Metabolomics approach to understand mycobacterial adaptation to stress and identify disease markers

A Thesis

Submitted to the University of Hyderabad for the award of a Ph.D degree

in Department of Biochemistry, School of Life Sciences

By

Arshad Rizvi

(Regd No. 12LBPH07)

Supervisor

Prof. Sharmistha Banerjee
Department of Biochemistry
School of Life Sciences
University of Hyderabad-500046
India

University of Hyderabad
Hyderabad- 500 046, India

CERTIFICATE

This is to certify that this thesis entitled ***"Metabolomics approach to understand mycobacterial adaptation to stress and identify disease markers"*** submitted by **Mr. Arshad Rizvi** bearing registration number **12LBPH07** in partial fulfillment of the requirements for award of Doctor of Philosophy in the Department of Biochemistry, School of Life Sciences, is a bonafide work carried out by him under my supervision and guidance.

This thesis is free from plagiarism and has not been submitted previously in part or in full to this or any other University or Institution for award of any degree or diploma.

Parts of this thesis have been presented in the following **conferences:**

1. Oral presentation at symposium on "Accelerating biology 2018". 09th -11th January 2018 at C-DAC, Pune. Received 2nd prize as **'Young Speaker Award'**
2. International Conference on Innovations in Pharma and Biopharma Industry: Challenges and Opportunities for Academy and Industry. 20th – 22nd December 2017 at University of Hyderabad, Hyderabad.

Publications:

Published:

1. Ganji R, Dhali S, **Rizvi A**, Rapole S, Banerjee S. Understanding HIV-Mycobacteria synergism through comparative proteomics of intra-phagosomal mycobacteria during mono- and HIV co-infection. Sci Rep. 2016 Feb 26;6:22060.
2. Ganji R, Dhali S, **Rizvi A**, Sankati S, Vemula MH, Mahajan G, Rapole S, Banerjee S. Proteomics approach to understand reduced clearance of mycobacteria and high viral titers during HIV-mycobacteria co-infection. Cell Microbiol. 2016 Mar;18(3):355-68.
3. Pratima K, Goud N, Mohammed A, Ramamoorthy G, Ananthathatmula R, Doble M, **Rizvi A**, Banerjee S, Alvala R, Alvala M. An efficient and facile Green synthesis of Bisindole Methanes as potential Mtb FtsZ inhibitors. Chem Biol Drug Des. 2018 Jul 13.
4. Ramaprasad EV, **Rizvi A**, Banerjee S, Sasikala C, Ramana CV. Mycobacterium oryzae sp. nov., a scotochromogenic, rapidly growing species is able to infect human macrophage cell line. Int J Syst Evol Microbiol. 2016 Nov;66(11):4530-4536.

Under communication:

1. **Rizvi A**, Yousf S, Balakrishnan K, Chugh J, Patel A , Mande S, Banerjee S*. Early adaptive responses of Mycobacterium smegmatis suggest involvement of methylamine pathway earlier reported in obligate methylotrophs to accumulate osmolytes to overcome microbicidal environment stresses.
2. **Rizvi A** , Shankar J A, Chatterjee A, More TH, Bose T, Dutta A, Balakrishnan K, Rapole S, Mande S*, Banerjee S*, Mande SC*. Adaptation of *Mycobacterium tuberculosis* to different stress conditions reveals rewiring of metabolic networks.
3. **Rizvi A** , Matzapetakis M, Coelho AV , Fernandes P, Sivangala R, Puchades L, Pineda-Lucena A Patel A, Sitaramaraju P, Kumar R, Gaddam S, Velluri V, Mande S, Banerjee S*. Identification of tuberculosis disease indicators using NMR-based metabolomics.

Further, the student has passed the following courses towards fulfilment of coursework requirement for Ph.D.

Course code	Name	Pass/Fail
BC 801	Seminar 1	Pass
BC 802	Research Ethics and Management	Pass
BC 803	Biostatistics	Pass
BC 804	Analytical techniques	Pass
BC 805	Lab work	Pass

Supervisor **Head of Department** **Dean of School**

University of Hyderabad
Hyderabad- 500 046, India

DECLARATION

I, **Arshad Rizvi**, hereby declare that this thesis entitled ***"Metabolomics approach to understand mycobacterial adaptation to stress and identify disease markers"*** submitted by me under the guidance and supervision of **Professor Sharmistha Banerjee**, is an original and independent research work. I also declare that it has not been submitted previously in part or in full to this University or any other University or Institution for the award of any degree or diploma.

Date:

Signature of the student

Name: Arshad Rizvi

Regd No. 12LBPH07

Signature of the Supervisor

TABLE OF CONTENTS

ACKNOWLEDGMENT

I would like to express my deepest gratitude and sincere appreciation to my supervisor **Prof**. **Sharmistha Banerjee.** I feel privileged to have carried out my research under her guidance. She is kind hearted, admirable, honest, inspiring teacher and an ideal mentor. Her positive outlook and confidence in my research inspired me and elevated my confidence. She has given me every opportunity to grow and develop in my academic endeavor and has pushed me to better myself. This made me to identify my own strength and drawbacks, and particularly boosted my self confidence. Her approach to supervision has allowed me to become a confident and independent scientist. I am blessed to be one of those who had an opportunity to work with her.

I am always thankful to my doctoral committee members: **Dr. Ravi Gutti** and **Dr. Prakash P Prabhu** for their critical comments and invaluable suggestions on my PhD work.

I would like to thank Heads of Department of Biochemistry (**Prof. Mrinal Kanti Bhattacharyya**, Prof. N. Siva Kumar, Prof. O.H. Setty, and Prof. K.V.A Ramaiah) and Deans of School of Life Sciences (**Prof. K.V.A Ramaiah**, Prof. P. Reddanna, Prof. Aparna Dutta Gupta, Prof. M. Ramanadham, Prof. A.S Raghavendra and Prof. R.P. Sharma) for allowing me to use all the central facilities of department and school.

I thank all the non-teaching staff of Dept of Biochemistry, especially **Mr. Chary**, for their help with paper work related to administrative processes. I thank **Dr. Ravindra Kumar** and **Mr. Fahim**, Health Centre, UH for helping us to collect healthy control samples and non-specific infection controls. I would like to thank all the faculty members of school of life sciences for their timely help.

I would also like to thank our collaborators **Dr. Sumanlatha Gaddam,** of MHRC for obtaining the patient samples. I thank **Ramya Sivangala** and volunteers, clinicians, and non-clinical staff of MHRC for obtaining the patient samples.

I would also like to specially mention **Dr. Shekhar Mande** for his moral support, inspiration and encouragement during my tenure.

During My PhD, I got the opportunity to work at Instituto de Tecnologia Quimica e Biologica Antonio Xavier da (ITQB), Universidade Nova de Lisboa, Portugal. It was a privilege to work under the guidance of **Dr. Manolis Matzapetakis**. Right from training on how to work with NMR and analyzing NMR spectra, it was an experience made successful with the help of him. I express my gratitude to **Dr. Manolis** for investing so much of his time to discuss experiments and encouraging me to do experiments in his lab. My stay at Lisboa was made comfortable and pleasant, thanks to the company of **Anaisa, Tiago, Ulric, Imran, Dr. Ana V Coelho and Dr. Pedro Fernandes.**

I would also like to thank **Dr. Shekhar Mande, Dr. Srikanth, Dr. Ankita and Tushar** from NCCS, Pune for helping me with the *M. tb* metabolomic data acquisition, identification analyses. I thank **Dr. Jeetender Chug** and **Saleem** from IISER, Pune for timely help with *M. smegmatis* metabolomics data acquisition and identification analyses. I also would like to thank **Dr. Anant B Patel** from CCMB, Hyderabad for helping me with sera sample and *M. smegmatis* acquisition at his lab. I would also like to thank **Dr. Sharmila Mande, Arvind** and **Tungadri** from TCS, Pune for helping in metabolomics data analysis.

I also thank my Ex- lab mates **Dr. Ronald, Dr. Atoshi, Dr. Rakesh Ganji, Dr. Harika, Swetha, Madhuri** for making my stay in the lab wonderful. I would like to thank project students, Manish, Haider, Sourab, Syed Haque, Veeda Narharain, for their support.

I am thankful to my present colleagues **Dr. Krishnaveni Mohareer, Kiran, Kannan, Suman, Jayashankar, Kumar, Govind, Sriram and Harish**. I thank **Mahesh** and **Srikanth** for keeping the lab and culture room clean to the extent possible. I relish all those lab parties we had together. A special thanks to **Dr. Krishnaveni Mohareer** for helping me in my thesis writing.

I thank **Mr. Prasenjit Banerjee** for his moral support and encouragement during my tenure. I feel privileged to be associated with him during my tenure He is an inspiring human great advisor and motivator. Whenever I am feeling low I talk to him to get motivation tonic.

I have been fortunate to make new friends during my tenure: **Ali, Rakesh, Arif, Sohail, Vanaja, Nisha, Fareed, Ritika, Amit, Majid, Neetu, Shamsul, Shabbir, Mujahid** and **Gopesh**. They literally supported me all through my ups and downs during my Ph.D. I also thank all my SLS friends.

I would like to acknowledge the financial support provided by CSIR for my research fellowships. I also thank all the funding bodies (CSIR, UGC, ICMR, DBT, CREBB, DST, FIST, UPE, and PURSE) for their financial assistance to the department and School. I also acknowledge EMBO for providing me financial support during my tenure in ITQB, Lisboa, Portugal.

I would like to thank my school and college teachers **Dr. Awasthi, Simon, Andrew**, **Dr. Vijay Neema** and **Neeraj** for encouragement during my school and college life.

I would like to thank my childhood friends **Major Amit Yadav, K.K Solanki, Manoj and Shiron Haider** for their constant support and encouragement.

I would like to thank my family for all the love and support. The emotional support and affection of my brothers **Imran, Farhan** and **Rehan** is highly commendable. Their love and care always made me work happily amidst all negative results. Especially my wife Tarannum, kids: **Humaira** and **Abdullah**, my parents, parents-in-law, brothers-in-law, **Danish** and **Furrukh** who always support me and pray for my success in life and hereafter.

A special thanks to my wife **Tarannum** for her patience, encouragement & support. She has been the rhythm of my life.

Arshad Rizvi

ABBREVIATIONS

AEC	Adenylate Energy Charge
AMP	Adenosine monophosphate
ANOVA	Analysis of variance
AUC	Areas under the curve
DCR	2,4-dienoyl-coA reductase
DSS	4,4-dimethyl-4-silapentane-1-sulfonic acid
DST	Drug Susceptibility Test
E. coli	Escherichia coli
ELISA	Enzyme-Linked Immunosorbent Assay
ESAT6	6 kDa early secretory antigenic target
ESX	Type VII secretion systems
FC	Fold change
FDR	False Discovery Test
FP	Forward Primer
GO	Gene Ontology
H37Rv	*M. tuberculosis*
HIV	Human Immunodeficiency virus
HMDB	Human Metabolome Database
INH	Isoniazid
KO	Knock out
LC-MRM/MS	Liquid chromatography-mass spectrometric multiple reaction monitoring based analysis
LC-MS/MS	Liquid Chromatography Tandem Mass Spectrometry
M. bovis	*Mycobacterium bovis*
M. kansasii	*Mycobacterium kansasii*
M. leprae	*Mycobacterium leprae*
M. marinum	*Mycobacterium marinum*
M. phlei	*Mycobacterium phlei*
M. smegmatis	*Mycobacterium smegmatis*
M. tb	*Mycobacterium tuberculosis*
MDR	Multi-Drug Resistance
mL	Milli Litre
mM	Milli Molar
NMR	Nuclear Magnetic Resonance Spectroscopy
OADC	Oleic acid Albumin Dextrose Catalase
^{o}C	Degree Celsius
OD	Optical density
PC	Principal Components
PCA	Principal Component Analysis
PCR	Polymerase chain reaction
PDB	Protein databank
PLS-DA	Partial Least Squares - Discriminant Analysis

PPD	Purified protein derivative
RIF	Rifampicin
RNA	Ribonucleic acid
RNS	Reactive Nitrogen Species
RNTCP	Revised National Tuberculosis Control Programme
ROC	Receiver Operating Characteristic Curve
ROI	Reactive oxygen Intermediates
ROS	Reactive oxygen species
RP	Reverse Primer
RT	Room temperature
SMART	Simple Modular Architecture Research Tool
STREP	Streptomycin
TB	Tuberculosis
TCA cycle	Tricarboxylic Acid Cycle
TSP	Trimethylsilylpropanoic acid
TST	Tuberculin Skin Test
VIP	Variable Importance in Projection
WHO	World health organization
XDR	Extensive Drug Resistant
ZN	Ziehl-Neelsen staining
μg	micro gram
μL	micro litre
μm	micro meter
μM	micro Molar

LIST FIGURES AND TABLES

Chapter 1

Introduction

1. Introduction

Ever since Robert Koch discovered *Mycobacterium tuberculosis* (*M. tb*) as the causative agent of tuberculosis (TB) in 1882, TB remains one of the world's most widespread and fatal diseases. It continues to be a major cause of human affliction, causing around 10.8 million new infections and around 1.8 million deaths in 2016 (WHO, 2016). TB is caused by inhalation of *M. tb* as airborne droplets from air that is expelled from a person with active disease (Leemans, Juffermans *et al*. 2001; Wolf, Linas *et al*. 2007). The bacteria reach the respiratory tract airways of a new host, where they are engulfed by alveolar macrophages and encounter various anti-microbial attacks, including reactive oxygen species/ reactive nitrogen species, low pH, phagosome-lysosome fusion and subsequent attack by lytic enzymes and release of antimicrobial peptides (Haas 2007). The infected host macrophages induce immune response, culminating in granuloma formation that attempts to contain infected cells by inducing the recruitment of monocytes from blood (Tsai, Chakravarty *et al*. 2006; Wolf, Linas *et al*. 2007). However, if the host is in an immune compromised state either due to immuno-suppressive drugs, HIV co-infection, age or malnutrition, the containment fails (Kaufmann 2001; Russell 2001; Russell 2007) leading to another cycle of infection, followed by development of cough facilitating expulsion of aerosol drops containing bacilli to spread the infection.

1.1 *Mycobacterium tuberculosis*, the TB causing mycobacteria

TB or TB-like disease in human is caused by many mycobacterial species (order *Actinomycetales* under the family *Mycobacteriaceae*) which includes members of the *M. tb* complex such *M. tb*, *M. bovis*, *M. microti*, *M. africanum*, *M. ulcerans* and *M. canetti*, which are genetically related and can cause TB in humans (Huard *et al.*, 2006). However, the most predominant species is *M. tb. M. tb* are aerobic, obligate, intracellular, non-sporulating, and immotile rod shaped bacteria with 1 to 5 μm length and 0.3 to 0.6 μm width. They have waxy and hydrophobic cell wall that is thicker than most of other bacteria. The outer mycolic acid layer is connected to peptidoglycan through branched arabinogalactan, mycolyl-arabinogalactan-peptidoglycan complex, which besides being indispensable for the viability and pathogenicity of *M. tb*, is also implicated in antibiotic resistance (Sakamoto 2012). Whole genome sequencing of *M. tb* (H37Rv) revealed high GC content (61 to 71%) in a 4.41 Mb long genome that codes for

more than 4000 genes (Cole *et al.*, 1998);(Kapopoulou *et al.*, 2011). A large fraction of the genome (~40%) codes for proteins involved in lipid metabolism that may mostly contribute to the biogenesis of lipid rich membrane of mycobacteria and makes it impermeable to Gram stain. Mycobacteria are "acid fast" bacteria. This is on the basis of Ziehl-Neelsen staining, where carbolfuchsin, a basic dye, stains mycobacteria pink that cannot be decolorized with acid-alcohol, therefore also referred to as acid fast bacilli (AFB).

On the basis of the generation time, mycobacteria have been categorized into fast and slow growing mycobacteria. Fast growers (generation time 2-4 hours) are typically avirulent including *M. smegmatis, M. abscessus* while slow growers (generation time 20-24 hours) can cause human as well as animal diseases, such as, *M. leprae, M. bovis, M. tb* (Hartmans *et al*., 2006). Some of the recent evidences revealed that an ancestral lineage of fast growing mycobacteria gave rise to various lineages of slow growing mycobacteria.

1.2 Tuberculosis, the disease

Tuberculosis is an aerosolic contagious disease that primarily affects lungs. Lung TB is also called pulmonary TB; however, besides lungs TB can attack several extra-pulmonary regions of the body like bones, nervous tissue, spinal cord, lymph nodes etc. While pulmonary TB is relatively easy to diagnose due to symptoms of persistent cough, night sweats along with low grade persistent fever, symptoms of extra-pulmonary TB are varied, diffused, making early detection a real challenge.

1.2.1 Stages of TB infection

The pathogenesis of pulmonary tuberculosis can be categorized into four stages (Turner, Basaraba and Orme, 2003) as schematically illustrated in **Figure 1**.

Stage 1

Once *M. tb* is inhaled as airborne droplets from the atmosphere, it travels into respiratory tract, where they are engulfed by alveolar macrophages and exposed to various kinds of microbicidal stresses. This is the most critical stage for the successful establishment of early infection as most of the bacteria are cleared by the microbicidal stress of the host cell.

Stage 2

M. tb that survive primary intracellular defense replicate. This triggers the recruitment of monocytes such as macrophages and DCs to the site to limit infection by pro-inflammatory chemokines and cytokines. *M. tb* engulfed by naive or resting macrophages survive. Thus, at this stage active disease is evident by tissue damaged by mycobacterial growth.

Stage 3

After two to three weeks of infection, *M. tb* infected macrophages and DCs that prime T cells towards Th1 arm of the immune response, secrete pro-inflammatory cytokines such as IFN-γ and TNF-α, which further activate macrophages. At this stage, activated macrophages eliminate intracellular pathogens by T cell-mediated immunity and the disease becomes stationary.

Stage 4

The outcome of the infection depends upon the growth of mycobacteria, formation of granuloma and the initiation of host adaptive immune response, which is in turn regulated by the interaction of T cells with infected macrophages and pro-inflammatory cytokines such as IFN-γ and TNF-α. The progression of infection depends upon the response of innate and adaptive immune systems. The inability of innate immune system to clear bacteria leads to activation of adaptive immune system. Further, this may lead to latent infection for life time if bacteria sustain human adaptive immune system. When there is a perturbation in the host immune status, such as suppressed or compromised immunity, reactivation occurs allowing replication of bacteria.

During the various stages of infection, *M. tb* is subjected to a variety of microbicidal stresses and for establishing an infection; it not only adapts successfully to these stresses but also overcomes the same. One of the objectives of my PhD work was to understand the metabolic adaptation of *M. tb* in microbicidal environment. The various stress conditions faced by *M. tb* upon engulfment by alveolar macrophages are discussed in the following section.

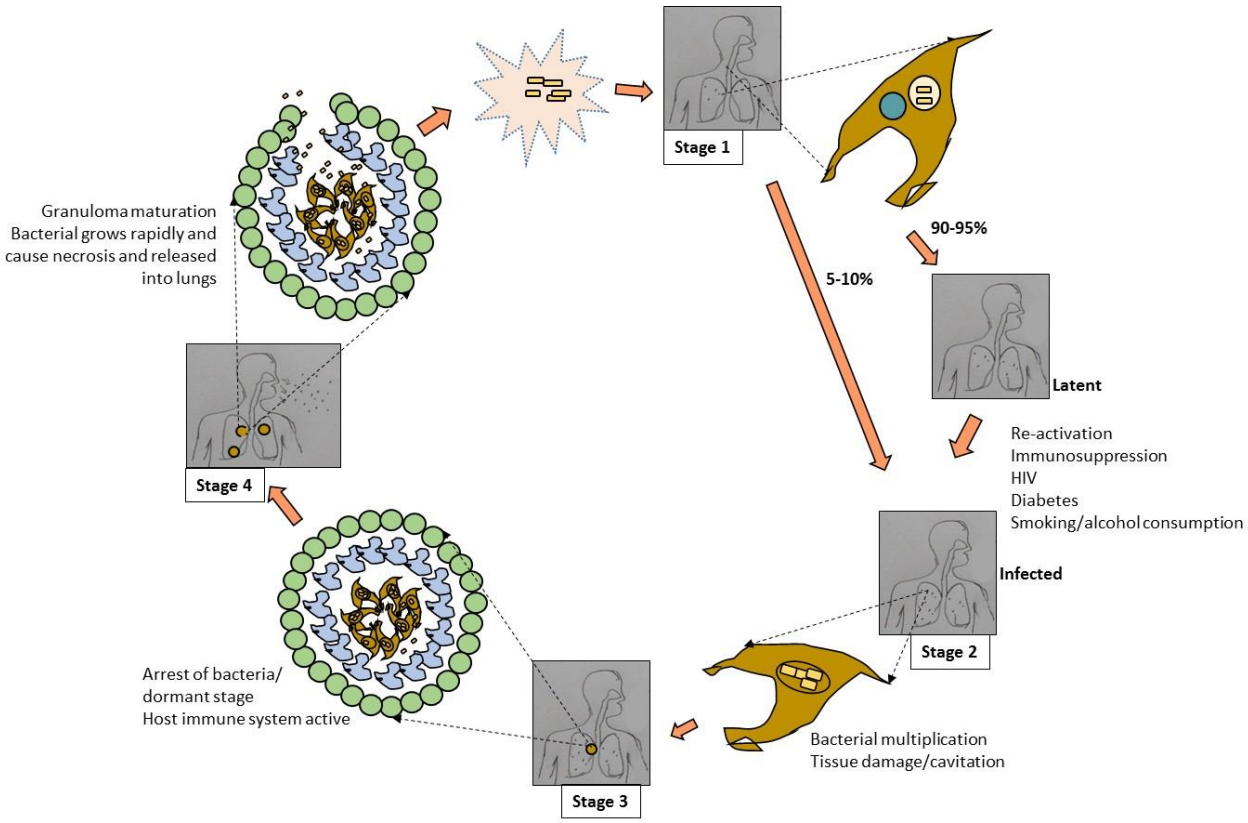

Figure 1: Schematic representation of the different stages of tuberculosis infection of human host
Tuberculosis is an infectious disease that spreads through air contaminated with mycobacteria by an actively infected TB patient. Mycobacteria have to cross several host barriers before they reach the lungs where they are engulfed by alveolar macrophages. In the alveolar macrophages, they encounter an arsenal of anti-microbial attacks (Stage 1). Mycobacteria that successfully evade host attack by a range of defense strategies, eventually multiply within the macrophages and cause tissue damage by necrosis (Stage 2). This causes release of several pro-inflammatory cytokines and subsequent infiltration of several types of immune cells, the activity of which causes clearance of most bacteria leading to a dormant stage (Stage 3). The development of organized structures of host immune cells, termed as granuloma, depends on host immune response. If mycobacteria circumvent the host immune responses, they multiply within the caseous necrotic matrix, cause necrosis and dissolve the caseous matrix with specific virulent proteins such as matrix metalloproteases and escape from the granuloma (Stage 4) into the extracellular space of the lungs and expelled out into the atmosphere.

