponding with increasing amplification. The higher the amount of double stranded DNA in the sample, the more intensity of the fluorescent detected earlier. These determined highly expressed target genes fluoresce earlier in the cycle whereas detectable fluorescent signal in less expressed gene was in the later cycles. *sig*A was used as a house keeping gene to normalize the amount of cDNA template used. The normalization was determined by calculating the expression deference between each target gene and *sig*A in separate well. The normalized data of each sample was used to compare the expression of genes in different conditions such as in iron treated and non-treated samples.

PerfeCTa SYBR Green FastMix, ROX kit was used for real time PCR master mix preparation. The reaction components and volume used was as follows

| Components of additives | Volume per reaction mix |
|---|---|
| PerfeCTa SYBR Green Fast Mix | 10 µl |
| 10 µM Forward Primer | 0.5 µl |
| 10 µM Reverse Primer | 0.5 µl |
| Nuclease free water | 4 µl |
| 2 ng/ µl cDNA | 5 µl |
| Total volume | 20 µl |

20 µl of Mastermix and cDNA template mix per reaction were loaded on 96 well PCR plate and triplicates were run for each cDNA sample. The plates were properly sealed and centrifuged at 1500 RPM for 3 minutes in order to prevent bubble or foam formation.

Applied biosysytems real-time PCR machine was programmed to amplify the samples according to the following standard thermo cycling steps which was one of the recommended programs by the kit. StepOne plus software, version 2.3 for 96 PCR well plates, was used with the following settings: 30 seconds initial denaturation followed by 40x PCR cycling for 5 seconds and 30 seconds final extension. Quantitation-comparative $C_T$ ($\Delta\Delta C_T$) for SYBR green reagents was chosen.

### 3.4.5 **Conventional PCR**

1 µM final concentration of each primer was mixed with 2X GO Taq green master mix (Promega) containing dNTPs, polymerase and buffer pre-mix per reaction in separate 1.5 ml tubes. Nuclease free water was added to adjust the final volume of master mix and 24 µl volumes were dispensed in each 0.2 ml PCR tube. 100 ng DNA template extracted from wild type *M. smegmatis mc*$^2$155 strain genome was added in to each labelled tubes to amplify *esp*$G_3$*,* *ecc*$B_3$ *and ecc*$C_3$. In separate reaction tubes, 100 ng promoter, esx-3 and reporter gene DNA was amplified. The Thermo cycler program for amplification was as follow: Denaturation 95 $^0$C for 30 sec, Annealing 55 $^0$C for 30 sec and elongation 72 $^0$C for 3 minutes and all 3 steps repeated for 25 cycles (BIO RAD). Initial denaturation and final elongation temperatures were 95 $^0$C for 3 min and 72 $^0$C for 5 minutes respectively.

### 3.4.6 **Gel Electrophoresis**

1% agarose powder (Lonza) was suspended in 1X Tris-Acetate EDTA (TAE) buffer and boiled in a microwave until fully dissolved. When the temperature cooled down to about 55 $^0$C, 50 ml agarose was mixed with 1 µl Gel red and casted on labelled flat tray. 1kb DNA ladder was used as a molecular weight marker. Samples were run at 80 volts for about 1 and ½ hours. The gel image was captured by Gel doc built in with camera (Carestream Gel Logic 212 PRO).

### 3.4.7 **Purification of PCR product**

If PCR product obtained was only from the target specific, Qiagen QIAquick PCR purification protocol was used to purify it. However, if multiple size PCR bands were observed in the gel electrophoresis, gel extraction protocol was followed. With the help of UV-light, the desired DNA band was excised and purified using Qiagen QIAquick Gel Extraction protocols.

### 3.4.8 **Plasmid isolation**

*E. coli* containing plasmids were grown in LB media with selective antibiotics overnight at 37 $^{0}$C shaking incubator. Isolation of plasmid was followed and done based on PureYield plasmid Miniprep system protocol (Promega kit). The isolation step was started from protocol for larger volume cultures as mentioned on the protocol.

## 3.5 Cloning

### 3.5.1 **Primer design**

attB/attBr flanked primers were designed based on multisite gateway cloning technology (clone manager , cmsuite9 software). Designed primers had 48-52 nucleotides in length which included 4 G's at 5' end, 22-25 attB or attBr reaction specific recombination sites followed by 18-25 nucleotides sequence of gene of interest without any stop codon. At the end of forward primers, 2 nucleotides were added to maintain the translated sequence codon in frame.

### 3.5.2 **Multisite Gateway cloning**

Multisite cloning technology is developed to clone multiple fragments of DNA by using specific reaction site sequences. It is based on principle of bacteriophage lambda site specific recombination system by two consecutive reaction pathways; the lysogenic and lytic pathway which are phage lambda integrase ( Int) and *E. coli* integration host factor (IHF) proteins ( BP clonase II enzyme mix) and phage lambda Int and excionase protein(xis) proteins and *E. coli* IHF protein ( LR clonase II plus enzyme mix) respectively (Invitrogen).

Three fragment Multi site Gateway cloning technology was applied to recombine promoter, target and reporter genes. The PCR product of the genes were purified by Qiagen kit (section 3.7) and stored at -20 $^{0}$C freezer until used for cloning. Specific recombination reaction flanked plasmids (attP flanked pDONR) were used to react with each purified PCR on 3 separate tubes. In the presence of enzyme BP clonase II, each unique attB flanked sites recombine between pDONR and PCR product (attB-attP) without any net gain or loss of nucleotides which created new reaction sites, attL. Invitrogen Multisite Gateway pro user guide was followed for primer design and cloning. In the BP recombination reaction,100 fmol of purified attB flanked PCR products from promoter, target genes and reporter gene were mixed with 150 ng of pDONR 221 p1-p4, pDONR p4r-p3r and pDONR p3-p4, respectively in separate tubes. 2 µl BP clonase II enzyme mix was added in to each reaction mix and incubated overnight at 25 $^{0}$C. Then, 1 µl of proteinase K was added and incubated at 37 $^{0}$C for 10 minutes which digests proteins in the reaction. The plasmids were transformed in to chemically competent DH5-α *E .coli* and then isolated (section 3.8) from successfully transformed bacteria. The quality and quantity was checked by NanoDrop and gel electrophoresis. In addition, to ensure the exact gene was inserted in to the entry clones that might be mutated during PCR, the entry clones were sequenced after BP reaction using the primers provided with the kit. Primers used for sequencing were:

**M13 Forward (-20):** 5'-GTAAAACGACGGCCAG-3'
**M13 Reverse:** 5'-CAGGAAACAGCTATGAC-3')

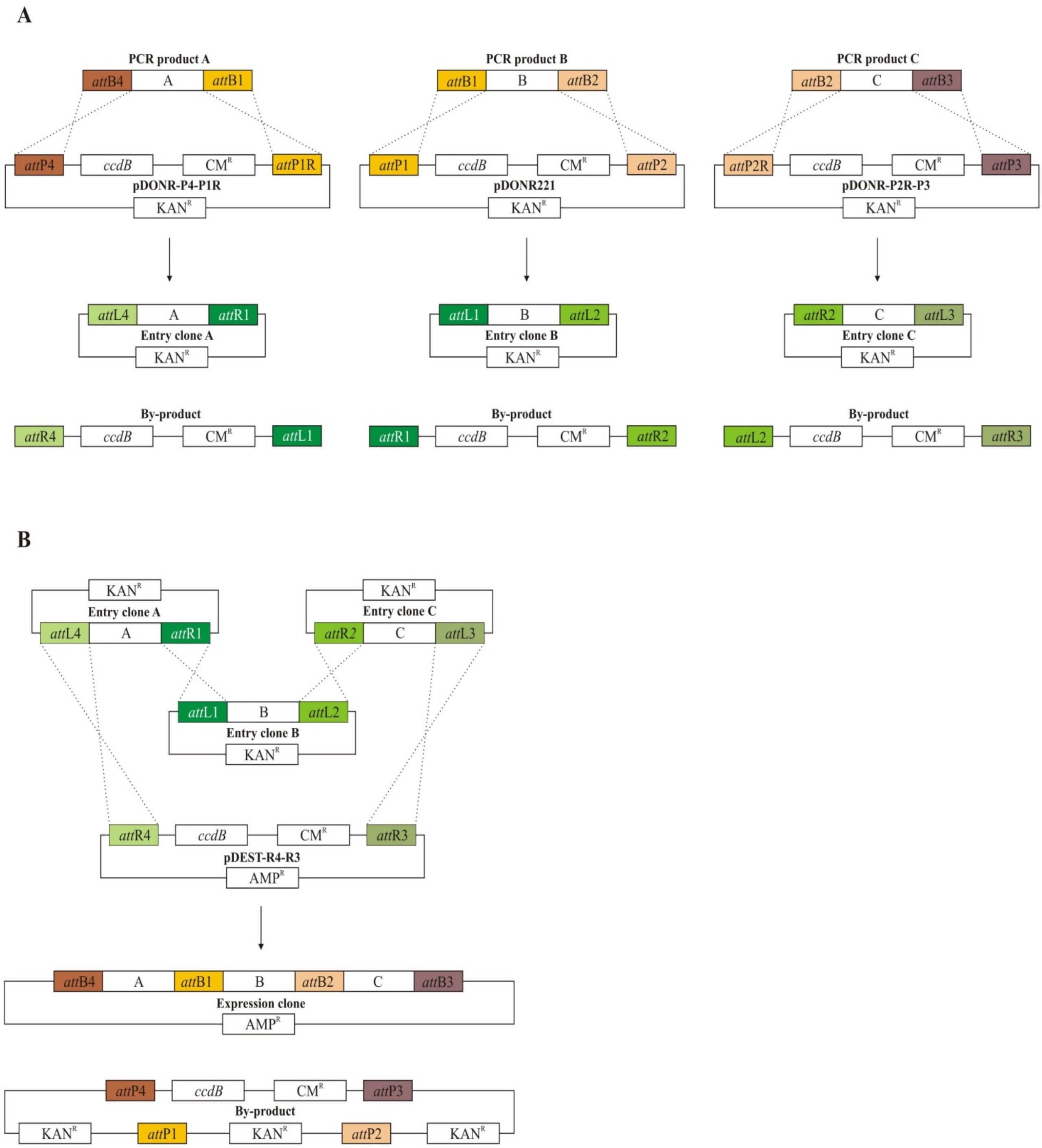



Figure 3.2. Schematic presentation of three fragment PCR products cloning using multi-Site Gateway cloning system. **a**) BP reaction. b) LR reaction. Target genes PCR products A, B and C amplified by attB flanked primers cloned into donor vectors (pDONR) and become entry vectors by which in vitro recombination of PCR products occurred during the BP reaction. In LR reaction, PCR products from the three entry vectors transferred into a Destination Vector (pDEST) by multisite sites specific recombination reaction [97].

All the 3 verified plasmids were then mixed with destination vector flanked with attR1 and attR2 reaction sites and LR clonase II plus enzyme. The corresponding attL sites of DONR vector recombine with attR sites of pENTR to recombine the promoter-target- reporter genes together without any stop codon in between. The promoter promotes expression of all fused gene at a time. Cloned kanamycin resistant pENTR expression plasmids were transformed in to competent *M .smegmatis* to express the target proteins. The pDONR and pENTR vectors had the ccdB gene cassette flanked by the recombination sites which encode a protein toxic for the enzyme gyrase in DNA replication. If the recombination was unsuccessful, the ccdB gene would remain in the plasmids and express the toxic protein and the bacteria would not survive. Transformation was done on competent *M. smegmatis* strains and expression of target gene was observed indirectly by exciting the reporter gene with the correct wavelength of beam of light.

### 3.5.3 **Transformation**

#### 3.5.3.1 **Electrocompetent *M. smegmatis* preparation**

*M. smegmatis* strains were sub-cultured in 7H9 from frozen stocks until mid-log phase $OD_{600}$ value reached 0.8 – 1.2. Bacterial cultures were spun down at 3500 RPM , 4 $^{0}$C for 10 minutes and the pellet washed with equal volume of 10% glycerol at 4 $^{0}$C. The wash step was repeated by adding 10% glycerol ½ $^{th}$ and 1/10$^{th}$ of original volume. The final pellet was re-suspended in 10% glycerol until it became thick (usually 1/100$^{th}$ original volume) and used for immediate electroporation or stored at -80 $^{0}$C for future work. 400 ng purified plasmid was mixed with washed bacterial pellet and left 5 minutes at room temperature in order to let the plasmid stick to the bacteria.

#### 3.5.3.2 **Electroporation**

The Electroporator apparatus setting was 2500 volts (2.5 kv) voltage, 25 µF capacitor and 1000Ὠ resistor. Just after electroporation, the transformants were recovered in 1 ml 7H9 at 37 $^{0}$C for about 4 hours and plated on7H10 agar plates containing antibiotic.

#### 3.5.3.3 **Heat shock transformation**

2 µl of each BP (Section 4.2) reaction mixture (Promoter, esx-3 and reporter) was added into 25 µl of ice thawed competent DH5-α *E. coli* cells on separate tubes. The reaction was incubated for 30 minutes on ice and followed by heat shock transformation at 42 $^{0}$C for 30 seconds followed by 2 minutes incubation on ice. Then, 250 µL of pre warmed, room temperature, S.O.C. medium was added and incubated at 37 $^{0}$C for 1 hour in constantly horizontal shaking incubator adjusted to 300 rpm. From each reaction mixture, 20 µl and 100 µl volumes were seeded in Luria Agar plates containing 50 µg/ml kanamycin for overnight

incubation at 37 $^{0}$C. Water with DH5α was also included as a negative transformation control.

## 3.6 Microscopy

### 3.6.1 Sample preparation

2 ml *M. smegmatis* culture with $OD_{600}$ between 0.6- 0.8 grown under kanamycin containing 7H9 was spun down. After discarding the supernatant, the pellet was fixed with 200 µl of 4% paraformaldehyde (PFA) and incubated at room temperature for 10 minutes to prevent degradation of proteins and keeping the morphology and components of the bacteria intact[98]. Then PFA was removed and the bacterial pellet was washed 2 times with 500 µl Phosphate buffer saline (PBS) by spinning down in each steps. The pellet was again re-suspended with 200 µl quenching buffer (Mixture of PBS and $NH_4Cl$). After discarding the final supernatant, 1ml PBS was added in to the bacteria pellet and sonicated for 15 minutes to weaken the cell wall. Finally, the bacterial suspension was passed through narrow bore needle (24 G) 4-5 times. Equal volume of bacterial suspension and Mowiol (Mowiol 4-88) was mixed and loaded on confocal dishes. Mowiol was used to immobilize the bacteria between glass slides and cover slip. In addition, Mowiol embedded samples solidified after several hours and stabilized fluorescent material for longer period when preserved at 4 $^{0}$C.

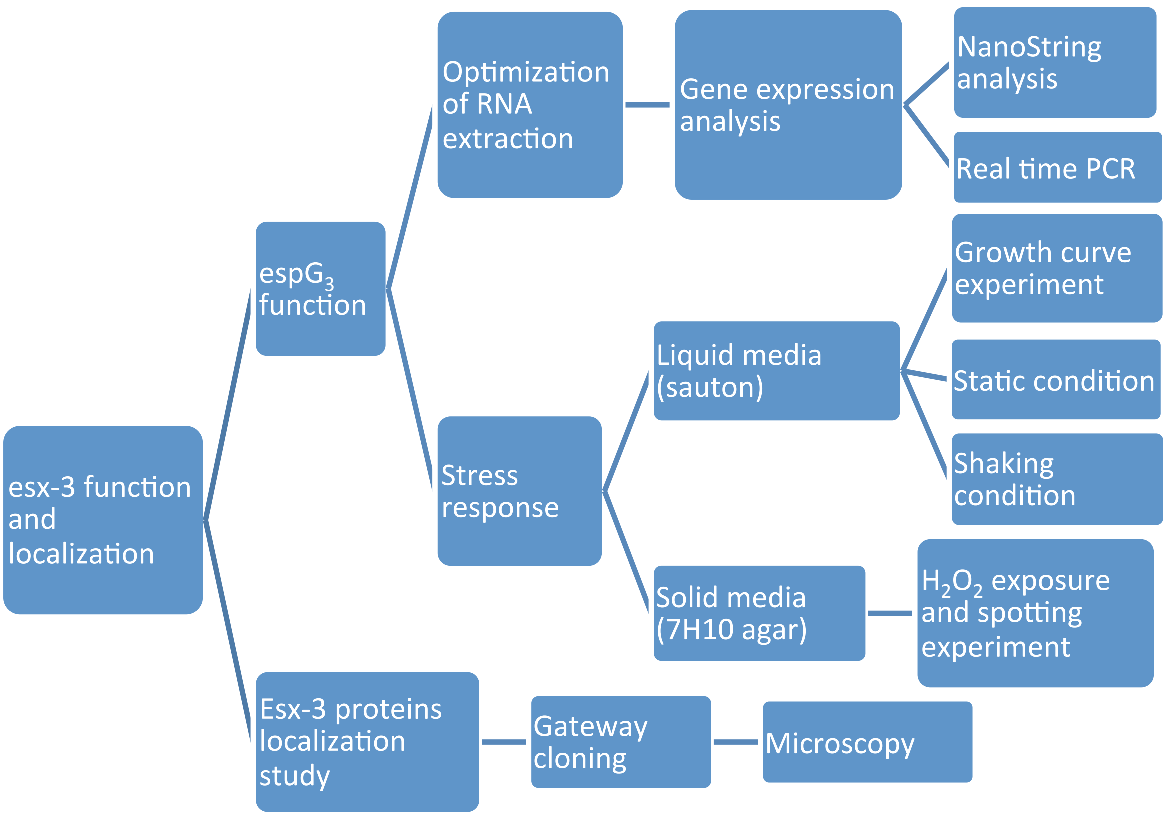
### 3.6.2 Hoechst DNA staining

Nucleic acid of the bacteria was stained with double strand intercalated Hoechst 33342 dye (Life Technologies). The dye is permeable through the complex mycobacteria cell membrane and stain DNA of the bacteria preferentially A-T reach region and minor groves of nucleic acids. Working concentration was prepared by diluting 5mg/ml stock solution with water 1 in 1000 dilution. Bacterial suspension prepared for microscopy was mixed with equal volume of Mowiol (Mowiol 4-88) and 0.1 µl working Hoechst dye was added per 100 µl samples. Mowiol embedded samples stabilized fluorescent stained material for longer period when preserved at 4 $^{0}$C.The mixture was kept 10 minutes on ice in light protected condition. 50 µl mixtures was dropped on microscopic slide and covered with cover glass. Excess volume was soaked by a piece of clean tissue paper and ready for Microscopy. 315 nm filter was used for excitation of blue fluorescent dye under confocal microscope.

### 3.6.3 Confocal Microscopy

Fluorescent protein tagged Esx-3 protein were assessed in *M. smegmatis* strains using a ZEISS PECON confocal microscopy with a GFP filter (490nm), Cherry (587nm). Bacteria were examined under confocal microscope and images were taken by high quality camera system built in the microscope. Captured images were analyzed by TIFF or ZEISS TIRF softwares. The brightness, resolution and contrast were normalized to avoid background noise and maximize the quality of images.

## 3.7 Overview of work flow



# 4 Results

## 4.1 Growth of esx-3 mutant *M. smegmatis* was suppressed under low iron condition

Previously, it was reported that Esx-3 is essential for survival of pathogenic mycobacteria whereas the saprophytic *M. smegmatis* need this system to survive under low iron condition[74,75]. Although Esx-3 was studied and required under low iron condition in *M. smegmatis*, the reason how it is function for growth under low iron is still not understood completely. Previous studies have used mutants of all components of Esx-3. We wanted to use a mutant of *M. smegmatis* mutated in only one gene, *esp*$G_3$, to make complementation possible.

In the first experiment, wild type (WT) and two mutant strains of *M. smegmatis* were used. The mutant used was a strain mutated in *esp*$G_3$, a gene that has been shown to be an essential component of Esx-3. The Δ*esp*$G_3$ mutant of *M. smegmatis* strains will from now on in this study be referred to as esx-3 mutant. Similarly we used *ΔmbtD*, *mbt*D is a gene in the *mbt* gene cluster that encodes proteins involved in mycobactin synthesis. This mutant will hereafter called mycobctin mutant.

WT, esx-3 mutant and mycobactin mutant *M. smegmatis* strains were cultivated in low and high iron (150µM $FeCl_3$) Sauton media. Under low iron (LI) condition, no iron was added with the assumption that trace metals from the water used and from glassware could fulfill the minimum iron requirement for growth. Growth curve studies (Figure 1) showed similar results as reported previously[29]. Growth of esx-3 mutant and mycobactin mutant *M. smegmatis* was suppressed in LI condition (Figure 1a) while in high iron (HI) condition suppression was less (Figure 1b).

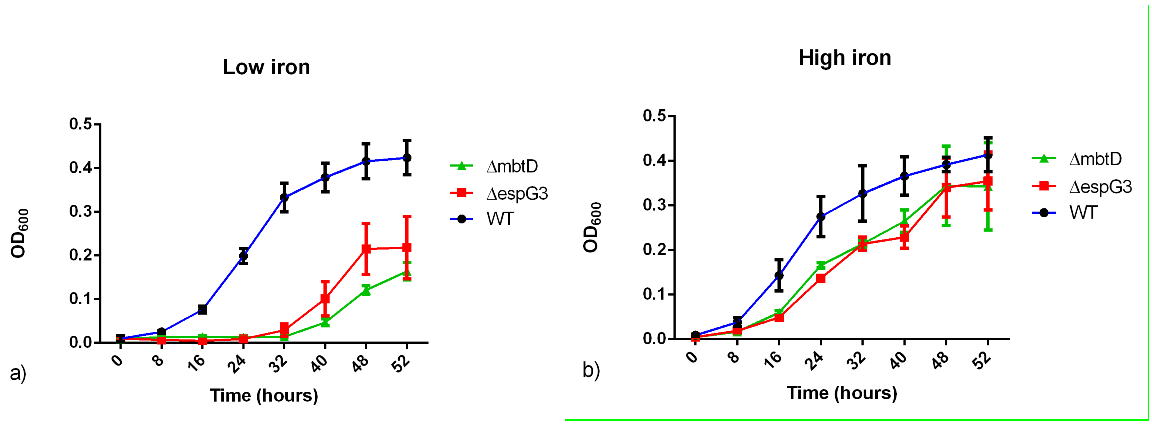



Figure 4.1. Growth curve of WT, esx-3 mutant and mycobactin mutant M. smegmatis grown under low and high iron conditions. (a) The growth of esx-3 and *mycobactin* mutant *M. smegmatis* strains was severely suppressed under LI growth condition. (b) In HI media, esx-3 mutant and mycobactin mutant *M. smegmatis* were less suppressed and didn't show big difference in growth curve pattern compared with WT.