1.3 Microbicidal stresses and mycobacteria

Pathogenic mycobacteria like *M. tb*, immediately upon engulfment by alveolar macrophages predominantly faces acidic, oxidative, nutrient starvation and iron deprivation in the hostile environment of phagolysosomes (as depicted in **Figure 2**), while *M. smegmatis*, being a soil bacteria may primarily face acidic, oxidative and nutrient starvation, but rarely iron deprivation. It is assumed that even if iron is limited in soil, it may not be completely absent or sequestered away from *M. smegmatis* (as can be the case for *M. tb*) hence iron deprivation may not be a primary environmental stress for *M. smegmatis*. Several studies have been carried out to understand mechanism of adaptation to these stresses, using both *M. tb* and *M. smegmatis* models, some of which are discussed below.

1.3.1 Acidic (low pH) stress

Following engulfment of bacilli by macrophage, lysosomes fuse with bacilli-laden phagosomes. The phagosome-lysosome fusion causes acidification of phagosome, activates microbicidal mechanisms by exposing bacteria to lysosomal hydrolases and produces reactive oxygen species (ROS) (Jackett et al., 1978; Stuehr and Nathan, 1989; Turk et al., 2001). It has been long established that *M. tb* can inhibit the fusion of phagosomes with lysosomes of resting macrophages, to survive inside in acidic environment of pH ~6.2 (MacMicking *et al*., 2003). Activation of the macrophage by IFN-γ relieves the phagosome–lysosome fusion block, thus, decreasing pH to 4.5-5.0 (MacMicking *et al*., 2003). Thus pathogenic bacteria such as *M. tb*(Armstrong and Hart, 1971; MacMicking et al., 2003), *M. leprae* (Sibley and Krahenbuhl, 1987), and *M. avium* (Sturgill-Koszycki *et al.*, 1994), have the ability to prevent maturation of phagosomes and thus infection progresses. *M. tb* can survive within the resting macrophages, but macrophages or DCs stimulated with the IFN-γ and TNF-α can actively clear the mycobacterium. Recently it was shown that a guanosine triphosphatase, host factor (LRG47) associated with phagosomes is induced by IFN-γ, which promotes maturation of phagosomes (MacMicking *et al*., 2003). Various studies have shown that increased expression of acid response genes and increased susceptibility of acid sensitive strains are important to maintain infection for example by expression of glutamine synthetase, aspartate transporter which would increase the intra-phagosomal pH by producing ammonia and thus arresting the maturation (Cheruvu et al., 2007;Rohde et al., 2007;Vandal et al., 2008). Several of these studies have used

M. smegmatis as a surrogate system for *M. tb* to understand the changes at RNA or protein levels (Wang and Marcotte, 2008), but changes at metabolite level have not been studied. Other factors that can increase the ability of bacilli to survive acid stress and to maintain pH homeostasis in bacteria are the secreted proteins (Vergne *et al.*, 2005); (Walburger *et al.*, 2004), MgtC, a potential Mg^{2+} transporter (Buchmeier *et al.*, 2000), cell wall glycolipids (Axelrod et al., 2008;Katti et al., 2008;Robinson et al., 2008;Vergne et al., 2005) and the outer membrane protein OmpATb (Raynaud *et al.*, 2002; Molle *et al.*, 2006).

1.3.2 Oxidative stress

One of the major defense mechanism employed by macrophages and phagocytic cells is the production of reactive oxygen species (ROS), Reactive nitrogen (RNI) species and peroxynitrite against invading bacteria (Akaki et al., 2000; Darrah et al., 2000; Karupiah et al., 2000; Nathan and Shiloh, 2000;Paziak-Domanska et al., 2000; Shiloh and Nathan, 2000; Yu et al., 1999). These strategies can directly damage DNA and protein lead to lethal effects which are irreversible. Thus *M. tb* has a diverse range of mechanisms to defend against these reactive species (Ehrt and Schnappinger, 2009). *M. tb* has a thick cell wall that harbors potent oxygen scavengers such as cyclo-propanated mycolic acids, lipoarabinomannan and phenolic glycolipid I (PGL-1) which enables bacilli to resist ROS. *M. tb* has a higher basal expression of ROS detoxification genes due to presence of a disrupted copy of oxyR gene. In *M. tb* complex, *M. smegmatis* and *M. avium*, oxyR gene is either inactivated or a pseudogene, whereas only in *M. leprae* it is unmutated (Sreevatsan *et al.*, 1997). Various *in-vitro* models were generated to understand the response of *M. tb* towards oxidative stress by generating reactive species using cumene hydroperoxide (Buchmeier *et al*., 2003; Springer *et al*., 2001), hydroperoxide (Gebhard *et al*., 2008). Hydroperoxide can easily cross membrane; its toxicity is due to generation of highly reactive hydroxyl radical by Fenton's reaction. Despite inactivation of oxyR gene in *M. tb*, it prevents the damage produced by reactive oxygen species by several other oxidative stress response genes that are preserved in the tubercle bacilli, such as superoxide dismutases (SodA and SodC*);(*Jackett *et al*., 1978; Piddington *et al*., 2001; Tullius *et al*., 2001), the peroxidase and peroxynitrite reductase (complex of AhpC, AhpD, SucB (DlaT), and Lpd)*;*(Bryk *et al*., 2002) and catalase (KatG*);(*Manca *et al*., 1999*);(*Ng *et al*., 2004). Besides this, *M. tb* has developed a redox-couple system that can monitor intracellular redox environment that acts as an antioxidant,

for instance, mycothiol, to resist oxidative damage (Buchmeier *et al*., 2006; Zhang *et al*., 1999). Thus, *M. tb* resists damage at protein and DNA level from reactive species by constitutive expression of genes that resist oxidative stress. As stated above, *M. smegmatis* has been quite often used as a model system for *M. tb* studies including proteomics and transcriptomics under oxidative stress, however metabolomics studies are limited.

1.3.3 Nutrient starvation

M. tb experiences a hostile environment upon phagocytosis by macrophage, thus its survival depends upon its ability to modulate cellular metabolism. There is increasing evidence that suggests that persistent *M. tb* suffers from nutrient deprivation in the lung lesions. Recent transcriptional studies in *M. tb* by genome wide microarray analyses suggest the nature of phagosomal environment such as low oxygen tension, oxidative-stress, nitrosative-stress and nutrient deprivation (Schnappinger *et al*. 2003,). It is observed that nutrient deprived *M. tb* from lung lesions displays altered morphology and staining properties as compared to *in-vitro* grown *M. tb* (Nyka, 1974). Similar properties were displayed when *M. tb* was allowed to grow in distilled water that remained in non-replicating state. Even after two years of nutrient starvation stress, when bacteria were resuspended in nutrient-rich media, they started replicating (Nyka, 1974). Nutrient starvation model was used by Loebel to study the metabolic changes of nutrient starvation on bacilli. In Loebel starvation model, bacilli were grown in nutrient rich media followed by bacillary harvest, washing and resuspension in phosphate buffered saline. It has been demonstrated that nutrient starvation results in decrease of respiration rate and gradual shut down (Loebel 1933). However, bacteria remain viable and upon transfer to nutrient rich media they would recover. To study the protein expression of nutrient starved bacilli, Betts *et al* performed 2D SDS-PAGE to compare expression profile of nutrient deprived to normal log phase cultures. Protein spots revealing differential expression were cut from the gel and identified either by MALDI, MS or LC-MS/MS. The result revealed high expression of HspX (α-crystallin homologue), Tig involved in protein export and DnaK (Betts *et al*., 2002). Further, they performed microarray analysis to elucidate adaptation of *M. tb* to nutrient starvation. Their study suggested that after 96 hours of nutrient starvation stress, expression of around 323 genes are significantly lower as compared to normal. The genes involved include prosthetic groups and carriers, DNA replication, repair, biosynthesis of cofactors, energy metabolism (15% of the 281

energy metabolism genes), post-translational modification (59% of the 51 genes of ribosomal protein synthesis), lipid biosynthesis (16% of the 63 lipid biosynthesis genes) and amino acid biosynthesis. Thus, this nutrient starvation model mimics *M. tb* persistent state.

1.3.4 Iron deprivation

The ability of bacilli to survive within macrophage is not only determined by its ability to evade immune response and oxidative stress, but also on its ability to acquire nutrients from the host. Growing body of evidence suggests that there is a preferential utilization of lipids as a carbon source by pathogenic *M. tb* (Keating *et al.*, 2005). In the metabolism of *M. tb*, iron plays an important role. Iron is the second most abundant metal and fourth most abundant element on earth. Among the micronutrients, iron is an essential micronutrient for all bacteria except for lactobacilli (Sritharan, 2000). In *M. tb*, iron acts as a co-factor for forty different enzymes including hydroxylases, ribonucleotide reductases, oxygenases and superoxide dismutases. Iron is involved in vital cellular functions such as haem biosynthesis, electron transport, DNA replication, pyrimidine biogenesis and amino acid biosynthesis (De Voss *et al.*, 1999)(Sritharan, 2000). Iron serves as a catalyst in biological redox reactions, as it has reduction potential (Fe3+ / Fe2+) of +110. Thus, it is important to maintain optimal levels of iron, as under high iron conditions, iron generates cellular toxicity by catalyzing Fentons' reaction in the presence of ROS. Upon infection, the host usually lowers the level of free iron below the required level for intracellular bacilli by holding iron by proteins such as transferrin, ferritin, lactoferrin (Ratledge and Dover, 2000; Bullen, Griffiths and Edmiston, 1999; Sritharan, 2000). They are regulated by hormone hepcidin, which initiates an innate immune response to further reduce availability of iron to cease *M. tb* growth (Drakesmith and Prentice, 2012). This is known as 'nutritional immunity' (Kochan *et al*, 1977). *M. tb* can acquire iron from the host by siderophore mediated iron acquisition machinery such as mycobactins (Mb), carboxymycobactins (CMb) and exochelins secreted by pathogenic and non-pathogenic mycobacteria respectively (De Voss et al., 1999; Sritharan, 2000). While iron deprivation can be a stress for intracellular mycobacteria, soil mycobacteria like *M. smegmatis* may rarely face the same.

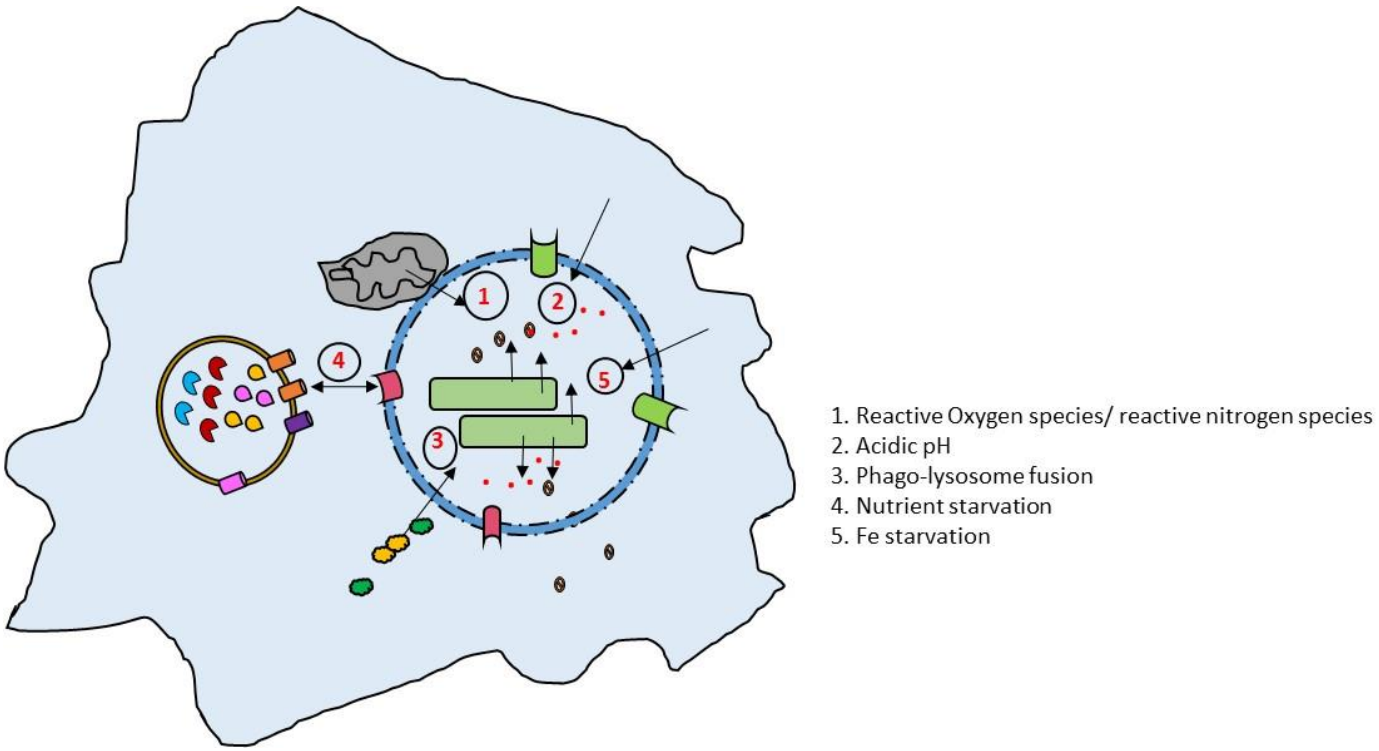

Figure 2: Microbicidal stress faced by mycobacteria in host macrophages: The mycobacteria that enter host macrophages face hostile conditions within the phagosome such as acidic pH, hypoxia, starvation, Fe deprivation, oxidative stress and so on. A successful mycobacterial infection overcomes these stress conditions by a variety of proteins and glycolipids and inhibits phago-lysosome fusion. Besides, several kinases and phosphatases dampen the host immune signaling mechanisms and help in successful establishment of infection.

1.4 Challenges of TB disease management

Recognizing the contagious nature of the disease, Hermann Brehmer started the concept of secluding the patients from healthy population leading to the concept of 'Sanatorium' with opening of the first TB sanatorium at Görbersdorf in 1854. The management of TB treatment started changing with the discovery of drugs like para-amino salicylic acid in 1943 (PAS) and thiosemicarbazone in 1945 though both were bacteriostatic, and disappointingly ineffective. The first bactericidal anti-mycobacterial drug used in 1944 was streptomycin, followed by isoniazid in 1952 and rifamycin in 1957 which led to the end of the 'Sanatorium' era and marked a beginning to a new era of TB disease management. Further TB disease management and control was tremendously supported by the commencement of anti-tuberculous treatment module including first and second line drugs**.** The first line drugs (primary agents) include isoniazid, ethambutol, rifampicin, pyrazinamide, and streptomycin and the second line drugs (retreatment agents) include ethanolamine, cycloserine, kanamycin, para-aminosalicylic acid (PAS) ofloxacin, ciprofloxacin and other fluoroquinolones (Kanabus *et al.,* 2016). This advancement had transformed TB disease from a fatal to a curable disease. However, TB treatment faces the problems of emergence of drug resistant strains, where strains are resistant to first line drugs [multi-drug resistance (MDR)] and sometimes resistant to second line drugs [extreme drug resistance (XDR)]. Dramatic outbreaks caused by multi-drug resistant strains, recently by XDR, extensive drug resistant strains; HIV epidemic; failure to identify asymptomatic yet infective cases, are possibly are some of the major reasons for this resurgence. Several efforts are being made to identify new targets and small molecule inhibitors (Chiarelli *et al*., 2016; Laughon and Nacy 2017; Pratima *et al*., 2018). Recently, bedaquiline (BDQ) had been approved for the treatment of TB; it is the first FDA-approved drug in the last four decades. It has bactericidal effect on *M. tb* and drug-resistant *M. tb*. But still, TB management requires novel medications, drug target identification by new strategies and vaccine along with new biomarkers to monitor disease progression (Global tuberculosis report) (2015).

Controlling TB through vaccination came into picture with Albert Calmette and Camille Guerin designing the first successful vaccine against TB in 1921, commonly known as Bacillus Calmette Guerin (BCG) vaccine. Though this vaccine effectively helps in preventing TB in children, it fails to do so in adults (Luca and Mihaescu 2013; Hawn, Day *et al*. 2014).

One of the main challenges is to detect TB early and accurately. As we understand, TB diagnosis still relies on medical history, X-ray of chest, Mantoux test, microscopy and culturing, which take considerable time. The sputum microscopy and culture method are still regarded as "gold standard" in TB diagnostics (Woods 2000). Getting pure mycobacterial samples from sputum itself is challenging and requires laborious processes. The much required drug sensitivity is assessed by calculating minimal inhibitory concentration (MIC) of different anti-TB drugs tested on cultures, taking at the least 3-4 weeks for the result. The window is sufficient for elevating the severity of the disease. Methods such as BACTEC MGIT 960 system, Bactec 460, Versa Trek system and MB/BacT (Caulfield and Wengenack, 2016) based on bacterial metabolism with radio-isotopic detection or fluorescent dyes in liquid culture are in use, which has tremendously reduced the delay in detection. However, these methods still need pure cultures, besides additional issues like radioactive waste disposal, high cost, sophisticated infrastructure and highly trained technicians. Diagnosing latent TB is still challenging and is not practiced in general. Most of the diagnostic methodologies do not distinguish between *M. tb* and other mycobacterial species.

In India, tuberculin skin test (TST) with purified protein derivative (PPD) is still the standard test for TB. Efficacy of the same is arguable as BCG vaccinated individual often cross reacts and may result in false positive, while immune compromised individuals, such as those co-infected with HIV, respond negatively to TST. Because of the fact that people with various backgrounds can have different degrees of immune response to TB, immunological tests, such as ELISA based assays for TB diagnosis, even with specific antigens such as ESAT-6, CFP-10, antigen-38, Hsp60, antigen-85 etc are not recommended by WHO as confirmatory (Andersen *et al.*, 2000). Another commonly used serological test, Quantiferon-TB gold, is a T-cell based assay that quantifies secreted IFN-γ in response to PPD and MPB64 patch test, which uses a 64-kDa antigen specific to *M. tb* complex. The molecular methods for TB diagnosis include PCR amplification of various markers *e.g* insertion element IS6110, 16S–23S rRNA spacer, *recA, rpoB*, and *gyrB* that allow rapid and sensitive detection. The 16S sequences of some species cannot be differentiated due to the presence of polymorphic sites in 16S rRNA gene. However, molecular assays such as line probe assay have been developed by researchers to perform rapid drug susceptibility test (DST) over conventional method by extracting DNA from mycobacterial

isolates or directly from clinical samples. The WHO recently endorsed molecular methods such as real-time PCR (GeneXpert system) as a diagnostic method for TB suspected of MDR-TB or HIV-associated TB (Boehme *et al.*, 2010);(WHO 2011). However, high cost makes it practically less feasible. Therefore continuous efforts to identify early disease indicators with different approaches are imperative.

Presently, although TB is curable, 8 million individuals develop active tuberculosis disease resulting in two million deaths per year, which is unacceptable (WHO 2012). Hence multiple approaches across the world have been initiated to identify new drug targets, new drug molecules and new disease indicators.

1.5 Omics' approaches to understanding cellular physiology to identify new drug targets

The whole-genome sequencing of *M. tb* in 1998 (Cole *et al*. 1998) led to a significant advancement in understanding the molecular biology of *M. tb.* Further, Sassetti and colleagues developed a random mutagenesis approach to identify a list of essential genes which are critical for mycobacterial physiology (Sassetti and Rubin, 2003). The essentiality of a gene can be used to explore it as a potential drug target. Genome-wide DNA microarrays, as performed in various studies, gave insights into the pattern of gene expression under various growth conditions and in identification of new targets (Waddell and Butcher, 2007). Many of these potential drug targets have been validated by gene knockdown techniques *in-vitro* and *in-vivo* (Mohr *et al.*, 2014). Genetic engineering studies including various techniques like transposon mutagenesis (Chen, 2010), site-directed mutagenesis (Rouse *et al.*, 1996), counter-selectable marker sacB (Barkan *et al*., 2011), recombineering systems (van Kessel and Hatfull, 2007) have immensely contributed to our understanding of the mechanisms implicated with virulence, resistance and new vaccines. Comparative genomics reveal that virulent mycobacteria have smaller genomes as compared to its non-virulent counterparts. Genomic analyses of *M. tb* clinical cohorts suggests that since its origin in Africa, *M. tb* has evolved to yield seven lineages of strains which can infect the human host (Yimer *et al.*, 2015). In *M. tb* about two thirds of the SNPs are non-synonymous and more than 50% fall in conserved positions of the coding regions, indicating phenotypic diversity of the pathogen is due to functional mutations (Otchere *et al.*, 2017). Genome-wide association studies (GWAS) along with standard methods identified 16 different drugs for which resistance has been

generated in *M. tb* as a result of mutation (Phelan *et al.*, 2016). Recently, a genomic study showed that across all lineages of *M. tb*, resistance to isoniazid (katG mutation encoding p.S315T) arises overwhelmingly before resistance to rifampicin (Dookie *et al.*, 2018). Despite the success, data generated by genomic studies alone is not sufficient to determine the molecular mechanisms of infection and persistence. This led researchers to elucidate functional genomics of *M. tb* to increase systems level understanding by using primarily mRNA analysis by microarrays or RNA sequencing approaches. For example, the transcriptomics study performed by Sherman and colleagues showed that the transcription factor, DosR/ DevR (Rv3133c) regulates ~50 genes of *M. tb* in dormancy (non-replicating) state on exposure to hypoxia (Sherman *et al.*, 2001). Another independent study from Tyagi and group demonstrated the differential regulation of dosR operon during hypoxia (Saini *et al.*, 2004). Her group further documented around 47 binding sites (including 24 new sites) which are temporally regulated (Chauhan *et al.*, 2011). ChIPSeq combined with expression data of *M. tb* induced with the same transcription factors revealed Rv0081 as a hub of transcriptional network of *M. tb* in hypoxia (Galagan *et al.*, 2013). However, some regulations are not at transcript level. For example, translation of persistence genes that exhibit a codon-bias for mycobacteria selectively occurs by 40 modified ribonucleosides in tRNA in response to hypoxia in BCG vaccine strain (Chionh *et al.*, 2016). Upon nitric oxide stress exposure, *M. tb* reacts promptly at the transcriptome level, however it takes a long time to observe changes at proteomics level as it is linked to protein degradation. Thus, suggesting that proteomics level data is important to elucidate complex mechanisms. Proteomics, besides identifying and quantifying proteins present at any given time in a sample, also can identify their post-translational modifications (PTMs), change in conformations or protein-protein interactions. Proteomics study using virulent and avirulent strains of *M. tb* identified 29 differential proteins associated with membrane such as SecF and three ABC-transporter proteins that were up regulated in the virulent strain of *M. tb*, suggesting virulence of the bacilli can be determined by its secretion and transport system (Målen *et al.*, 2011). Proteomics study performed with ancient and modern Beijing strains, explains that a higher mutation rate and drug resistance in modern Beijing strain could be a result of increased DNA damage by oxidation due to low level of SseA (Bespyatykh *et al.*, 2016). Proteomic data generated earlier in our laboratory revealed 92 differential proteins belonging to various

functional categories such as type VII (Esx) secretion systems, cation transporters and toxin–antitoxin (TA) modules. Proteomic studies have empowered researchers to gain in-depth knowledge on pathogenicity of *M. tb* by elucidating regulatory pathways such as that by PhoP, which regulates small non-coding RNA (Mcr7) and indirectly affects Twin Arginine Translocation (Tat), secretion of BlaC and the antigen 85 complex (Ag85) (Solans *et al.*, 2014). Proteomic studies with *M. tb* culture filtrate revealed that secretion of EsxG and EsxH co-dependently influence the secretion of PE/PPE family protein for example PE5 (Tufariello *et al.*, 2016).

Omics have now graduated to yet another level of studying cellular metabolites as signatures of functional proteins of a cell, the approach called 'metabolomics' helps a great deal in understanding the regulation of cell metabolism, metabolic fluxes and metabolic integration. For example, studies on riboflavin-producing *Bacillus subtilis* (Sauer *et al.* 1997; Sauer *et al.*, 2006) and use of ^{13}C-based metabolic flux analysis (Zamboni *et al.* 2009);(Gerosa and Sauer, 2011). These studies help in understanding regulatory network, multiple pathways and metabolic fluxes of bacteria metabolism. Thus, to understand the intricate details of cellular physiology, multi-omics data integration can prove to be a useful tool. Integration of genomics, transcriptomics, proteomics and metabolomics data will elucidate how genomic variation /stress/ stimuli alter quantitative biomolecular profiles that eventually determine cellular physiology (Hawkins *et al.*, 2010);(Gehlenborg *et al.*, 2010). The present study is a small step towards understanding the metabolic adaptation of mycobacteria under microbicidal stress it faces during infection. Metabolomics will be discussed in detail in following sections.

The 'omics', has now expanded, to lipidomics, glycomics, metabonomics and sometimes more specific functional congregation of cellular factors, such as 'secretomics' which is a the study of the sum total of all the factors secreted by a cell, including components of secretory pathway. These studies, though relatively recent, have added valuable information towards understanding of mycobacterial physiology, adaptability, antibiotic resistance and pathobiology, which can be further explored for identifying new drug targets.

1.6 Metabolomics: a new approach to target and disease marker identification

The term metabolite is usually restricted to small organic compounds, which are produced in or required for metabolism. Metabolites are typically micro molecules, which are engaged with numerous cellular functions, for example energy, stimulation, inhibition, protection, etc. Based on their cellular activities metabolites are divided into two types-primary and secondary metabolites. Primary metabolites are directly involved in physiological function and are required for the growth and maintenance of cellular function, while secondary metabolites are not are not required for the growth and maintenance of the cellular functions. Comprehensive and quantitative study of all metabolites in a biological system can be defined as metabolomics. Thus metabolomics is a study to evaluate all metabolites within a cell and observe changes in chemicals or their concentration in response to stimuli or genetic alterations (Barsch *et al*., 2004)**;**(Rabinowitz and Kimball, 2007)**.** The common samples for metabolomics study are bio-fluids such as blood and urine including fecal extracts, breast milk, amniotic fluid, umbilical cord blood and various cells and tissues. These samples are easy to collect, where often the aim is to identify biomarkers for a disease. Recent development in identification techniques like NMR (Harrigan *et al*., 2004), Fourier transform-infrared spectroscopy (FT-IR) (Sauer *et al*. 1997; Johnson *et al*. 2004) and mass spectrometry (MS) have been used for metabolomic applications. NMR is a non-discriminating and a non-destructive technique. It requires minimal steps in sample preparation. The samples used for collecting NMR spectra can be used for other experiments since it does not destroy the samples. NMR offers a convenient method wherein the individual metabolites are identified from a complex mixture based on chemical shift signals. CHENOMX suite software helps to identify and quantify metabolites by spectral deconvolution method. However, mass spectrometry-based techniques offer to identify and quantify with high selectivity and sensitivity. It usually requires a sample preparation step, which can cause metabolite loss, and specific metabolite classes may be discriminated based on the sample introduction system and the ionization technique used. The changes resulted due to environmental perturbations, inherited and gut microbiome pressure can be studied by metabolomics. The fine regulation of cellular metabolism is governed by local metabolite concentration thus metabolomics studies acts as a linker between genotype and phenotype. Thus metabolomics studies provide complete picture on how a cell functions (Fiehn, 2002)**.**

1.7 Why study metabolites and use metabolomics approach?

Basic requirement for every living system is its ability to adapt to its changing environment and metabolic needs. Thus cells continuously monitor their requirements and cellular metabolic state at different levels, these changes can be quick and mild or slow but more drastic. These regulations can be at DNA sequences (genome), mRNA (transcriptome) and proteins (proteome) levels. Since all these changes are in response to long term environmental change these are long term regulation. However, quick adjustment or fine regulation of cellular metabolism is triggered by changes in local concentrations of metabolites leading to an allosteric or post-translational regulation of enzyme activity. The aim of studying metabolomics usually is to get biological data to facilitate the understanding of cellular functions. In general, metabolomics include both extracellular and intracellular metabolites. Metabolomics of intracellular metabolites are vital to study the regulation of cell metabolism, metabolic fluxes and metabolic integration for example, riboflavin-producing *Bacillus subtilis* (Sauer *et al.* 1997; Sauer 2006), ^{13C}-based metabolic flux analysis (Zamboni, Fendt *et al.* 2009; Gerosa and Sauer 2011). These studies help to understand the regulatory networks, multiple pathways and metabolic fluxes of bacterial metabolism.

To build metabolite network from metabolite profile, one requires to know the direction of reaction, models, algorithms, flux analysis etc. One of the challenges in modeling the rates of metabolic systems is to determine the intracellular metabolites concentration, as most intracellular metabolites are in low concentration i.e. below 1 mM (Holms 1996; Hiller *et al.* 2007). Thus for analyzing intracellular metabolites cells are separated from media and metabolic activity is quenched immediately after sampling. Metabolomic studies not only provide qualitative but also quantitative data. Quantitative measurements can be used to define physiological state of biological system. Quantification of intracellular metabolites can be achieved by the following steps: 1) harvesting 2) ceasing metabolic activity by quenching 3) extraction and 4) finally quantitative analysis (Dobson *et al.* 2010). Once the metabolites and their concentrations are defined, these voluminous data could make sense only if this information could be built-in to a network or direct the metabolites to a specific metabolic flux. Recent surge in systems biology approach has made available advanced tools and methods to deal with the huge loads of data and churn them to meaningful revelations and insights into cellular processes.

1.8 Recent studies involving metabolomics approach towards understanding basic biology, drug discovery and biomarker identification

Recently, the application of metabolomics has made significant impact in disease, toxicity or genetic manipulation research field. Models for metabolome measurements for central metabolism of different organisms include *Saccharomyces cerevisiae*, *Escherichia coli*, *Aspergillus niger* and *Lactococcus lactis* (Chassagnole *et al*. 2002; Oliveira *et al*. 2005; Andersen *et al*., 2008; Usuda *et al*. 2010), *Streptococcus pneumonia* for shikimate pathway and *E. coli* and *Corynebacterium glutamicum* for amino acid metabolism (Yang *et al*. 1999; Chassagnole *et al*. 2002; Magnus *et al*. 2006; Noble *et al*. 2006) have been described.

The metabolomics of *E. coli* has been investigated extensively using qualitative analysis (Simão *et al.*, 2005);(Dwivedi *et al.*, 2010). Metabolomics approach has emerged as a new tool to investigate *M. tb* physiology. Given the ability to respond to dynamic environments and regulate entry and exit from cell cycle, intense research on the metabolism of *M. tb* is need of the hour (Muñoz-Elías and McKinney, 2006);(Beste and McFadden, 2010). As we understand, the 'omics' cascade has come a long way from genomics, transcriptomics and proteomics to elucidate the essential components and their dynamic interactions. In spite of these studies we have limited understanding of *M. tb* metabolic network. Thus metabolomics of products from this complex interaction will provide a better understanding of cellular metabolism of *M. tb* adaptive physiology. Recently researchers have started to use metabolomics approach to study *M. tb* metabolic adaptations. Rhee and group have shown that *M. tb* has the ability to co-catabolize different nutrients simultaneously by distinct metabolic pathways to maximize the growth. In their study, they had also shown that mycobacteria have the ability to metabolize carbon source by same pathway and also reversibly. For this study, they used ^{13}C isotope-labeled carbon sources (i.e. dextrose, acetate, or glycerol). In most bacteria, the regulatory trait of carbon utilization is "catabolite repression", but mycobacteria shows unprecedented regulatory metabolic network, where it can co-catabolize substrate. In response to environmental stresses, *M. tb* remodels its central carbon metabolism which has been studied by several groups. Beste and group have shown a novel pathway for pyruvate dissimilation, termed as Glyoxylate, Anapleurotic and Succinyl CoA (GAS) pathway that appears to be activated in response to carbon starved condition by using ^{13}C isotope tracing metabolomics studies. Another study by

Baek and group showed that the growth-limiting stresses induced the biosynthesis of triacylglycerols, by remodeling flux away from the TCA cycle (Baek *et al*., 2011). Thus, this remodeling helped the bacteria to resist antibiotics. All these observations showed the unprecedented ability of *M. tb* to utilize carbon and that its availability affects its growth rate. A recent study using MS-based metabolomics approach revealed the possible utilization of macromolecules by *M. tb* infected macrophages as reflected by low levels of carbohydrates, nucleotides and amino acid (Zimmermann *et al*., 2017). Meissner and group performed metabolomic analysis of hyper- and hypo-virulent Beijing strains of *M. tb* suggesting virulence in *M. tb* may be influenced by ESX-1 gene cluster (Meissner-Roloff *et al.*, 2012). Recently, Loots and group confirmed the association of ESX-1 with virulence using metabolomics approach by comparing *M. smegmatis* with *M. smegmatis* ESX-1 knock-out strain (Loots *et al.*, 2016). Metabolomics has significantly contributed to our understanding disease-associated metabolomic adaptations of both the microbe and the host by identifying biomarkers originating from it.

1.9 Metabonomics approach to disease biomarker identification

While metabolomics refers to metabolites involved in normal endogenous metabolism, metabonomics is preferentially referred to altered metabolite profiling in response to environmental factors or a diseased state (Jeremy and Ramsden, 2017). Metabolite biomarkers can serve as indicators for improved diagnostic strategies and may serve as indicator for prognostic indicators for improving TB treatment strategies. Metabolomics has been successfully used to develop TB treatment strategies in identifying new metabolite biomarkers contributing to an improved understanding of the mechanisms of TB drug action, its resistance mechanism and their side effects (Luies *et al*., 2017). A recent metabolomics study by Li and his group on urinary metabolite profiles of healthy controls and INH-administered healthy volunteers identified 7 new INH-derived metabolites which belong to hydroxylated and hydrazone class of metabolites (Li *et al.*, 2011). These hydrazones are intermediates of essential amino acid metabolism such as tryptophan, phenylalanine, tyrosine, lysine, leucine and isoleucine, indicating additional targets of INH inhibits growth of *M. tb*. Another metabolomics study carried out by Nandakumar and group identified a common subset of metabolome of *M. tb* exposed to sublethal dose ofthree anti-mycobacterials Isoniazid (INH), rifampicin (RIF) and

Samples	Biomarker identified	Study
Serum	5-oxoproline	(Che *et al.*, 2013)
	cysteine-glutathione disulfide;threonine; citrulline;Histidine; urea; inosine; cysteine; glycocholenate sulfate;tryptophan; mannoseacetylneuraminate; 4-methyl-5-propyl-2-furanpropanoate; phenylalanine; pyroglutamine. octadecanedioate; glutamine; taurocholenate sulfate; asparate; glycylvaline; g-glutamylglutamine;	Weiner *et al*. 2012
Urine	glucose; norepinephrine; 4- hydroxybenzoic acid; Lactic acid;hydroquinone.	(Das *et al.*, 2015)
	5-hydroxyhydantoin; indole-3-carboxylic acid;2-C-methylglycerol; glycerol monostearate; oxalic acid; 2-octenoic acid quinolinic acid; 5-hydroxyhexanoic acid ribitol; tryptophan; 5-hydroxyindoleacetic acid;;phenylacetic acid; l-rhamnulose;homovanillic acid; N-acetyltyrosine ; kynurenic acid;phenyllactic acid;	Luies and Loots (2016)
Breath	Methyl nicotinate	(Syhre and Chambers, 2008)
	1,4-dichloro-benzene;1-methyl- naphthalene; camphene; 1,3-dimethyl-trans-cyclohexane; 2-butyl-1-octanol; 1,4-dimethyl-cyclohexane;	(Phillips *et al.*, 2007)
	4-methyl-dodecane; Bis-(3,5,5-trimethylhexyl) phthalate;3-(1-methylethyl)-oxetane; hexyl- cyclohexane; 1,3,5-trimethyl-benzene; tridecane; 5-ethyl-2-methyl- heptane3,7-dimethyl-decane; 4-methyl-1-hexene; 4, 6, 8-trimethyl-1-nonene.	(Phillips *et al.*, 2010)
Sputum	eicosanoic acid ; TBSA; myo-inositol;Nonadecanoic acid; hexacosanoic acid;Schoeman acetohydroxamic acid; palmitoleic acid; D-glucopyranoside; furan;d-glucose;propane; d-glycero-l-manno-heptanoic acid;Lignoceric acid, arachidonate; octadecanoic acid; d-mannose;heptadecanoic acid; 9-octadecenoic acid; d-galactose; d-glucosamine; glycerol; uridine; d-fructose; 2-O-glycerol-à-d-galactopyranoside; cadaverine; docosanoic acid;á-d-xylopyranose; arabinofuranose; á-d galactofuranose; d-erythro-pentitol; á-d-galactopyranoside;l-threonine; and phenylethanolamine	(Schoeman, du Preez and Loots, 2012)
	à-d-glucopyranoside; à-d-mannopyranoside; à-d-glucopyranose-2-acetylamino-2-deoxy; N-acetyl-glucosamine;á-l-mannopyranose; lactone; d-galactose-6-deoxy;d-gluconic acid; d-citramalic acid;à-d-galactopyranose; 17-methyl-octadecanoic acid; d-glucosamine; 3,4-dihydroxybutanoic acid; 10-heptadecenoic acid; onadecanoic acid; glutaric acid; sebacic acid; ethane; butanal; 2-deoxy-d-erythro-pentitol; gamma-aminobutyric acid;normetanephrine;9-octadecenoic acid;	(Du Preez and Loots, 2013)

Table 1: Previous studies identified differential metabolites as biomarkers for disease prognosis and diagnosis of TB

streptomycin (STREP), respectively (Nandakumar, Nathan and Rhee, 2014). These identified metabolites represent biosynthetic pathways of various amino acids, glyoxylate pathway and tricarboxylicacid (TCA) cycle. These results suggest that INH, RIF and STREP apart from having their known antibacterial targets, they activate isocitrate lyases. These findings were in agreement with the study showing that *M. tb* strains deficient in isocitrate lyase are significantly susceptible to anti-mycobacterial drugs. Further, **Table 1** shows previous studies done to identify differential metabolite to use them as biomarker. These studies provide a better understanding of TB disease and also provide better insights in to *M. tb* adaptive physiology by providing information on the fine regulation of cell under the given stress conditions as compared to broad regulation at genomics and proteomics levels.

As ongoing efforts to understand mycobacterial adaptations during infection, both as single infection and as co-infection with HIV, I was a part of study that used proteomics approach to study differential proteomics of intra-phagosomal mycobacteria where we could observe changes in major pathways of intermediary metabolism, lipid metabolism, virulence and adaptation (Ganji *et al*. 2016a; Ganji *et al*. 2016b). The proteomics study strongly indicated metabolic adaptation of mycobacteria, both during mono- and co-infection. These observations prompted us to take up metabolomics approach to understand early metabolic adaptation of mycobacteria to microbicidal stress during infection. However, unlike proteins, it is difficult to distinguish metabolites of the pathogen from that of the host; therefore I studied the comparative metabolomics applying *in-vitro* stress models for mycobacterial growth. I focused on using comparative metabolomics approach to understand immediate adaptation of *M. tb* to microbicidal stress along with studying differential metabolites from the sera of host (TB patient) as compared to healthy individuals for identification of disease indicators for my thesis work.

1.10 Objectives, hypotheses and key questions

1.10.1 Objective 1: Comparative and quantitative metabolite profiling of pathogenic mycobacteria H37Rv and non-pathogenic soil mycobacteria, *M. smegmatis* MC^2155, under microbicidal stresses

Hypothesis: The key to the survival of mycobacteria in microbicidal environment is a two phase adaptation process, where subtle plastic phenotypic changes through alterations in metabolic flux will precede the long term adaptation through changes at protein, transcription or genetic level. Therefore studying the changes in metabolic profile of mycobacteria soon after exposure to microbicidal stresses will reveal early adaptation processes of these mycobacterial species.

Key questions

- What are the early adaptive metabolic changes in mycobacteria in response to microbicidal stresses like low pH, oxidative stress, iron deficiency and nutrient deprivation?
- Are there overlaps in metabolic adaptations to these stresses in pathogenic and non-pathogenic mycobacteria?

1.10.2 Objective 2: Comparative metabolic profiling of sera from TB patients, their clinically healthy household contacts and unrelated clinically healthy volunteers

Hypothesis: A disease state brings about a plethora of changes in metabolism of human host that gets reflected in the alterations in the sera composition including levels of these intermediate products of metabolic reactions. Comparing sera from diseased (here TB patients) *vs* healthy (here comprising of two groups; household healthy contacts of patients and unrelated healthy volunteers) will lead to identification of metabolite profiles that can distinguish diseased from healthy state, serving as biological indicators for TB disease.

Key questions

- Does the impact on metabolism during TB disease reflect in the peripheral sera of TB patients?
- Will the quantitative metabolite profiles differ in sera from TB patients, their household clinically healthy contacts and unrelated clinically healthy volunteers?

- Will the differences in metabolites in these groups able us to identify a TB disease indicator?

1.11 Organization of the thesis

The objectives drawn are discussed in the ensuing chapters. Each chapter has a chapter overview. Methodologies have been compiled in chapter (Chapter 2). The results and discussion relevant toobjective 1 is written in chapter 3. Chapter 3 has two parts: (A) Metabolic profiling of *M. smegmatis* (MC^2155) under stress conditions and major observations and (B) Metabolic profiling of *M. tb* (H37Rv) under stress conditions and major observations. Results of objective 2 are compiled and discussed in chapter 4 followed by an overall summarization of the observations, outcome and future leads from the thesis as chapter 5.

Chapter 2

Material and Methods

2. Material and Method

This chapter details the methodology used for achieving the objectives under my thesis work. The methodology has been detailed objective-wise for ease of reference. All experiments involving mycobacteria were performed inside BSL-3 lab. Institutional Biosafety Committee approval (SB/BSL-F-60) was taken. Prior ethical committee clearances (ECR/450/Inst/AP2013, ECR/450/Inst/AP/20131RR-16 and UH/IEC/2014/36) were taken and written consents obtained from the subjects recruited for collection of sera samples for objective 2.

2.1 Objective 1: Comparative and quantitative metabolite profiling of pathogenic and non-pathogenic mycobacteria under microbicidal stresses

2.1.1 Growth media

1) Lowenstein-Jensen (LJ) medium slants were purchased from HiMedia, India
2) Middlebrook 7H9 broth medium

   Media was prepared as per manufacturer protocol. In brief, 2.35 g of 7H9 Middlebrook base (HiMedia, India), 0.05% tylaxapol or 0.05% Tween and 0.2% glycerol was dissolved in 450 ml of MilliQ water in 2L flask and autoclaved at 121°C, 15 lbs/sq. inch pressure for 10 min. Before inoculating, 10% of OADC enrichment supplements were added to cool the media under aseptic conditions.

2.1.1 Growth media used to study stress responses Stress media

2.1.1.1 Sauton's media (Normal / Control)

Media was prepared dissolving 0.25g of monopotassium phosphate (KH_2PO_4), 2g of L-asparagine, 0.25g of magnesium sulphate ($MgSO_4$), 1g of citric acid, 5mL of glycerol, and 0.025g of ferric Ammonium Citrate with 0.05% tyloxapolin 500mL of water. pH was adjusted to 7.2 with sodium hydroxide (NaOH). After autoclaving, 0.1% w/v concentration of sterile zinc sulfate ($ZnSO_4$) was added just before using media.

2.1.1.2 Acidic stress media

Acid stress media was prepared as given above for Sauton's media with modification i.e the pH was adjusted 5.5 (Piddington, Kashkouli *et al*. 2000).

2.1.1.3 Oxidative stress

Acid stress media was prepared as given for Sauton's media with modification i.e hydrogen peroxide (10mM) was added, just prior starting of experiment (Voskuil, Bartek *et al*. 2011).

2.1.1.4 Iron deficiency or depletion stress (Iron stress)

For iron deprivation, the glasswares and the media were made iron-free as described earlier (Hall and Ratledge 1982).

Preparation of iron free glassware

- ✓ Glassware were soaked in 2% methanolic KOH over night.
- ✓ Wash glassware thrice with double-distilled water.
- ✓ Glasswares were soaked in 8 N HNO_3 overnight.
- ✓ Wash glassware again thrice with double-distilled water.

2.1.1.5 Preparation of iron free Sauton's media

While preparing iron deprived media, all constituents were added to as mentioned for sautons media except ferric ammonium citrate. Apart from ingredients, 3 g of alumina was added to media and then autoclaved for 20 min. After autoclaving, media was cooled and then Whatman grade 541 filter was used to filter media and autoclaved before using for experiment (Hall and Ratledge 1982).

2.1.1.6 Nutrient starvation

For nutrient starvation 1X phosphate buffer saline [8g of sodium chloride (NaCl), 1.16g of disodium phosphate (Na_2HPO_4), 0.2g of monopotassium phosphate (KH_2PO_4) and 0.2g of potassium chloride (KCL) in 1L of double-distilled water] was used along with 0.05% tyloxapol (Loebel *et al* 1933).

2.1.2 Bacterial Strains and its maintenance

M. smegmatis (mc^2155) and *M. tuberculosis* (*M. tb);(*H37Rv) were grown on LJ media. The slant was incubated at 37^0C until colonies appeared. A single colony was picked and inoculated in Middlebrook 7H9 media supplemented with 10% OADC and 0.05% tyloxapol at 37^0C until the OD600 nm reached 0.8 to 1. Prior to making glycerol stocks, Ziehl-Neelsen (ZN) staining was performed to rule out any contamination, for all set of experiments these glycerol stocks were used as inoculum for primary culture.