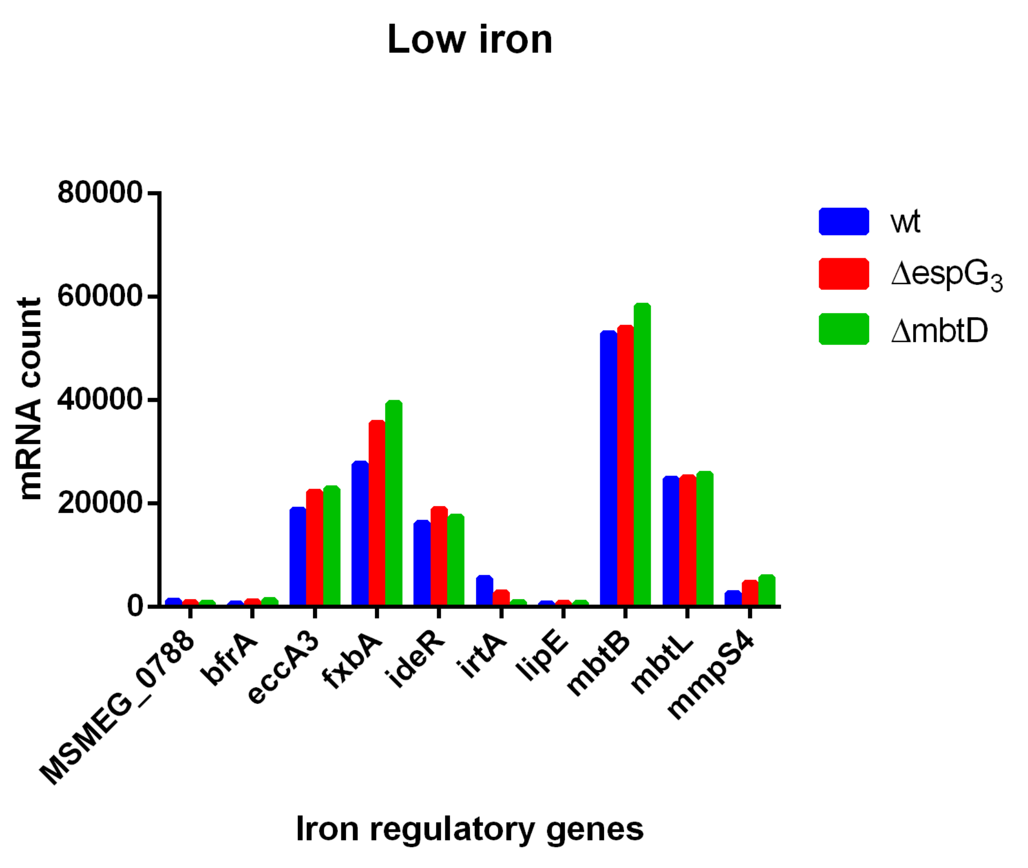
## 4.2 High quality RNA extraction method was optimized for gene expression analysis

To understand Mycobacteria pathogenesis, gene expression profiles are an important tool. Traditionally relative quantification using real-time PCR (RT-PCR) and microarrays have been used. More recently novel technology like NanoString technology has become available [99]. For reliable data, gene expression analysis like NanoString and RT-PCR experiment require high quality RNA with very little interference from DNA, proteins and other organic solvents. However, extraction of high quality RNA from mycobacteria is difficult unless a combination of mechanical, chemical and enzymatic methods used [99]. Isolation of RNA from mycobacteria can be challenging due to the special cell wall that prevent efficient lysis. To facilitate disruption methods like bead beating is common, a lysis method that also increase DNA contamination [100]. We therefore decided to develop a method for obtaining good quality yield of stable RNA for our expression analysis with NanoString technology. We also wanted the method that provided undetectable DNA levels and avoid organic solvents. Both these factors could affect later experiments in our gene expression analysis. To evaluate the RNA extraction method we grew *M. smegmatis* in 7H9 and Sautons medium until an OD of 0.8-1.2. At this specific OD we harvested the cells by centrifugation and resuspended the bacterium in Trizol which is a mixture of phenol and guanidium isothiocynate that solubilizes organic material and proteins while maintaining the integrity of RNA (Life technologies) and added the cell suspension to tubes containing beads for bead beating. We combined the mechanical bead beating methods to weaken and crack the waxy cell wall in the presence of TRIZOL and enzymatic DNA degradation step was followed in later steps to minimize DNA contamination as little as possible. At the beginning of optimization, large yellow precipitate pellet was frequently observed after isopropanol step. In spite of yellow pellet, extraction of RNA steps were followed until double DNAse digest and at the end we found bad quality RNA with absorbance ratio of both 260/280 and 260/230 below one which supposed to be around 2. We suspected that while care was taken only to pipette the aqueous phase containing the RNA, TRIZOL might still be a possible contaminant. To ensure less contamination only the upper part of supernatant was taken in the next experiment. Despite the change in procedure and several changes were made to eliminate the yellow pellet, the problem persisted and only low quality RNA was harvested. Finally, we found that changing the isopropanol from brand PRIMA to TEKNISK dramatically affected the result and we concluded that the large yellow precipitate was because of isopropanol from the brand PRIMA. Then after, we optimized a consistent high quality RNA extraction protocols for *M. smegmatis*. The quality of RNA was checked by NanoDrop spectrophotometer and real time PCR and in all extraction good quality RNA was isolated consistently. Good quality RNA was indicated when the ratio of the absorbance at 260/280 and 260/230 was nearly 2 or more (supplementary table 1) in addition to a sigmoidal curve shown in the spectrophotometric reading report, as indicated as a sign of good quality RNA by the NanoDrop manual ( ND-1000 Spectrophotometer V3.3 User's Manual ). Moreover, the quality of RNA extracted from *M. smegmatis* grown in LI and HI condition was also checked by RT-PCR before actual NanoString expression experiment. In the RT-PCR experiment we wanted to investigate the cDNA quality and the level of DNA contamination in the sample. *irt*A and *mbt*L targeted primers were used for amplification test for both RNA and corresponding cDNA from the same sample. None of RNA samples were amplified to the detectable levels, indicating that no DNA was contaminating the RNA samples, while the corresponding cDNA amplified the target at the expected levels (supplementary figure 1). The melting curve also showed, no any non-specific amplification peaks were observed which indicated nonspecific DNA was undetectable in the RNA sample (Data not shown).

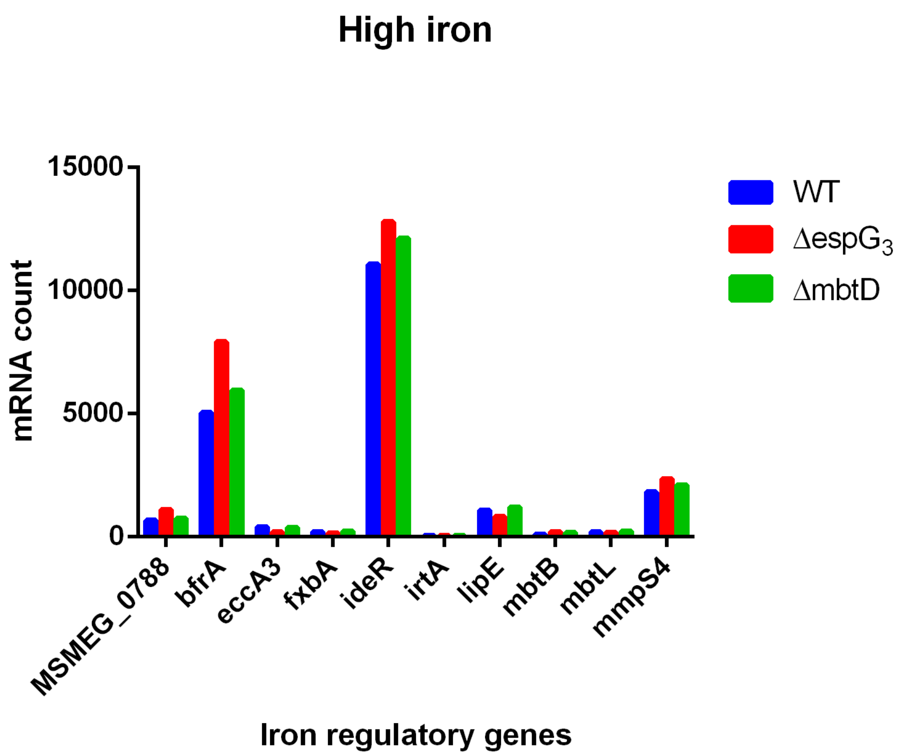
## 4.3 Gene expression analysis

### 4.3.1 NanoString analysis showed expression of *M. smegmatis* iron regulatory and iron dependent regulon was affected by esx-3 mutation

To study *esp*$G_3$'s role in Esx-3 secretion system, mRNA expression levels of iron regulatory genes were compared when WT, *esp*$G_3$ mutant and mycobactin mutant *M. smegmatis* were grown under LI and HI growth condition (Figure 2ab). Among tested functional categorized genes, the expressions of genes for protein involved in iron uptake were upregulated in all the strains we investigated in LI condition as expected (Figure 2a). Those upregulated genes were *ecc*$A_3$, *fxb*A, *irt*A, *mbt*B, *mbt*L and *ide*R. On the other hand, *bfr*A which involved in iron storage and activated by *ide*R [101] was upregulated in esx-3 mutant under HI relative to WT and mycobactin mutant. In similar condition i*deR* was also slightly over expressed in esx-3 mutant. The expression pattern in *bfr*A and *ide*R gene in the 3 *M. smegmatis* strains showed similar which indicated that *bfr*A is positively regulated by *ide*R (Figure 2b). Previously it was reported that esx-3 gene cluster is regulated by *ide*R [74]. This result showed that the iron uptake defect in esx-3 mutant *M. smegmatis* could affect the activity of iron dependent regulon *ide*R.




a)




b)

Figure 4.2. Normalized NanoString gene expression profile for iron regulatory genes expressed in WT, Δ*esp*$G_3$ esx-3 mutant and Δ*mbt*D mutant *M. smegmatis*. a) Grown in LI condition. b) Grown in HI condition.

Since the expression of iron regulatory genes in LI were very high compared to HI condition, it was difficult to directly compare the expression level by simple graph as shown in Figure 2ab. Thus, the ratio of mRNA expression in LI to HI was calculated by using nSolver software and effect of iron in gene expression was looked for each strain separately and compared. In esx-3 mutant, we observed that expression of growth in LI to HI ratio in *ecc*$A_3$, *fxb*A, *irt*A and MSMEG_0788 genes were different compared to wild type and mycobactin mutant *M. smegmatis* (Table 1). The ratio of *ecc*$A_3$ in *Δesp*$G_3$ mutant was much higher compared to WT and *mbtD* mutant. Similarly, expression of *irt*A in *esx-3* mutant was lower than WT but about 4 fold higher than mycobactin mutant *M. smegmatis*. On the contrary, MSMEG_0788 which encoded putative conserved membrane proteins was under expressed in the absence of *esp*$G_3$ compared to both WT and mycobactin mutant *M. smegmatis*.

Table 4.1 mRNA ratio of selected iron regulatory genes expressed in WT, Δ*esp*$G_3$ and Δ*mbt*D mutant *M. smegmatis* grown under LI versus HI.

**mRNA count ratio**

| genes | WT LI/HI | Δ*esp*$G_3$ LI/HI | Δ*mbt*D LI/HI |
|---|---|---|---|
| **MSMEG_0788** | 1.85 | **-1.22** | 1.14 |
| ***bfr*A** | -8.02 | -6.99 | -4.62 |
| ***ecc*$A_3$** | 46.42 | **112.86** | 60.72 |
| ***fxb*A** | 137.12 | 219.75 | 173.89 |
| ***ide*R** | 1.47 | 1.48 | 1.44 |
| ***irt*A** | 102.44 | **64.3** | 17.21 |
| ***lip*E** | -1.51 | -1.07 | -1.6 |
| ***mbt*B** | 502.1 | 259.4 | 356.47 |
| ***mbt*L** | 126.44 | 134.82 | 108.66 |
| ***mmp*$S_4$** | 1.5 | 1.99 | 2.73 |

In addition, to compare and see iron regulatory gene expression difference in mutant and WT *M. smegmatis*, ratio of mRNA for each iron regulatory genes in mutant against WT was analyzed for LI and HI condition separately as shown in Table 2. In esx-3 mutant *M. smegmatis,* expression of *irt*A was low in both LI and HI condition which supported that esx-3 involvement in iron acquisition could be through IrtA membrane protein. In the absence of mycobactin, *irtA* was much lower expressed in LI condition which also corroborates esx-3 is involved in iron acquisition through mycobactin pathway [75]. Similarly, MSMEG_0788 was down regulated in the two mutants under LI condition which could function similar to *irt*A. On the contrary, *bfr*A and *mmp*$S_4$ were upregulated in both esx-3 and mycobactin mutant in LI compared to WT in similar condition. In *M. tb*, Mmp$S_4$ protein is required in siderohpore synthesis and transport under LI condition [52] which supported that mutation of mycobactin and esx-3 in *M. smegmatis* could lower cytosolic iron that the mutants interpret as a further need to export siderophores and therefore increase expression of *mmp*$S_4$.

Table 4.2 mRNA ratio for selected iron regulatory genes expressed in esx-3 and mycobactin mutant *M. smegmatis* against WT grown under low LI and HI condition and analyzed by nSolver separately.

| | | **mRNA count ratio** | | | | | |
|---|---|---|---|---|---|---|---|
| **gene name** | **Accession number** | **In low iron condition against /WT** | | | **In high iron condition against/WT** | | |
| | | **WT** | **Δ*esp*$G_3$** | **Δ*mbt*D** | **WT** | **Δ*esp*$G_3$** | **Δ*mbt*D** |
| **MSMEG_0788** | MSMEG_0788 | 1 | -1.35 | -1.44 | 1 | 1.67 | 1.13 |
| ***bfr*A** | MSMEG_3564 | 1 | 1.8 | 2.05 | 1 | 1.57 | 1.18 |
| ***ecc*$A_3$** | MSMEG_0615 | 1 | 1.18 | 1.22 | 1 | -2.05 | -1.07 |
| ***fxb*A** | MSMEG_0014 | 1 | 1.29 | 1.42 | 1 | -1.25 | 1.12 |
| ***ide*R** | MSMEG_2750 | 1 | 1.16 | 1.07 | 1 | 1.16 | 1.1 |
| ***irt*A** | MSMEG_6554 | 1 | -2.01 | -5.75 | 1 | -1.26 | 1.04 |
| ***lip*E** | MSMEG_6575 | 1 | 1.07 | 1.05 | 1 | -1.31 | 1.12 |
| ***mbt*B** | MSMEG_4515 | 1 | 1.02 | 1.1 | 1 | 1.97 | 1.55 |
| ***mbt*L** | MSMEG_2132 | 1 | 1.01 | 1.04 | 1 | -1.06 | 1.21 |
| ***mmp*$S_4$** | MSMEG_0380 | 1 | 1.72 | 2.1 | 1 | 1.3 | 1.15 |

### 4.3.2 NanoString gene expression analysis showed, majority of *M. smegmatis* genes involved in different metabolic pathway were unaffected by esx-3 mutation at transcription level

The function of about 84 % mycobacterial gene encoded proteins which annotated in databases was identified from prediction and sequence similarity while, around 16% of the proteins which identified as unique to mycobacteria had unknown function yet [11]. To test whether a metabolic pathway is affected by a mutation in a given strain, it is possible to investigate the transcription of genes for proteins involved in the process. Esx-3 has been shown to be involved in iron uptake via the mycobactin pathway. Recent work in our lab however gives us reason to believe that Esx-3 also have some other function [75]. Thus, to understand more on function of Esx-3 we wanted to investigate whether any metabolic pathway in the bacterium was altered using NanoString technology. For doing the analysis we had selected 107 mycobacterial genes for proteins involved in many different metabolic processes (supplementary table 3). *M. smegmatis* strains were grown under 2 different growth conditions LI and HI. RNA was harvested at $OD_{600}$ 0.8-1.2 and followed NanoString expression analysis. In addition, after cDNA synthesis the samples were also analyzed by qPCR for few representative *M. smegmatis* genes that were differently expressed in esx-3 mutant strain. Screening result by NanoString gene expression profile showed that the majority of genes were unaffected in *esx-3* mutant compared to wild type (supplementary table 2). However, there were some genes that affected in esx-3 mutant compared to wild type *M. smegmatis*. If Esx-3 only has a role in mycobactin-mediated iron uptake it would be expected that the mycobactin mutants expression pattern would be identical to the esx-3 mutant expression pattern. NanoString profile showed few genes were 2 fold or more differently expressed in esx-3 mutant compared to WT. In esx-3 mutant *M. smegmatis*, the expression of *fur*A was 2 fold upregulated in LI but more than 4 fold down regulated in HI condition. *atp*A was down regulated in esx-3 mutant *M. smegmatis* independent of media iron level whereas, under LI condition *irt*A and *tgs*1 were down regulated in both esx-3 and mycobactin mutant. *rip*A was also upregulated in esx-3 mutant and slightly more upregulated in siderophore mutant. This shows esx-3 could be directly involved in *atp*A gene transcription. esx-3 mutant could have also defect in iron transport through mycobactin pathway which indirectly affected the expression of *irt*A, tgs1 and *rip*A (Table 4).

Under high iron growth condition, expression of *dos*R*, kst*D and *xer*C were upregulated while *ahp*C*,atp*A*,atp*B*,ecc*$A_3$*,fur*A*,mur*A and one of the 3 *M. smegmatis kat*G gene (MSMEG_6384) were under expressed in esx-3 mutant (Table 4).

Table 4.3. mRNA ratio of esx-3 mutant and mycobactin mutant versus WT *M. smegmatis* for those differently expressed genes in mutant strains grown under low and high iron condition (analyzed by nSolver v1.1 software). Yellow and red indicate over and under expression respectively compared to WT.

| **Ref. gene *sig*A Normalization with nSolver v1.1** | **Low iron (LI)** | | **Ref. gene *sig*A Normalization with nSolver v1.1** | **High iron (HI)** | |
|---|---|---|---|---|---|
| | **mRNA ratio of mutant vs WT** | | | **mRNA ratio of mutant vs WT** | |
| **gene name/accession no.** | **Δ*esp*$G_3$** | **Δ*mbt*D** | **gene name/accession no.** | **Δ*esp*$G_3$** | **Δ*mbt*D** |
| MSMEG_0219 | -1.51 | -2.57 | *ahpC* (MSMEG_4891) | -2.88 | -1.02 |
| *atpA* (MSMEG_4938) | -2.04 | -1.02 | *atpA* (MSMEG_4938) | -2.06 | -1.09 |
| *bfrA (MSMEG_3564)* | 1.8 | 2.05 | *atpB* (MSMEG_4942) | -2.05 | -1.1 |
| *ctaE* (MSMEG_4260) | 2.62 | 2.27 | *dosR* (MSMEG_3944) | 2.28 | 1.88 |
| *fas* (MSMEG_4757) | -1.43 | -2.07 | *eccA3* (MSMEG_0615) | -2.05 | -1.07 |
| *furA* (MSMEG_6383) | 2.03 | 1.38 | *furA* (MSMEG_6383) | -4.57 | 1.06 |
| *groS* (MSMEG_1582) | 1.1 | 1.19 | *katG* (MSMEG_6384) | -6.35 | -1.18 |
| *irtA* (MSMEG_6554) | -2.01 | -5.75 | *kstD* (MSMEG_5941) | 2.17 | 1.52 |
| *katG* (MSMEG_6384) | 1.65 | 1.35 | *murA* (MSMEG_4932) | -2.19 | -1.16 |
| *mmpS4* (MSMEG_0380) | 1.72 | 2.1 | *xerC* (MSMEG_2515) | 2.47 | 1.37 |
| *mshA* (MSMEG_0933) | -1.84 | -2.04 | | | |
| *recB* (MSMEG_1327) | -1.5 | -2.4 | | | |
| *ripA* (MSMEG_3145) | 2.31 | 3.67 | | | |
| *ssb* (MSMEG_4701) | -1.23 | -2.18 | | | |
| *tgs1* (MSMEG_5242) | -2.59 | -2.25 | | | |

### 4.3.3 Esx-3 mutation affected expression of *M. smegmatis* redox regulatory genes

In addition to iron regulatory genes, the expression of redox regulatory genes was also affected in esx-3 mutant *M. smegmatis*. Almost all screened redox regulatory genes expression were less in esx-3 mutant compared with WT and mycobactin mutant *M. smegmatis* grown in HI condition (Figure 3b). This would indicate that the esx-3 mutant has an additional defect affecting the red-ox system of the bacterium compared to the mycobactin mutant. Under LI conditions the gene expression of most red-ox genes seemed unaffected by any of the mutations. However, *msh*A which is one of mycothiol encoding essential mycobacteria gene involved in antioxidant activity [102] was upregulated in only WT strain. The reason could be HI in WT expose the bacteria for oxidative stress. Whereas *ahp*C and *whi*B7 were slightly upregulated in esx-3 mutant compared with WT and mycobactin mutant strains in LI media (Figure 3a). LI affect the repression activity of iron dependent repressor genes thus slightly higher expression in *ahp*C and *whi*B7 could be less repression by iron dependent regulon due to low iron.

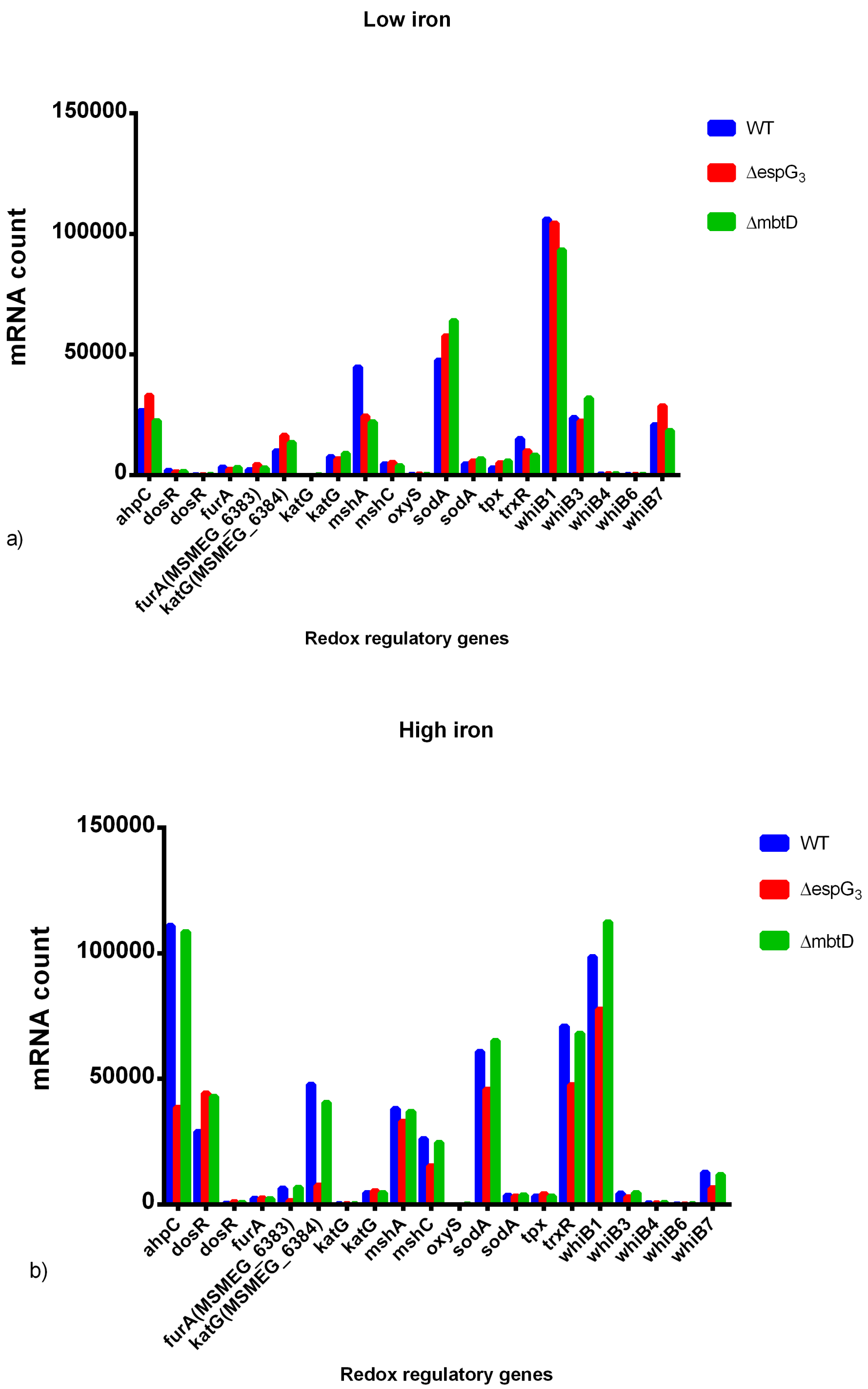



Figure 4.3. NanoString gene expression profile for redox regulatory gene expression in esx-3 mutant M. smegmatis. *sig*A and WT was taken as housekeeping genes and reference strain respectively a). Majority of redox regulatory gene expression in LI were unaffected in esx-3 mutant compared to wild type *M. smegmatis* but *msh*A gene was under expressed in mutant strains b) under HI condition, many

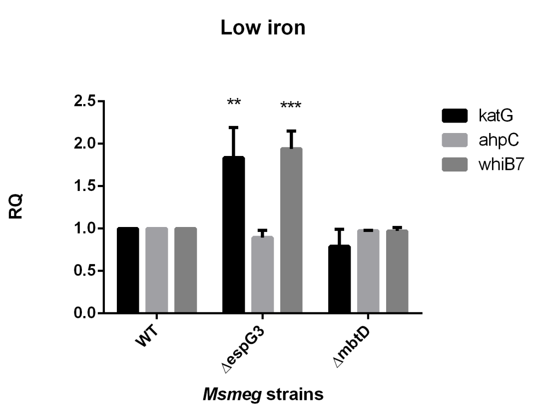
of redox regulatory genes were down regulated in esx-3 mutant compared to WT and mycobactin mutant *M. smegmatis*.