2.1.3 *In vitro* growth under stress conditions

Primary cultures were grown by inoculation of respective strain (*M. smegmatis and M. tb*) glycerol stocks in 7H9 media (consisting 0.4% glycerol, 0.05% tylaxapol and supplemented with10% OADC) till its optical density ($O.D_{600}$) reached 1.0. Further 1% primary culture was inoculated in 500 ml and 100 ml 7H9 media in 2000 ml and 500 ml size culture screw-capped conical flask for *M. smegmatis* and *M. tb* respectively. Cells were grown at 37 °C with shaking at 180 rpm till OD 600nm reached to 0.6 - 0.7. Cultures were then harvested, washed and resuspended in respective stress media for 4 hours (*M. smegmatis*) and 36 hours (*M. tb*). The study was performed by using well-established in-vitro models mimicking microbicidal stress conditions such as acidic stress (pH 5.5) (Piddington *et al..*, 2000), oxidative stress (10 mM H2O2) (Voskuil *et al*., 2011), iron deprivation (no iron supplemented*);(*Hall and Ratledge 1982) in Sauton's minimal medium and for nutrient starvation stress 1x phosphate buffer saline was used as described by Loebel (Loebel *et al* 1933). The culture was ruled out for any possible contamination by using Ziehl-Neelsen (ZN) staining before and after stress. Each of the stress condition was replicated 10 times to obtain statistically significant number of samples while averaging out data within a stress condition.

2.1.4 Identification and quantification of metabolite from non-pathogenic bacteria

2.1.4.1 Extraction of intracellular metabolites from non-pathogenic bacteria

M. smegmatis cells were extracted using a modified version of an extraction previously described (Nicolas *et al* 2013). After 4 hours of stress, cultures were harvested and quenched in liquid

nitrogen. The cell pellets were thawed on ice then extracted with a Methanol/chloroform method i.e. 1ml of methanol/chloroform (2:1) was added along with 0.1mm of zirconia beads to lyse the cells for 10 times using bead beater with an interval of 1min on ice. Supernatant was collected after centrifuging at 1000 rpm for 45 secs, to this supernatant 500 µl of MilliQ water and 500 µl of chloroform was added and vortexed for 30 sec. The culture was then centrifuged at 12000 rpm for 30 min. The upper aqueous layer and lower organic phase were collected in 1.5 ml vial separately and upper aqueous layer phase samples were lyophilized, for further analysis

2.1.4.2 NMR Spectroscopy Analysis

2.1.4.2.1 Sample preparation

NMR spectroscopy was done for non-pathogenic bacteria (*M. smegmatis*). Briefly, 580µl of 100% D_20 NMR buffer [20 mM sodium phosphate pH 7.4 (Composition for 1litre PBS is 0.623 g Na_2HPO_4. H2O; 4.15g NaH_2PO_4. $7H_2O$ and the pH of the buffer was adjusted by addition of either dilute HCl or NaOH)]were added to above lyophilized samples, containing 0.4 mM of 4,4-dimethyl-4-silapentane-1-sulfonic acid (DSS) for 2 min at RT and vortexed. Then the supernatant was collected in NMR tube after spinning at 4000g for 2 min.

2.1.4.2.2 NMR Spectroscopy

The above M. smegmatis samples were transferred into a NMR tube and ^{1}D NOESY spectra were collected in a 600 MHz NMR spectrometer (In collaboration with Dr. Chugh, IISER, Pune), each sample (n=40) was measured at 298K using a Bruker Avance III HD Ascend NMR spectrometer equipped with a quad-channel (HCNP-D) cryo-probe operating at a proton frequency of 600.13 MHz. Water suppression pulse sequence noesygppr1d was used to record ^{1}H NMR spectra, which uses water presaturation and spoiler gradients during the relaxation delay and is of the form –RD-G1-90°-t-90°-tm-G2-90°-ACQ, where RD is the relaxation delay of 5 secs, t is a short delay typically of ~3µs, 90° represents a 90° RF pulse, tm is the mixing time of 100 ms and ACQ is the data acquisition period. For a given sample, a total of 64 transients were collected into 32k data points for each spectrum with a spectral width of 12 ppm. Pulse width, receiver gain and water suppression parameters were kept identical among all the ^{1}H experiments recorded for various samples to rule out intensity variation while recording the

NMR data. For ^{1}H-^{1}H TOCSY experiment (mixing time = 80 ms), a total of 2048 × 1024 data points with 40 transients per increment in the indirect dimension were recorded spanning a spectral width of 10 ppm in both the dimensions. Additional homonuclear and heteronuclear spectra were also collected to assist with metabolite identification. These were ^{1}H J-resolved, ^{1}H-1H COSY and ^{1}H-^{1}H TOCSY, ^{1}H DOSY and ^{1}H-^{13}C HSQC spectra. Spectra were processed and analyzed by using TOPSPIN (Bruker). This includes, automatic and manual phase correction, spectra referencing, and baseline correction. These processed spectra were used for further metabolomic analysis. Chenomx (Chenomx, Inc) suite was used to identification and quantification. Each metabolite identified in respective samples were picked and converted to relative concerntration (micromoles) using Chenomx NMR suite 8.1 by comparing with internal standard (DSS). All the concentration data such obtained from each 10 replicated of each condition were exported in matrix format to a spreadsheet file for univariate and multivariate analysis.

2.1.5 Identification and quantification of metabolite from pathogenic bacteria

2.1.5.1 Extraction of intracellular metabolites from pathogenic bacteria.

M. tb cells were extracted after 36 hours of stress, cultures were harvested and quenched in liquid nitrogen. The cell pellets were thawed on ice then extracted with 1 ml of Methanol along with 0.1mm of zirconia beads to lyse the cells for 10 cycles using mini bead beater with an interval of 1min on ice. The supernatant was collected after centrifuging at 1000 rpm for 45 sec. The supernatant was then centrifuged at 12,000 rpm for 30 min. The upper layer phase was collected in 1.5ml vial separately and the solvent was evaporated by speed vac. The dried powder was used for further analysis.

2.1.5.2 Liquid chromatography-mass spectrometric multiple reaction monitoring (LC-MRM/MS) based analysis

2.1.5.2.1 Sample preparation

The dried metabolite extract from *M. tb* were dissolved in 50 µl sample buffer (6.5:2.5:1 acetonitrile: methanol: water) and used for positive ionization mode (HILIC Chromatography). In case of negative ionization mode, the dried metabolite extract from *M. tb* was dissolved in

50µl ultrapure water (T3 RPLC Chromatography). For both modes, 10 µl of the sample was injected using an auto sampler.

2.1.5.2.2 LC-MRM/MS analysis for *M. tb*

Targeted metabolomic analysis was performed using multiple reaction monitoring (MRM) based approach i.e LC-MRM/MS analysis (Tushar H. More *et al* 2016). 108 metabolite standards were purchased from Sigma Aldrich to build an in-house MRM method (In collaboration with Dr.Srikanth, NCCS, Pune). For each metabolite standard, parent ion to daughter ion transitions was selected using MS/MS fragmentation. Based on the fragmentation pattern of metabolites, 108 metabolite standards were further divided into positive and negative ionization mode. For each MRM transition, collision energy (CE) and declustering potentials (DP) were optimised. The information obtained was exported to build the acquisition methods for positive and negative ionisation modes. In order to maximise the metabolomic coverage, chromatographic separation was performed using hydrophilic and reversed phase chromatography columns for positive and negative ionisation mode respectively. For both modes, 10 µl of sample was injected into the mass spectrometer using Shimadzu Prominence HPLC autosampler.

MS data was acquired using a 4000 QTRAP triple quadrupole mass spectrometer (AB SCIEX, Foster City, CA) equipped with Shimadzu Prominence binary HPLC pump (Shimadzu Corporation, Japan). For positive ionisation mode, the chromatographic separation was achieved using XBridge HILIC column (Waters, Milford, MA) that was eluted at 700 µl/min with a 32 min linear gradient starting from 5% mobile phase A (10 mM ammonium formate with 0.1% formic acid) increasing to 60% mobile phase B (acetonitrile with 0.1% formic acid). The column was kept at 60% mobile phase B for 3 min then returned to 5% mobile phase A for equilibration. For negative ionization mode, the chromatographic separation was achieved using ATLANTIS T3 column (Waters, Milford, MA) that was eluted at 500 µL/min with a 40 min linear gradient starting with 100% mobile phase A (10 mM ammonium hydroxide with 0.1% acetic acid) increasing to 98% mobile phase B (100% MeOH). The column was kept at 98% mobile phase B for 5 min then returned to 100% mobile phase A for equilibration. MRM was used to acquire targeted MS data for specific metabolites in the positive and negative ion modes. Transitions, dwell time and collision energies were set by using Analyst 1.5 software (SCIEX, Foster City,

CA). MS conditions were set as follows, source temperature: 400°C, interface heater: on, curtain gas: 30, declustering potential: 90, entrance and exit potential: 10, and the two ion source gases were set at 45 (arbitrary units). Analyst 1.5 software (Sciex, Foster City, CA) was used to analyse LC-MRM/MS data by manual inspection of chromatograms and for the detection of the compounds.Analyst quantitation wizard was used for integration of peak areas. For quality measures, samples order was randomized at the time of analysis and integration of peaks was performed in a blinded manner. Metabolites with a minimum of 15% of base peak intensity were considered for quantitation. The peak areas obtained after integration were exported in a matrix formatto a spreadsheet file for univariate and multivariate analysis.

2.1.6 Statistical analysis

The data matrix file obtained from either NMR (*M.smegmatis*) or LC-MS/MS (*M. tb*) were used as input files to perform statistical analysis by using online tool MetaboAnalyst (www.metaboanalyst.ca/MetaboAnalyst). All the required and recommended statistical analyses were performed (Xia *et al*., 2012), to arrive at sets of metabolites that have the most distinguishing abilities between the groups. To begin with, the data obtained for each condition was normalized using the respective number of cells at the point of metabolite extraction. Each set of data was then pre-processed to achieve normal distribution (i.e. they follow a Gaussian or 'normal' distribution), the methods applied for each condition is briefed below (**Table 2**). Such normalization is an essential pre-requisite for any statistical analysis as otherwise most of the standard statistical tests become unreliable (Xia *et al*., 2012). Univariant and multivariant analysis like fold change, t-test, principal component analysis (PCA), partial least-squares discrimination analysis (PLS-DA) were performed to identify metabolites differentially regulated between control and stress. Further pathway analysis was used to identify altered pathway in both pathogenic and non-pathogenic mycobacteria.

	Pre-Processing Method	Acidic - Control	Oxidative-Control	Iron-Control	Starvation-Control	Acidic-Oxidative-Control	Acidic-Oxidative-Iron-Starvation
M. smegmatis	Normalization	----	Sum	NA	----	Sum	NA
	Transformation	----	Cube Root	NA	----	----	NA
	scaling	Pareto	Pareto	NA	Pareto	Auto	NA
M. tb	Normalization	Sum	Sum	Median	Median	NA	Sum
	Transformation	Cube Root	------	Cube Root	Cube Root	NA	Cube Root
	Scaling	Auto	Range	Pareto	Auto	NA	Range

Table 2: Pre-processing method to perform standard statistical analysis

2.1.7 Protein Structure Homology-Modelling

The primary amino acid sequence of 2,4-dienoyl-coA reductase [Mycobacterium smegmatis str. MC2 155] MSMEG_5124 was obtained from Mycobrowser (Kapopoulou *et al.*, 2011) and subjected to domain analysis tools to predict conserved domains in the 2,4-dienoyl-coA reductase [Mycobacterium smegmatis str. MC2 155] MSMEG_5124. These include MOTIF (http://www.genome.jp/tools/motif/) and SMART (http://smart.embl-heidelberg.de/). To further confirm the catalytic site of 2,4-dienoyl-coA reductase [Mycobacterium smegmatis str. MC2 155] MSMEG_5124, we have generated a model using SWISS-MODEL (protein structure homology-modelling server) (Biasini *et al.*, 2014) and GENO3D (automatic modelling of proteins three-dimensional structure) (Combet *et al.*, 2002) . The predicted protein structure homology was checked by using Chimera version 1.10.2 and PyMOL viewer.

2.1.8 Urease activity

The urease activity was measured by colorimetric method at 670nm using urease activity assay kit (ab204697, Abcam Plc., UK). In summary, after respective stress, cells were washed with ice cold 1XPBS and then resuspended in 300 µl of 1XPBS with protease inhibitor. The cells were lysed by bead beating, after adding 0.1mm zirconia beads to each sample. Cell lysate was centrifuged at 10,000 rpm at 4 ºC for 20 min. The supernatant was filtered by 0.22micron filter. The protein concentration of the resultant supernatant was measured. Equal amount of protein from each sample was taken and the volume was made up to 10µl. To each 10µl of sample, 90µl of reaction mix (urease buffer and urea). The samples were homogenised and incubated at 37ºC for 30 min. Then, 80µl of reagent1 followed by 40µl of reagent 2 were added. Samples were incubated for 30 min (room temperature). Absorbance was measured at 670 nm in multi-well plate reader (Biotek). All the experiments were performed at the least three times.

2.1.9 Estimation of ammonia released

The ammonia released was estimated by colorimetric method at 430nm by using Nessler's reagent as described (Gouzy *et al.*, 2014) with modifications in a 96 well plate reader. In summary, after giving respective stress, samples were harvested at 3,500 rpm and lysed as described. The resultant supernatant was filtered by 0.22 micron. To 10 µl of supernatant, 90 µl of Nessler's reagent was added and incubated at room temperature for 15 min. Then absorbance was measured at 430 nm in a multi-well plate reader (Biotek). For respective stress, media was taken as control. All the experiments were performed at the least three times.

2.1.10 Extracellular pH measurement

pH measurement was done as describe (Gouzy *et al*. 2014). In summary, after giving respective stress, the samples were harvested at 3,500 rpm. The resultant supernatant was collected in 50 ml falcon. The pH of the supernatant was measured inside the biosafety cabinet type II B2. For respective stress, media was taken as control. All the experiments were performed at the least three times.

2.1.11 RNA Extraction

RNA extraction was performed by using Trizol method with some modification. After giving respective stress, the samples were harvested at 3,500 rpm. The resultant pellet was snap freed in liquid nitrogen and stored at -80 ºC until processed for RNA. The pellet were resuspended in Trizol reagent (Invitrogen, CA, USA) along with 0.1 mm glass beads and subjected to lysis by by bead beating with pulse on: 1 min and pulse off: 2 min on ice. Final concentration of 200 µg/mL of glycogen was added and kept for 10 min at RT. Vortex vigorously after adding chloroform and incubate at RT for 10 min. Spin at 10,000 rpm at 4 ºC for 20 min. The upper aqueous layer is collected. Isopropanol along with of glycogen was used to precipitate RNA. 75% ethanol was used to wash pellet. The pellet was air dried at RT and dissolved in RNase-free water (Qiagen, Hilden, Germany). Any residual DNA contamination was removed by DNase treatment prior to reverse transcription the RNA. This DNase treated RNA was further reverse transcribed with random hexamers as primer using Superscript III Reverse Transcriptase (Invitrogen). The reverse transcribed RNA was used for RT-PCR. RT-PCR was carried out with 1: 10 diluted cDNA for 30 cycles details of it is given with respective primers with initial denaturation at 95 ºC (3 min), cycles of 95° C 15 s/ appropriate temperature mentioned in the Table3 for *Mycobacterium smegmatis* (Table 3A, 3B, 3C, 3D) and Table 4 for *M. tb* (Table 4A, 4B, 4C)3 ºC 20 s/72 ºC 20 s, final extension at 72 ºC for 10 min.

Table 3A: Trehalose biosynthesis and utilization pathway in *Mycobacterium smegmatis*

Gene name	Sequence (5'-3')	Annealing Temperature (°C)
MSMEG_6515treSRTFp	TCGGTATCGACGGTTTCC	57
MSMEG_6515treSRTRp	GGAAGTGGAACGCCATGT	57
MSMEG_4916glgERTFp	AGATCCTGCAGATGTCGAA	57
MSMEG_4916glgERTRp	AACTGCTGCGCGGCCTCCA	59
MSMEG_4918glgBRTFp	AACCTCATCGACTACCGC	59
MSMEG_4918glgBRTRp	AGTGGTTGAACTCGCCGAT	57
MSMEG_6514Mak-Pep2RTFp	ACGCCATCTTCAAGCTCTT	57
MSMEG_6514Mak-Pep2RTRp	TGCTGGTGGTCGCCATGT	57
MSMEG_4535TreHRTFP	CAGTGGATCAACGTCGGTG	57
MSMEG_4535TreHRTRP	GACAGGTGGTGCAGCTCTTC	57

Table 3B: Betaine Biosynthesis synthesis pathways in *Mycobacterium smegmatis*

Gene name	Sequence (5'-3')	Annealing Temperature (°C)
MSMEG_5815BBADFP	GGTGGCTGCGTCGTGTGT	57
MSMEG_5815BBADRP	CGGCGTCAGGGAAGATGATGTT	57
MSMEG_3444BCDFP	CAATCTCGACGTGATCTCCGA	57
BCDMSMEG_3444BCDRP	CTCACCGTGGTTGTGCTC	57

Table 3C: *Trimethylamine* biosynthesis synthesis pathways in *Mycobacterium smegmatis*

Gene name	Sequence (5'-3')	Annealing Temperature (°C)
2,4-DCA MSMEG5124FP	AGATCGTCTGTTTGGCAAC	58
2,4-DCAMSMEG5124RP	GATCACTCAGCAGCGGAC	58
YeaXMSMEG_4371RTFP	AGAGATCCACGAGCAGTT	57
YeaXMSMEG_4371RTRP	CTGCGGGAGCACCTCGAC	57
TR2Fe-2SMSMEG_0657FP	TGCTCGTCCATCCATCCCG	57
TR2Fe-2SMSMEG_0657RP	GACGGCCAGCGGTGTCAT	57

Table 3D: Stress control *Mycobacterium smegmatis*

Gene name	Sequence (5'-3')	Annealing Temperature (°C)
MSMEG_4325FabDRTFP	CTGCTGGCGCACGAAGAACT	59
MSMEG_4325FabDRTRP	TTGGCCGGGACGAGGTCGA	59
MSMEG_6383 FurARTFP	CGTGTACGCACACCCGCA	57
MSMEG_6383 FurARTRP	CAGTCCACGTCGGCGATGT	57
MSMEG_1804sigFFP	GCTCAAGGAACTCCACTTGC	55
MSMEG_1804sigFRP	GATGGACAGCGTGTTGTACG	55
MSMEG_2758 sigAFP	GAAGACACCGACCTGGAACT	55
MSMEG_2758 sigA RP	GACTCTTCCTCGTCCCACAC	55

Table 4A: Stress control *Mycobacterium tuberculos*is

Gene name	Sequence (5'-3')	Annealing Temperature (°C)
FabD Rv2243RTFP	CGGCGCAGCGGACCAGAT	58
FabD Rv2243RTRP	CGCCGCCGAGCACCGCAGACAT	58
FurARv1909cRTFP	GCCGTGTACGACGTGCTGC	58
FurARv1909cRTRP	GTCCAACAGGAAGCCGTTAT	58
Rv3286cSigFRTFP	AGGTCCGACGACACTTCCG	58
Rv3286cSigFRTRP	ATCTGGTCAAGACCCGCATC	58
16SrRNARTFP	TAGGCGTTCCCTTGTGGC	58
16SrRNARTRP	CAGTCTCTCACGAGTCCC	58
Rv1193ACSF3RTFP	TGATCACCCTATCGACGC	57
Rv1193ACSF3RTRP	ATCAAGTCGACCGACTCG	57

Table 4B: Glutamate-Glutamine pathway *Mycobacterium tuberculosis*

Gene name	Sequence (5'-3')	Annealing Temperature (°C)
Rv2220GSRTFP	CAACGGCTCCTTCTACGAG	55
Rv2220GSRTRP	TGCAGCAGCGAATTGAACT	55
Rv2476cGDHRTFP	GGAACTGTTGCGTTGGCTG	55
Rv2476cGDHRTRP	CGATGGCATAGGGATAGGC	55
Rv3859cGOGATRTFP	GCAGGTGTTCTTGGCTGGC	55
Rv3859cGOGATRTRP	CGAGTGCACGATGCCTAGC	55

Table 4C: ADI pathway *Mycobacterium tuberculosis*

Gene name	Sequence (5'-3')	Annealing Temperature (°C)
ArcARv1001RTFP	CCTCATCTATGCTCATCA	57
ArcARv1001RTRP	CATCGTGCACACCGTGTC	57

Table 4C: Urease pathway *Mycobacterium tuberculosis*

Gene name	Sequence (5'-3')	Annealing Temperature (°C)
Rv1850 ureCFP	GACGTTCGATTCACGCCTA	57
Rv1850 ureCRP	CATCTTCTGCCGCAATGGT	57

2.2 Objective 2: Comparative metabolic profiling of sera from TB patients, their household clinically healthy contacts and unrelated clinically healthy volunteers.

2.2.1 Study cohort

A total of 140 subjects in the age group of 25-45 years were recruited under the study after taking prior ethical committee clearances (ECR/450/Inst/AP2013, ECR/450/Inst/AP/20131RR-16 and UH/IEC/2014/36) and written consents from the subjects. The patient and house hold contacts were recruited at Mahavir Hospital and Research Centre (MHRC) in collaboration with Dr.VijaylakshmiVelluri and Dr.Sumanlatha Gaddam. The healthy volunteers were recruited at UoH under the guidance of Dr.Ravindra Kumar. The study population of 140 subjects comprised of three categories, namely, clinically healthy donors (n=60), primary TB patients (n=40) and their respective household contacts (n=40). TB patients were identified as per Revised National Tuberculosis Control Programme (RNTCP), Government of India, guidelines with confirmed diagnosis from sputum, culture, Mantoux test and chest X-ray in patients at Mahavirhospital, Hyderabad, India.

Household contacts of the respective patients were those who resided in-house of the TB patient during 3months period for at least seven consecutive days prior to the diagnosis of tuberculosis. Clinically healthy donors were screened for TB infection by using QuantiFERON®-TB Gold (QFT®) ELISA kit (Reference# 0594-0201) and the healthy donors had no symptoms of any other diseaseat the time of collection. For QuantiFERON-TB Gold, results were analysed by QuantiFERON-TB Gold Analysis software (Version 2.62) as per the manufacturer's instructions. Before collecting samples from subjects, written informed consent was obtained. All standard procedures were followed to perform the experiments.

Criteria for selection of subjects for the study

Inclusion criteria

1. Age 25-45 years
2. Patients willing to abide by study procedures
3 With confirmed diagnosis from sputum, culture, Mantoux test and chest X-ray

Exclusion criteria

4. Patients on other medication (i.e. steroid therapy,)
5. Terminally ill patients as per treating clinician's judgment
6. HIV
7. Patients with no other respiratory infections
8. Patients with previous history of smoking, alcohol or drug abuse
9. Transplant patients, diabetes mellitus, liver disease, renal failure or, malignancy etc

2.2.1.1 Serum collection

About 2-5 ml blood was collected from subjects in a vacutainer tube free of anticoagulant (no additive). Stabilizers such as EDTA, citrate and collection tubes with gels were avoided as their presence generate additional signal in the NMR spectra. The following protocol was used for serum collection.