### 4.3.4 Quantitative PCR expression analysis confirmed *ahp*C, *kat*G and *whi*B7 was differently expressed in esx-3 mutant *M. smegmatis*

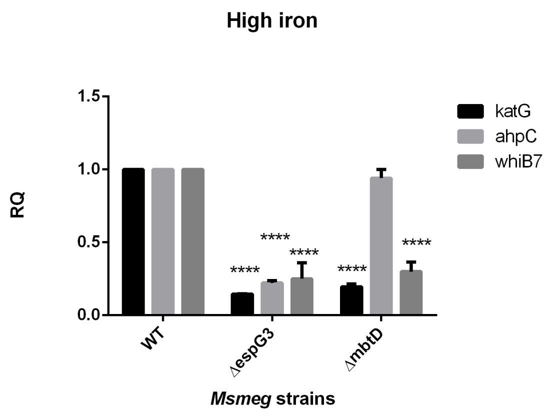
NanoString expression experiment was used only as screening tool to see the expression difference among *M. smegmatis* strains (Supplementary table 3). In order to confirm and check the reproducibility of expression results from NanoString, comparative gene expression experiment was done by qPCR for some selected genes. We decided to choose *ahp*C and *kat*G genes that were highly affected in esx-3 mutants and one of putative transcription regulatory genes *whi*B7 which was also moderately affected in both LI and HI condition (Figure 3ab).

qPCR comparative expression analysis also confirmed that expression of *kat*G, *ahp*C and *whiB*7 were significantly down regulated in esx-3 mutant *M. smegmatis* grown in HI condition compared to WT grown in similar condition ( Figure 4b). On the other hand, only expression of *kat*G and *whi*B7 genes were upregulated in esx-3 mutant grown in LI condition compared with WT grown in similar way (Figure 4a).

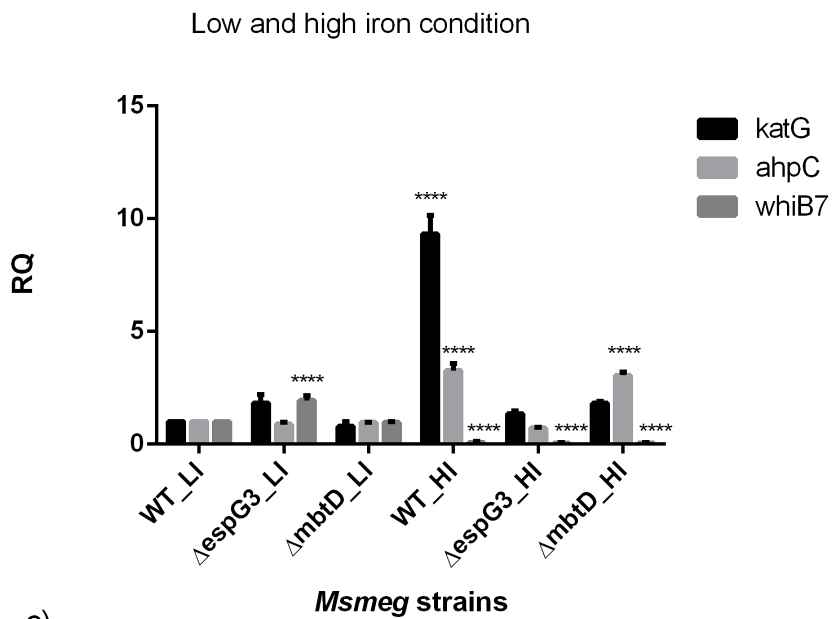
To check whether the expression difference we observed was a down regulation or an inability to upregulate, we analyzed all the samples against WT in LI condition as a reference. We found that expression of *kat*G was significantly high in WT_HI condition compared with the reference. Similarly, *ahp*C expression was high in both WT and mycobactin mutant in HI compared with WT_LI condition. *whiB*7 expression was raised only in esx-3 mutant in LI condition similar to NanoString result (Figure 3a) whereas low expression in all strains in HI condition compared to WT_LI as shown in Figure 4c. Based on these findings it seemed that the esx-3 mutant is unable to turn on expression of *ahp*C under HI conditions, while the mycobactin mutant has retained this ability.




a)




b)



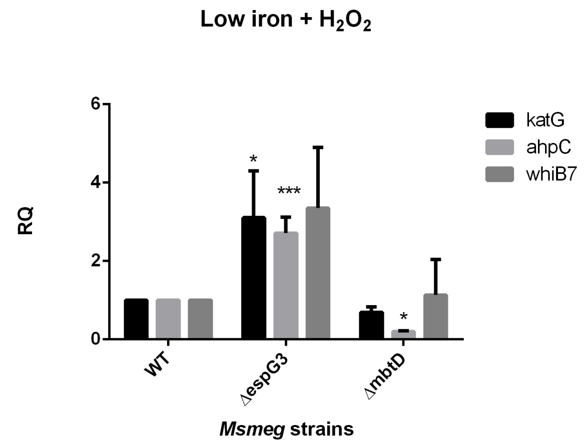

c)

Figure 4.4. qPCR expression analysis and relative quantification (RQ) results among selected genes from *M. smegmatis* strains grown in different growth condition. Figure (a) and (b) indicated relative expression of *kat*G*, ahp*C and *whi*B7 genes between *M. smegmatis* mutant strains compared with WT as a reference strain that were grown in similar condition. a) LI growth condition. b) HI growth condition. c) Expression levels of *M. smegmatis strains* in both conditions (LI and HI) relative to WT in LI (WT_LI) as a reference strain. RQ mean was calculated from three technical triplicates and normalized by expression value of a housekeeping gene *sig*A. Average RQ values with SEM error bars were plotted. Statistical significance over the bars shows there is a significant difference in expression of wild type genes compared with mutant strains of *M. smegmatis.* Gene expression differences were determined by one way ANOVA statistical tests for mean of RQ values with a p value cutoff of 0.05. Difference in expression of each gene in between WT versus mutant strains was compared and showed statistically significant which is indicated by asterisk over the bars. The significance differences were observed as * = P < 0.05, **= P < 0.05, *** = P < 0.0001, **** = P < 0.0001.

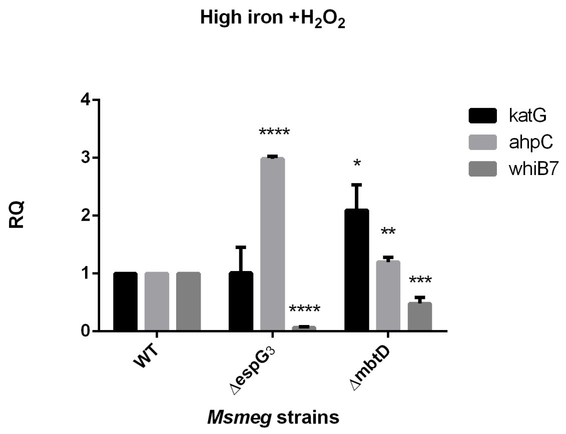
In a study by Dussurget *et al.,* the level of *sod*A and *kat*G encoded proteins was reduced in *ide*R mutant *M. smegmatis* and susceptible to the anti-tuberculosis drug isoniazid (INH) which is a potent oxidative intermediate inducer [47] and we hypothesis that esx-3 mutant strains are much more susceptible to stress inducer hydrogen peroxide ($H_2O_2$) than wild type. From NanoString expression profile (Supplementary Table 2) and (Figure 3b) almost all screened redox regularly genes were down regulated in esx-3 mutant compared with WT in HI condition. In this experiment WT, esx-3 mutant and mycobactin mutant *M. smegmatis* were challenged with 15 mM $H_2O_2$ for 1 hour prior to RNA extraction. qPCR experiment and analysis were done as described earlier for expression comparison in low and high iron condition except here bacteria were challenged by $H_2O_2$ before RNA extraction. qPCR analysis showed expression of *ahp*C in esx-3 mutant was significantly higher when exposed to $H_2O_2$ both in high and low iron condition whereas low expression was observed in mycobactin mutant *M. smegmatis* grown in LI condition. In general; *ahp*C, *kat*G and *whi*B7 seem upregulated only in esx-3 mutant under LI condition (Figure 5a and c). Under HI condition, *ahp*C was upregulated in both esx-3 mutant and Δ*mbt*D mutant *M. smegmatis* whereas *kat*G was slightly upregulated only in Δ*mbt*D mutant (Figure 5b).

To check whether the expression difference we observed was a down regulation or an inability to upregulate, we analyzed all the samples with WT_ LI challenged with $H_2O_2$ as a reference (Figure 5c). We found that all the 3 genes were up regulated in esx-3 mutant grown in LI with $H_2O_2$ despite *whi*B7 expression was not statistically significant compared with the reference WT_LI with $H_2O_2$. The expression of *ahp*C after $H_2O_2$ challenge was upregulated in both low and high iron condition only in *esx-3* mutant strain. On the other hand, *kat*G was upregulated only in LI condition in esx-3 mutant *M .smegmatis* (Figure 5c).

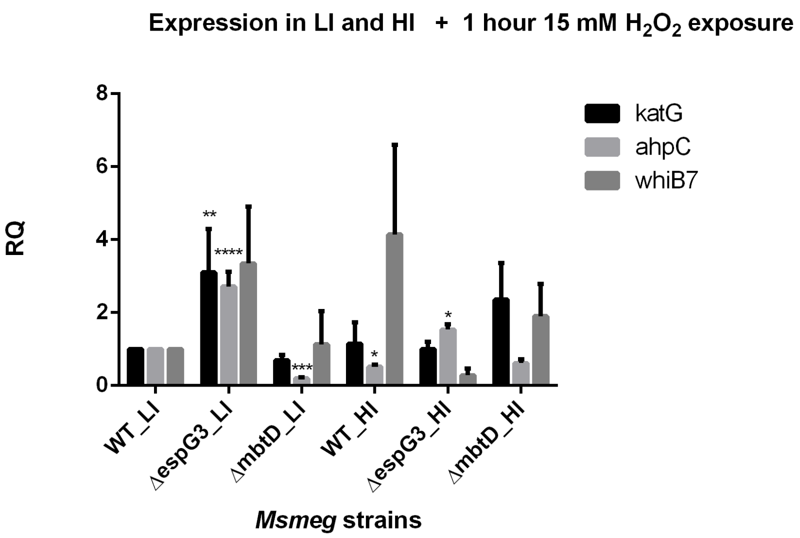
Based on these findings it seemed that the esx-3 mutant is able to turn on expression of *ahp*C under both high and low iron conditions*,* while the WT and mycobactin mutant has unable to retain this ability. Thus, we for the first time explore that *M. smegmatis esp*$G_3$ encoded protein involve in *ahp*C gene or *esp*$G_3$ could affect the transcription of *ahp*C in response to stress. But further study is required to investigate how $EspG_3$ act on AhpC function.




a)




b)




c)

Figure 4.5. Graphs depicted to show qPCR analysis of relative quantification for ahpC, katG and whiB7 genes expression. The relative expression of WT as a reference was compared with ΔespG3 esx-3 mutant or ΔmbtD mycobactin mutant M. smegmatis. a) Gene expression in LI growth condition + 1 hour 15mM H2O2 exposure prior to RNA extraction. b) Gene expression in HI growth condition + 1 hour 15mM H2O2 exposure prior to RNA extraction. c) Expression levels of ahpC, katG and whiB7 in M. smegmatis strains grown under LI and HI + 1 hour 15mM H2O2 exposure prior to RNA extraction relative to WT in LI + 1 hour 15mM H2O2 exposure prior to RNA extraction. RQ mean was calculated from three technical triplicates and normalized by expression level of a housekeeping gene sigA. Average RQ values with SEM error bars were plotted. Difference in expression of each gene in between WT_LI versus mutant strains was compared and showed statistically significant which is indicated by asterisk over the bars. The significance differences were observed as * = $P < 0.05$, **= $P < 0.05$, *** = $P < 0.0001$, **** = $P < 0.0001$.

## 4.4 Esx-3 mutant *M. smegmatis* had better tolerance against $H_2O_2$ stress than WT in both LI and HI condition

In order to investigate if stress tolerance and viability of esx-3 mutant strains was correlated with redox regulatory gene expression; WT, esx-3 mutant and mycobactin mutant *M. smegmatis* strains were cultivated in Sauton with low and high iron media containing 15mM, 10mM and 5mM hydrogen peroxide ($H_2O_2$) incubated at 37 $^0$C both in static and shaking condition. In all conditions, none of the strains tolerated the exposure of 5mM – 15 mM $H_2O_2$ concentration in Sauton media and no change in optical density of media was observed (data not shown). In addition, we also cultivated WT, esx-3 mutant and mycobactin mutant *M. smegmatis* in the presence of 15, 10, 5, 3, 1, 0.5, 0.1 and 0.01 mM $H_2O_2$ Sauton media with high and low iron under constantly shaking growth curve machine for 3 – 5 days. Among the tested concentration of $H_2O_2$, the maximum tolerable limit in liquid Sauton media for *M. smegmatis* was 0.1 mM $H_2O_2$. We observed that esx-3 mutant *M. smegmatis* had better tolerated when exposed to $H_2O_2$ in both high and low iron condition. This result showed the upregulation of *ahp*C gene in esx-3 mutant *M. smegmatis* could protect from oxidative stress compared to WT (Figure 6ab).

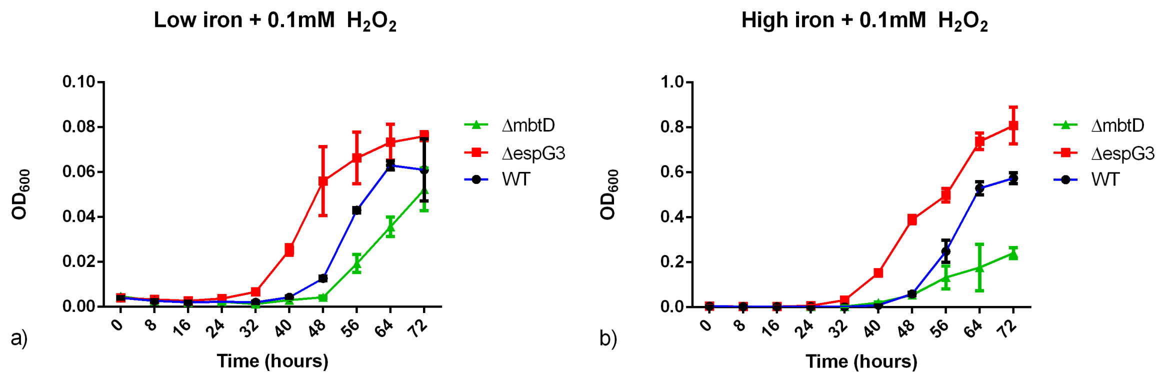



Figure 4.6. Growth curve for WT, ΔespG3 mutant and ΔmbtD mycobactin mutant M. smegmatis treated with H2O2 under low and high iron condition. a) Growth curve in $H_2O_2$ and low iron condition Sauton media. b) Growth curve in $H_2O_2$ and high iron Sauton media.

Similarly to investigate esx-3 mutant stress tolerance in solid media and to identify the involvement of *esp*$G_3$'s in stress response; first we exposed WT, esx-3 mutant and mycobactin mutant *M. smegmatis* (OD 0.01) to 15mM, 10mM and 5mM $H_2O_2$ for 1 hour at 37 $^0$C in high and low iron Sauton media. Then, after 1 hour the dilution (from $10^{-2}$ – $10^{-6}$) was spotted on nutritionally enriched 7H10 agar media and incubated for 3 - 5 days. After spotting on 7H10 media, anti-oxidative substances from the media such as catalase was able to reduce the effect of $H_2O_2$ .Thus, phenotypic difference observed in WT and mutant strains of *M. smegmatis* reflect, the effect of $H_2O_2$ and iron from previous media. Here, we observed that esx-3 mutant strain was better tolerated to $H_2O_2$ in both low and high iron condition compared to both WT and mycobactin mutant *M. smegmatis* (Figure 7).

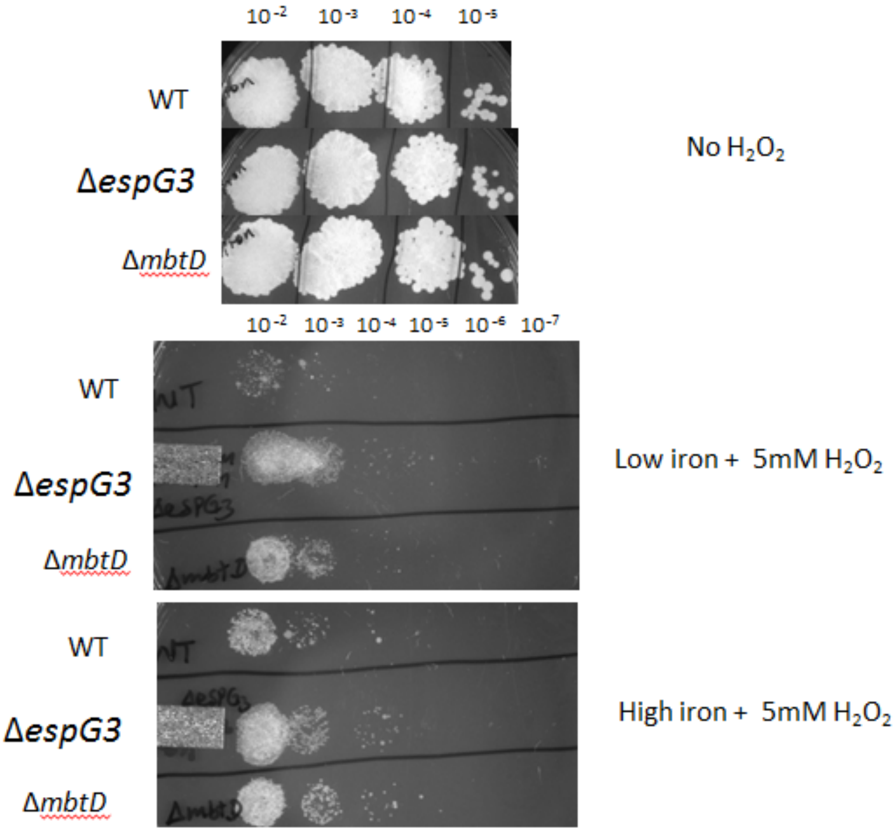
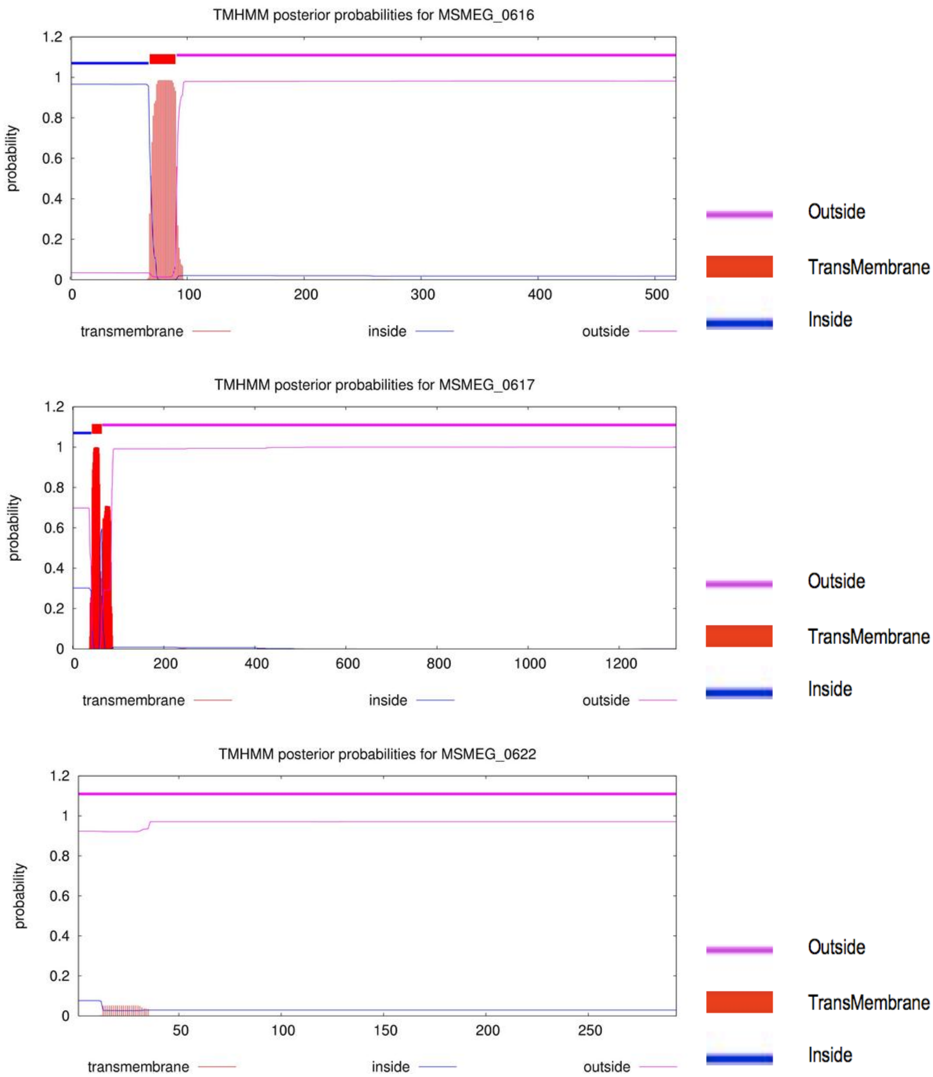



Figure 4.7. Growth of M.smegmatis strains after treatment with H2O2. Spotting of WT, esx-3 mutant and mycobactin mutant M. smegmatis in 7H10 agar media prior to spotting on 7H10, M. smegmatis strains were treated with H2O2 for 1 hour at 370C under low and high iron condition.

To see the effect of esx-3 mutation in *M. smegmatis* growth at lower pH stress; WT, esx-3 mutant and mycobactin mutant strains were cultivated in pH=6 Sauton media at $37^0$C both in low and high iron media under static and shaking condition. All tested strains were grown in pH=6 Sauton both under static and shaking condition in low and high iron similar to growth in neutral pH Sauton media. However, it took longer to observe visible turbidity in static low iron growth condition (data not shown). Future experiments should use lower pH for investigating the pH-stress effect in esx-3 mutant *M. smegmatis* both in shaking and static condition.

## 4.5 *M. smegmatis ecc*$B_3$ and *esp*$G_3$ genes were cloned using 3 fragment PCR gateway cloning system

An important part of defining appropriate protein function is protein localization. Studying localization of esx-3 proteins has not previously been performed and by identifying localization we anticipated to get clues to the actual function of the system. Esx-3 proteins were predicted as components of proteins encoded by 11 genes that form core membrane components and non-membrane soluble proteins[66]. In our study, we wanted to include at least one membrane bound proteins and one non-membrane (cytosolic) proteins. Based on HMM prediction information, $EccB_3$ and $EccC_3$ which were predicted as core membrane proteins and the non-membrane protein $EspG_3$ was selected (Figure 8)




a)

b)

c)

Figure 4.8. Representative M. smegmatis Esx-3 Trans Membrane Hidden Markov Model (TMHMM) predicted proteins. a) EccC3, b) EccB3 and c) EspG3 proteins. The eccB3(MSMEG_0616) and eccC3(MSMEG_0617) genes encoded proteins showed only few amino acids reside in transmembrane while the majority lay either cytosolic or extracytosolic sites. c) espG3

(MSMEG_0622) gene encodes either secreted or cytosolic proteins and none of amino acids shown trans membrane path 103-105.

In order to visualize the localization of esx-3 encoded proteins with microscope, fluorescent proteins encoding genes were fused to esx-3 genes by using MultiSite Gateway cloning system. Using this system we could clone a strong constitutive mycobacterial promoter (p750), the esx-3 genes (*ecc*$B_3$, *ecc*$C_3$ and *esp*$G_3$) and reporter genes (gfp and cherry). For expression in mycobacteria, we also selected an integrating plasmid vector that would increase the stability of the system. Briefly, at first primers for individual genes were designed based on the MultiSite Gateway cloning system for 3 fragment PCR product. MultiSite Gateway cloning for 3 fragments recombination system (promoter, esx-3 gene and reporter gene) was designed using clone manager 9 software and Invitrogen user manual guideline was followed. attB flanked primers were designed for promoter target and reporter gene. PCR was amplified from DNA of promoter (p750), esx-3(*esp*$G_3$, *ecc*$B_3$ and *ecc*$C_3$) and reporter (gfp and cherry). Initially, we designed the primers only for Esx-3 protein coding region without including ribosome binding site (RBS) which in mycobacteria is expected to be up to -10 nucleotides upstream of start codon[106]. We tested the designed primers using the respective DNA sample and all esx-3 genes including reporter were amplified (Figure 9).

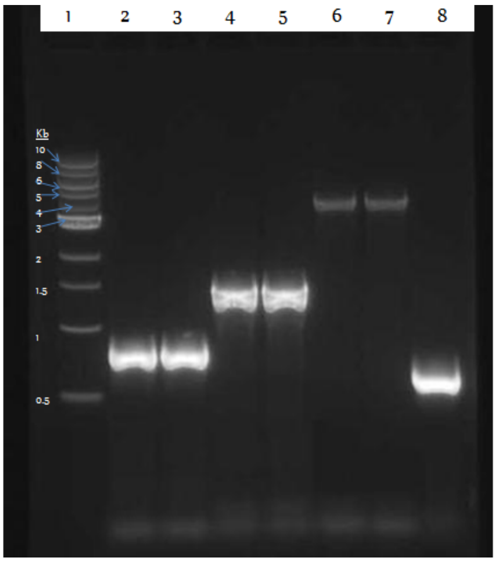



Figure 4.9. PCR product amplified from MultiSite Gateway for 3 fragments cloning system design. Lane 2 and 3= espG3 (882 bp), lane 4 and 5 = eccB3 (1,557 bp), lane 6 and 7 = eccC3 (3957bp) and lane 8 = gfp(714 bp)

After testing whether the methods were working or not, we decided to include RBS and new primers were designed in similar ways. This way, all the designed primers amplified their target DNA except *ecc*$C_3$. Then, all the amplified PCR products were purified and cloned

(Figure 9). The cloned entry plasmids were transformed in to chemically competent DH5-α *E. coli* and then isolated from successfully transformed bacteria. The quality and quantity was checked by NanoDrop and PCR gel electrophoresis which showed correct size DNA bands (Figure 10). In addition, to ensure the exact gene was inserted in to the entry clones that might be mutated during PCR, the entry clones were sequenced and found to be correct after BP reaction using the primers provided with the kit.

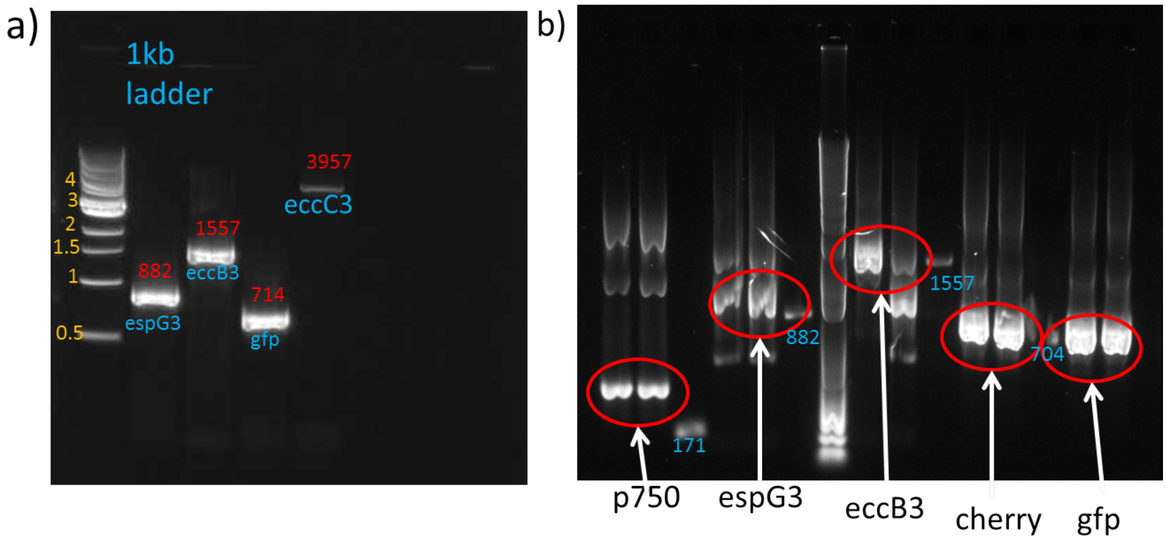



Figure 4.10. PCR products (a) before cloning (b) after BP reaction. The controls 171, 882, 1557 and 704 bp were PCR products directly amplified from M. smegmatis DNA. The difference in PCR product size that observed between the controls and sample was due to additional sequence from attB flanked cloning construct included in the PCR product after cloning in BP reaction.

All the 3 verified plasmids were then mixed with destination plasmids pDE43 flanked with attR1 and attR2 reaction sites and LR clonase II plus enzyme. The corresponding attL sites of DONR vector recombine with attR sites of pDE43 plasmid to fuse the p750-esx3- reporter gene (gfp or cherry) together without any stop codon in between. The promoter promotes expression of both fused gene as one gene. Cloned kanamycin resistant pDE43 expression plasmids were transformed in to competent *M. smegmatis* to express the target proteins. Transformation was done on chemically competent *M. smegmatis* strains where *esp*$G_3$ and *ecc*$B_3$ was tagged with reporter gene gfp and cherry respectively. In this study, in order to test the predicted locations of proteins, we successfully cloned *esp*$G_3$ fused with gfp and *ecc*$B_3$ with cherry and constitutively expressed in competent *M. smegmatis* by using p750 promoter as mentioned in gateway cloning section (Figure 11).

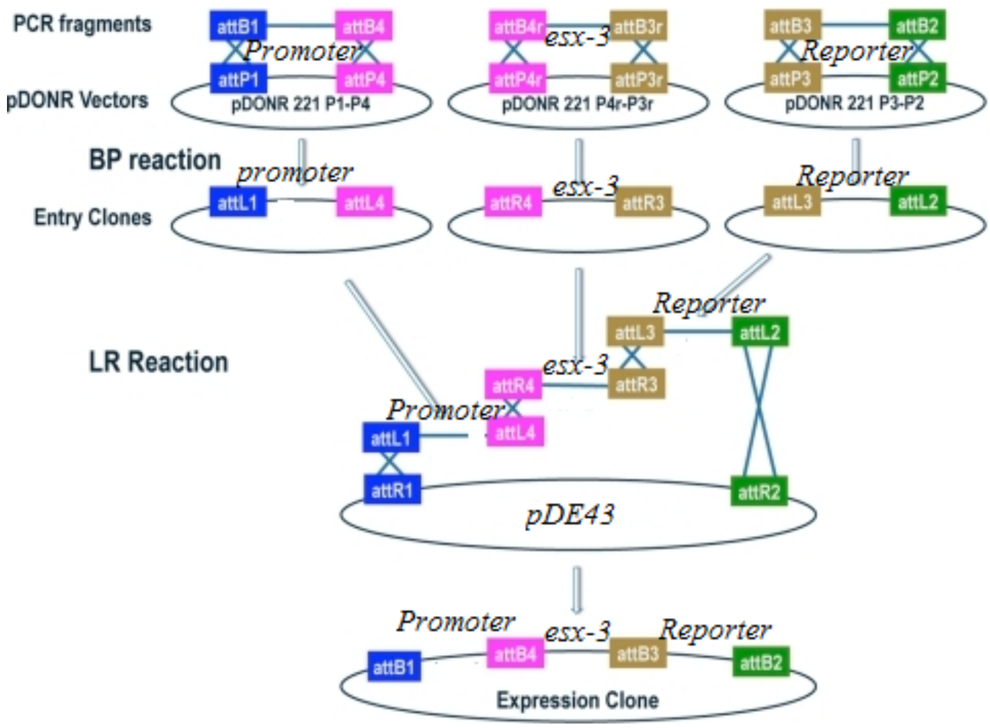
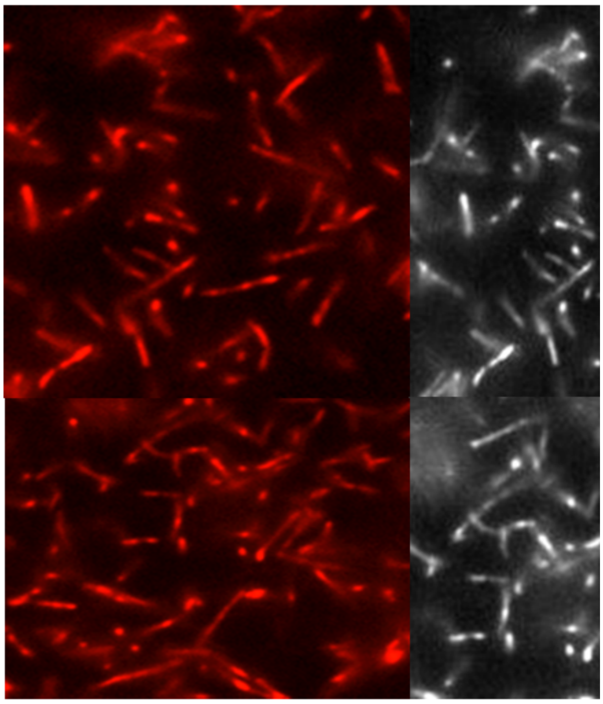


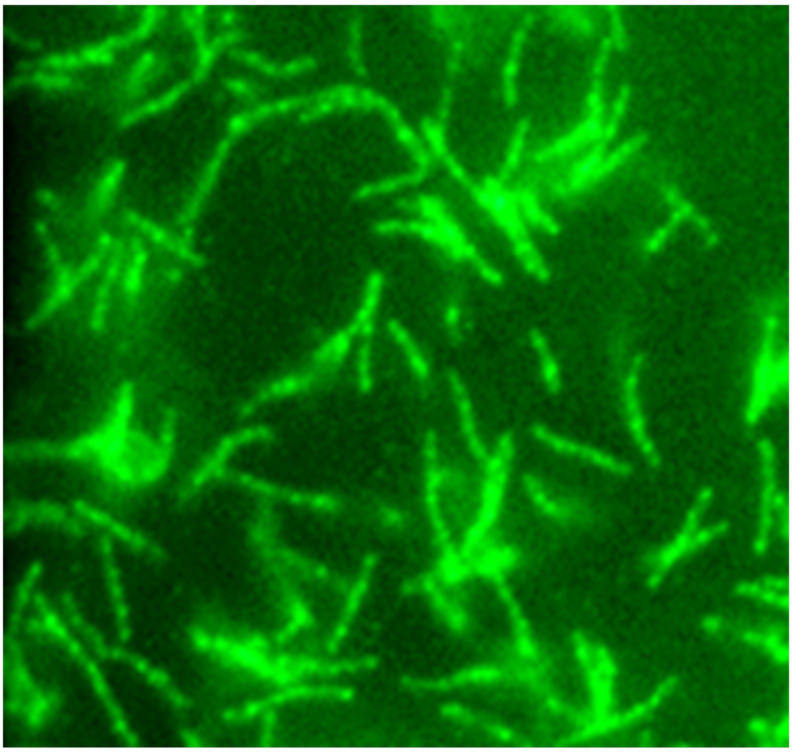

Figure 4.11. Schematic presentation of gateway cloning construct for recombination with 3 fragments PCR product. attB flanked PCR was amplified using primers designed for 3 fragment gateway cloning system. At the beginning attB1 and attB4 sites were flanked in p750 promoter, in the middle attB4r and attB3r sites flanked in one of esx-3 protein coding gene and at the end attB3 and attB2 sites were flanked in reporter gene (either cherry or gfp). The fragment attB1- p750- attB4 was recombined with pDONR 221 P1-P4, attB4r-esx-3-attB3r recombined with pDONR 221 P4r-P3r, and attB3-cherry/gfp-attB2 recombined with pDONR 221 P2-P2 in separate BP reaction tube. The BP recombination reaction of donor plasmids and PCR products were pENTR attL1-p750-attL4, pENTR attR4-esx-3-attR3, and pENTR attL3-cherry/gfp-attL2. In LR reaction the 3 BP reaction products, entry clones were recombined with the destination vector, pDE-43-MCK(Addgene) to produce an expression vector containing promoter (p750) gene of interest(esx-3) and reporter gene( cherry or gfp) in a defined sequential position. The stop codon was removed between esx-3 and reporter gene during the gate way primers design and the promoter promote the expression of esx-3 and reporter gene as operon. (Adapted from Invitrogen MultiSite Gateway pro user manual).

### 4.5.1 *M. smegmatis* EspG$_3$ and EccB$_3$ proteins localization under Confocal Microscopy

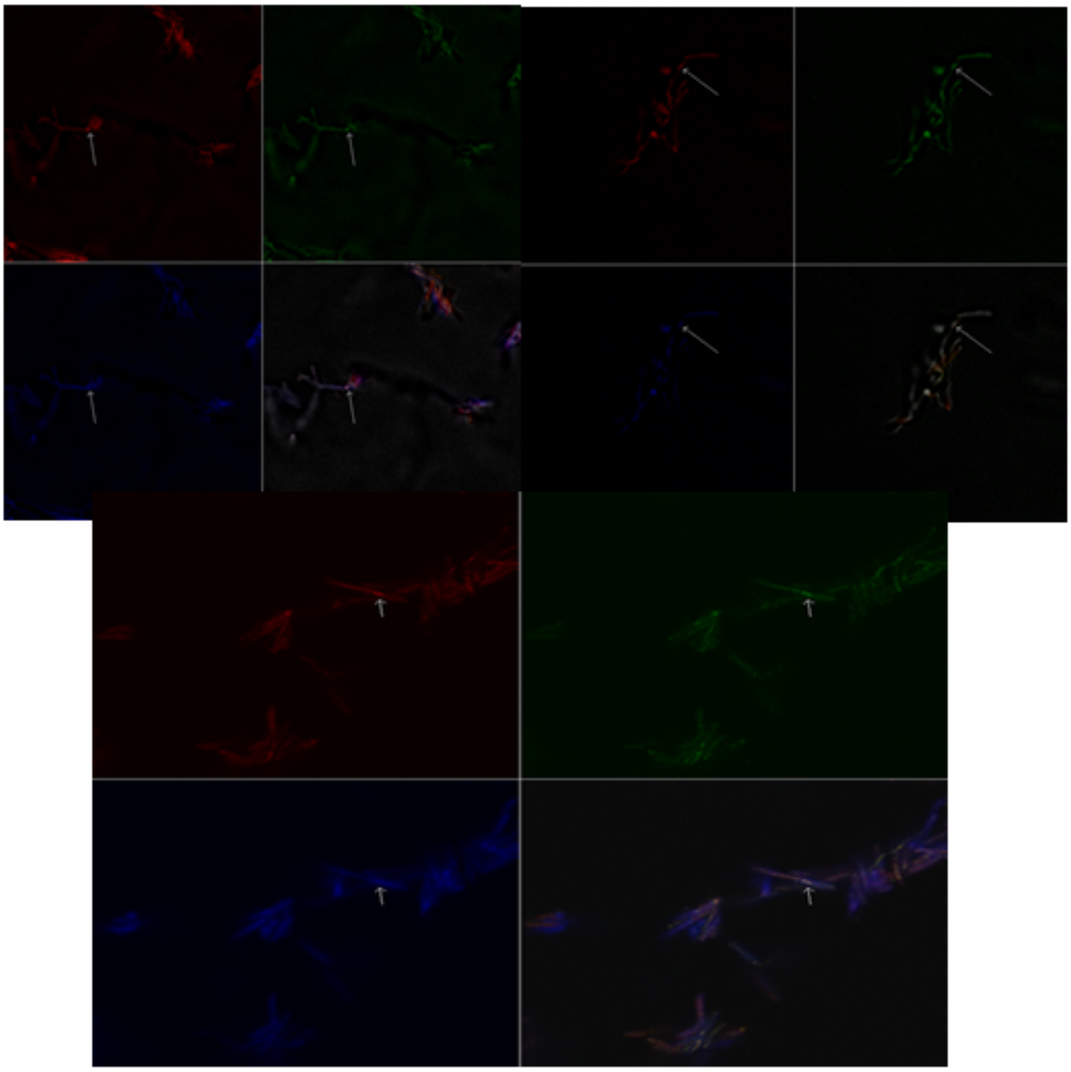
In order to see the cloned Esx-3 proteins localization, confocal microscopy was used as described in method 5.3 section. We found that *ecc*B$_3$ was strongly expressed and observed by confocal microscopy as bright large beads on average 2 – 3 per bacteria with predominantly polar localization (Figure 12a). EspG$_3$ on the other hand, was seen as small green fluorescent dots scattered throughout the bacterium, with more dots per cell than in EccB$_3$ (Figure 12b). We for the first time probably identified that two esx encoded proteins seem not co-localized as shown in figure11ab. However, when the two cloned genes expressed with in the same bacterium EccB$_3$ and EspG$_3$ proteins seem to co-localize as shown in Figure 12c. This could be an artifact because *esp*G$_3$ and *ecc*B$_3$ were cloned in a vector having similar antibiotic resistance gene. In order to see the real co-localization, the two genes must be cloned in vectors with different antibiotic resistance gene.



a)



b)



c)

Figure 4.12. Localization and colocalization of cloned Esx-3 component proteins in M. smegmatis. a) EccB3 proteins cloned together with Cherry was localized as bright large red beads in the bacteria predominantly localized towards pole. b) EspG3 proteins cloned with GFP localized as small green dots scattered along the rode shape bacteria and the number of beads was more and fine compared to EccB3. c) Arrows showing a probable co-localization of EccB3 and EspG3 proteins in WT M. smegmatis and blue staining showed Nucleic acid staining by Hoechst dye.

$EspG_3$ staining shown in figure 11b resembles DNA-staining previously shown in our group and by others [100]. $EspG_3$ is also predicted to be a DNA binding proteins [107]. To test this prediction an *M. smegmatis* WT strain containing $EspG_3$ tagged with GFP was stained with Hoechst dye which binds double stranded DNA. However, it was difficult to see staining of nucleic acids as the bacteria was uniformly stained with the Hoechst dye. This was because of scattered DNA material throughout bacterial cytosol (Figure 12c). Therefore, in this study we were unable to determine whether $EspG_3$ was a DNA binding proteins or not.

# 5 Discussion

## 5.1 Nanostring gene expression technology can be used as a tool for screening mycobacteria gene expression

RNA based gene expression studies are mainly relying on microarray, sequencing and more recently by Nanostring nCounter assay. Using NanoSting gene expression screening is advantageous over the other methods that Nanostring nCounter directly quantitate the level of specific RNA, measures up to 800 gene expression with in a single reaction tube, has high specificity and sensitivity and no amplification required unlike PCR. However, there are some limitations that, NanoString probes are expensive and RNA count is affected by impurity of RNA [95,108]. In addition, RNA is unstable and degraded during extraction process and storage which affected the real expression level by NanoString RNA quantification. Thus, for confirmation of NanoString expression results, other less expensive but high sensitive and specific expression analysis methods are required. Despite limitation, quantitative PCR (qPCR) comparative gene expression analysis is one of those expression analyses which were done for confirmation of few NanoString screened samples. qPCR gene expression analysis could also affected by several factors, such as non-specific amplification which might give falsely elevated expression due to DNA contamination. Mycobacteria have high GC DNA that also affects qPCR primer efficiency. High GC DNA requires high annealing temperature which in turn is good for specificity and excludes most non-specific amplification but sensitivity is affected. Due to these factors, expression of specific gene might seem falsely under expressed. Thus, for confirmation by qPCR high quality RNA should be extracted for cDNA synthesis, correct annealing temperature should be chosen, efficient primer should be designed and several repeat might be important to avoid false upregulation and down regulation result.

## 5.2 Nanostring gene expression analysis showed the expression of few *M. smegmatis* genes involved in different metabolic pathway were affected by esx-3 mutation

From NanoString expression profile we observed that most of the genes involved indifferent metabolic pathway was unaffected. However, some genes had different expression profiles between the WT and mutations. Among which *atp*A and *atp*B genes were down regulated in esx-3 mutant *M.smegmatis* both under low and high iron condition suggesting that iron starvation could be the reason due to iron transport defect (Table 3). Iron is required in respiratory enzyme function and in a study both genes suggested to be involved in respiratory ATP synthesis and down regulated in nutrient starvation[109]. The other possibility could be that esx-3 might be involved in regulation of *atp*A and *atp*B gene as both genes down regulated in esx-3 mutant regardless of iron. Further study could require to show that how esx-3 affect the expression of *atp*A and *atp*B. Under HI, expression of iron regulatory, redox regulatory and *mur*A, genes were down regulated while *xer*C, *Dos*R and *kst*D were upregulated in *esx-3* mutant *M. smegmatis* compared to WT .Which shows the difference in cytosolic iron between esx-3 mutant and WT *M. smegmatis* could make the expression difference (Table 3). Briefly, this could be because despite the presence of high iron in the media, iron couldn't transported efficiently which in turn affected the level of cytosolic iron. On the other hand, under LI condition, *cta*E, *fur*A, *bfr*A and *rip*A were upregulated which also directly or indirectly affected by the level of iron (Table 3). Iron could directly be involve as cofactor for enzymatic metabolic process or indirectly affects the function of iron dependent gene regulator or regulon such as *ide*R and *fur*A.