1. The sample was left on ice for 30 minto coagulate
2. The sampleswere centrifuged at 4,200 rpm at 4 ° C for 15 min
3. The supernatant was removed and snap freezed in liquid nitrogen to cease any enzymatic or chemical reactions

Stored the supernatant at -80ºC until analysis

2.2.2 Identification and quantification of metabolite from sera samples

2.2.2.1 Sample preparation

- To 500µl of serum, 150 µl of Trimethylsilylpropionic acid (TSP) buffer was added and mixed thoroughly (prepared in D_2O).
- 650 µl of the above sample was then passed through 0.22 µm filter, the filtrate was collected in to a fresh 1.5ml tube
- The filtrate was further passed through a centrifugal filter (cut off 10 KDa) (centrifuged at 10,000 g at 4 ºC) to remove macromolecules (the centrifugal filter was washed four times with de-ionised water for 5 min at 5000 g before use.
- The filtrate is then transferred to NMR tubes and analysed.

2.2.2.2 Data acquisition

The supernatant was transferred into a NMR tube and ^{1}D NOESY, CPMG and J-Res ^{1}H NMR spectra were collected in a 600 MHz NMR spectrometer equipped with a room temperature HCN inverse Z-gradient probe. The pulse sequence used was ^{1}D gradient NOESY with water pre-saturation using the Bruker standard profiling parameters. Additional homonuclear and heteronuclear spectra were also collected to assist with metabolite identification. These were^{1}H J-resolved, ^{1}H-^{1}H COSY and ^{1}H-^{1}H TOCSY, ^{1}H DOSY and ^{1}H-^{13}C HSQC spectra. Spectra were processed and analysed by using TOPSPIN (Bruker). This includes, automatic and manual phase correction, spectra referencing, and baseline correction. These processed spectra were used for further metabolomic analysis. Chenomx (Chenomx, Inc) suite was used for identification and quantification. Intensities of the identified metabolites were converted to absolute concentrations (micromoles) using Chenomx NMR suite 8.1 by comparing with the signal intensity from internal standard (TSP) of known concentration. All the concentration data obtained from each sample of each group were exported to a comma separated values (.csv) or tab delimited text (.txt) file format for univariate and multivariate analysis.

2.2.3 Statistical analysis

To identify differential metabolomic signatures associated with groups, uni-variate and multi-variate statistical analysis, pathway and biomarker analysis were carried out by using the online tool MetaboAnalyst (www.metaboanalyst.ca/MetaboAnalyst). All required and recommended statistical analyses were performed (Xia *et al*., 2012; Xia *et al.*, 2013; More *et al.*, 2018) to arrive at sets of metabolites that have the most distinguishing abilities between the groups. Each set of data was then pre-processed to achieve normal distribution (i.e. they follow a Gaussian or 'normal' distribution), by using median normalization cube root transformation and range scaling methods. Such normalization is an essential pre-requisite for any statistical analysis as otherwise most of the standard statistical tests become unreliable (Xia *et al.*, 2012). Initially, principal component analysis (PCA) was performed to examine the intrinsic variation in groups and then partial least squarediscrimination (PLS-DA) was used to maximize the separation between healthy, contacts and patients subjects.

The quality of the models was described by R^2 and Q^2 values. The variable importance of projection (VIP) scores obtained from the PLS-DA model was used to identify key metabolic features significant for each group of patients, contact and healthy subjects. The metabolites identified by PLS-DA were subjected to biomarker analysis tool, to evaluate them as distinctly different metabolites between groups. Quantitative pathway enrichment module on MetaboAnalyst online platform was used to identify enriched pathway for each category. For biomarker analysis, biomarker analysis module was selected in MetaboAnalyst online platform, For each metabolite module generate receiver-operating-characteristic curves (ROC) was generated, this were calculated and expressed as areas under the curve (AUC), with an asymptotic 95% confidence interval (CI). ROC curve based model creation and evaluation was done to create biomarker model using selected features by using PLS-DA algorithm. Further, validation was carried out with the selected features, to assess whether they could be used as potential biomarker for active TB or latent TB.

Chapter 3

Comparative and quantitative metabolite profiling of non-pathogenic and pathogenic mycobacteria under microbicidal stresses

3.1 Overview of the chapter

Mycobacterial species, especially a successful pathogen like *Mycobacterium tuberculosis*, is known for its adaptability to various microbicidal environments. I hypothesized that the key to the survival of mycobacteria in microbicidal environment is a two phase adaptation process, where subtle plastic phenotypic changes through alterations in metabolic flux will precede the long term adaptation through changes at protein, transcription or genetic level. Therefore studying the changes in metabolic profile of mycobacteria soon after exposure to microbicidal stresses will reveal early adaptation processes of these mycobacterial species.

Under this objective, I studied the metabolic changes in *Mycobacterium smegmatis* (soil mycobacteria, non-pathogenic) and *Mycobacterium tuberculosis* (pathogenic) through steady-state metabolic profiling to understand the early adaptive changes in mycobacteria in response to four different microbicidal stresses. The stress conditions included under the study are acidic, oxidative, iron deprivation and nutrient starvation. Each of these stress conditions was studied using well-established *in-vitro* models (detailed in Chapter 2).

The early adaptive responses to stresses by *Mycobacterium smegmatis* and *Mycobacterium tuberculosis* (*M. tb*) have been studied independently keeping in mind the specific nature of environmental stresses that these bacilli face and differences in the time duration that may define 'early adaptation' for these two mycobacterial species. The stress was given for 4 hours for *M. smegmatis* and 36 hours for *M. tb*. Each of the stress condition was replicated as independent experiments for 8-10 times to obtain statistically significant number of samples while averaging out data within a stress condition and ensure reproducibility. Every sample was checked for contamination using Ziehl-Neelsen (ZN) staining. Following stress, the intracellular metabolites were extracted and subjected to Nuclear magnetic resonance spectroscopy (NMR) or liquid chromatography–tandem mass spectrometry (LC-MS/MS) with multiple reaction monitoring (MRM) (LC-MRM) for identification and quantification in combination with multivariate data analysis. The methodology / protocols have been detailed in Chapter 2.

The chapter has two parts:

A. Metabolic profiling of *M. smegmatis* (MC^2155)and major observations

B. Metabolic profiling of *M. tb* (H37Rv) and major observations

Some of the recent metabolomics studies on *M. smegmatis* include the impact of drugs like rifampin and capreomycin (Man *et al.,* 2018) or pretomanid (Baptista *et al*. 2018). Yet another study used comparative metabolomics of egt1/2 (involved in oxidative stress responsive Ergothioneine and Selenoneine biosynthetic pathway) mutants (Pluskal *et al*. 2014) or changes in the metabolism of early lag to late stationary phase/beginning of non-replicating phases (Drapal *et al*. 2014). Recent metabolomics studies on *M. tb* explored the effect of drugs colist in methanesulfonate and its mechanism (Koen *et al*. 2018). Another study showed that inhibition of glutamate synthtase (GltB/D) enhances the level of propionate toxicity in *M. tb* (Lee *et al*., 2018). While another study used an *mpr* mutant (persistence regulator) to investigate the role of iron deprivation in the persistence of *M. tb* (Kurthkoti *et al*. 2017). A recent study by Zimmermann *et al* explored the metabolic interaction between the host and *M .tb* by integrating dynamic metabolomics and dual RNA-seq data during early infection where they identified subnetworks that were differentially active during infection and also revealed a multiple-nutrient strategy adopted by bacilli during early infection (Zimmermann *et al*. 2017). The present study is the first comprehensive *in vitro* cataloguing of differential metabolites spanning four microbicidal stresses for *M. tb* (H37Rv) and three microbicidal stresses for *M. smegmatis* to understand the early adaptive changes.

PART 3A

Metabolic profiling of *M. smegmatis* (MC^2155) and major observations

3.2 PART A: Metabolic profiling of *M. smegmatis* (MC^2155) and major observations

The study was performed by subjecting *M. smegmatis* (MC^2155) to three stresses; acidic stress (pH 5.5) (Piddington *et al*. 2000), oxidative stress (10 mM H_2O_2) (Voskuil *et al*., 2011) in Sauton's minimal medium and nutrient starvation stress in 1X phosphate buffered saline (Loebel *et al.*, 1933). The stress was given for 4 hours. Each of the stress condition was performed 10 independent times. Every sample was checked for contamination using Ziehl-Neelsen (ZN) staining. Following stress, intracellular metabolites were extracted and subjected to untargeted metabolomics using Nuclear Magnetic Resonance spectroscopy (NMR). Please refer Chapter 2 for detailed methodology and protocols.

3.2.1 Results

3.2.1.1 Metabolite identification and chemical shift assignment

All the ^{1}H NMR spectra were manually phased and baseline-corrected using Topspin (v3.5) software (www.bruker.com/bruker/topspin). ^{1}H chemical shift dimension was directly referenced to the DSS (4, 4- dimethyl-4-silapentane-1-sulfonic acid) resonance. For TOCSY, prior to Fourier transform, the FIDs were weighted in both dimensions by a sine-bell function and zero-filled to 2048 and 1024 data points in F1 and F2 dimensions respectively. The distinguishing chemical shifts in the frequency domain that originated from the sets of spectral measurements were allocated to particular metabolites. Metabolite resonances present in the ^{1}H NMR spectrum (**Figure 3**) were identified using Chenomx NMR Suite 8.1 software and confirmed with biological magnetic resonance data bank (BMRB) database (Ulrich *et al*., 2008) and human metabolome database (HMDB) (Wishart *et al*., 2007). Further confirmation of metabolites was achieved using 2D total correlation spectroscopy (TOCSY) *via* semi-automated software – MetaboMiner (Xia *et al*., 2008). About 56 abundant metabolites were assigned using above strategy and are marked on the ^{1}H NMR spectrum as shown in **Figure 3.** Five metabolites (GTP, CDP, tryptophan, fructose-1-6 biphosphate, fumarate,) were excluded as their representative peaks were not clear in all spectra**.** All the 50 (excluding the internal standard, DSS) metabolites and their respective ^{1}H chemical shifts (in reference to DSS) are listed in **Table 5.**

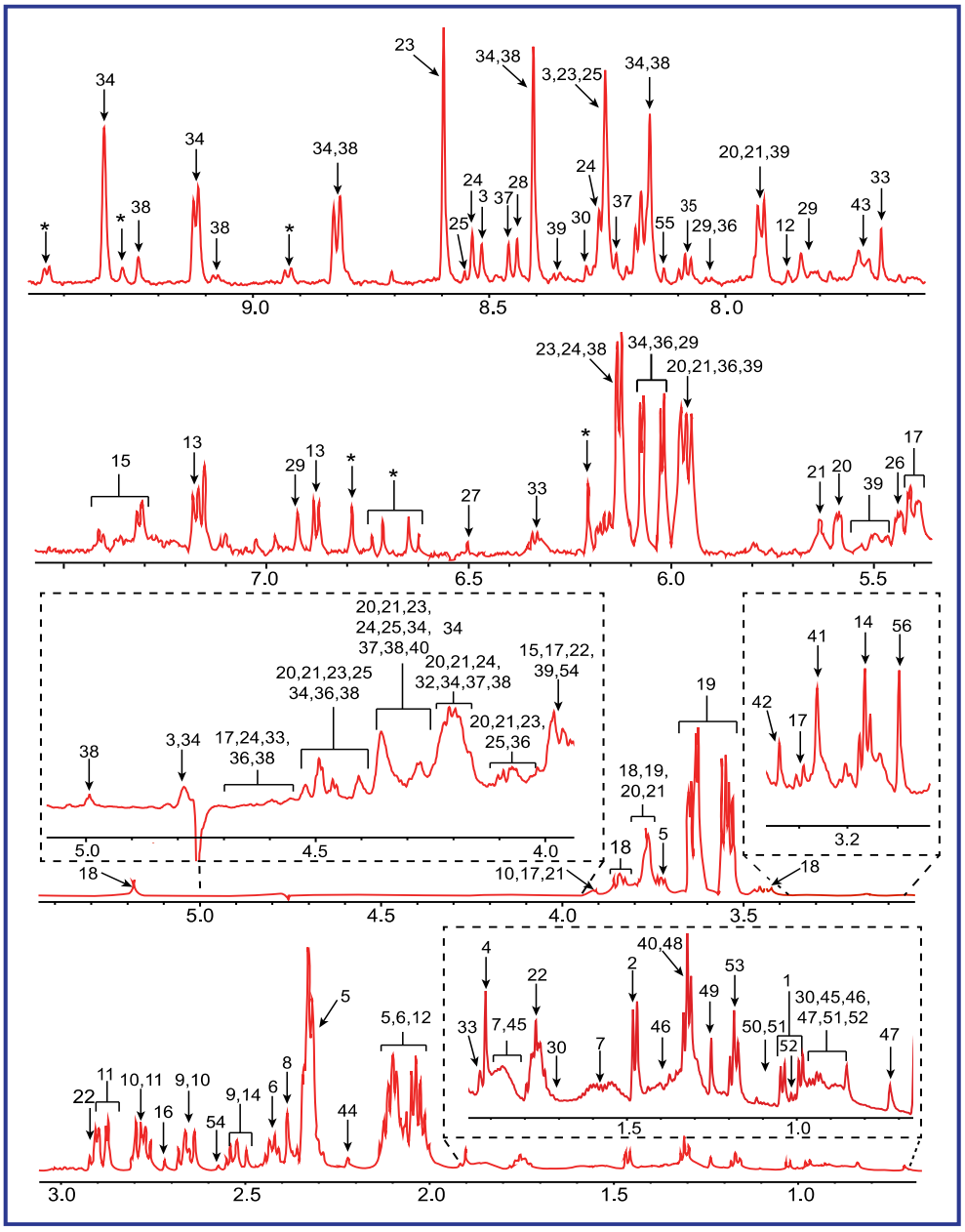

Figure 3: **Representative ^{1}H NMR spectrum of *Mycobacterium smegmatis.*** The ^{1}H NMR spectrum was obtained for each sample as described.Key: 1, Valine; 2, Alanine; 3,ATP; 4, Acetate; 5, Glutamate; 6, Glutamine; 7, Citrulline; 8, Succinate; 9, Citrate; 10, Aspartate; 11, Asparagine; 12, Homoserine; 13, Tyrosine; 14, beta-alanine; 15, Phenylalanine; 16, Dimethylamine; 17, Maltose; 18, Trehalose ; 19, Glycerol; 20, UDP-glucose; 21, UDP-galactose; 22, DSS; 23, AMP; 24, ADP; 25, IMP; 26, Glucose-1-phosphate; 27, fumarate ; 28, Formate; 29, CDP; 30, Leucine; 31, Lysine; 32, Fructose 1-6, bisphosphate; 33, dTTP; 34, NAD+; 35, N-acetyl glucosamine; 36, UMP; 37, NADPH; 38, NADP+; 39, UDP-N-acetylglucosamine; 40, Threonine; 41, Betaine; 42, Methanol; 43, Tryptophan; 44, Acetone; 45, 2-aminobutyrate; 46, Caprate; 47, Cholate; 48, Lactate 49, 3-hydroxyisovalerate; 50, 3-methyl-2-oxovalerate; 51, 2-hydroxy-3-methylvalerate; 52, Isoleucine; 53, Ethanol; 54, Methylamine; 55, GTP; and 56, Malonate.

3.2.1.2 Statistical analysis to identify metabolic signatures associated with respective stress (PCA and PLS-DA modeling)

To assess the influence of acidic, oxidative and nutrient starvation stresses on *M. smegmatis*, the data matrix obtained from each of the 10 experimental replicates of each condition were analyzed using univariate and multivariate analysis (Xia *et al.*, 2012). The unsupervised multivariate analysis was done by principal component analysis (PCA). For the univariate analysis, fold changes were calculated and t-test was performed to identify the differences between all the three stresses as compared to control. The values of each metabolite are tabulated in **Table 6.** A heat map was generated to depict the graphical representation of individual metabolite levels in different stress conditions in replicates as indicated in **Figure 4**, suggesting that metabolite levels in different samples change that in turn, are visualized as a color spectrum.

Overall differential metabolites observed were 31, 20 and 47 for acidic, oxidative and starvation stress respectively as compared to normal growth conditions with cutoff of fold change (FC >1.2), p-value ($P<0.05$) and false discovery rate ($FDR<0.05$) as shown in **Table 6** (A, B and C for acidic, oxidative and nutrient starvation stresses respectively), highlighting the differences at metabolite levels among the respective stresses. Principal component analysis (PCA) was then performed to examine the intrinsic variation in groups. The total variance explained by five components PCA analyses were 94.1% for acidic stress, 80.1% for oxidative stress and 94.4% for nutrient starvation stress. The figure 2 represents divergent separation on the score plot of the first two principal components PC1 and PC2. Percentage variance explained by PC1 and PC2 for respective stresses are mentioned in their respective plots **(Figure 5).** While the 3D plot **(Figure 5D**) showed stresses exposed samples were clearly separated from each other.Segregation indicates the metabolic differences in control and stress exposed samples. The 2D-PCA score plot revealed that the maximum discriminatory features are in control vs starvation (**Figure 5C)** compare to control vs acidic stress and control vs oxidative stress. PCA score plot represents a specific individual sample, and samples with similar metabolic profiles are grouped together in clusters.

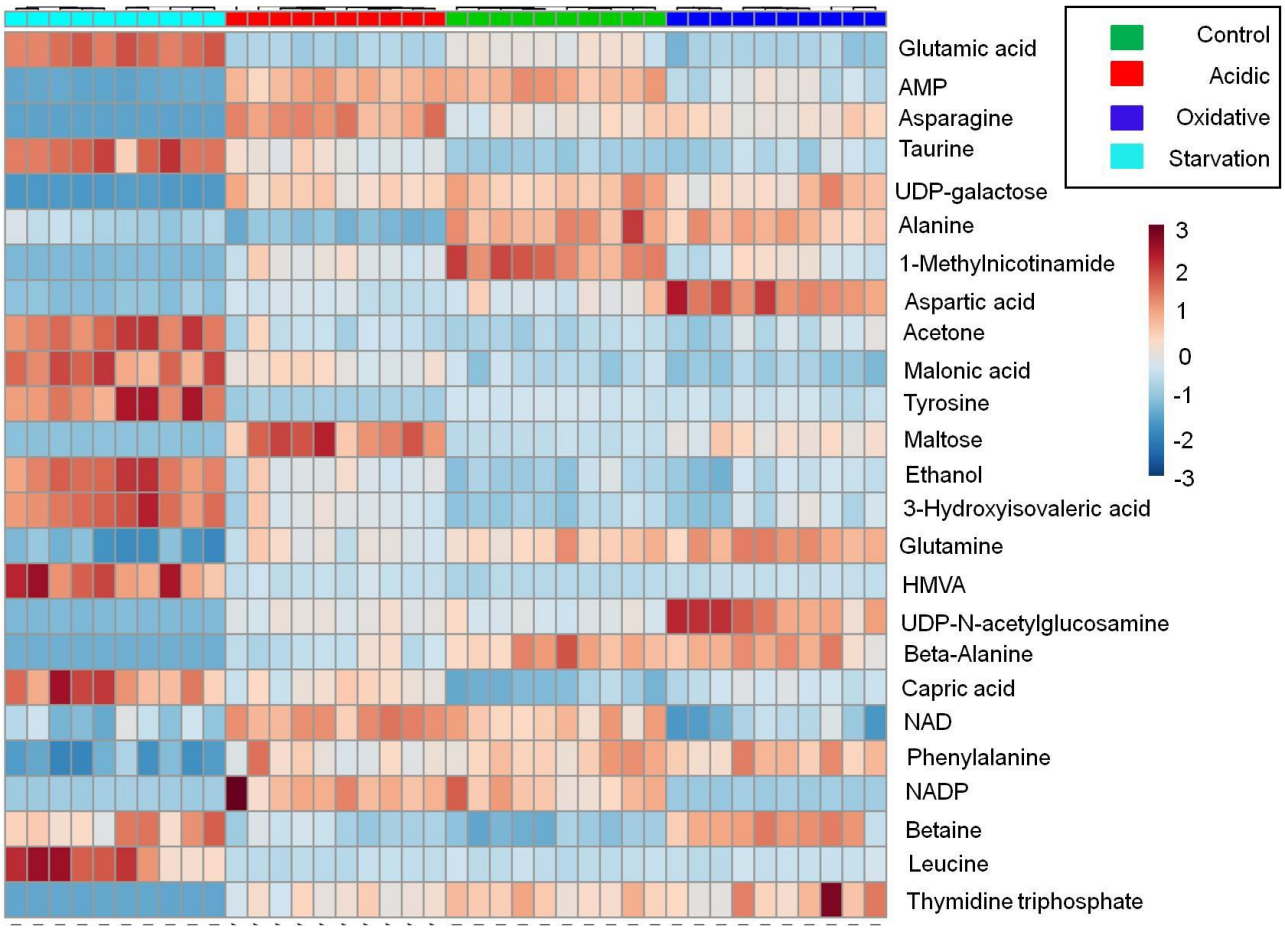

Figure 4: Heat map representation of the metabolic changes in *Mycobacterium smegmatis* upon exposure to microbicidal stress

Microbicidal stress induced by acidic, oxidative and nutrition starvation stress cause metabolic changes in *Mycobacterium smegmatis*. Each row represents a single metabolite detected in the study. Colordifferences demonstrate the relative concentration of metabolites acrossthe different conditions and experiment groups. From each group, 8 to 10 biological samples were analyzed.Individual samples are placed on horizontal axis and metabolites were placed on vertical axis.