## 5.3 Expression of iron regulatory and iron dependent repressor/activator genes were affected by esx-3 mutation

It is known fact that esx-3 is important in mycobacterial iron acquisition system. To see whether there is any transcriptional difference in iron regulatory genes due to esx-3 mutation, we analyzed expression of iron regulatory genes in esx-3 intact and mutant *M. smegmatis* strains. In this study under HI condition, *ecc*$A_3$ was down regulated in esx-3 mutant *M. smegmatis* compared to WT (Table 1 and Table 2) which indicates that despite surplus iron present in media, esx-3 mutant possess lower cytosolic iron while WT *M. smegmatis* acquire iron from all possible pathway which makes the expression of *ecc*$A_3$ higher in WT. This could because the activity of iron dependent repressor genes such as *ide*R is affected by the level of iron which could negatively regulates *ecc*$A_3$ under high cytosolic iron in WT *M. smegmatis*. Previously, it was reported that esx-3 gene cluster is regulated by ideR [74].

Iron regulated transporter gene, *irt*A was under expressed in esx-3 mutant *M. smegmatis* both in high and low iron condition (Table 2). This could be because, esx-3 is important in iron acquisition through mycobactin pathway[29,74] and possibly through other unknown system. WT *M. smegmatis* strain acquires iron both through mycobactin and esx-3 independent exochelin. This shows the IrtA protein channel might be busy and overexpressed in WT. In another studies it was reported that *irt*A is an ABC transporter involved in utilization of iron through siderohore pathway [110] and important in homeostasis of iron bound siderophores [51]. However, the expression difference between esx-3 mutant and mycobactin mutant showed (Table 1), esx-3 could not be the only system involved in iron acquisition through mycobactin pathway but also other esx-3 independent iron uptake via mycobactin could be possible which indicted that in esx-3 mutant mycobactin was intact and could still transported through irtA channel independent of esx-3.

In this study, both mutant and WT *M. smegmatis ide*R was upregulated in LI compared to HI condition (Figure 2a and Table 1) and this could be because, many of iron regulatory genes were upregulated under LI and *ide*R negatively regulate the upregulation of iron regulatory genes. On the contrary, *bfr*A was down regulated under LI which previously reported as it was positively regulated by *ide*R [101]. In microarray expression study, about 153 mycobacteria genes was shown to be iron dependent and 1/3 of them were regulated by iron dependent repressor (ideR) [79] among which transcription of genes encoding for esx-3 secretion system clusters and other iron regulatory genes were controlled by ideR [111]. Our result indicated that the expression of genes involved in siderophore synthesis was down regulated under HI and upregulated in LI. This evidence shows that under LI condition, the bacteria use all the possible iron acquisition mechanism and one of the pathways of iron acquisition is through esx-3 dependent mycobactin pathway. In a study under LI condition, acquisition of iron is mainly through affinity transport mechanism which depends on iron bound siderophores. Under HI media, affinity independent iron transport mechanism like Msp porins also contribute in iron uptake in addition to affinity dependent siderophore pathway in *M. smegmatis* [112]. This data supports that the regulatory function of *Ide*R is dependent on the level of iron which in turn the activation of *ide*R is affected by defect in iron acquisition system. We conclude that esx-3 mutation affects the level of cytosolic iron which directly affects iron dependent and iron regulatory gene expression in *M. smegmatis*. On the other hand, the level of iron could also affect esx-3 gene expression repressor and activator gene.

## 5.4 Genes involved in oxidative stress response were down regulated in esx-3 mutant compared to WT and mycobactin mutant *M. smegmatis* under high iron condition.

Among the screened genes that have been shown to be involved in oxidative stress response, most of them were differently expressed in esx-3 mutant compared to WT and mycobactin mutant (Figure 3b). Here, a mycobactin mutant showed, expression trend of redox regulatory genes that was similar to that of the wild type. These data supports that Esx-3 has a direct functional role in oxidative stress response. Since it was difficult to conclude based only on the nanostring data, we decided to confirm some of representative genes by real time PCR. We chose *kat*G and *ahp*C that are known mycobacteria enzyme encoding genes involved in oxidative stress response [78]. The expression of *kat*G could be dependent on the level of cytosolic iron which was significantly high in only WT. This shows *kat*G could positively be regulated in surplus iron by *ide*R as mutation of esx-3 and mycobactin affects the level of cytosolic iron while unaffected in WT. Previous report showed that activation of *ide*R increases activity of *kat*G and was also suggested redox regulatory enzymes such as KatG and SodA response was reduced in *ide*R mutant strains[113]. Very high expression levels of *katG* in high iron reference strain (WT) masked the real expression level of mutant *M.smegmatis* strains and looked all genes under expressed in esx-3 mutant grown under HI condition. From combination of both high and low iron analysis; it was only the expression of *katG* in WT grown under high iron was upregulated and statistically significant (Figure 4c). Similarly, *ahp*C expression in esx-3 mutant was the lowest under high iron compared with both WT and mycobactin mutant in similar condition. This data supported that *kat*G expression could positively regulated by iron dependent *ide*R or the expression of *kat*G was directly proportional with bacterial iron uptake. esx-3 mutant and mycobactin mutant *M. smegmatis* lack optimal utilization of iron acquisition. The expression of *ahp*C was only switched on in WT and mycobactin mutant *M. smegmatis* under HI condition. This could indicate that, esx-3 could involve in regulation of *ahp*C gene transcription. On the other hand, esx-3 in turn could be regulated by iron dependent repressor genes like *fur*A and *ide*R. Therefore, *M. smegmatis* $EspG_3$ protein potentially involve in oxidative stress response activity of AhpC enzyme. Lack of the notable bacterial antioxidant OxyR in *M. smegmatis* could be substituted by AhpC to respond against oxidative stressors.

## 5.5 Upregulation of *ahp*C expression in esx-3 mutant *M. smegmatis* could be cumulative effect of iron, iron dependent repressor genes and $H_2O_2$ stress and responsible for stress tolerance.

Despite expression of *ahp*C was not able to switch on in esx-3 mutant under both LI and HI condition, $H_2O_2$ able to induce the expression. However, unlike esx-3 mutant, both WT and mycobactin mutant were unable to switch on the expression of *ahp*C. This shows esx-3 could directly affect the expression of *ahp*C in the presence of iron dependent repressor or activator genes. Iron indirectly affects the activity of esx-3 by iron dependent repressor like *fur*A but *fur*A might be inactivated by $H_2O_2$. In *M. smegmatis* expression of *ahp*C was raised in response to hydrogen peroxide exposure [82] and hydrogen peroxide exposure increases the level of cAMP which in turn increases expression of *ahp*C. *fur*A could also be inactivated by oxidative stress [78]. Both *M. tb* and *M. smegmatis* lack functional stress response regulator OxyR and in this species FurA could replace function of OxyR in stress response [78]. The presence of *kat*G downstream of *fur*A supported the involvement in oxidative response regulator. In mycobacteria the main oxidative stress response, *kat*G and *ahp*C are negatively regulated by *fur*A [78,114] and *fur*A conserved in most mycobacterial species [114,115]. Our result showed, the involvement of *fur*A in oxidative response could be by regulating esx-3 genes or both *ahp*C and esx-3 could be controlled by *fur*A. However *fur*A could be inactivated by $H_2O_2$ and unable to repress esx-3 in WT *M. smegematis. ide*R could also indirectly responsible in the regulation of oxidative stress response by controlling siderophore synthesis and other iron regulatory genes which in turn control the level of iron [79]. Microarray expression study by Rodriguez shows, *esp*$G_3$ negatively regulated by iron and *ide*R [79].

EspG3 protein also predicted as a chaperon for core T7SS membrane protein and secreted substrate. Lack of this protein could affect *M. smegmatis* membrane integrity which in turn increases diffusion of $H_2O_2$ in to the bacteria and induce expression of *ahp*C by directly increasing the level of cAMP or indirectly by inactivating its negative regulator *fur*A which has protective effect in esx-3 mutant *M. smegmatis*. EspG3 also predicted as putative DNA binding protein which could involve in regulation of redox regulatory genes like *ahp*C.

High expression of ahpC gene in esx-3 mutant treated with $H_2O_2$ had better tolerance against hydrogen peroxide stress which clearly supported that in *M. smegmatis*, *ahp*C could be the main oxidative stress response which affected by esx-3. The involvement of esx-3 function in oxidative stress response could be by interacting with AhpC. From all comparative qPCR analysis, liquid and solid media experiment shown in Figure 5, 6 and 7 respectively supported that esx-3 mutant has better tolerance against $H_2O_2$ in both low and high iron indicating esx-3 could have a role in oxidative stress response by regulating *ahp*C gene. The function of oxidative stress regulatory genes was affected by esx-3 probably by iron dependent repressor genes such as *fur*A and i*de*R and $H_2O_2$ induced gene expression. In general, esx-3 could be regulated by both *fur*A and *ide*R and involved in *ahp*C oxidative protection. The defect in *fur*A and *ide*R affect esx-3 regulon which in turn oxidative stress regulatory response through *ahp*C is affected. However, further study could be required to investigate how esx-3 interacts with *ahp*C in oxidative stress response.

## 5.6 Localization of $EspG_3$ and $EccB_3$ proteins in *M. smegmatis*

In order to visualized the localization of esx-3 proteins under microscope, fluorescent protein encoding gene was tagged with esx-3 genes together with appropriate promoter using gateway cloning technology (Figure 10). The gateway cloning system was working perfectly for esx-3 reporter tagging (Figure 8) and compatible for 3 fragment PCR gateway cloning system despite unable to amplify *ecc*$C_3$ DNA. We observed that 80% of -10 region of *ecc*$C_3$ DNA sequence was identified as G/C. This high GC region might affected annealing temperature and hamper primer efficiency that we couldn't amplify ecc$C_3$ gene. In order to alleviate the problem, we did gradient annealing temperature PCR (data not shown) from 50 – 65 $^0$C .But, we were not able to amplify *ecc*$C_3$ DNA. Next time, new primers should be designed by including sequence having recommended GC content usually 40 – 60 % (Life technologies) for efficient PCR amplification.

We established getaway cloning for 3 PCR fragment products to fuse promoter esx-3 and reporter proteins sequentially in a defined position. Despite successful cloning was made for esx-3 fused with reporter genes and expressed in competent *M. smegmatis*, the exact localization was unable to be determined as the size of the bacteria was small compared to eukaryotic cells and was difficult to discriminate between membrane and cytosolic proteins by Confocal microscopy. Initially, we were also unable to get a clear image as the bacteria were moving in suspended saline. Later we embedded the bacteria in Mowiol which immobilized the bacteria and preserve morphology for longer periods. The resolution of the camera could also affect to see compartments of the bacteria. From our confocal imaging, it was difficult to determine the exact localization and whether cytosolic or membrane proteins (Figure 11ab). However, we wanted to try to transform both *esp*$G_3$-gfp and *ecc*$B_3$-cherry construct in to a single competent *M. smegmatis*. This procedure is complicated because of the problems with having two plasmids of the same type in one cell. The only way of identifying successful transformants is to take kanamycin positive clones and screen them for expression of both GFP and Cherry. In the strains identified as double positive, the expression of *esp*$G_3$-gfp and *ecc*$B_3$-cherry were weak compared with single transformation. Thus, co-localization of EspG3 and $EccB_3$ proteins in competent *M. smegmatis* could be affected by low expression or indeed be false positives. Confocal microscope imaging for co-localization was observed weak fluorescent protein to see the definite co-localization of $EspG_3$ and $EccB_3$ protein (Figure 11c). Although it was difficult to determine, $EccB_3$ and

$EspG_3$ proteins seem co-localized proteins (as shown in figure 11c) without telling cytosolic or membrane compartment proteins. $EccC_3$ and $EccB_3$ were predicted as membrane proteins [66] but unfortunately, we couldn't clone $EccC_3$ to see the localization.

$EspG_3$ predicted as a DNA binding proteins [116] and to determine whether this protein is DNA binding or not, nucleic acid of *M. smegmatis* was stained with Hoechst dye (Figure 11c). Since the nucleic acid of *M. smegmatis* was scattered throughout the cytosol [100], majority portion of the bacteria was stained with the Hoechst dye and thus we were unable to determine whether $EspG_3$ is DNA binding protein or not.

### 5.6.1 **Steric effect and promoter strength could affect the localization of Esx-3 proteins.**

Expression of target ESX-3 tagged to florescent proteins reporter could affect the native localization of $EspG_3$ and $EccB_3$ proteins. This could be due to steric effect that abundance of large proteins could hinder the translocation of proteins to its native location. Besides, the effect of promoter p750 strength could also affect abundance of proteins in the cell which could interfere to see real localization and co-localization of proteins. In a study, the overexpression of fluorescent proteins tagged proteins interfere natural localization of target proteins [117]. Structural folding of $EspG_3$ and $EccB_3$ proteins cloned with florescent reporter proteins could also different compared to the expression of native target ESX-3 protein alone. This shows tagging with fluorescent protein contribute for mislocalization of target proteins. In previous study reported that tagging protein with fluorescent protein could affect folding and localization of translated proteins due to steric effect[117]. Despite we suggested that $EspG_3$ seem to localize to the center of the bacterium and $EccB_3$ proteins to the poles of the bacterium. The strength of promoters, steric effects due to tagged proteins could affect natural localization and co-localization of proteins and further studies are needed to confirm localization.

# 6 Conclusion and recommendation

Our NanoString gene expression screen on 107 selected *M. smegmatis* genes was promising in further understanding novel role of Esx-3 secretion system and in other metabolic pathway. Optimization of high quality mycobacteria RNA extraction method will enhance the credibility of gene expression results and expand to other important mycobacteria gene expression. From our expression experiment, esx-3 system has a role in oxidative response.

Esx-3 could be negatively regulated by both *fur*A and *ide*R. Esx-3 could also involve in *ahp*C oxidative stress response, possibly through gene regulation. Defects in one of the repressor *fur*A or *ide*R will indirectly affects the function of Esx-3 and increase *M. smegmatis* oxidative stressor susceptibility. We also established that gateway cloning system is an efficient and simple method that is compatible with esx-3 genes if appropriate promoter, strong reporter and compatible entry and destination plasmids is used. Thus, not only for esx-3, but also other essential gene could be cloned and the localization of proteins could be investigated.

In the future, expression profile of pathogenic mycobacteria oxidative regulatory and other related genes such as sigma factors, two component system should be screened by NanoString gene expression analysis under different condition. Since Esx-3 is essential and not possible to knockout from pathogenic mycobacteria, cloning *M. tb* genes in esx-3 mutant *M. smegmatis* strains are indispensable to study the overall function of Esx-3 system in virulence, pathogenicity and metabolic function. Despite it is not possible to directly knockout esx-3 in *M. tb*, knocking down could be possible to study Esx-3 function in directly in virulent *M. tb* strain.

Further work should also be done to investigate how mycobacteria *ahp*C gene interacts with esx-3 system in oxidative stress response and role of *fur*A in esx-3 function. Non pathogenic mycobacteria stress response could be different from pathogenic mycobacteria. Thus, *M. tb* strain *ahp*C gene should be complemented in *ahp*C mutant *M. smegmatis* to confirm and see whether there is difference in expression pattern on pathogenic *ahp*C against hydrogen peroxide exposure. The growth of esx-3 mutant strains also should be evaluated against other stress inducers like cumene hydroperoxide (CHP), t-butyl hydroperoxide (t-BHP), SDS, Low pH and others.

*M. tb* esx-3 gene should also be cloned and expressed in competent pathogenic strains to identify the actual localization and co-localization of proteins. Since protein localization in pathogenic and nonpathogenic mycobacteria might be different. For co-localization study, antibiotic resistant gene should be different for each entry vector. So, in future cloning of esx-3 gene, different antibiotic resistance gene constructed in different entry vector should be used to express 2 or more proteins in the same destination vector for co-localization study. As the size of the bacteria is small compared with eukaryotic cells, visualization of localized and co-localized protein could be difficult and thus high resolution camera system integrated microscopy should be used. It is difficult to determine the exact localization and co-localization of esx-3 protein in mycobacteria by relying only on confocal microscopy. It should be supported by electron microscopy, membrane dissection, fluorescent dye staining and other technologies. $EspG_3$ Protein which was predicted as a DNA binding protein should be tested by Electrophoretic Mobility Shift Assay (EMSA) or by DNase I footprinting. Not only esx-3 but also other T7SS putatively predicted function by bioinformatics tool should be tested in wet lab by cloning using gateway cloning strategy. Promising finding on saprophytic *M. smegmatis* should be confirmed by experiments on pathogenic mycobacteria before results can be extrapolated. NanoString gene expression could be used as a mycobacterial gene expression screening tool which can be broaden to study other essential mycobacteria gene function by evaluating expression pattern at different condition. The same is true for localization and co-localization study that essential or virulent mycobacteria genes cloned with reporter gene supports the possible function of unknown mycobacterial proteins if supported with high resolution imaging.

# 7 References


1. Tullius, M.V., *et al.* Discovery and characterization of a unique mycobacterial heme acquisition system. *Proceedings of the National Academy of Sciences of the United States of America* **108**, 5051-5056 (2011).
2. Yimer, S.A., *et al.* Primary drug resistance to anti-tuberculosis drugs in major towns of Amhara region, Ethiopia. *APMIS : acta pathologica, microbiologica, et immunologica Scandinavica* **120**, 503-509 (2012).
3. Economou, A., *et al.* Secretion by numbers: Protein traffic in prokaryotes. *Molecular microbiology* **62**, 308-319 (2006).
4. WHO. Global tuberculosis report 2013. http://www.who.int/tb/publications/global_report/en/.
5. Guo, H., Naser, S.A., Ghobrial, G. & Phanstiel, O.t. Synthesis and biological evaluation of new citrate-based siderophores as potential probes for the mechanism of iron uptake in mycobacteria. *Journal of medicinal chemistry* **45**, 2056-2063 (2002).
6. Havlir, D.V. & Barnes, P.F. Tuberculosis in patients with human immunodeficiency virus infection. *The New England journal of medicine* **340**, 367-373 (1999).
7. Zumla, A., *et al.* The WHO 2014 global tuberculosis report--further to go. *The Lancet. Global health* **3**, e10-12 (2015).
8. Green, K.D. & Garneau-Tsodikova, S. Resistance in tuberculosis: what do we know and where can we go? *Frontiers in microbiology* **4**, 208 (2013).
9. Mukhopadhyay, S., Nair, S. & Ghosh, S. Pathogenesis in tuberculosis: transcriptomic approaches to unraveling virulence mechanisms and finding new drug targets. *FEMS microbiology reviews* **36**, 463-485 (2012).
10. Kim, J.H., *et al.* A genetic strategy to identify targets for the development of drugs that prevent bacterial persistence. *Proceedings of the National Academy of Sciences of the United States of America* **110**, 19095-19100 (2013).
11. Cole, S.T., *et al.* Deciphering the biology of Mycobacterium tuberculosis from the complete genome sequence. *Nature* **393**, 537-544 (1998).
12. Orme, I.M., Robinson, R.T. & Cooper, A.M. The balance between protective and pathogenic immune responses in the TB-infected lung. *Nature immunology* **16**, 57-63 (2015).
13. Andersen, P. & Doherty, T.M. The success and failure of BCG - implications for a novel tuberculosis vaccine. *Nature reviews. Microbiology* **3**, 656-662 (2005).
14. Behr, M.A. & Small, P.M. Has BCG attenuated to impotence? *Nature* **389**, 133-134 (1997).
15. Pym, A.S., *et al.* Recombinant BCG exporting ESAT-6 confers enhanced protection against tuberculosis. *Nature medicine* **9**, 533-539 (2003).
16. Grode, L., Kursar, M., Fensterle, J., Kaufmann, S.H. & Hess, J. Cell-mediated immunity induced by recombinant Mycobacterium bovis Bacille Calmette-Guerin strains against an intracellular bacterial pathogen: importance of antigen secretion or membrane-targeted antigen display as lipoprotein for vaccine efficacy. *Journal of immunology (Baltimore, Md. : 1950)* **168**, 1869-1876 (2002).
17. Liang, J., *et al.* Enhanced and durable protective immune responses induced by a cocktail of recombinant BCG strains expressing antigens of multistage of Mycobacterium tuberculosis. *Molecular immunology* **66**, 392-401 (2015).
18. Conradt, P., Hess, J. & Kaufmann, S.H. Cytolytic T-cell responses to human dendritic cells and macrophages infected with Mycobacterium bovis BCG and recombinant BCG secreting listeriolysin. *Microbes and infection / Institut Pasteur* **1**, 753-764 (1999).
19. Koul, A., Arnoult, E., Lounis, N., Guillemont, J. & Andries, K. The challenge of new drug discovery for tuberculosis. *Nature* **469**, 483-490 (2011).
20. Gold, M.C., *et al.* Human mucosal associated invariant T cells detect bacterially infected cells. *PLoS biology* **8**, e1000407 (2010).