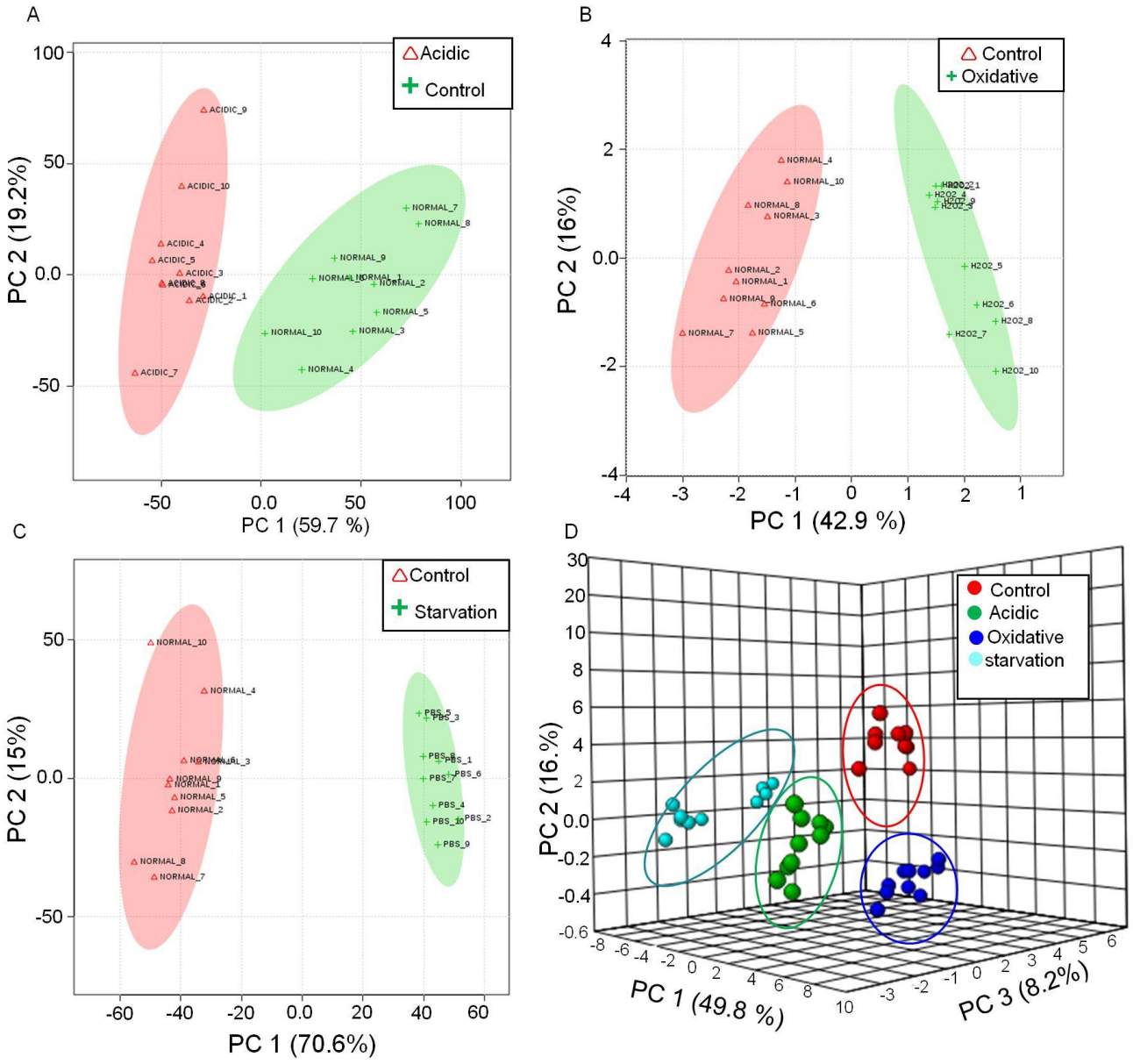

Figure 5: Principal Component Analysis segregates metabolites from different stress conditions in *M.smegmatis.* PCA analysis differentially segregates metabolites from acidic stress, oxidative stress, nutrient starvation and control: Principal Component Analysis (PCA) 2D score plot of A) Acidic *vs* Control. B) Oxidative *vs* Control. C) Nutrient starvation *vs* Control. D) 3D score plot of acidic stress, oxidative stress, nutrient starvation and control, suggesting all groups could be distinctly categorized.The dots inside the all the plots correspond to biological replicates under each category.

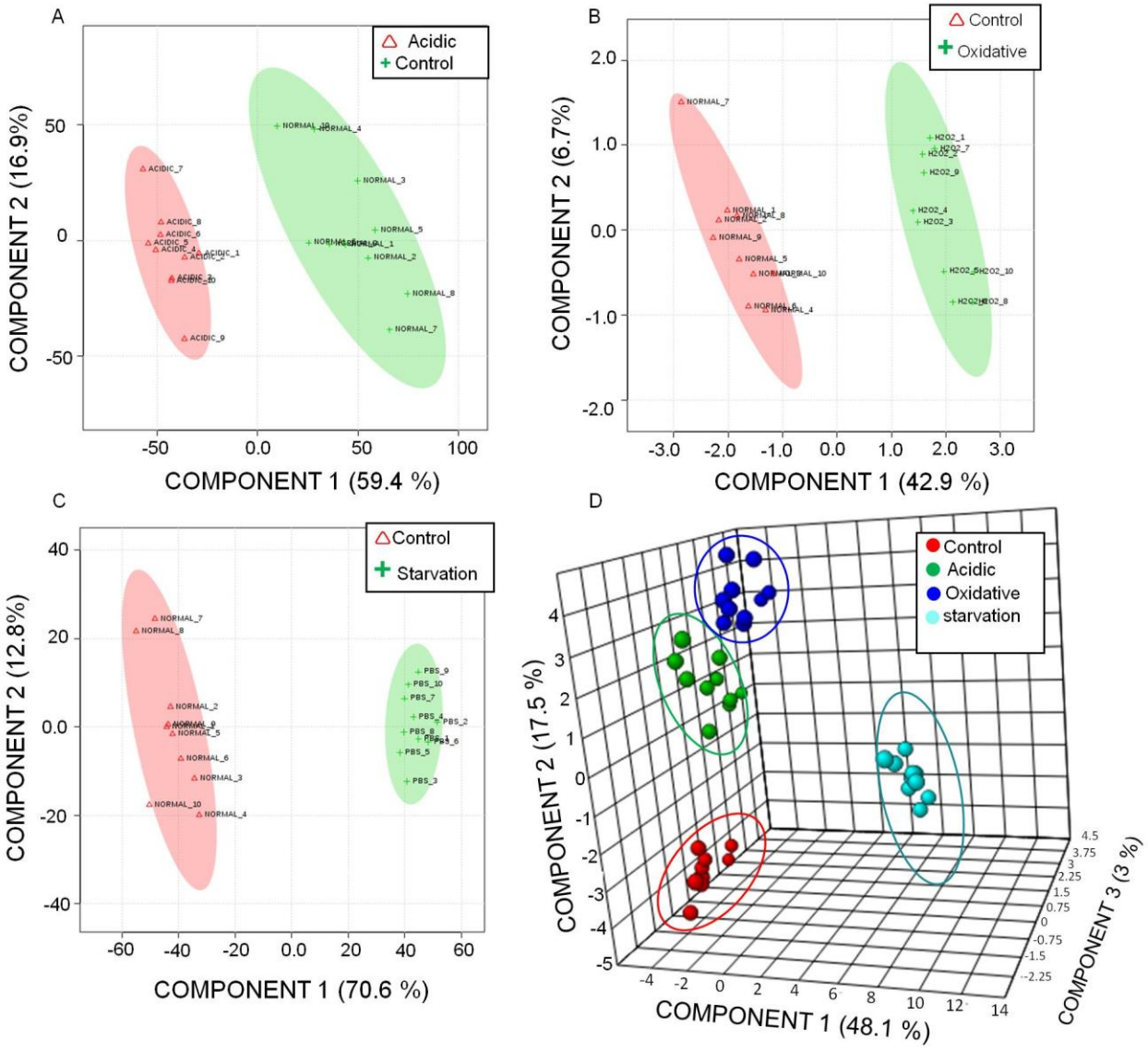

Figure 6: Partial Least Squares - Discriminant Analysis (PLS-DA) 2D score plot in *M.smegmatis*. A) Acidic *vs* Control. B) Oxidative *vs* Control. C) Nutrient starvation *vs* Control. D) 3D score plot ofacidic stress, oxidative stress, nutrient starvation and control.

PLS-DA shows further segregation between the groups analyzed under A, B and C as compared to PCA analysis. When all the groups were analyzed together (D) showed maximum segregation of nutrient starvation from acidic and oxidative stress. The dots inside the all the plots correspond to samples numbers under each category.The separation between control and stress groups were maximized when partial least square discrimination (PLS-DA) was subsequently performed (**Figure 6**). Unlike PCA, PLS considers the groups and seeks to reduce the dimension while maximizing the separation of the groups, it has higher discriminatory potential ideal for classification of groups. Clear discrimination between control, acidic stress, oxidative stress and nutrient starvation, could be observed (**Figure 6D).** The quality (goodness of fit) of the PLS-DA models was described by R^2 and Q^2 values which are given in table (**Table 7).** The R^2 value indicated goodness of fit, and Q^2 which indicates the goodness of predictability. R^2 measures the strength of the least-squares fit to the training set activities, while Q^2 is the R^2 value that one gets from applying Quantitative structure–activity relationship (QSAR) model to the test set instead of the training set (Szymanska *et al*. 2012). The model has good discriminatory abilities of the model if the values are closer to 1. The **table 7** clearly shows all stress conditions can be distinctly discriminated from control.

Conditions	**Measure**	**1 comps**	**2 comps**	**3 comps**	**4 comps**	**5 comps**
	R2	0.90	0.96	0.97	0.98	0.99
Acidic-Control	Q2	0.87	0.92	0.92	0.86	0.78
	R2	0.95	0.98	0.99	0.99	0.99
Oxidative-Control	Q2	0.92	0.92	0.93	0.92	0.88
	R2	0.99	0.99	1.00	1.00	1.00
Starvation-control	Q2	0.98	0.97	0.98	0.98	0.98
	R2	0.80	0.95	0.98	0.99	0.99
Acidic-Oxidative-Starvation-Control	Q2	0.77	0.93	0.95	0.97	0.97

Table 7: R^2 and Q^2 values of Partial Least Squares - Discriminant Analysis (PLS-DA)
Comps: component

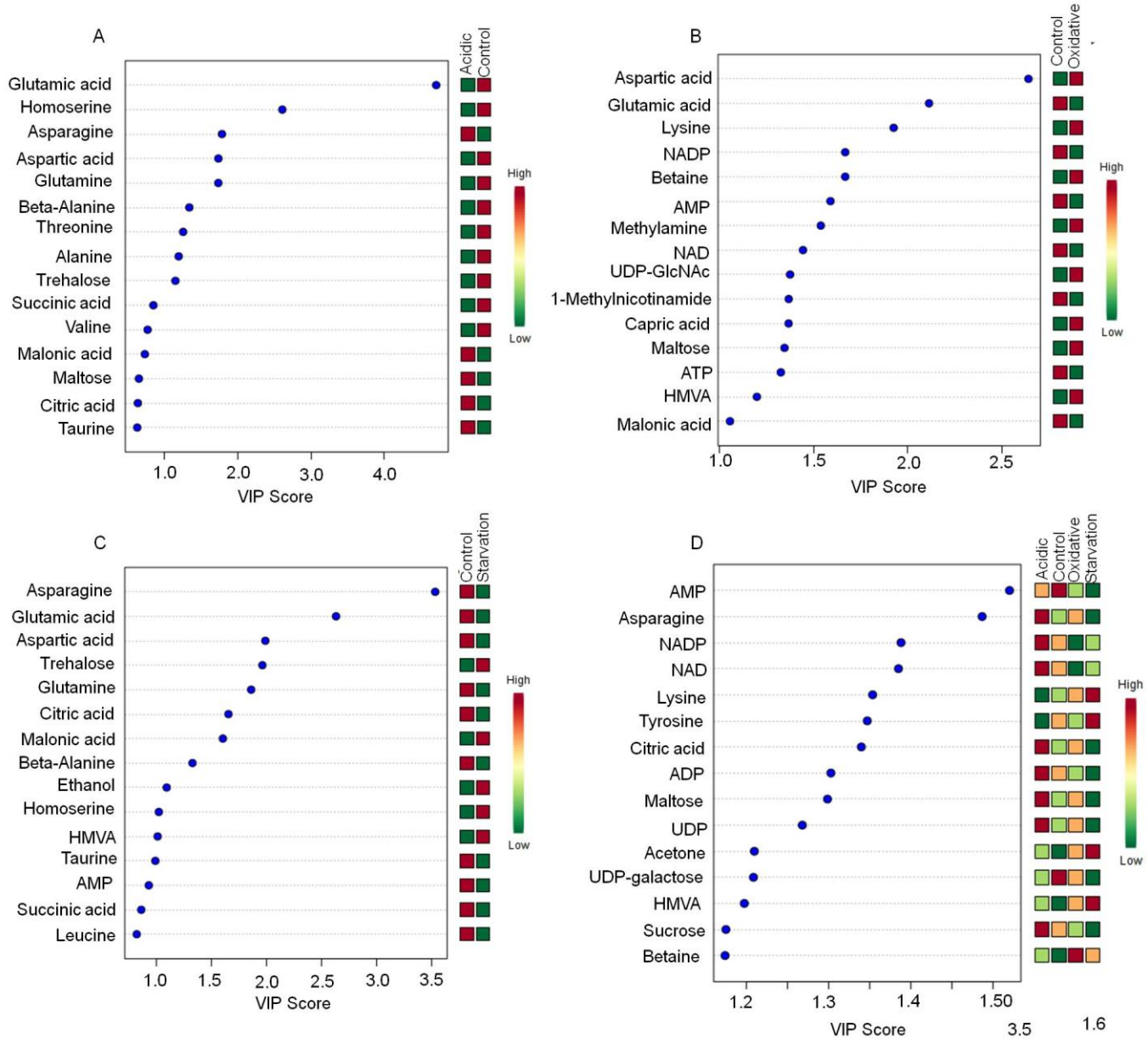

Figure 7: Variable Importance in Projection (VIP) displays the most discriminating metabolites in *M.smegmatis*. VIP plots indicating the most discriminating metabolitewere identified through PLS-DA analyses, indescending order of importance for A) Acidic *vs* Control. B) Oxidative *vs* Control. C) Nutrient starvation *vs* Control. These important variables are responsible for the segregation among the groups as indicated.

The variable importance of projection (VIP) scores obtained from the PLS-DA model was used to identify key metabolic features significantly distinct in each stress condition acidic*vs* control (**Figure 7A**), oxidative *vs* control (**Figure 7B**) and nutrient starvation *vs* control (**Figure 7C**). Metabolites differentiating various conditions with their VIP scores are tabulated as **Tables 6 (A, B and C)**.The analyses showed distinct differences in control as compared acidic, oxidative stress and nutrient starvation. Based on VIP score of more than 1, common metabolites distinguishing acidic stress, oxidative stress, nutrient starvation and control are AMP, asparagines, NADP, NAD, lysine, tyrosine, citric acid, ADP, maltose, UDP, acetone, UDP-galactose, 2-Hydroxy-3-methylpentanoic acid (HMVA), sucrose and betaine. It was observed that metabolites that scored high for discriminating control from acidic, oxidative stress and nutrient starvation majorly fall in energy metabolism and amino acid metabolism. Not surprisingly, metabolites ADP, AMP, NAD and NADP scored high as for distinguishing stress conditions from control. As per these scores, distinguishing metabolites for acidic *vs* control were glutamic acid, homoserine, asparagine, aspartic acid, glutamine, beta-alanine, threonine, alanine and trehalose, suggesting major perturbations in amino acid metabolism.This was similar to nutrient starvation vs control, where the major differential metabolites were asparagine, glutamic acid, aspartic acid,trehalose, glutamine,citric acid, malonic acid and beta-alanine.Important differential metabolites specific for oxidative *vs* control included aspartic acid, glutamic acid, lysine, NADP, betaine, AMP, methylamine, NAD, UDP-GlcNAc,1-Methylnicotinamide, capric acid, maltose, ATP and HMVA.

3.2.1.3 Pathway analysis

While the above section scores some of the critical metabolites that are distinctly differential between control and stresses, the metabolic pathway impact analysis was conducted considering all the metabolites to identify relevant pathways perturbed under each stress conditions. List of pathways perturbed in each condition was based on impact factor (> 0.1) p-value ($p<0.05$) and false discovery rate (FDR) (FDR<0.05) **(Table 8)**. It could be observed that across the different stress conditions, certain pathways appear in common. However, the significance of each pathway on the given stress condition seems to be different as reflected by p-value. In the acidic vs control,most significant pathway perturbed based on highest -log (P) (the negative natural log of the P value) are glycine-serine-threonine metabolism and glutamine-glutamate metabolism.In

the oxidative vs control, pathways perturbed with significant scoreswere nicotinate-nicotinamide metabolism andglycine-serine-threonine metabolism. For nutrient starvation condition the major pathway perturbed was alanine-aspartate-glutamate metabolism and beta-alanine metabolism. In acidic and oxidative stresses,bacilli adapt to maintain homeostasis of pH and redox state respectively, as it is essential for bacilli survival.Thus glycine-serine-threonine metabolism can counter both the stress by utilizing threonine that can be metabolized to glycine and further it can move through glycine cleavage system where glycine gets catabolized to release carbon dioxide, ammonia and reducing equivalents (NADH) (Sharma *et al*., 2016). While glutamate-glutamine metabolism pathway counters acidic and oxidative stress respectively by metabolizing glutamine to glutamate and releasing ammonia and by glutamate synthatase by metabolizing glutamate to oxoglutarate to release ammoniaandreducing equivalents (NADH) (Gallant *et al*., 2016), thereby helping bacteria adapt to both acidic and oxidative stress. During nutrient starvation, bacilli utilize intracellular amino acids as a source of alternate to carbon through anapleurotic reactions. Beta-alanine metabolism suggest utilization of β-alanine to form malonate (our data shows accumulation of malonate) which can be converted to malonyl-CoA to and enter fatty acid metabolism.

All tables referred in this section are annexed at the end of this section before the results and discussions on data obtained for H37Rv (Part B). Some major observations / inferences drawn from the above metabolomics analysis are discussed below followed by enlisting some of the leads from this study.

3.2.2 Discussion

3.2.2.1 Adenylate energy charge (AEC)

With biochemical energy as the central constituent to maintain appropriate turnover of biomolecules and functional metabolic viability, we began with looking at ATP, ADP and AMP levels during various stress conditions and calculating their ratios to reflect adenylate energy charge (AEC). Atkinson defined this ratio as ([ATP] X 0.5 [ADP]) / ([ATP] + [ADP] + [AMP])(Atkinson and Walton 1967). A living cell is known to uphold AEC within fine physiological values, inspite of large fluctuations in intracellular adenine nucleotides concentrations (Hochachka and McClelland 1997; Lim *et al*., 2010; Ozalp *et al*., 2010; Boender

et al., 2011; Edwards *et al*., 2012; Ytting *et al*., 2012). The AEC values were about 0.560 when grown in Sauton's media, 0.496, 0.476 and 0.647 for acidic, oxidative and nutrient starvation stresses respectively. It was clear that though large fluctuations in the adenosine nucleotide concentrations were observed during microbicidal stresses in *M. smegmatis* (**Figure 8 A,B,C**), the AEC values were maintained close to that of normal, reflecting the organisms' ability to adapt and maintain homeostasis despite stress by changing energy dynamics. This implied that the conditions in which total intracellular metabolites of *M. smegmatis* were measured indeed represented metabolic adaptation to stress. AEC value for *M. smegmatis* grown in Sauton's media is not reported. Though we would have preferred observing an AEC value of atleast 0.7 during normal growth, we considered 0.560 reflecting basal homeostasis in our condition. Further, a low AEC for *M. smegmatis* grown in Sauton's media may also be because the *in vitro* growth conditions are much different from soil growth conditions for which it has evolved.

Low levels of ATP production during nutrient starvation was apparent and expected which also resulted in low levels of ADP and AMP accumulation **(Figure 8A, B and C)**. High ratios of ATP/ADP **(Figure 8D)** and ATP/AMP **(Figure 8E)** in all stresses as normalized to normal growth conditions mean that ATP generation is lower during all stresses, while utilization is more.

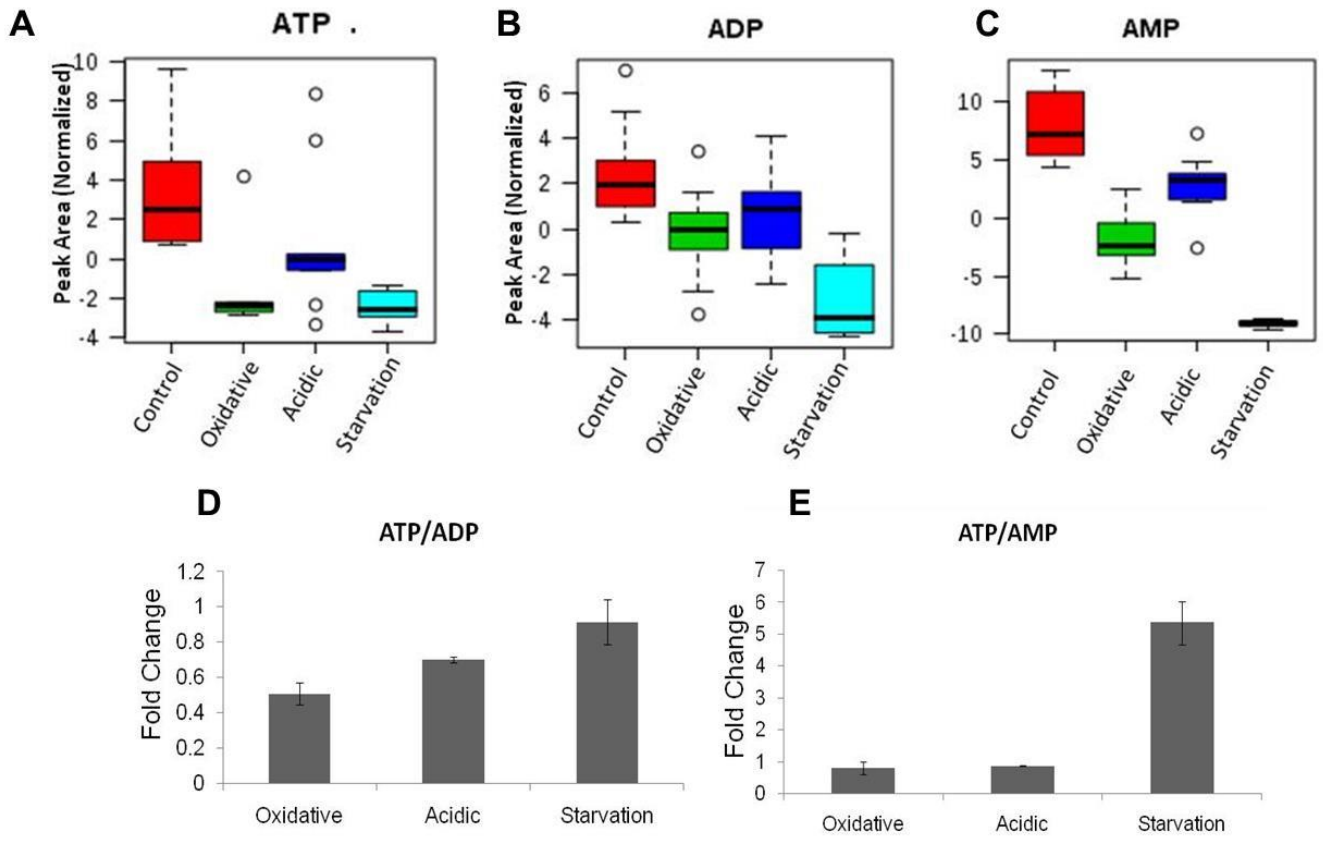

Figure 8: Adenosine nucleotide levels in *M. smegmatis* during microbicidal stresses. (A) ATP, (B) ADP and (C) AMP. (D) ATP/ADP and (E) ATP/AMP ratios in fold change normalized to the ratio in control growth conditions.

3.2.2.2 Osmolytes accumulation during oxidative stress and nutrient starvation: tracing the presence of a possible pathway of biosynthesis of methylated amines in *M. smegmatis* earlier reported in obligate methylotrophs

Metabolomics data showed high intercellular concentrations of osmolytes like betaine, methylamine and dimethylamine when *M. smegmatis* was subjected to stresses (**Figure 5A**). These osmolytes are known to play a role in adaptation to various stresses besides ionic imbalance (Whatmore *et al*., 1990; Holmstrom *et al*., 1994; Park *et al*., 1995; Yancey *et al*. 2002; Burg and Ferraris 2008). *M. tb* is reported to acquire betaine from macrophages to gain osmoregulatory advantage while colonizing the host (Price *et al*. 2008). Though methylamine, dimethylamine and betaine were identified in all the conditions, it was significantly ($P<0.05$) high when *M. smegmatis* was under oxidative and starvation stresses (**Figure 5A**). While many bacteria, including that in human microbiota can produce trimethylamines from carnitine (Zhu *et al*. 2014), the biosynthesis pathway of methylamines is not reported in *M. smegmatis*. With no direct supplements of Tri- or di-, or methylamine, we asked the question if orthologs of these genes are present in *M. smegmatis* and if they are expressed during stress conditions.

In methylotrophic bacteria, dimethylamine can be produced from trimethylamine by trimethylamine dehydrogenase and methylamine from dimethylamine using dimethylamine dehydrogenase. In certain bacteria, both the steps are mediated by single dehydrogenase alone whereas in others, each step is mediated by a distinct dehydrogenase. It is to be noted that both the dehydrogenases have similar physical, chemical, spectral and kinetic properties (Meiberg and Harder, 1978). An ORF/gene by this name could not be located within the annotated *M. smegmatis* genome (https://mycobrowser.epfl.ch/). A protein sequence similarity search using PSI-BLAST was then performed using the protein sequence of trimethylamine dehydrogenase (TMD) from methylotroph *Methylophilus methylotrophus* [(Bacterium W3A1), UniProt ID P16099] as the query sequence and MC^2155 genome as the subject. A sequence identity of 26% with query coverage of 72% was observed with MSMEG_5124 of *M. smeg*. MSMEG_5124 is annotated as 2,4-dienoyl-coA reductase (DCR), identified on the basis of protein family similarity (https://mycobrowser.epfl.ch/genes/MSMEG_5124). Although a limited similarity, a thorough literature search showed a reported case of similarity of TMD of *Methylophilus*

methylotrophus with DCR of *E. coli* (Hubbard *et al*. 2003). The authors solved the crystal structure of *E.coli* DCR and observed domain arrangement and overall polypeptide fold resembledthat of TMD. They also mention that superimposition of DCR with a monomer of TMD showed r.m.s deviation of ~1.5 Å over 467 Cα carbon atoms in primarily middle and N-terminal domains, concluding the high degree of structural similarity between DCR and TMD (Hubbard *et al*. 2003). With these clues, MSMEG_5124 was then aligned with *E.coli* 2, 4-dienoyl-CoA reductase (P42593) which showed 100% query coverage and an identity of 54%. This suggested that there is a possibility that though annotated as a DCR, MSMEG_5124 may also function as TMD for *M. smegmatis*. Homology modeling of MSMEG_5124 using TMD of *Methylophilus methylotrophus*, Bacterium W3A1 (Mm-TMD) showed overlapping catalytic sites as illustrated diagrammatically (**Figure 9C**). Cysteine residues for 4Fe-4S in Mm-TMD (CYS315, 351, 364) and MSMEG_5124 (CYS 335, 342, 354) are conserved. With this supporting information, I next checked if this probable ORF is expressed in *M. smegmatis* during both normal and stress conditions for which RT-PCR for MSMEG_5124 was performed using specific primers **(Figure 9B)**. The results showed that MSMEG_5124 is expressed in *M. smegmatis* during all conditions. However, no differences in the transcript levels amongst all the conditions also suggested that MSMEG_5124 is not regulated at transcription level for early adaptation to stress. With the confirmation of expression of MSMEG_5124, a possible TMD in *M. smegmatis*, and evidence of intermediate metabolites, di-methylamine and methylamine from the metabolomics data, suggested the presence of the pathway converting tri-methylamine to di-methylamine and methylamine.

Next, I wanted to know the possible intracellular source of tri-methylamine (TMA). Metabolomics data showed that betaine is accumulated during oxidative and starvation stresses. Reported in many bacteria, also those associated with gut microbiota, carnitine monoxygenase reductase subunit (YeaX) as complex with an oxygenase component [Rieske (2Fe-2S) (YeaW) region] of *E. coli* K-12 can use, carnithine, γ-butyrobetain, choline and betaine as substrates to produce trimethylamine (Koeth *et al*. 2014). I next looked for these genes in the genome of *M. smegmatis* MC^2155. With *E. coli* (P76254 and P0ABR7) genes as references, a scanning of MC^2155 using PSI-BLAST two orthologs, MSMEG_4371 and MSMEG_0657, with similarity to YeaX (38% identity) and YeaW (37% identity) respectively. MSMEG_4371 and

MSMEG_0657 are hypothetical ORFs with no experimental evidence of their expression in *M. smegmatis*. The expressions of MSMEG_4371 and MSMEG_0657 in *M. smegmatis* during both normal and stress conditions were checked using RT-PCR. Once again, these were observed to be expressed in all conditions and the differences in the intermediate metabolites of the pathway did not reflect in expression levels of these genes (**Figure 9B**) (Koeth *et al*. 2014) had further demonstrated that YeaW and YeaX interact with each other and participate as a complex in assimilation to TMA. We checked the same for MSMEG_4371 and MSMEG_0657 using STRING, which predicted their interaction (**Figure 9D**). With this, we have a possible pathway that can convert betaine into TMA and TMA to methylamines in *M. smegmatis*. It is to be noted that TMA can also be produced from carnitine, which can be synthesized from N^{ε}-trimethyllysine which comes from methylation of lysine (Borum and Broquist 1977). Hence, betaine, which is accumulated during oxidative and starvation stresses, is a possible intracellular source for synthesis of TMA, catalyzed by MSMEG_4371 and MSMEG_0657.

In addition, betaine can be synthesized from choline *via* betaine aldehyde (Landfald and Strom 1986). The possible annotated ORFs for the synthesis of betaine are MSMEG_5815, an ortholog of betaine aldehyde dehydrogenase and ORF MSMEG_3444, an ortholog of choline dehydrogenase. The expressions of these ORFs during normal and stress conditions were confirmed by RT-PCR. MSMEG_4371 and MSMEG_0657 in *M. smegmatis* during both normal and stress conditions were checked using RT-PCR. The existence of these ORFs at the transcript levels was evident by RT-PCR (**Figure 9B**).

With the evidence of expression all these hypothetical ORFs from *M. smegmatis* genome and presence of intermediate metabolites that are catalyzed by these ORFs, we have the evidence of existence of a pathway that can utilize choline to betaine to TMA to methylamines. Choline can be fed into the pathway from lipid metabolism (Barksdale and Kim 1977; Dhariwal *et al*. 1978; Landfald and Strom 1986). Methylamines have been shown to be utilized to synthesize formaldehyde and ammonia(Kim, Bae and Lee, 2001), which may serve as an advantage to pathogenic mycobacteria, especially during acidic stress. Formaldehyde can further be converted to formate. While released ammonia was separately measured we could not detect formaldehyde in our metabolomics data. However, we could detect formate which reflects utilization of

formaldehyde under stress conditions. This unreported pathway can be studied in detail for its significance in physiological adaptation to stress in *M. smegmatis*.

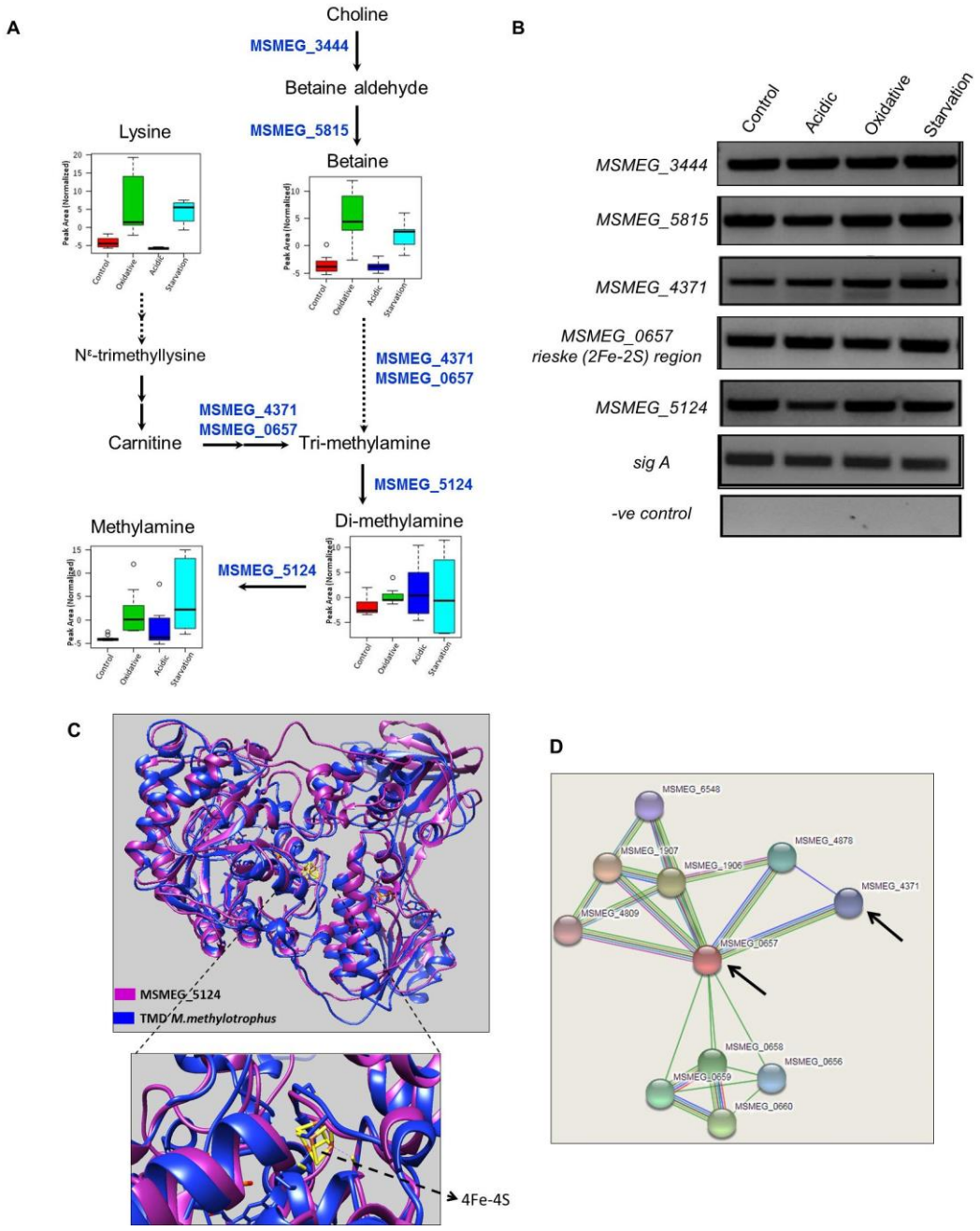

Figure 9: A possible pathway of biosynthesis of methylated amines in *Mycobacterium smegmatis* supported by *in silico* analyses and RT-PCR A) Possible methylamine pathway catalyzed by factors homologous to methylotrophic bacteria. B) RT_PCR oforthologue ORFs and their existence at transcript level in *M. smegmatis*. C) Homology modeling of MSMEG_5124 using TMD of Bacterium W3A1 (Mm-TMD) showed overlapping catalytic sites as illustrated. D) STRING analysis predicts that MSMEG_4371 and MSMEG_0657 interact (indicated with arrows).

3.2.2.3 Impact on α-Glucan biosynthesis: regulation at transcription level as an early adaptation event during stresses

Metabolomics data showed that metabolites associated with α-Glucan biosynthesis, such as trehalose, UDP-glucose, Glucose-1-phosphate, D-maltose etc, were differential during different stress conditions. We observed low levels of trehalose in all stresses as compared to control whereas an accumulation of maltose was observed in acidic and oxidative stresses. Additionally, it was observed thatboth trehalose and maltose levels were lowest during starvation as compared to all other conditions, which is suggestive of utilization of these metabolites during nutrient deprivation. The levels of these metabolites indicated perturbations in glucan biosynthesis. While the perturbations definitely can be due to differential enzymatic efficiencies as an early event during these stresses, the possible regulation at transcription levels was checked by RT-PCR. It was very apparent that the expression of genes for maltose kinase [*mak* (*pep2*)] and α-1,4-glucan branching enzyme (*glgB*) remained unchanged,while that for trehalose synthetase (*treS*) was downregulated in acidic stress and maltose transferase (*glgE*) was upregulated only under oxidative stress (**Figure 10B**). Since for trehalase (*treH*), the differences were very refined, we performed densitometric analysis which showed that its expression levels are lesser in oxidative stress as compared to others while in starvation it was marginally higher. With *glgE* levels higher and *treH* lower, it could be inferred that this pathway is oriented for biosynthesis of α-glucan towards capsule formation **(Figure 10C)**. In starvation stress, marginal increase in expression of *treH* points to utilization of trehalose to glucose as bacteria is deprived of nutrition. Acidic stress showed poor expression of *treS* and no increase in *glgE*, which may explain the accumulated maltose levels. *M. smegmatis treS* has been reported to have the property catalyzing reversible conversion of trehalose to maltose (Pan *et al*. 2004). With implication of these pathways in intracellular and capsular α-glucan synthesis (Sambou *et al*. 2008) their regulation may be critical in adaptation to different stress. This was the first time a regulation at transcription level was observed as an early adaptation event in these set of experiments.

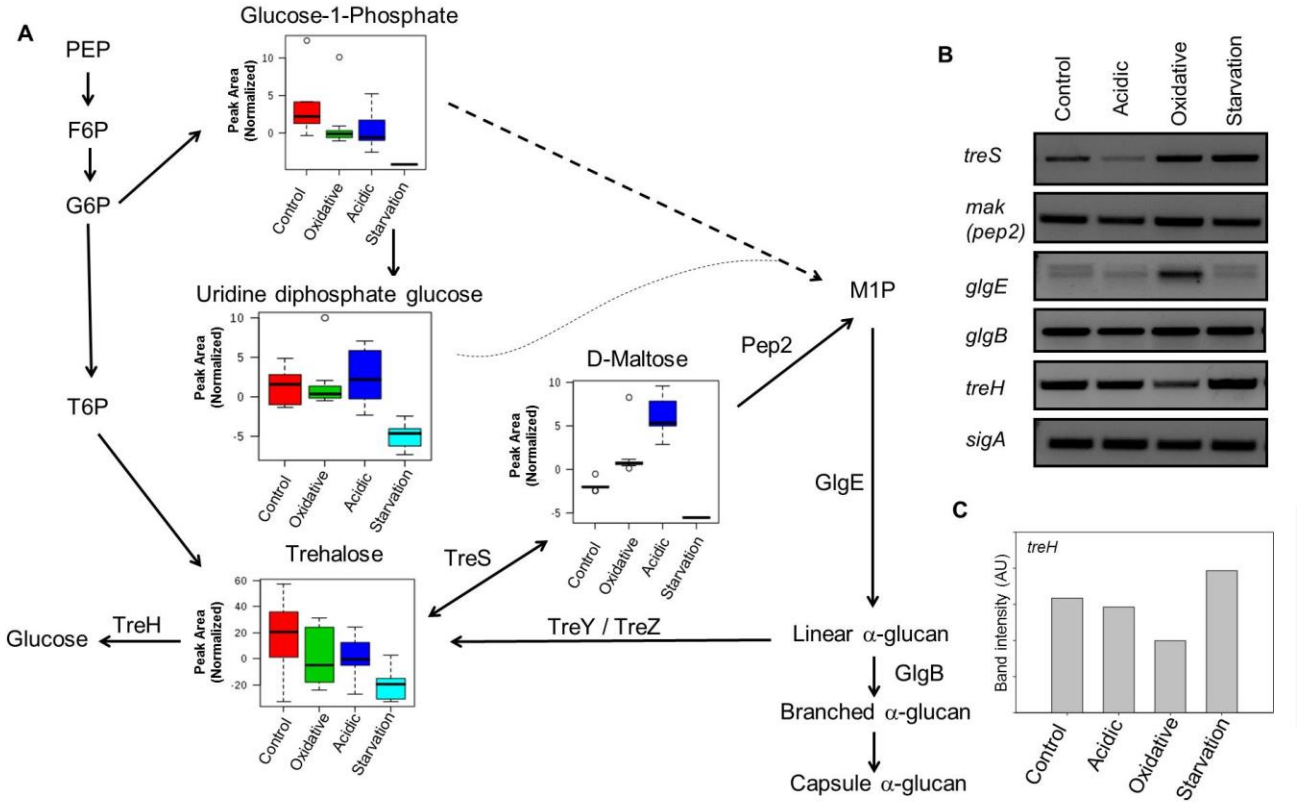

Figure 10: Metabolic pathways associated with α-glucan, GlgE and glycogen metabolism in *M.smegmatis*. A) α-glucan and GlgE pathway metabolites identified in the study. B) RT-PCR of genes taking part in α-glucan pathway and their possible regulation at transcription levels. C) Densitometric analysis of TreH.

3.2.2.4 Perturbations in amino acid metabolism: Asparagine accumulation correlated with ammonia release and *M. smegmatis* growth during acidic stress

The quantitative measurement of amino acids accumulated under various growth conditions clearly showed that amino acid metabolism and nitrogen assimilation pathways are perturbed during stresses **(Table 6 A, B and C for acidic, oxidative and nutrient starvation stresses)**. The figure below is a representation of amino acid biosynthesis pathway for amino acids that showed differential concentrations during different stresses in *M. smegmatis* **(Figure 11)**.

Asparagine was the only amino acid whose levels were high amongst different amino acids detected during early metabolic adaptation to acidic stress **(Figure 11)**. Branched chain amino acids (BCAA), including proteinogenic amino acids Leucine, iso-Leucine and Valine, were low in acidic stress condition. Though not reported in M.smeg, asparagine has been implicated in supporting *M. tb*survive acid stress (Song *et al*. 2011). One of the suggested mechanisms is that it gets assimilated by enzyme asparaginase, releasing ammonia to neutralize the acidic environment (Gouzy *et al*. 2014). Taking cues from these reported studies, the concentration of released ammonia in culture media **(Figure 12A)** and the changes in the pH of media post-stress as compared to pre-stress were measured (**Figure 12B**). While some basal level of NH_3 was released in control, it was ~ 2 fold more during acidic stress. NH_3 levels were negligible in oxidative and starvation conditions, which points to the possibly of its utilization as nitrogen source during these conditions. Interestingly we observed a rise in pH of the media post-stress only during acidic stress and in no other condition. With this, it could be postulated that *M. smegmatis* overcomes acidic stress by neutralizing the environment by releasing molecules like NH_3, a mechanism followed by many bacteria. One of the early adaptations to acidic stress may be through enzymatic regulation of asparagine catabolism. ORF MSMEG_3173 that codes for probable L-asparginase, annotated on the basis of identification by match to this protein family (https://mycobrowser.epfl.ch/) is a promising candidate for further experimental validation through KO studies to understand basic mycobacterial physiology. Such studies will be continued in lab.

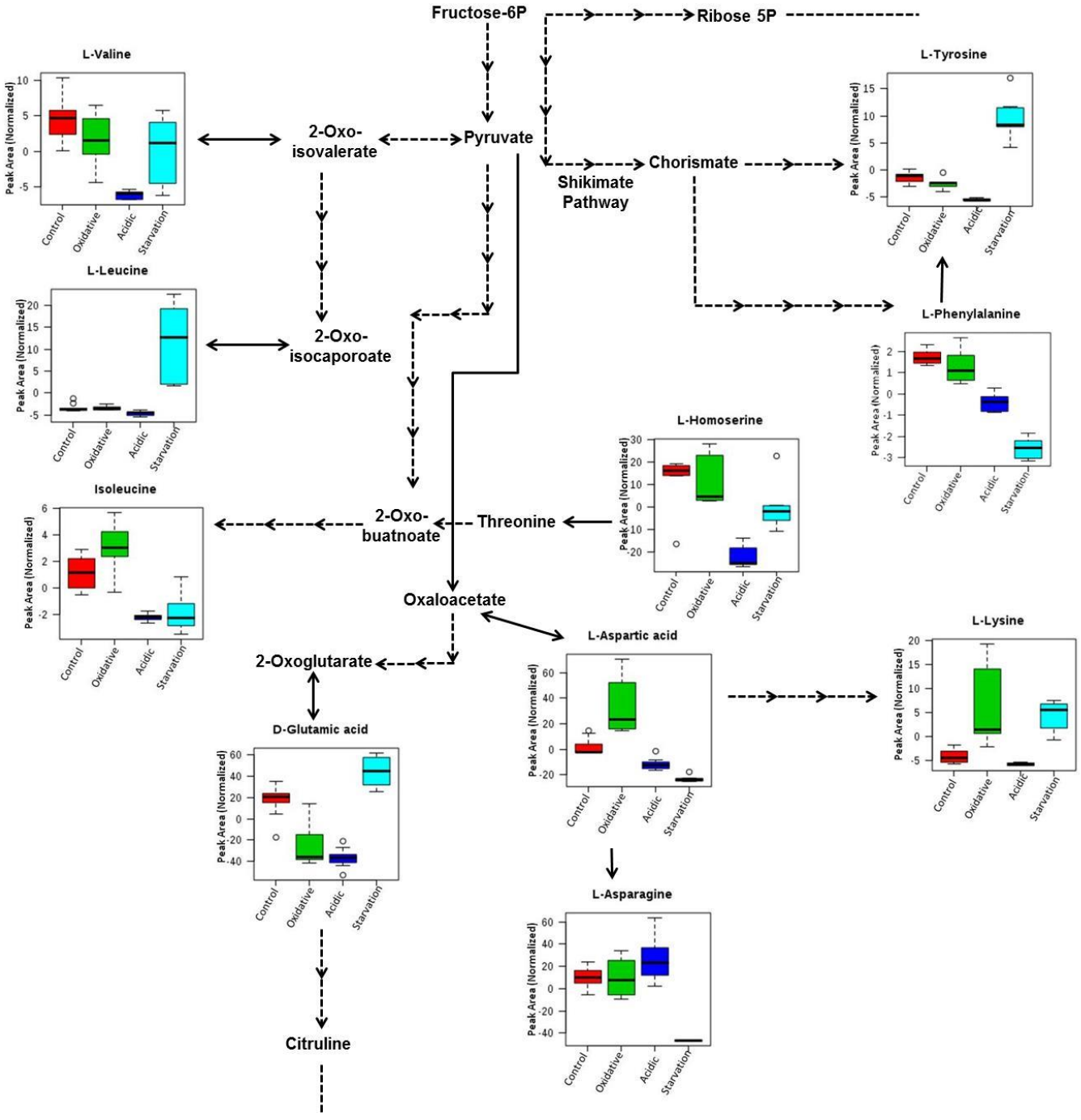

Figure 11: Network map (partial) of altered amino acids in the amino acid biosynthesis pathway during the different stresses in *M. smegmatis*. The amino acids identified in this study are displayed in the aminoacid biosynthesis pathway along with box plots illustrating their differential levels during stresses as normalized peak area.