21. Arcos, J., *et al.* Human lung hydrolases delineate Mycobacterium tuberculosis-macrophage interactions and the capacity to control infection. *Journal of immunology (Baltimore, Md. : 1950)* **187**, 372-381 (2011).
22. Leemans, J.C., *et al.* Depletion of alveolar macrophages exerts protective effects in pulmonary tuberculosis in mice. *Journal of immunology (Baltimore, Md. : 1950)* **166**, 4604-4611 (2001).
23. Dorhoi, A. & Kaufmann, S.H. Perspectives on host adaptation in response to Mycobacterium tuberculosis: modulation of inflammation. *Seminars in immunology* **26**, 533-542 (2014).
24. Schlesinger, L.S. Macrophage phagocytosis of virulent but not attenuated strains of Mycobacterium tuberculosis is mediated by mannose receptors in addition to complement receptors. *Journal of immunology (Baltimore, Md. : 1950)* **150**, 2920-2930 (1993).
25. Le Cabec, V., Carreno, S., Moisand, A., Bordier, C. & Maridonneau-Parini, I. Complement receptor 3 (CD11b/CD18) mediates type I and type II phagocytosis during nonopsonic and opsonic phagocytosis, respectively. *Journal of immunology (Baltimore, Md. : 1950)* **169**, 2003-2009 (2002).
26. Ishikawa, E., *et al.* Direct recognition of the mycobacterial glycolipid, trehalose dimycolate, by C-type lectin Mincle. *The Journal of experimental medicine* **206**, 2879-2888 (2009).
27. Sweeney, K.A., *et al.* A recombinant Mycobacterium smegmatis induces potent bactericidal immunity against Mycobacterium tuberculosis. *Nature medicine* **17**, 1261-1268 (2011).
28. Lindestam Arlehamn, C.S., *et al.* Memory T cells in latent Mycobacterium tuberculosis infection are directed against three antigenic islands and largely contained in a CXCR3+CCR6+ Th1 subset. *PLoS pathogens* **9**, e1003130 (2013).
29. Siegrist, M.S., *et al.* Mycobacterial Esx-3 is required for mycobactin-mediated iron acquisition. *Proceedings of the National Academy of Sciences of the United States of America* **106**, 18792-18797 (2009).
30. Deshayes, C., *et al.* MmpS4 promotes glycopeptidolipids biosynthesis and export in Mycobacterium smegmatis. *Molecular microbiology* **78**, 989-1003 (2010).
31. Forrellad, M.A., *et al.* Virulence factors of the Mycobacterium tuberculosis complex. *Virulence* **4**, 3-66 (2013).
32. Reyrat, J.M. & Kahn, D. Mycobacterium smegmatis: an absurd model for tuberculosis? *Trends in microbiology* **9**, 472-474 (2001).
33. Mohan, A., Padiadpu, J., Baloni, P. & Chandra, N. Complete Genome Sequences of a Mycobacterium smegmatis Laboratory Strain (MC2 155) and Isoniazid-Resistant (4XR1/R2) Mutant Strains. *Genome announcements* **3**(2015).
34. Abdallah, A.M., *et al.* Type VII secretion--mycobacteria show the way. *Nature reviews. Microbiology* **5**, 883-891 (2007).
35. Houben, E.N., Korotkov, K.V. & Bitter, W. Take five - Type VII secretion systems of Mycobacteria. *Biochimica et biophysica acta* **1843**, 1707-1716 (2014).
36. Koo, I.C., *et al.* ESX-1-dependent cytolysis in lysosome secretion and inflammasome activation during mycobacterial infection. *Cellular microbiology* **10**, 1866-1878 (2008).
37. Bayan, N., Houssin, C., Chami, M. & Leblon, G. Mycomembrane and S-layer: two important structures of Corynebacterium glutamicum cell envelope with promising biotechnology applications. *Journal of biotechnology* **104**, 55-67 (2003).
38. Zuber, B., *et al.* Direct visualization of the outer membrane of mycobacteria and corynebacteria in their native state. *Journal of bacteriology* **190**, 5672-5680 (2008).
39. Lemassu, A. & Daffe, M. Structural features of the exocellular polysaccharides of Mycobacterium tuberculosis. *The Biochemical journal* **297 ( Pt 2)**, 351-357 (1994).
40. Feltcher, M.E., Sullivan, J.T. & Braunstein, M. Protein export systems of Mycobacterium tuberculosis: novel targets for drug development? *Future microbiology* **5**, 1581-1597 (2010).

41. Bottai, D., Serafini, A., Cascioferro, A., Brosch, R. & Manganelli, R. Targeting type VII/ESX secretion systems for development of novel antimycobacterial drugs. *Current pharmaceutical design* **20**, 4346-4356 (2014).
42. Champion, P.A. & Cox, J.S. Protein secretion systems in Mycobacteria. *Cellular microbiology* **9**, 1376-1384 (2007).
43. Jones, C.M. & Niederweis, M. Role of porins in iron uptake by Mycobacterium smegmatis. *Journal of bacteriology* **192**, 6411-6417 (2010).
44. Minnikin, D.E., Kremer, L., Dover, L.G. & Besra, G.S. The methyl-branched fortifications of Mycobacterium tuberculosis. *Chemistry & biology* **9**, 545-553 (2002).
45. Abdallah, A.M.*, et al.* Type VII secretion [mdash] mycobacteria show the way. *Nat Rev Micro* **5**, 883-891 (2007).
46. Chou, C.J.*, et al.* Functional studies of the Mycobacterium tuberculosis iron-dependent regulator. *The Journal of biological chemistry* **279**, 53554-53561 (2004).
47. Dussurget, O. & Smith, I. Interdependence of mycobacterial iron regulation, oxidative-stress response and isoniazid resistance. *Trends in microbiology* **6**, 354-358 (1998).
48. Olakanmi, O., Schlesinger, L.S., Ahmed, A. & Britigan, B.E. The nature of extracellular iron influences iron acquisition by Mycobacterium tuberculosis residing within human macrophages. *Infection and immunity* **72**, 2022-2028 (2004).
49. Olakanmi, O., Schlesinger, L.S., Ahmed, A. & Britigan, B.E. Intraphagosomal Mycobacterium tuberculosis acquires iron from both extracellular transferrin and intracellular iron pools. Impact of interferon-gamma and hemochromatosis. *The Journal of biological chemistry* **277**, 49727-49734 (2002).
50. Schaible, U.E., Collins, H.L., Priem, F. & Kaufmann, S.H. Correction of the iron overload defect in beta-2-microglobulin knockout mice by lactoferrin abolishes their increased susceptibility to tuberculosis. *The Journal of experimental medicine* **196**, 1507-1513 (2002).
51. Farhana, A.*, et al.* Mechanistic insights into a novel exporter-importer system of Mycobacterium tuberculosis unravel its role in trafficking of iron. *PloS one* **3**, e2087 (2008).
52. Wells, R.M.*, et al.* Discovery of a siderophore export system essential for virulence of Mycobacterium tuberculosis. *PLoS pathogens* **9**, e1003120 (2013).
53. Owens, C.P., Du, J., Dawson, J.H. & Goulding, C.W. Characterization of heme ligation properties of Rv0203, a secreted heme binding protein involved in Mycobacterium tuberculosis heme uptake. *Biochemistry* **51**, 1518-1531 (2012).
54. Quadri, L.E., Sello, J., Keating, T.A., Weinreb, P.H. & Walsh, C.T. Identification of a Mycobacterium tuberculosis gene cluster encoding the biosynthetic enzymes for assembly of the virulence-conferring siderophore mycobactin. *Chemistry & biology* **5**, 631-645 (1998).
55. Li, W., He, J., Xie, L., Chen, T. & Xie, J. Comparative genomic insights into the biosynthesis and regulation of mycobacterial siderophores. *Cellular physiology and biochemistry : international journal of experimental cellular physiology, biochemistry, and pharmacology* **31**, 1-13 (2013).
56. Rodriguez, G.M. Control of iron metabolism in Mycobacterium tuberculosis. *Trends in microbiology* **14**, 320-327 (2006).
57. Mangan, J.A., Sole, K.M., Mitchison, D.A. & Butcher, P.D. An effective method of RNA extraction from bacteria refractory to disruption, including mycobacteria. *Nucleic acids research* **25**, 675-676 (1997).
58. Fang, Z., Sampson, S.L., Warren, R.M., Gey van Pittius, N.C. & Newton-Foot, M. Iron acquisition strategies in mycobacteria. *Tuberculosis (Edinburgh, Scotland)* **95**, 123-130 (2015).
59. Chatzi, K.E., Sardis, M.F., Karamanou, S. & Economou, A. Breaking on through to the other side: protein export through the bacterial Sec system. *The Biochemical journal* **449**, 25-37 (2013).

60. Papanikou, E., Karamanou, S. & Economou, A. Bacterial protein secretion through the translocase nanomachine. *Nature reviews. Microbiology* **5**, 839-851 (2007).
61. Sato, K., *et al.* A protein secretion system linked to bacteroidete gliding motility and pathogenesis. *Proceedings of the National Academy of Sciences of the United States of America* **107**, 276-281 (2010).
62. Ates, L.S., *et al.* Essential Role of the ESX-5 Secretion System in Outer Membrane Permeability of Pathogenic Mycobacteria. *PLoS genetics* **11**, e1005190 (2015).
63. Dhandayuthapani, S., Zhang, Y., Mudd, M.H. & Deretic, V. Oxidative stress response and its role in sensitivity to isoniazid in mycobacteria: characterization and inducibility of ahpC by peroxides in Mycobacterium smegmatis and lack of expression in M. aurum and M. tuberculosis. *Journal of bacteriology* **178**, 3641-3649 (1996).
64. van der Woude, A.D., Luirink, J. & Bitter, W. Getting across the cell envelope: mycobacterial protein secretion. *Current topics in microbiology and immunology* **374**, 109-134 (2013).
65. Tseng, T.T., Tyler, B.M. & Setubal, J.C. Protein secretion systems in bacterial-host associations, and their description in the Gene Ontology. *BMC microbiology* **9 Suppl 1**, S2 (2009).
66. Bitter, W., *et al.* Systematic genetic nomenclature for type VII secretion systems. *PLoS pathogens* **5**, e1000507 (2009).
67. Deng, W., Xiang, X. & Xie, J. Comparative genomic and proteomic anatomy of Mycobacterium ubiquitous Esx family proteins: implications in pathogenicity and virulence. *Current microbiology* **68**, 558-567 (2014).
68. Delogu, G., Sali, M. & Fadda, G. THE BIOLOGY OF MYCOBACTERIUM TUBERCULOSIS INFECTION. *2013* **5**(2013).
69. Guinn, K.M., *et al.* Individual RD1-region genes are required for export of ESAT-6/CFP-10 and for virulence of Mycobacterium tuberculosis. *Molecular microbiology* **51**, 359-370 (2004).
70. Coros, A., Callahan, B., Battaglioli, E. & Derbyshire, K.M. The specialized secretory apparatus ESX-1 is essential for DNA transfer in Mycobacterium smegmatis. *Molecular microbiology* **69**, 794-808 (2008).
71. Solomonson, M., *et al.* Structure of EspB from the ESX-1 type VII secretion system and insights into its export mechanism. *Structure (London, England : 1993)* **23**, 571-583 (2015).
72. Romagnoli, A., *et al.* ESX-1 dependent impairment of autophagic flux by Mycobacterium tuberculosis in human dendritic cells. *Autophagy* **8**, 1357-1370 (2012).
73. Houben, D., *et al.* ESX-1-mediated translocation to the cytosol controls virulence of mycobacteria. *Cellular microbiology* **14**, 1287-1298 (2012).
74. Serafini, A., Pisu, D., Palu, G., Rodriguez, G.M. & Manganelli, R. The ESX-3 secretion system is necessary for iron and zinc homeostasis in Mycobacterium tuberculosis. *PloS one* **8**, e78351 (2013).
75. Siegrist, M.S., *et al.* Mycobacterial Esx-3 requires multiple components for iron acquisition. *mBio* **5**, e01073-01014 (2014).
76. Di Luca, M., *et al.* The ESX-5 associated eccB-EccC locus is essential for Mycobacterium tuberculosis viability. *PloS one* **7**, e52059 (2012).
77. Winterbourn, C.C. Toxicity of iron and hydrogen peroxide: the Fenton reaction. *Toxicology letters* **82-83**, 969-974 (1995).
78. Lee, H.N., Lee, N.O., Han, S.J., Ko, I.J. & Oh, J.I. Regulation of the ahpC gene encoding alkyl hydroperoxide reductase in Mycobacterium smegmatis. *PloS one* **9**, e111680 (2014).
79. Rodriguez, G.M., Voskuil, M.I., Gold, B., Schoolnik, G.K. & Smith, I. ideR, An essential gene in mycobacterium tuberculosis: role of IdeR in iron-dependent gene expression, iron metabolism, and oxidative stress response. *Infection and immunity* **70**, 3371-3381 (2002).
80. Milano, A., Forti, F., Sala, C., Riccardi, G. & Ghisotti, D. Transcriptional Regulation of furA and katG upon Oxidative Stress in Mycobacterium smegmatis. *Journal of bacteriology* **183**, 6801-6806 (2001).

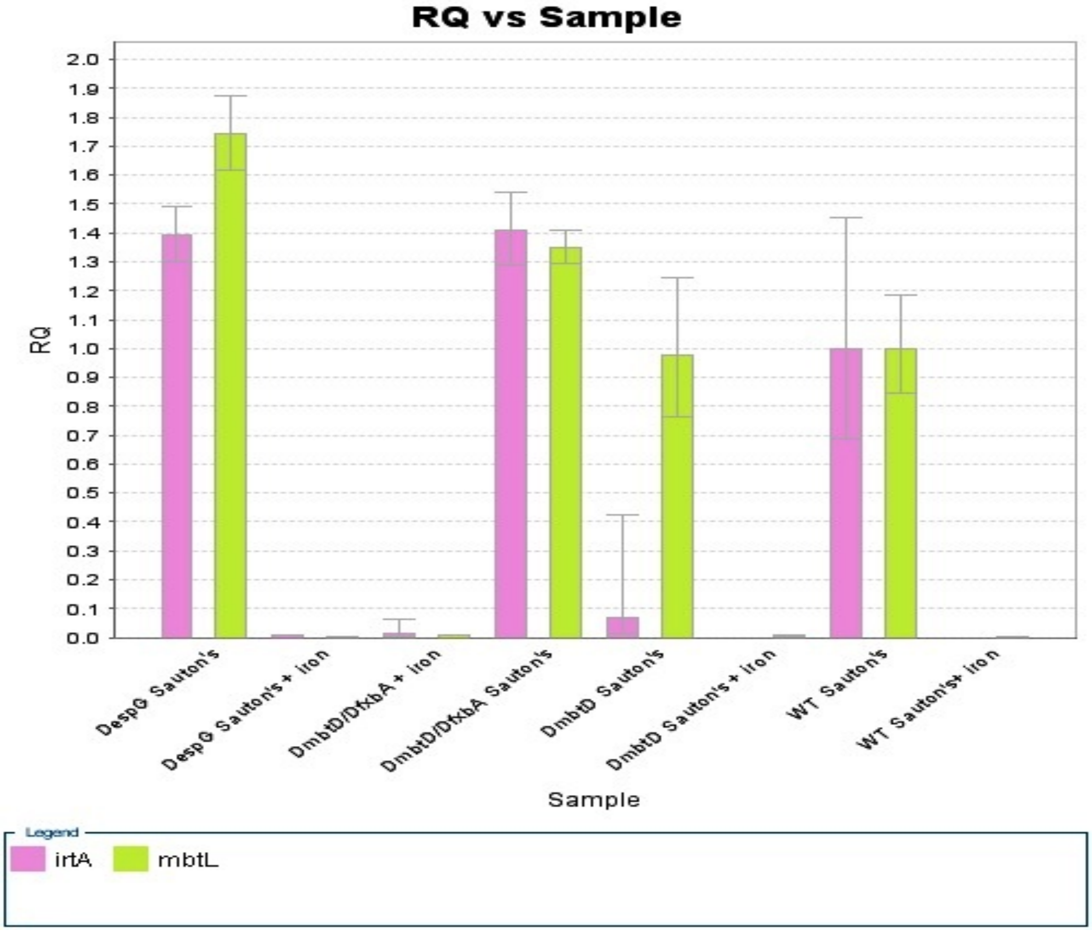
81. Teixeira, V.H., *et al.* Molecular details of INH-C10 binding to wt KatG and Its S315T mutant. *Molecular pharmaceutics* **12**, 898-909 (2015).
82. Sherman, D.R., *et al.* Disparate responses to oxidative stress in saprophytic and pathogenic mycobacteria. *Proceedings of the National Academy of Sciences of the United States of America* **92**, 6625-6629 (1995).
83. Hillas, P.J., del Alba, F.S., Oyarzabal, J., Wilks, A. & Ortiz De Montellano, P.R. The AhpC and AhpD antioxidant defense system of Mycobacterium tuberculosis. *The Journal of biological chemistry* **275**, 18801-18809 (2000).
84. Koshkin, A., Knudsen, G.M. & Ortiz De Montellano, P.R. Intermolecular interactions in the AhpC/AhpD antioxidant defense system of Mycobacterium tuberculosis. *Archives of biochemistry and biophysics* **427**, 41-47 (2004).
85. Guimaraes, B.G., *et al.* Structure and mechanism of the alkyl hydroperoxidase AhpC, a key element of the Mycobacterium tuberculosis defense system against oxidative stress. *The Journal of biological chemistry* **280**, 25735-25742 (2005).
86. Daugherty, A., Powers, K.M., Standley, M.S., Kim, C.S. & Purdy, G.E. Mycobacterium smegmatis RoxY is a repressor of oxyS and contributes to resistance to oxidative stress and bactericidal ubiquitin-derived peptides. *Journal of bacteriology* **193**, 6824-6833 (2011).
87. Pagan-Ramos, E., Song, J., McFalone, M., Mudd, M.H. & Deretic, V. Oxidative stress response and characterization of the oxyR-ahpC and furA-katG loci in Mycobacterium marinum. *Journal of bacteriology* **180**, 4856-4864 (1998).
88. Dosanjh, N.S., Rawat, M., Chung, J.H. & Av-Gay, Y. Thiol specific oxidative stress response in Mycobacteria. *FEMS microbiology letters* **249**, 87-94 (2005).
89. Saikolappan, S., Das, K., Sasindran, S.J., Jagannath, C. & Dhandayuthapani, S. OsmC proteins of Mycobacterium tuberculosis and Mycobacterium smegmatis protect against organic hydroperoxide stress. *Tuberculosis (Edinburgh, Scotland)* **91 Suppl 1**, S119-127 (2011).
90. Rickman, L., *et al.* A member of the cAMP receptor protein family of transcription regulators in Mycobacterium tuberculosis is required for virulence in mice and controls transcription of the rpfA gene coding for a resuscitation promoting factor. *Molecular microbiology* **56**, 1274-1286 (2005).
91. Dona, V., *et al.* Evidence of complex transcriptional, translational, and posttranslational regulation of the extracytoplasmic function sigma factor sigmaE in Mycobacterium tuberculosis. *Journal of bacteriology* **190**, 5963-5971 (2008).
92. Bhaskar, A., *et al.* Reengineering redox sensitive GFP to measure mycothiol redox potential of Mycobacterium tuberculosis during infection. *PLoS pathogens* **10**, e1003902 (2014).
93. Nunoshiba, T., Hidalgo, E., Amabile Cuevas, C.F. & Demple, B. Two-stage control of an oxidative stress regulon: the Escherichia coli SoxR protein triggers redox-inducible expression of the soxS regulatory gene. *Journal of bacteriology* **174**, 6054-6060 (1992).
94. Gupta, S. & Chatterji, D. Stress responses in mycobacteria. *IUBMB life* **57**, 149-159 (2005).
95. Waggott, D., *et al.* NanoStringNorm: an extensible R package for the pre-processing of NanoString mRNA and miRNA data. *Bioinformatics (Oxford, England)* **28**, 1546-1548 (2012).
96. Ramakers, C., Ruijter, J.M., Deprez, R.H. & Moorman, A.F. Assumption-free analysis of quantitative real-time polymerase chain reaction (PCR) data. *Neuroscience letters* **339**, 62-66 (2003).
97. Magnani, E., Bartling, L. & Hake, S. From Gateway to MultiSite Gateway in one recombination event. *BMC Molecular Biology* **7**, 46-46 (2006).
98. Beaume, M., *et al.* Orientation and expression of methicillin-resistant Staphylococcus aureus small RNAs by direct multiplexed measurements using the nCounter of NanoString technology. *Journal of microbiological methods* **84**, 327-334 (2011).
99. Stephan, J., Bail, J.G., Titgemeyer, F. & Niederweis, M. DNA-free RNA preparations from mycobacteria. *BMC microbiology* **4**, 45-45 (2004).
100. Wang, N. Norwegian University of Science and Technology (NTNU) (2014).

101. Gold, B., Rodriguez, G.M., Marras, S.A., Pentecost, M. & Smith, I. The Mycobacterium tuberculosis IdeR is a dual functional regulator that controls transcription of genes involved in iron acquisition, iron storage and survival in macrophages. *Molecular microbiology* **42**, 851-865 (2001).
102. Buchmeier, N. & Fahey, R.C. The mshA gene encoding the glycosyltransferase of mycothiol biosynthesis is essential in Mycobacterium tuberculosis Erdman. *FEMS microbiology letters* **264**, 74-79 (2006).
103. Smegmalist. http://mycobrowser.epfl.ch/tmhmm_smegma/MSMEG_0616.html. Vol. 2015 (2015).
104. Smegmalist. http://mycobrowser.epfl.ch/tmhmm_smegma/MSMEG_0617.html. Vol. 2015 (2015).
105. Smegmalist. http://mycobrowser.epfl.ch/tmhmm_smegma/MSMEG_0622.html. Vol. 2015 (2015).
106. Newton-Foot, M. & Gey van Pittius, N.C. The complex architecture of mycobacterial promoters. *Tuberculosis (Edinburgh, Scotland)* **93**, 60-74 (2013).
107. Smegmalist. http://mycobrowser.epfl.ch/smegmasearch.php?gene+name=espg3&submit=Search. Vol. 2015 (2015).
108. Brumbaugh, C.D., Kim, H.J., Giovacchini, M. & Pourmand, N. NanoStriDE: normalization and differential expression analysis of NanoString nCounter data. *BMC bioinformatics* **12**, 479 (2011).
109. Betts, J.C., Lukey, P.T., Robb, L.C., McAdam, R.A. & Duncan, K. Evaluation of a nutrient starvation model of Mycobacterium tuberculosis persistence by gene and protein expression profiling. *Molecular microbiology* **43**, 717-731 (2002).
110. Rodriguez, G.M. & Smith, I. Identification of an ABC transporter required for iron acquisition and virulence in Mycobacterium tuberculosis. *Journal of bacteriology* **188**, 424-430 (2006).
111. Serafini, A., Boldrin, F., Palu, G. & Manganelli, R. Characterization of a Mycobacterium tuberculosis ESX-3 conditional mutant: essentiality and rescue by iron and zinc. *Journal of bacteriology* **191**, 6340-6344 (2009).
112. Gao, L., *et al.* Pulmonary Langerhans cell histiocytosis with cervical lymph node involvement, and coexistence with pulmonary tuberculosis and right pneumothorax: a case report and review of literature. *International journal of clinical and experimental pathology* **8**, 2146-2152 (2015).
113. Dussurget, O., Rodriguez, M. & Smith, I. Protective role of the Mycobacterium smegmatis IdeR against reactive oxygen species and isoniazid toxicity. *Tubercle and lung disease : the official journal of the International Union against Tuberculosis and Lung Disease* **79**, 99-106 (1998).
114. Zahrt, T.C., Song, J., Siple, J. & Deretic, V. Mycobacterial FurA is a negative regulator of catalase-peroxidase gene katG. *Molecular microbiology* **39**, 1174-1185 (2001).
115. Milano, A., Forti, F., Sala, C., Riccardi, G. & Ghisotti, D. Transcriptional regulation of furA and katG upon oxidative stress in Mycobacterium smegmatis. *Journal of bacteriology* **183**, 6801-6806 (2001).
116. http://mycobrowser.epfl.ch/smegmasearch.php?gene+name=MSMEG_0622&submit=Search. (2015).
117. Snapp, E. Design and use of fluorescent fusion proteins in cell biology. *Current protocols in cell biology / editorial board, Juan S. Bonifacino ... [et al.]* **Chapter 21**, Unit 21.24 (2005).

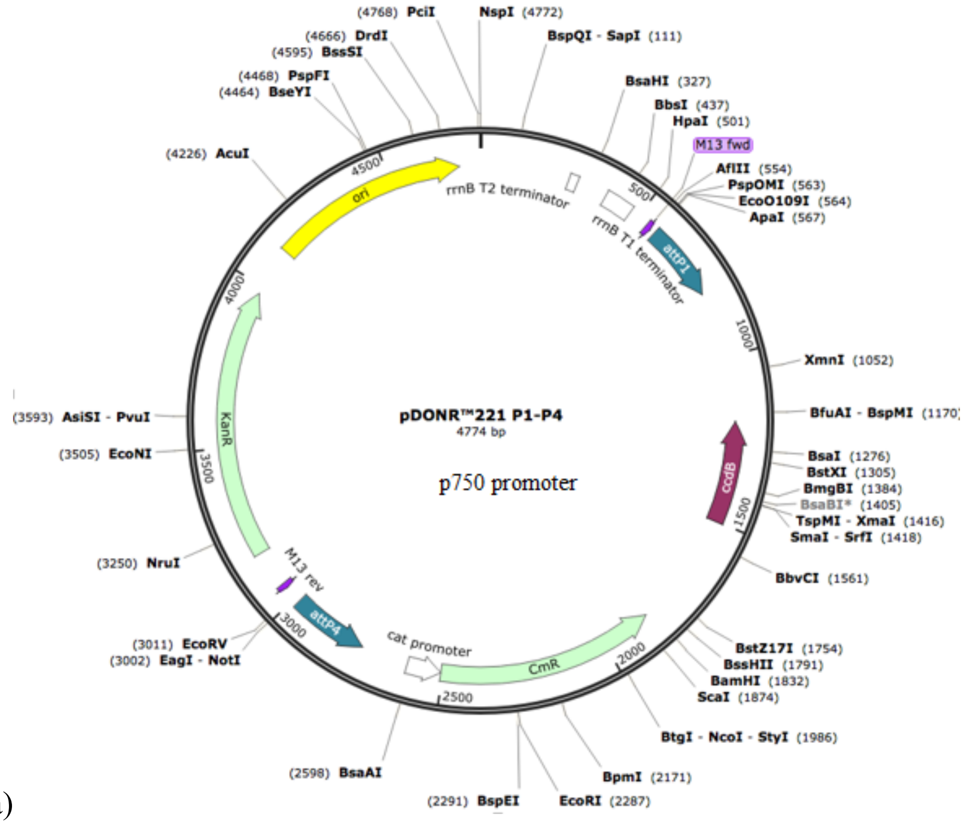
# 8 Appendix

Supplementary table 1. List of primers designed and used in this study

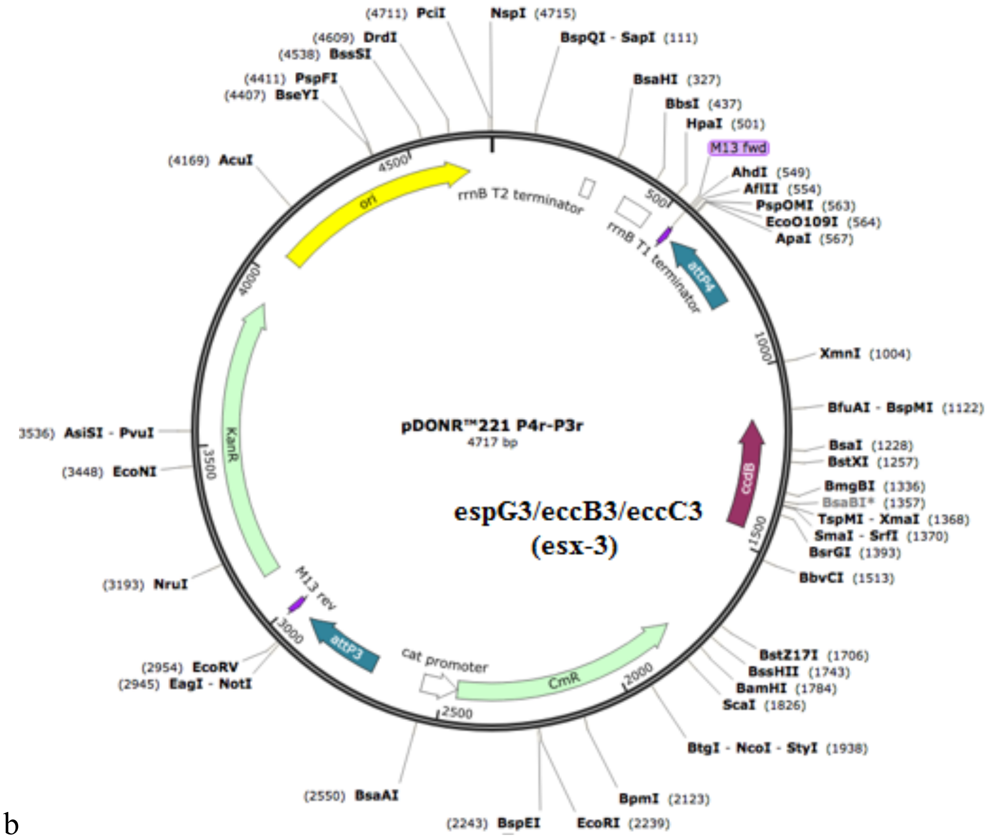
| **Name** | **Primer sequence 5' -----> 3'** |
|---|---|
| *SigA* forward | AAGACACCGACCTGGAACTC |
| *SigA* reverse | AGCTTCTTCTTCTTCCTCGTCCTC |
| *irtA* forward | ACACGATGTCCGAGATGTCA |
| *irtA* reverse | GTCGTTCTCGTCGTGCTCT |
| *katG(MSMEG_6384)*forward | TGGCCCAATCAGCTCAATC |
| *katG(MSMEG_6384)*reverse | GACAGGTTCTTGCCGTACTTC |
| *ahpC* forward | CAGCCCGATGACTACTTCACC |
| *ahpC* reverse | CTCACGCTTGAGATCCGAGAC |
| *whib7* forward | TGACATGCGAGACGCGACTG |
| *whib7* reverse | TTGCCGCCGGAATCCTTACG |
| *oxyS* forward | GGCCAACGTCTCCATATCCTC |
| *oxyS* reverse | TCGATGAACTGCCGGAACC |
| *whib3* forward | CAACTACCCGGTCCGAATG |
| *whib3* reverse | AATTCCGCGCTTGAGCAG |
| **gateway primers** | |
| B1-p750 promoter ( forward) | GGGGACAAGTTTGTACAAAAAAGCAGGCTGGACCAGGCCTAGATCTGGGGAC |
| B4-p750 promoter ( reverse) | GGGGACAACTTTGTATAGAAAAGTTGGGTGGGTGCATGCGGTTGTGAGC |
| B4r-espG3 (forward) | GGGGACAACTTTTCTATACAAAGTTGGGCGGATCACCGTGGGGCCTAA |
| B3r-espG3 ( reverse) | GGGGACAACTTTATTATACAAAGTTGTTGCTTTCTGGGTTCTTCTCTGGAG |
| B4r-eccB3 ( forward) | GGGGACAACTTTTCTATACAAAGTTGGGGGTGCCGGCATGACCGGCCCCGTCA |
| B3r-eccB3 ( reverse) | GGGGACAACTTTATTATACAAAGTTGTTCGGGAGGCCTCCATACGCG |
| B4r-eccC3 ( forward) | GGGGACAACTTTTCTATACAAAGTTGTAGGCCTCCCGATGAGCCGGCTCATCT |
| B3r-eccC3 ( reverse) | GGGGACAACTTTATTATACAAAGTTGTTCGGTATTCCCCTCCTCGGTTGTG |
| B3-gfp ( forward) | GGGGACAACTTTGTATAATAAAGTTGGGGTGAGCAAGGGCGAGGAGC |
| B2-gfp ( reverse) | GGGGACCACTTTGTACAAGAAAGCTGGGTACTTGTACAGCTCGTCCATGCC |
| B3-cherry ( forward) | GGGGACAACTTTGTATAATAAAGTTGGGATGGTGAGCAAGGGCGAGGAG |
| B2-cherry ( reverse) | GGGGACCACTTTGTACAAGAAAGCTGGGTATTACTTGTACAGCTCGTCCATGCCG |



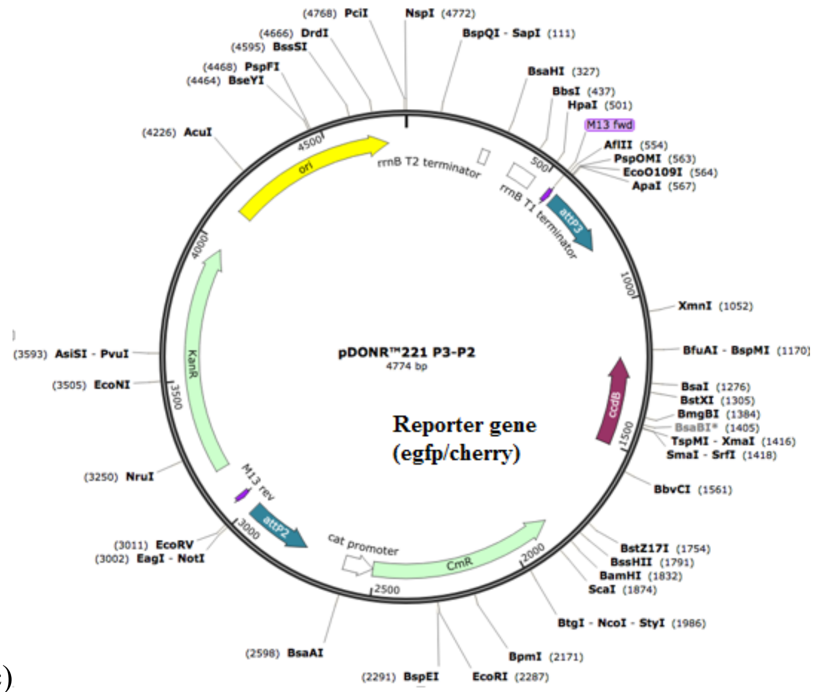

Supplementary figure 1. qPCR test to check the quality of RNA which extracted from bacteria grown in low iron (LI) and high iron (HI) condition. RNA extract was tested by real time PCR using irtA and mbtL primers with both cDNA and RNA sample. None of RNA samples was amplified to the detectable levels (undetermined CT value) while the corresponding cDNA amplified the target. The melting curve also showed (not shown here), nonspecific amplification peaks were not observed which indicated nonspecific DNA was undetectable in the RNA sample.


pDONR™221 P1-P4
4774 bp
p750 promoter

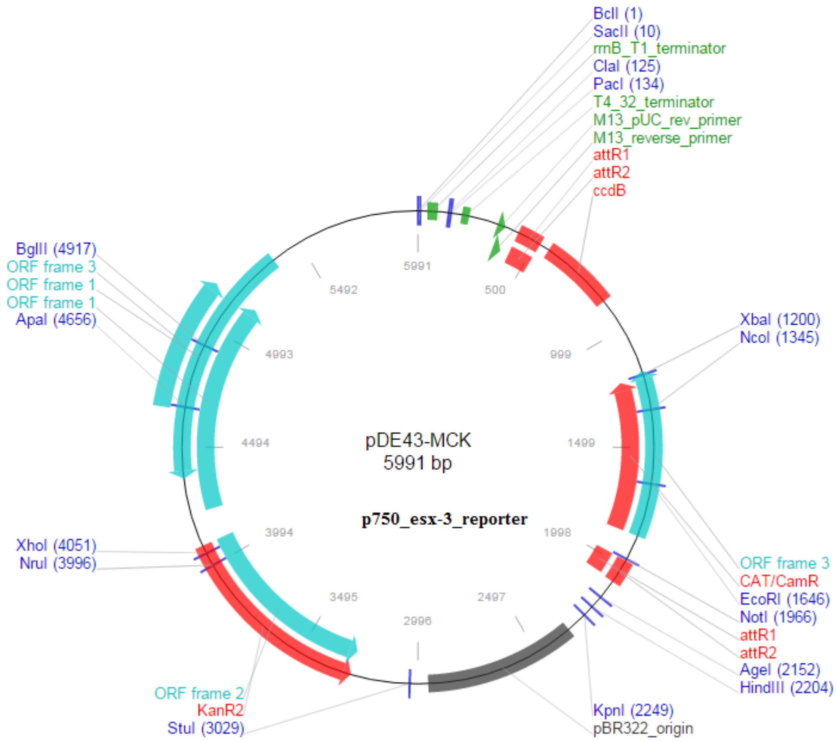
a)


pDONR™221 P4r-P3r
4717 bp
espG3/eccB3/eccC3
(esx-3)

b




c)




d)

Supplementary figure 2. Map of vectors used for multisite gateway cloning system. a, b and c were kanamycin resistant donor vectors for promoter( p750),espG3/eccB3/eccC3(esx-3) and egfp/cherry(reporter gene) respectively. Each of the attB flanked PCR products recombine with attB sites of compatible vector and replaced a casette containing ccdB gene to make an entry clones.(source from invitrogen life technlogies) d) pDE43-MCK, kanamycin resistant destination vector with attR1 and attR2 reaction sites that was engineered to compatible with Mycobacteria and 3 fragment PCR gatewaycloning system. The 3 pDONR vectors cloned with PCR products of interested genes were recombine with the destination vector pDE43 and reaction was facilitated by LR clonase plus enzyme. The cloned pDE43 then transformed in to M. smegmatis and succesful cloned bacteria was isolated from kanamycin containing media. (map adopted from www.addgene.org/49523/and 10)

## 1.3 Supplementary tables

Supplementary table 2. M. smegmatis RNA quantity and quality check by NanoDrop spectrophotometer measurement.

Plots | Report | Test type: Nucleic Acid | 30.01.2015 21:41 | Exit

Report Name | Max Buffer Size 200 | Buffer Mode Save Report & Clear

| Sample ID | User ID | Date | Time | ng/ul | A260 | A280 | 260/280 | 260/230 | Constant | Cursor Pos. | Cursor abs. | 340 raw |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| water | Default | 30.01.2015 | 21:21 | 0,30 | 0,007 | 0,001 | 10,79 | 1,22 | 40,00 | 230 | 0,006 | 0,001 |
| 1 | Default | 30.01.2015 | 21:22 | 153,17 | 3,829 | 1,943 | 1,97 | 1,17 | 40,00 | 230 | 3,272 | 0,038 |
| 2 | Default | 30.01.2015 | 21:22 | 103,26 | 2,582 | 1,301 | 1,98 | 1,66 | 40,00 | 230 | 1,556 | 0,024 |
| 3 | Default | 30.01.2015 | 21:23 | 314,55 | 7,864 | 3,891 | 2,02 | 2,13 | 40,00 | 230 | 3,689 | 0,030 |
| 4 | Default | 30.01.2015 | 21:24 | 310,57 | 7,764 | 3,862 | 2,01 | 2,01 | 40,00 | 230 | 3,854 | 0,060 |
| 5 | Default | 30.01.2015 | 21:25 | 323,24 | 8,081 | 4,016 | 2,01 | 1,79 | 40,00 | 230 | 4,510 | 0,045 |
| 6 | Default | 30.01.2015 | 21:27 | 159,94 | 3,999 | 2,000 | 2,00 | 1,84 | 40,00 | 230 | 2,173 | 0,010 |
| 7 | Default | 30.01.2015 | 21:29 | 245,83 | 6,146 | 3,023 | 2,03 | 1,96 | 40,00 | 230 | 3,132 | 0,040 |
| 8 | Default | 30.01.2015 | 21:30 | 240,54 | 6,014 | 2,941 | 2,04 | 2,06 | 40,00 | 230 | 2,913 | 0,037 |
| 9 | Default | 30.01.2015 | 21:31 | 351,07 | 8,777 | 4,432 | 1,98 | 2,05 | 40,00 | 230 | 4,272 | 0,053 |
| 10 | Default | 30.01.2015 | 21:31 | 219,10 | 5,478 | 2,711 | 2,02 | 1,90 | 40,00 | 230 | 2,882 | 0,012 |
| 11 | Default | 30.01.2015 | 21:32 | 416,47 | 10,412 | 5,259 | 1,98 | 1,36 | 40,00 | 230 | 7,642 | 0,062 |
| 12 | Default | 30.01.2015 | 21:33 | 336,36 | 8,409 | 4,192 | 2,01 | 2,09 | 40,00 | 230 | 4,019 | 0,024 |
| 13 | Default | 30.01.2015 | 21:34 | 308,03 | 7,701 | 3,821 | 2,02 | 2,06 | 40,00 | 230 | 3,746 | 0,034 |
| 14 | Default | 30.01.2015 | 21:35 | 271,28 | 6,782 | 3,389 | 2,00 | 2,07 | 40,00 | 230 | 3,273 | 0,026 |
| 15 | Default | 30.01.2015 | 21:36 | 261,96 | 6,549 | 3,269 | 2,00 | 1,98 | 40,00 | 230 | 3,299 | 0,025 |
| 16 | Default | 30.01.2015 | 21:36 | 291,85 | 7,296 | 3,648 | 2,00 | 1,91 | 40,00 | 230 | 3,829 | 0,035 |

Supplementary table 3. Result from NanoString mRNA count ratio for 107 M. smegmatis genes. All the following Mycobacteria smegmatis genes listed were selected to screen out the expression levels in wild type, ΔespG3 and ΔmbtD mutant M. smegmatis under low and high iron growth conditions by NanoString nCounter gene expression analysis and normalized by housekeeping gene sigA. The mRNA ratio of each mutant strains versus WT gene in low and high iron condition were calculated and analyzed by nSolver version 1.1 software.

| **gene name** | **SAUTONS** | | | **SAUTONS + IRON** | | |
|---|---|---|---|---|---|---|
| | **mRNA Ratios against WT** | | | **mRNA Ratios against WT** | | |
| | **WT** | **Δ*espG*$_3$** | **Δ*mbt*D** | **WT** | **Δ*espG*$_3$** | **Δ*mbt*D** |
| MSMEG_0219 | 1 | -1.51 | -2.57 | 1 | -1.11 | 1.02 |
| MSMEG_0409 | 1 | 1.5 | 1.96 | 1 | -1.14 | 1.12 |
| MSMEG_0788 | 1 | -1.35 | -1.44 | 1 | 1.67 | 1.13 |
| MSMEG_1906 | 1 | 1.27 | 1.06 | 1 | -1.28 | -1.17 |
| MSMEG_3117 | 1 | 1.55 | 1.35 | 1 | 1.21 | -1.06 |
| MSMEG_3519 | 1 | 1.14 | 1.26 | 1 | 1.04 | -1.07 |
| MSMEG_3628 | 1 | -1.02 | -1.11 | 1 | 1.24 | 1.12 |
| MSMEG_4279 | 1 | 1.07 | 1.08 | 1 | -1.02 | 1.07 |
| MSMEG_5018 | 1 | 1.2 | 1.28 | 1 | -1.21 | 1.08 |

| | | | | | | |
|---|---|---|---|---|---|---|
| MSMEG_5900 | 1 | 1.14 | 1.45 | 1 | 1.17 | -1.03 |
| MSMEG_5994 | 1 | 1.28 | 1.19 | 1 | 1.01 | -1.05 |
| MSMEG_6014 | 1 | 1.21 | -1 | 1 | -1.14 | -1.37 |
| MSMEG_6069 | 1 | 1.16 | 1.03 | 1 | 1.02 | 1.08 |
| MSMEG_6942 | 1 | 1.12 | 1.12 | 1 | -1.26 | 1.07 |
| *mpr*B (MSMEG_5487) | 1 | 1.55 | 1.77 | 1 | 1.45 | 1.09 |
| *acn* (MSMEG_3143) | 1 | -1.02 | -1.08 | 1 | 1.16 | 1.1 |
| *aft*A (MSMEG_6386) | 1 | 1.09 | 1.07 | 1 | -1.12 | 1.07 |
| *ahp*C (MSMEG_4891) | 1 | 1.24 | -1.18 | 1 | -2.88 | -1.02 |
| *alr* (MSMEG_1575) | 1 | 1.43 | 1.32 | 1 | 1.38 | 1.28 |
| *atp*A (MSMEG_4938) | 1 | -2.04 | -1.02 | 1 | -2.06 | -1.09 |
| *atp*B (MSMEG_4942) | 1 | -1.67 | 1.1 | 1 | -2.05 | -1.1 |
| *bfr*A (MSMEG_3564) | 1 | 1.8 | 2.05 | 1 | 1.57 | 1.18 |
| *bla*I (MSMEG_3630) | 1 | 1.48 | 1.31 | 1 | 1.37 | 1.22 |
| *bla*R (MSMEG_3631) | 1 | 1.26 | 1.25 | 1 | 1.26 | -1.01 |
| *cat*A (MSMEG_1911) | 1 | 1.41 | 1.29 | 1 | 1.47 | -1.34 |
| *cta*E (MSMEG_4260) | 1 | 2.62 | 2.27 | 1 | 1.77 | 1.18 |
| *cyd*D (MSMEG_3233 | 1 | 1.48 | 1.38 | 1 | 1.28 | 1.2 |
| *dev*B (MSMEG_3099) | 1 | 1.34 | 1.34 | 1 | -1.43 | -1.06 |
| *din*X (MSMEG_3172) | 1 | -1.13 | -1.4 | 1 | -1.38 | -1.13 |
| *dla*T (MSMEG_4283) | 1 | 1.24 | 1.43 | 1 | 1.11 | 1.08 |
| *dna*A (MSMEG_6947) | 1 | 1.24 | 1.48 | 1 | -1.21 | -1.09 |
| *dna*N (MSMEG_0001) | 1 | -1.32 | -1.66 | 1 | -1.06 | 1.07 |
| *dos*R (MSMEG_5245) | 1 | -1.44 | -1.3 | 1 | 1.53 | 1.48 |
| *dos*R (MSMEG_3944) | 1 | -1.04 | 1.32 | 1 | 2.28 | 1.88 |
| *ecc*A3 (MSMEG_0615) | 1 | 1.18 | 1.22 | 1 | -2.05 | -1.07 |
| *fab*H (MSMEG_3953) | 1 | -1.16 | 1.23 | 1 | 1.28 | 1.18 |
| *fad*B (MSMEG_5720) | 1 | -1.19 | -1.24 | 1 | 1.44 | -1 |
| *fad*E34 (MSMEG_6041) | 1 | 1.31 | 1.3 | 1 | 1.23 | 1.53 |
| *fad*E5 (MSMEG_0406) | 1 | -1.21 | -1.17 | 1 | 1.35 | 1.11 |
| *fas* (MSMEG_4757) | 1 | -1.43 | -2.07 | 1 | -1.12 | -1.07 |
| *fts*Z (MSMEG_4222) | 1 | 1.04 | 1.08 | 1 | 1.35 | 1.11 |
| *fur*A (MSMEG_3460) | 1 | -1.37 | -1.07 | 1 | 1.04 | -1.04 |
| f*ur*A (MSMEG_6383) | 1 | 2.03 | 1.38 | 1 | -4.57 | 1.06 |
| *fxb*A (MSMEG_0014) | 1 | 1.29 | 1.42 | 1 | -1.25 | 1.12 |
| *glc*B (MSMEG_3640) | 1 | -1 | -1.04 | 1 | 1.13 | -1.05 |
| *gln*A1 (MSMEG_4290) | 1 | 1.1 | 1.34 | 1 | 1.17 | 1.1 |
| *glp*X (MSMEG_5239) | 1 | 1.31 | 1.33 | 1 | 1.09 | 1.07 |
| *gro*S (MSMEG_1582) | 1 | 1.1 | 1.19 | 1 | 1.12 | 1.08 |
| *hsa*A (MSMEG_6038) | 1 | 1.47 | 1.42 | 1 | 1.22 | 1.01 |
| *ide*R (MSMEG_2750) | 1 | 1.16 | 1.07 | 1 | 1.16 | 1.1 |
| *irt*A (MSMEG_6554) | 1 | -2.01 | -5.75 | 1 | -1.26 | 1.04 |
| *kat*G (MSMEG_6384) | 1 | 1.65 | 1.35 | 1 | -6.35 | -1.18 |
| *kat*G (MSMEG_3729) | 1 | -1.13 | 1.4 | 1 | -1.15 | -1.15 |
| *kat*G (MSMEG_3461) | 1 | -1.14 | 1.18 | 1 | 1.18 | 1.03 |
| *kgd* (MSMEG_5049) | 1 | 1.18 | 1.1 | 1 | -1.13 | 1.09 |

| | | | | | | |
|---|---|---|---|---|---|---|
| *kst*D (MSMEG_5941) | 1 | 1.5 | 1.73 | 1 | 2.17 | 1.52 |
| *kst*R (MSMEG_6042) | 1 | 1.23 | 1.25 | 1 | -1.02 | 1.06 |
| *lip*E (MSMEG_6575) | 1 | 1.07 | 1.05 | 1 | -1.31 | 1.12 |
| *mbt*B (MSMEG_4515) | 1 | 1.02 | 1.1 | 1 | 1.97 | 1.55 |
| *mbt*L (MSMEG_2132) | 1 | 1.01 | 1.04 | 1 | -1.06 | 1.21 |
| *mmp*S4 (MSMEG_0380) | 1 | 1.72 | 2.1 | 1 | 1.3 | 1.15 |
| *msh*A (MSMEG_0933) | 1 | -1.84 | -2.04 | 1 | -1.15 | -1.03 |
| *msh*C (MSMEG_4189) | 1 | 1.14 | -1.19 | 1 | -1.7 | -1.06 |
| *mur*A (MSMEG_4932) | 1 | -1.62 | 1 | 1 | -2.19 | -1.16 |
| *mur*E (MSMEG_4232) | 1 | 1.41 | 1.38 | 1 | -1.07 | 1.03 |
| *mur*I (MSMEG_4903) | 1 | 1.03 | 1.06 | 1 | -1.38 | 1.01 |
| *nuo*A (MSMEG_2063) | 1 | 1.52 | 1.37 | 1 | 1.5 | 1.05 |
| *nuo*N (MSMEG_2050) | 1 | 1.54 | 1.35 | 1 | 1.01 | -1.05 |
| *oxy*S (MSMEG_0156) | 1 | 1.42 | 1.15 | 1 | 1.09 | 1.31 |
| *par*A (MSMEG_6939) | 1 | 1.09 | 1.14 | 1 | -1.11 | 1.01 |
| *par*B (MSMEG_6938) | 1 | 1.23 | 1.28 | 1 | 1.02 | 1.03 |
| *pck*A (MSMEG_0255) | 1 | -1 | -1.04 | 1 | -1.04 | 1.01 |
| *pfk*B (MSMEG_3947) | 1 | -1.03 | -1.17 | 1 | 1.99 | 1.43 |
| *pho*T (MSMEG_5779) | 1 | -1.02 | 1.11 | 1 | 1.02 | 1.01 |
| *pkn*A (MSMEG_0030) | 1 | 1.33 | 1.39 | 1 | 1.08 | 1.01 |
| *pks*2 (MSMEG_4727) | 1 | 1.74 | 1.91 | 1 | 1.15 | -1.01 |
| *pon*A1 (MSMEG_6900) | 1 | -1.13 | -1.25 | 1 | -1.58 | -1.07 |
| *pon*A2 (MSMEG_6201) | 1 | 1.01 | 1.06 | 1 | -1.24 | 1.03 |
| *ppa* (MSMEG_6114) | 1 | 1.44 | 1.63 | 1 | -1.08 | 1.03 |
| *ppg*K (MSMEG_2760) | 1 | 1.52 | 1.77 | 1 | -1.04 | 1.02 |
| *ppr*B (MSMEG_5663) | 1 | 1.29 | 1.46 | 1 | 1.24 | 1.03 |
| *prp*C (MSMEG_6647) | 1 | 1.34 | 1.3 | 1 | -1.01 | -1.08 |
| *pst*C (MSMEG_5781) | 1 | 1.13 | 1.18 | 1 | 1.06 | -1 |
| *rec*B (MSMEG_1327) | 1 | -1.5 | -2.4 | 1 | -1.23 | 1.12 |
| *rip*A (MSMEG_3145) | 1 | 2.31 | 3.67 | 1 | 1.44 | 1.18 |
| *rod*A (MSMEG_0032) | 1 | 1.09 | 1.27 | 1 | 1.12 | 1.07 |
| *rpm*B2 (MSMEG_6068) | 1 | 1.46 | 1.36 | 1 | 1.34 | 1.1 |
| *scp*A (MSMEG_3742) | 1 | 1.1 | 1.24 | 1 | 1.32 | 1.01 |
| *sdh*C (MSMEG_1672) | 1 | 1.81 | 1.87 | 1 | 1.34 | 1.16 |
| *senx*3 (MSMEG_0936) | 1 | 1.07 | 1.18 | 1 | -1.04 | 1.04 |
| *sig*A (MSMEG_2758) | 1 | -1 | -1 | 1 | -1 | -1 |
| *sig*B (MSMEG_2752) | 1 | 1.67 | 1.28 | 1 | 1.61 | 1.18 |
| *sig*F (MSMEG_1804) | 1 | 1.09 | 1.27 | 1 | 1.49 | -1.16 |
| *sod*A (MSMEG_6427) | 1 | 1.21 | 1.34 | 1 | -1.33 | 1.07 |
| *sod*A (MSMEG_6636) | 1 | 1.27 | 1.47 | 1 | -1.11 | 1.04 |
| *ssb* (MSMEG_4701) | 1 | -1.23 | -2.18 | 1 | -1.15 | 1.12 |
| *tgs*1 (MSMEG_5242) | 1 | -2.59 | -2.25 | 1 | 1.94 | 1.64 |
| *tkt* (MSMEG_3103) | 1 | 1.59 | 1.59 | 1 | 1.14 | 1.01 |
| *tpx* (MSMEG_3479) | 1 | 1.77 | 2 | 1 | 1.28 | -1.01 |
| *trx*R (MSMEG_6933) | 1 | -1.5 | -1.83 | 1 | -1.49 | -1.04 |
| *ure*C (MSMEG_3625) | 1 | -1.16 | 1.38 | 1 | 1.11 | 1.01 |

| | | | | | | |
|---|---|---|---|---|---|---|
| *whi*B1 (MSMEG_1919) | 1 | -1.02 | -1.14 | 1 | -1.27 | 1.14 |
| *whi*B3 (MSMEG_1597) | 1 | -1.06 | 1.34 | 1 | -1.52 | 1.06 |
| *whi*B4 (MSMEG_6199) | 1 | 1.26 | 1.28 | 1 | -1.07 | 1.21 |
| *whi*B6 (MSMEG_0051) | 1 | 1.47 | 1.42 | 1 | -1.1 | 1.31 |
| *whi*B7 (MSMEG_1953) | 1 | 1.37 | -1.13 | 1 | -1.96 | -1.07 |
| *xer*C (MSMEG_2515) | 1 | 1.02 | 1.42 | 1 | 2.47 | 1.37 |

Supplementary table 4. List of mRNA count for selected M. smegmatis genes from Nanostring gene expression profile. RNA was isolated from WT, esx-3 mutant and mycobactin mutant M. smegmatis grown under low and high iron condition. Normalization was made by housekeeping gene sigA and known concentration of negative and positive control.

| **Function** | **Gene** | **Accession no.** | **mRNA count** | | | | | |
|---|---|---|---|---|---|---|---|---|
| | | | **Low iron** | | | **High iron** | | |
| | | | **WT** | **Δ*espG*$_3$** | **Δ*mbt*D** | **WT** | **Δ*espG*$_3$** | **Δ*mbt*D** |
| **Iron metabolism** | MSMEG_0788 | MSMEG_0788 | 1230.71 | 909.76 | 854.11 | 665.11 | 1111.93 | 748.39 |
| | *bfrA* | MSMEG_3564 | 628.35 | 1132.94 | 1289.04 | 5040.53 | 7920.03 | 5961.79 |
| | *eccA3* | MSMEG_0615 | 18807.78 | 22283.27 | 22929.57 | 405.19 | 197.44 | 377.65 |
| | *fxbA* | MSMEG_0014 | 27702.13 | 35615.95 | 39442.3 | 202.03 | 162.07 | 226.82 |
| | *ideR* | MSMEG_2750 | 16260.97 | 18934.35 | 17417.27 | 11069.65 | 12791.09 | 12122.77 |
| | *irtA* | MSMEG_6554 | 5580.89 | 2779.17 | 970.72 | 54.48 | 43.22 | 56.42 |
| | *lipE* | MSMEG_6575 | 713.74 | 764.04 | 752.2 | 1078.25 | 821.18 | 1204.33 |
| | *mbtB* | MSMEG_4515 | 52999.45 | 54017.3 | 58280.98 | 105.55 | 208.24 | 163.49 |
| | *mbtL* | MSMEG_2132 | 24827.68 | 25028.31 | 25772.39 | 196.35 | 185.65 | 237.18 |
| | *mmpS4* | MSMEG_0380 | 2730.58 | 4698.47 | 5738.17 | 1818.27 | 2357.44 | 2100.1 |
| **Zinc metabolism** | MSMEG_6069 | MSMEG_6069 | 55769.94 | 64656.14 | 57382.75 | 47265.93 | 48036.82 | 50875.59 |
| | *rpmB2* | MSMEG_6068 | 108769.39 | 158667.54 | 147889.71 | 120488.18 | 161161.66 | 133105.3 |
| ***BlaR*** (Possible sensor transducer) | MSMEG_3117 | MSMEG_3117 | 483.56 | 749.6 | 651.35 | 954.53 | 1155.15 | 903.83 |
| | MSMEG_6942 | MSMEG_6942 | 3234.56 | 3635.11 | 3627.59 | 3477.64 | 2767.05 | 3733.89 |
| | *atpA* | MSMEG_4938 | 4115.36 | 2020.38 | 4040.46 | 7962.02 | 3856.38 | 7313.5 |
| | *atpB* | MSMEG_4942 | 1045.08 | 626.2 | 1148.26 | 2373.28 | 1155.15 | 2154.21 |
| | *blaI* | MSMEG_3630 | 3020.16 | 4475.3 | 3963.77 | 1132.73 | 1554.93 | 1379.34 |
| | *blaR* | MSMEG_3631 | 5142.81 | 6460.23 | 6452.55 | 2658.17 | 3341.68 | 2643.55 |
| **Cell wall** | *aftA* | MSMEG_6386 | 1201.01 | 1314.1 | 1283.79 | 1375.62 | 1228.82 | 1469.15 |
| | *alr* | MSMEG_1575 | 774.07 | 1106.68 | 1020.1 | 946.59 | 1304.45 | 1208.94 |
| | *murA* | MSMEG_4932 | 302.57 | 186.42 | 303.61 | 475.56 | 217.08 | 408.74 |
| | *murE* | MSMEG_4232 | 922.57 | 1298.35 | 1272.23 | 1441.45 | 1349.64 | 1482.96 |
| | *murI* | MSMEG_4903 | 2344.48 | 2408.97 | 2475.12 | 5507.02 | 4001.76 | 5587.6 |
| | *ponA1* | MSMEG_6900 | 8086.86 | 7161.26 | 6483.02 | 12252.32 | 7755.99 | 11413.53 |
| | *ripA* | MSMEG_3145 | 153.14 | 354.45 | 562.05 | 136.2 | 195.47 | 160.04 |
| **Cell cycle** | *ftsZ* | MSMEG_4222 | 12471.39 | 12986.1 | 13481.86 | 7295.78 | 9881.61 | 8083.77 |
| | *parA* | MSMEG_6939 | 2425.22 | 2632.14 | 2754.57 | 2214.38 | 2002.84 | 2230.2 |
| | *pknA* | MSMEG_0030 | 732.3 | 975.4 | 1014.84 | 678.73 | 732.77 | 687.37 |
| | *ponA2* | MSMEG_6201 | 4473.62 | 4514.68 | 4754.84 | 5609.17 | 4526.29 | 5769.52 |
| | *rodA* | MSMEG_0032 | 3036.87 | 3302.98 | 3861.87 | 2708.11 | 3034.23 | 2887.64 |

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | *scpA* | MSMEG_3742 | 1743.04 | 1920.61 | 2157.85 | 1904.53 | 2506.75 | 1925.09 |
| | *sigA* | MSMEG_2758 | 17482.4 | 17482.4 | 17482.4 | 17482.4 | 17482.4 | 17482.4 |
| | *sigB* | MSMEG_2752 | 18507.07 | 30935.86 | 23700.68 | 36010.14 | 58149.27 | 42601.84 |
| | *sigF* | MSMEG_1804 | 2726.87 | 2964.28 | 3475.26 | 5333.36 | 7946.55 | 4612.39 |
| | *xerC* | MSMEG_2515 | 523.47 | 532.99 | 745.9 | 228.13 | 563.82 | 312.02 |
| **Red-ox** | *ahpC* | MSMEG_4891 | 26683.96 | 32983.81 | 22542.96 | 110998.44 | 38538.29 | 108497.1 |
| | *dosR* | MSMEG_5245 | 1893.4 | 1318.04 | 1460.28 | 28908.45 | 44218.75 | 42916.17 |
| | *dosR* | MSMEG_3944 | 123.44 | 118.15 | 162.84 | 396.11 | 904.67 | 744.94 |
| | *furA* | MSMEG_3460 | 3304.17 | 2405.03 | 3096.01 | 2329.02 | 2433.08 | 2235.96 |
| | *furA* | MSMEG_6383 | 2119.87 | 4296.76 | 2922.66 | 6348.05 | 1387.94 | 6745.88 |
| | *katG* | MSMEG_6384 | 10009.04 | 16464.99 | 13476.61 | 47814.14 | 7533.01 | 40507.5 |
| | *katG* | MSMEG_3729 | 74.25 | 65.64 | 104.01 | 188.41 | 164.04 | 163.49 |
| | *katG* | MSMEG_3461 | 7567.11 | 6621.7 | 8916.12 | 4534.32 | 5348.45 | 4657.29 |
| | *mshA* | MSMEG_0933 | 44749.23 | 24373.23 | 21981.96 | 38067.89 | 32961.96 | 36947.46 |
| | *mshC* | MSMEG_4189 | 4579.43 | 5234.09 | 3857.66 | 26098.19 | 15323.38 | 24624.35 |
| | *oxyS* | MSMEG_0156 | 232.96 | 330.82 | 266.84 | 62.42 | 67.78 | 81.75 |
| | *sodA* | MSMEG_6427 | 47622.75 | 57720.68 | 64034.91 | 60866.64 | 45846.37 | 65160.63 |
| | *sodA* | MSMEG_6636 | 4596.14 | 5858.97 | 6760.37 | 3548.01 | 3195.32 | 3700.5 |
| | *tpx* | MSMEG_3479 | 2882.8 | 5108.06 | 5753.93 | 3275.61 | 4194.28 | 3235.35 |
| | *trxR* | MSMEG_6933 | 15079.45 | 10033.64 | 8230.1 | 70977.21 | 47534.89 | 68049.42 |
| | *whiB1* | MSMEG_1919 | 106032.31 | 104449.38 | 93329.83 | 98528.2 | 77686.58 | 112339.2 |
| | *whiB3* | MSMEG_1597 | 23764.04 | 22352.85 | 31945.49 | 4398.12 | 2887.87 | 4651.53 |
| | *whiB4* | MSMEG_6199 | 373.11 | 468.67 | 476.96 | 545.93 | 508.81 | 663.19 |
| | *whiB6* | MSMEG_0051 | 199.55 | 294.06 | 282.6 | 122.58 | 111 | 161.19 |
| | *whiB7* | MSMEG_1953 | 20886.81 | 28715.93 | 18428.96 | 12659.79 | 6473.14 | 11798.09 |
| | *dlaT* | MSMEG_4283 | 12411.06 | 15357 | 17758.7 | 13660.86 | 15102.37 | 14715.66 |
| **pH** | MSMEG_0409 | MSMEG_0409 | 7527.2 | 11287.35 | 14774.05 | 3846.51 | 3368.2 | 4310.73 |
| | MSMEG_4279 | MSMEG_4279 | 177.27 | 190.35 | 192.25 | 180.46 | 177.79 | 192.28 |
| | MSMEG_5018 | MSMEG_5018 | 946.7 | 1136.87 | 1207.1 | 1213.31 | 1006.82 | 1311.41 |
| **Lipids** | MSMEG_3519 | MSMEG_3519 | 206.05 | 234.99 | 259.49 | 125.98 | 130.64 | 117.44 |
| | MSMEG_5900 | MSMEG_5900 | 1135.11 | 1298.35 | 1640.98 | 760.45 | 892.88 | 741.48 |
| | MSMEG_5994 | MSMEG_5994 | 111.38 | 143.09 | 132.37 | 74.91 | 75.63 | 71.38 |
| | MSMEG_6014 | MSMEG_6014 | 139.22 | 168.04 | 138.67 | 64.69 | 56.97 | 47.21 |
| | *fadE34* | MSMEG_6041 | 91.89 | 120.78 | 119.76 | 48.8 | 59.92 | 74.84 |
| | *hsaA* | MSMEG_6038 | 301.64 | 442.41 | 427.58 | 148.68 | 181.72 | 150.83 |
| | *kstD* | MSMEG_5941 | 295.15 | 442.41 | 511.62 | 86.26 | 187.61 | 131.26 |
| | *kstR* | MSMEG_6042 | 1303.11 | 1596.35 | 1625.22 | 2724 | 2683.56 | 2891.09 |
| **DNA-synthesis and repair** | *dinX* | MSMEG_3172 | 355.48 | 315.07 | 253.19 | 726.4 | 524.53 | 642.46 |
| | *dnaA* | MSMEG_6947 | 1904.54 | 2363.02 | 2823.91 | 1095.27 | 901.72 | 1002.84 |
| | *dnaN* | MSMEG_0001 | 27747.61 | 20946.85 | 16711.29 | 23712.42 | 22335.79 | 25460.25 |
| | *parB* | MSMEG_6938 | 746.22 | 918.95 | 952.86 | 715.05 | 725.9 | 736.88 |
| | *recB* | MSMEG_1327 | 6354.03 | 4238.99 | 2642.16 | 6578.46 | 5370.06 | 7381.43 |
| | *ssb* | MSMEG_4701 | 428.8 | 349.2 | 196.46 | 2997.53 | 2601.05 | 3351.64 |
| **Energy/Carbon** | *acn* | MSMEG_3143 | 3445.25 | 3388.31 | 3196.86 | 6897.39 | 7992.71 | 7592.13 |

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| **flux** | *ctaE* | MSMEG_4260 | 615.36 | 1613.42 | 1399.35 | 2003.27 | 3548.93 | 2364.91 |
| | *cydD* | MSMEG_3233 | 326.7 | 483.11 | 451.74 | 237.21 | 304.5 | 284.39 |
| | *devB* | MSMEG_3099 | 798.2 | 1069.92 | 1066.32 | 2230.27 | 1556.89 | 2107.01 |
| | *dlaT* | MSMEG_4283 | 12411.06 | 15357 | 17758.7 | 13660.86 | 15102.37 | 14715.66 |
| | *fabH* | MSMEG_3953 | 532.75 | 459.48 | 654.5 | 356.39 | 457.74 | 421.4 |
| | *fadB* | MSMEG_5720 | 2624.77 | 2200.23 | 2116.88 | 2245.03 | 3222.82 | 2237.11 |
| | *fadE5* | MSMEG_0406 | 9271.17 | 7668 | 7926.49 | 8047.15 | 10833.43 | 8911.6 |
| | *fas* | MSMEG_4757 | 3039.65 | 2121.47 | 1466.58 | 2407.33 | 2140.36 | 2254.38 |
| | *glcB* | MSMEG_3640 | 5737.75 | 5719.82 | 5534.36 | 5740.83 | 6469.22 | 5492.03 |
| | *glnA1* | MSMEG_4290 | 56567.21 | 62005.62 | 75821.16 | 23306.09 | 27381.7 | 25611.07 |
| | *glpX* | MSMEG_5239 | 3142.67 | 4116.9 | 4190.69 | 2453.87 | 2668.82 | 2637.79 |
| | *kgd* | MSMEG_5049 | 16968.21 | 19988.52 | 18629.62 | 24094.91 | 21288.69 | 26373.28 |
| | *nuoA* | MSMEG_2063 | 428.8 | 653.77 | 586.21 | 698.02 | 1048.08 | 732.27 |
| | *nuoN* | MSMEG_2050 | 52.9 | 81.39 | 71.44 | 55.61 | 55.99 | 52.96 |
| | *pckA* | MSMEG_0255 | 9628.5 | 9614.86 | 9291.17 | 5904.27 | 5655.9 | 5954.89 |
| | *pfkB* | MSMEG_3947 | 349.91 | 340.01 | 299.41 | 499.4 | 994.06 | 713.85 |
| | *phoT* | MSMEG_5779 | 1632.59 | 1597.66 | 1810.12 | 1375.62 | 1397.77 | 1382.8 |
| | *pks2* | MSMEG_4727 | 2716.66 | 4722.1 | 5193.98 | 978.37 | 1120.77 | 966 |
| | *ppa* | MSMEG_6114 | 11285.23 | 16266.76 | 18366.98 | 11470.31 | 10647.78 | 11834.93 |
| | *ppgK* | MSMEG_2760 | 671.97 | 1020.04 | 1186.08 | 351.85 | 337.9 | 358.08 |
| | *prpC* | MSMEG_6647 | 465.93 | 622.26 | 604.07 | 248.56 | 246.55 | 230.27 |
| | *pstC* | MSMEG_5781 | 2518.04 | 2844.81 | 2966.79 | 1937.44 | 2052.94 | 1934.3 |
| | *sdhC* | MSMEG_1672 | 401.88 | 728.6 | 751.15 | 1254.17 | 1684.59 | 1456.48 |
| | *tgs1* | MSMEG_5242 | 193.98 | 74.83 | 86.15 | 2157.63 | 4195.27 | 3533.55 |
| | *tkt* | MSMEG_3103 | 3583.54 | 5710.63 | 5709.8 | 3932.77 | 4474.23 | 3977.98 |
| **Two component systems** | *MprB* | MSMEG_5487 | 815.83 | 1264.22 | 1441.37 | 1055.55 | 1534.3 | 1147.92 |
| | *pprB* | MSMEG_5663 | 892.87 | 1148.69 | 1302.7 | 889.84 | 1106.03 | 919.94 |
| | *senx3* | MSMEG_0936 | 1821.93 | 1944.24 | 2153.65 | 1325.68 | 1279.89 | 1378.19 |
| **Others** | MSMEG_0219 | MSMEG_0219 | 16529.2 | 10911.9 | 6430.49 | 12815.28 | 11548.52 | 13112.95 |
| | MSMEG_1906 | MSMEG_1906 | 27.84 | 35.45 | 29.42 | 36.32 | 28.49 | 31.09 |
| | MSMEG_3628 | MSMEG_3628 | 7972.7 | 7797.96 | 7165.88 | 2989.59 | 3718.87 | 3345.88 |
| | *catA* | MSMEG_1911 | 100.24 | 141.78 | 129.22 | 170.25 | 249.5 | 126.65 |
| | *groS* | MSMEG_1582 | 112367.78 | 123512.38 | 134062.22 | 120515.42 | 135283.81 | 130019.6 |
| | *ureC* | MSMEG_3625 | 3207.64 | 2768.67 | 4416.56 | 1544.73 | 1707.18 | 1557.8 |