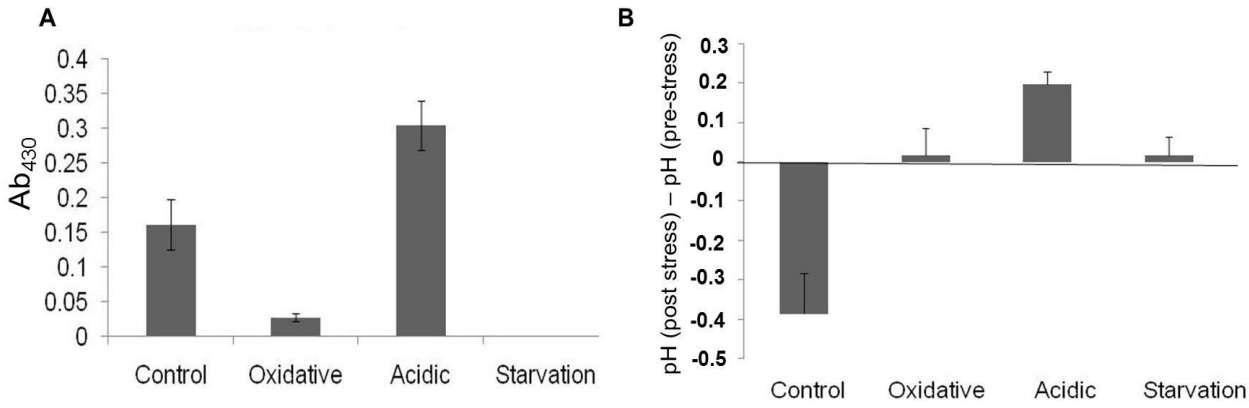

Figure 12: Estimation of ammonia release in *M. smegmatis* during different microbicidal stress conditions.
Determination of the levels of ammonia released in to culture media during different stress as measured by Nessler's reagent (refer chapter 2 for methodology) (A) and the changes in the pH of media post-stress as compared to pre-stress(B).

3.2.3 New observations and leads from this study-

1. The study has led to identification of a *hitherto* unreported possible pathway of intracellular biosynthesis of betaine, methylamine and dimethylamine.With the genes encoding the enzymes of these pathways transcribed during stresses, the physiological significance of this new pathway that can convert choline to betaine or carnitine to TMA to methylamines in adaptation to environmental microbicidal stresses in *M. smegmatis* can be studied further.

2. The study reports for the first time the differential regulation of α-Glucan biosynthesis at transcript levels of selected ORFs in *M. smegmatis* during different stress conditions. With implication of α-Glucan biosynthesis in capsule formation, the regulation of this pathway can be studied to understand critical adaptive nodes for different stress conditions in *M. smegmatis.*

Needless to say, the orthologues of these pathways can be traced in pathogenic mycobacteria and their significance may be studies to adapt to microbicidal stresses mimicking intracellular survival.

PART 3B

Metabolic profiling of *Mycobacterium tuberculosis* (H37Rv) and major observations

3.3 PART B: Metabolic profiling of *Mycobacterium tuberculosis* (H37Rv) and major observations

The study was performed by subjecting *M. tb* (H37Rv) to four stresses; acidic stress (pH 5.5) (Piddington *et al*. 2000), oxidative stress (10 mM H_2O_2) (Voskuil *et al*. 2011), iron deprivation (no iron supplemented) (Hall and Ratledge 1982) in Sauton's minimal medium and nutrient starvation stress in 1X phosphate buffered saline (Loebel *et al* 1933). The iron deprivation stress is referred to as iron stress throughout the text. The stress was given for 36 hours. Each of the stress condition was performed as 5 independent experiments with 2 technical replicates. Every sample was checked for contamination using Ziehl-Neelsen (ZN) staining. Following stress, intracellular metabolites were extracted and subjected to targeted metabolomics using LC-MRM/MS. Please refer Chapter 2 for detailed methodology and protocols.

The untargeted metabolomics study using NMR described in the previous section had two limitations (i) minimum culture volume required for effective metabolomics was 100ml for each experiment/replicate (a total volume of nearly 1 litre for each stress condition for 10 replicates was used), (ii) limited number of metabolites could be detected (about 55). Since H37Rv is a pathogenic strain, we had limitation in growing large scale cultures. For LC-MS, minimum culture volume required for effective metabolomics was only 10 ml for each experiment/replicate, and a total volume of 100 ml for each stress condition were sufficient (10 ml per experiment with 2 technical replicates). Each experiment was performed five times with two technical replicates. A standard of 108 known metabolites were used and we could identify nearly 88 metabolites. The metabolites used as standard with their retention time are tabulated as **Table 9**, which is annexed at the end of this chapter.

3.3.1 Results

3.3.1.1 Statistical analysis to identify metabolic signatures associated with respective stress (PCA and PLS-DA modelling)

To assess the influence of acidic, oxidative, iron and nutrient starvation stresses on H37Rv, all the data obtained from 5 replicates for each stress condition were analysed using univariate and

multivariate analysis. The unsupervised multi-variant analysis i.e principal component analysis (PCA) was performed for this set of data as well, which indicated metabolic differences in control *vs* respective stress exposed samples. For the univariate analysis, fold changes were calculated and t-test was performed, to identify differences between all the three stresses as compared to normal. The values of each metabolite are tabulated in **Table 10.**A heat map was generated to depict the graphical representation of individual metabolite levels in different stress conditions in replicates as indicated **(Figure 13)** suggesting that metabolite levels in different samples change that in turn, are visualized as a color spectrum. The total number of differential metabolites observed were 34, 6, 10 and 48 for acidic, oxidative, iron and nutrient starvation stress respectively as compared to normal growth conditions with a cut-off of fold change (FC >1.2), p-value (P<0.05) and false discovery rate (FDR<0.05) as shown in **Table 10 (Table 10A, 10B, 10C and 10D** for acidic, oxidative, iron and nutrition starvation stress respectively), highlighting the differences at metabolite levels among the respective stresses.

Principal component analysis (PCA) was then performed to examine the metabolic signatures associated with different stress conditions. The total variance explained by five components PCA analyses were 90.4% for acidic stress, 88.5% for oxidative stress, 91.8% for iron stress and 92.6% for nutrient starvation stress. **Figure 14** represents divergent separation on the score plot of the first two principal components PC1 and PC2. Percentage variance explained by PC1 and PC2 for respective stresses are mentioned in their respective plots. 3D visualization of PCA scores revealed that the maximum discriminatory features are in starvation and acidic stress; whereas metabolites from iron stress overlapped both with those of control, oxidative stress samples suggestion low discrimination between these groups.

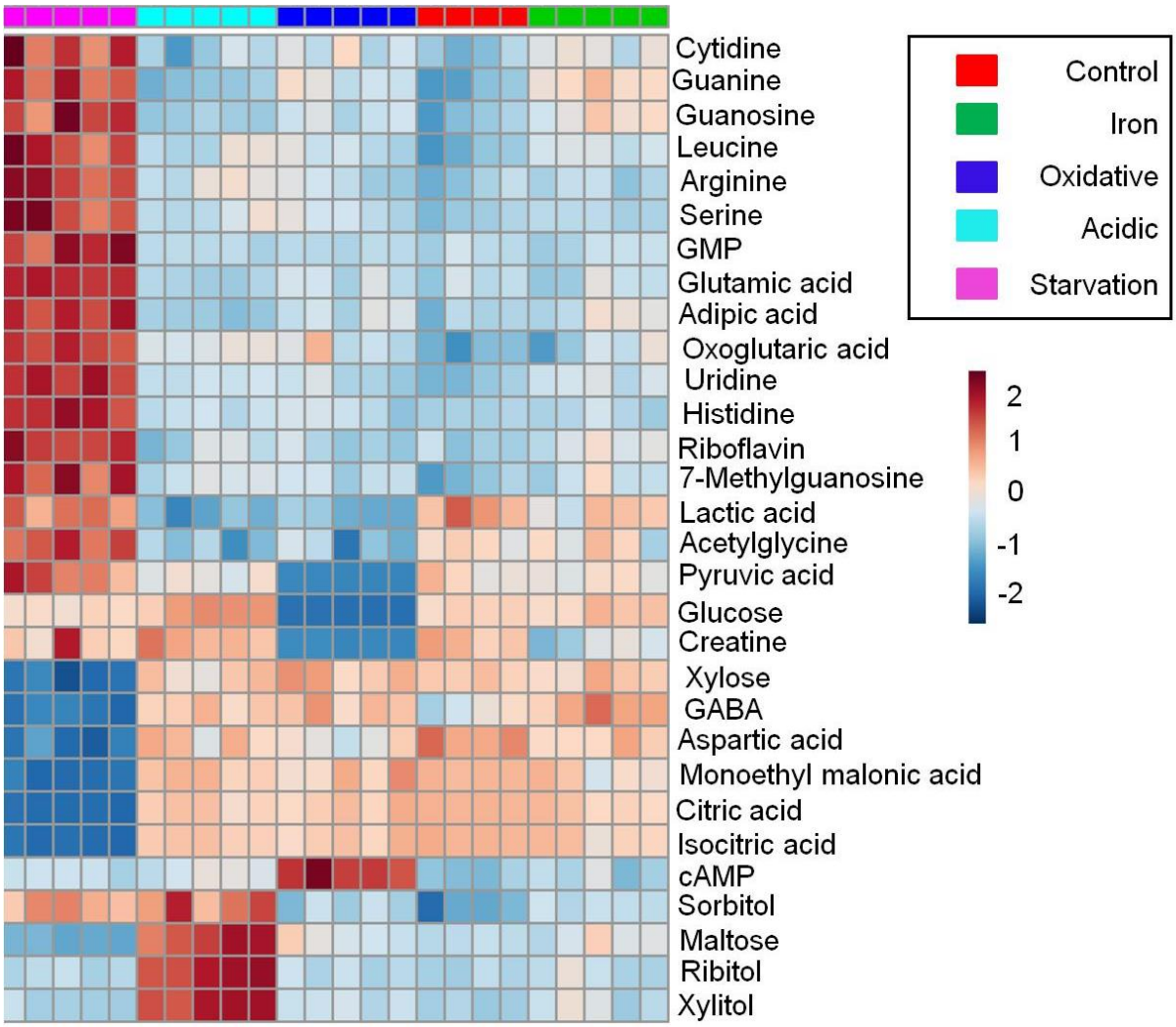

Figure 13: Heat map representation of the metabolic changes in *M. tb* induced by acidic, oxidative, iron and nutrition starvation stress.Each row represents a single metabolite detected in the study. Colour differences demonstrate the relative concentration of metabolites acrossthe different conditions and experiment groups. From each group, 4 to 5 biological samples were analysed. Individual samples are placed on horizontal axis and metabolites were placed on vertical axis.

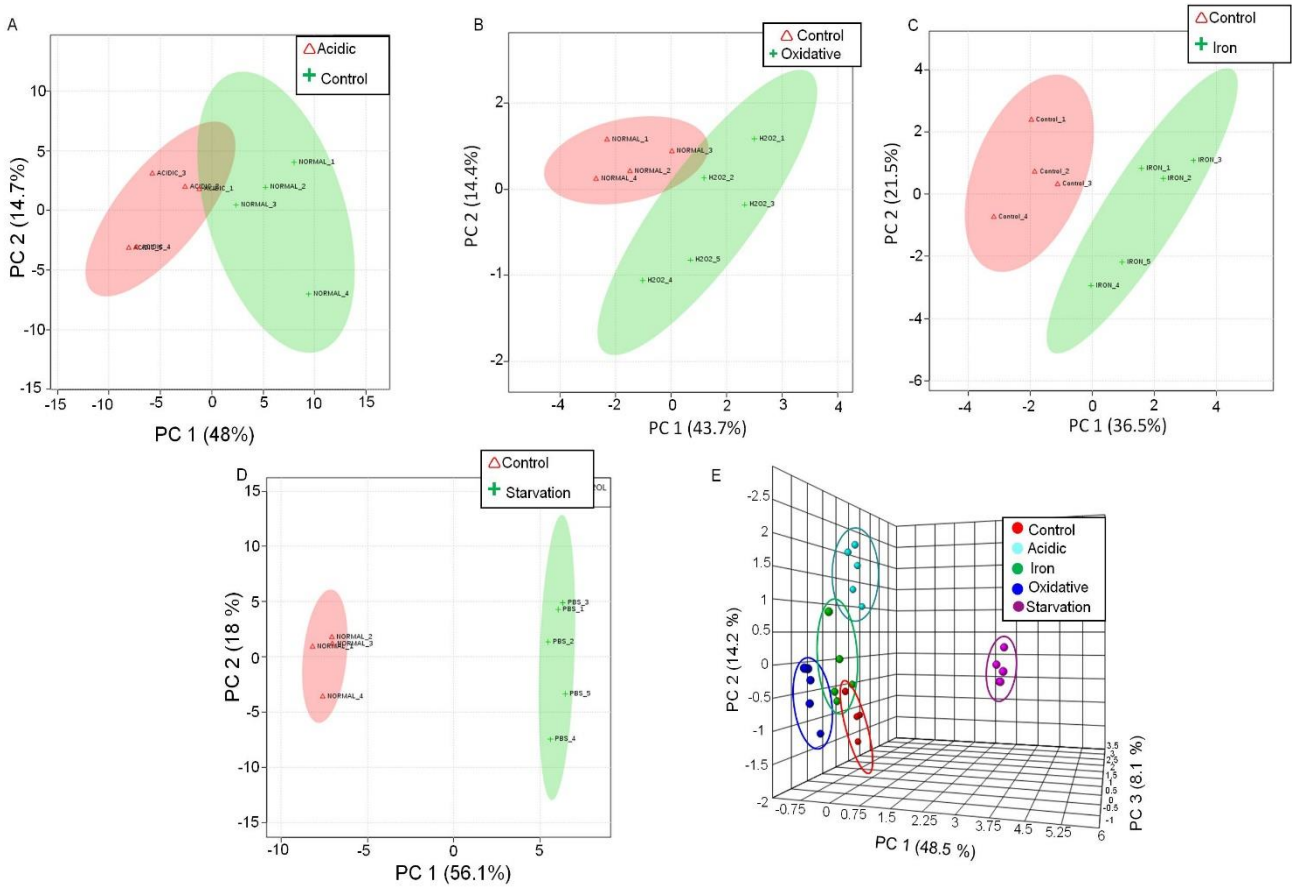

Figure 14: Principal Component Analysis differentially segregates metabolites from acidic stress, oxidative stress, iron stress, nutrient starvation and control in *M. tb*: Principal Component Analysis (PCA) 2D score plot of A) Acidic stress *vs* Control. B) Oxidative stress *vs* Control. C) Iron stress *vs* Control. D) Nutrient starvation stress *vs* Control. F) 3D score plot of acidic stress, oxidative stress, iron, nutrient starvation stress and control, showing acidic stress and nutrient starvation are distinctly categorized from control while iron and oxidative are showing subtle changes. The dots inside the plots correspond to biological replicates under each category.

Further, supervised analysis partial least squares discrimination (PLS-DA) was subsequently performed. It further maximized the separation between the control and stress exposed samples **(Figure 15)**, strongly suggesting metabolic variations between groups and identified additional metabolites than using PCA. PLS-DA shows further segregation between the groups analysed under A, B, C and D as compared to PCA analysis. When all the groups were analysed together (E) showed segregation of each stresses distinctly.

Amaximum segregation was observed in nutrient starvation and acidic stress, while Iron stress showed least segregation. A possible reason may be the presence of intracellular iron stores despite exogenous iron deprivation, could be utilized by mycobacteria. Since the stress duration was short, it is possible that mycobacteria continued to utilize its intracellular iron resources, thus showing metabolism nearly similar to control conditions. The quality (goodness of fit) of the PLS-DA models was described by R^2 and Q^2 values whichare given in **Table 11** below. The R^2 value and Q^2 for acidic, oxidative, iron and nutrient starvation stresses model were listed in table 3 where R^2 indicate goodness of fit and Q^2 which indicates the goodness of predictability.

Groups	Measure	1 comps	2 comps	3 comps	4 comps
Acidic-Control	R2	0.90	0.96	0.99	0.99
	Q2	0.85	0.88	0.91	0.92
Oxidative-Control	R2	0.80	0.99	1.00	1.00
	Q2	0.56	0.75	0.82	0.82
Iron-control	R2	0.93	0.98	0.99	1.00
	Q2	0.76	0.84	0.87	0.86
Starvation-control	R2	1.00	1.00	1.00	1.00
	Q2	0.99	1.00	1.00	1.00
Acidic-Oxidative-Iron-Starvation-Control	R2	0.70	0.92	0.98	0.99
	Q2	0.62	0.85	0.93	0.96

Table 11: R^2 and Q^2 values of Partial Least Squares - Discriminant Analysis (PLS-DA)

Comps: component

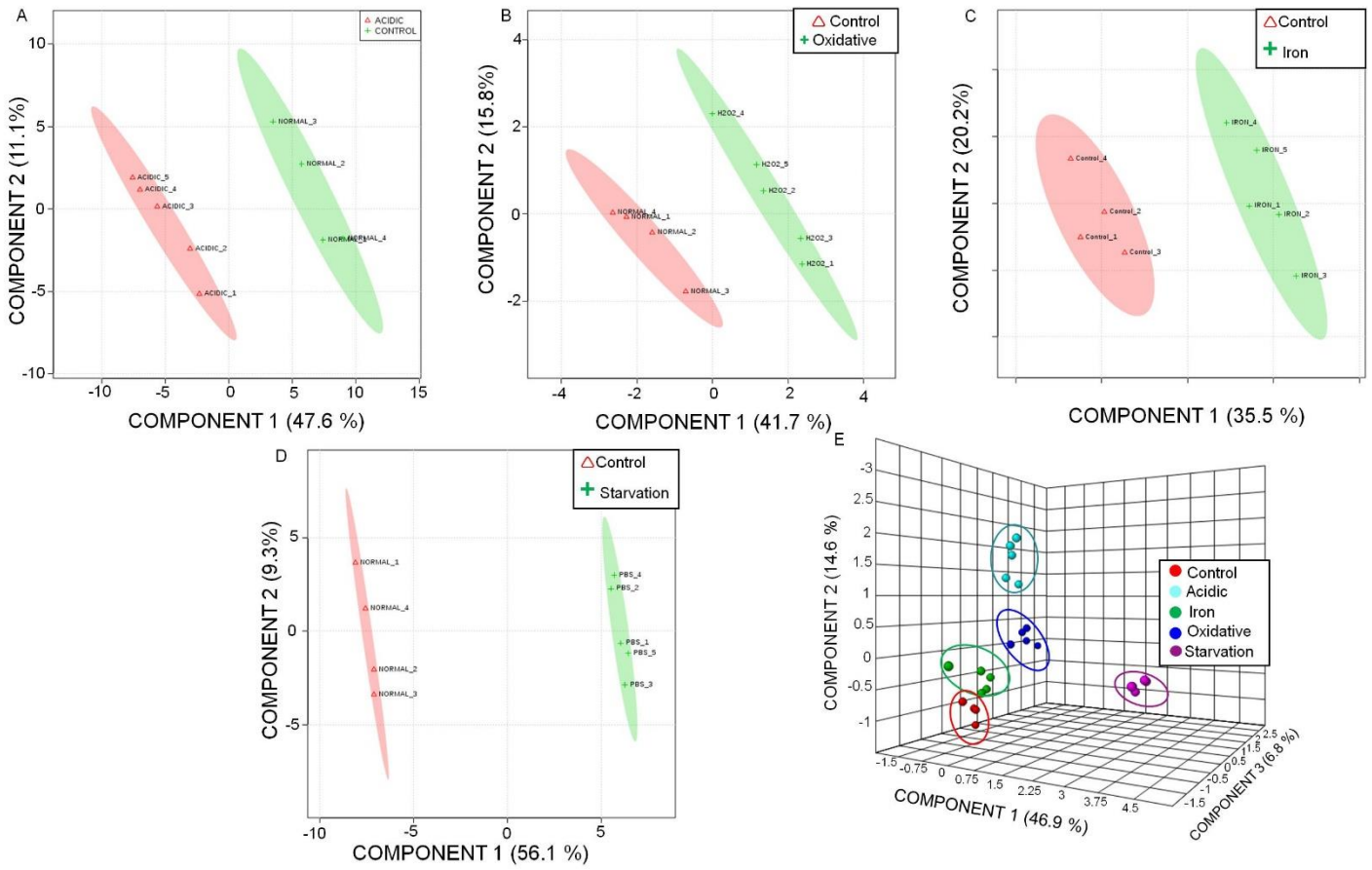

Figure 15: Partial Least Squares - Discriminant Analysis (PLS-DA) 2D score plot. A) Acidic stress *vs* Control. B) Oxidative stress *vs* Control. C) Iron stress *vs* Control D) Nutrient starvation stress *vs* Control. E) 3D score plot of acidic stress, oxidative stress, iron stress, nutrient starvation and control in *M. tb*. The dots inside the plots correspond to biological replicates under each category.

Potential metabolites under each stress condition were selected based on the VIP score obtained by PLS-DA model. In the model the metabolites with VIP score >1 wasbelieved to contribute significantly for group segregation; it shows 47, 34, 22 and 48 metabolites were majorly contributing for segregation in the model for acidic, oxidative, iron and nutrient starvation stresses respectively **(Figure 16)**. When all the conditions were analysed together, 39 metabolites were responsible for all group segregation. Some of these metabolites and their significance are discussed. Based on VIP score >1 **(Table 10)**, top fifteen common metabolites distinguishing

acidic stress, oxidative stress, iron stress, nutrient starvation stress and control **(Figure 16E)** are citric acid, isocitric acid,oxoglutaric acid, histidine, uridine, arginine, glutamic acid, kynurenic acid, serine, mono-ethylmalonate, aspartic acid, GMP, tryptophan, xylose, sorbitol. The metabolites discriminating control from acidic, oxidative stress and nutrient starvation majorly fall in the metabolism related to energy (TCA cycle), amino acid, nucleotide and sugar alcohol. Important metabolites to differentiate acidic *vs* control **(Figure 16A)** were xylitol, ribitol, maltose, lactic acid, oxoglutaric acid, sorbitol, homoserine, uridine, maltitol, lactose, malic acid, cAMP, methionine, acetlyglycine and isocitric acid suggesting majority of metabolites fall in sugar alcohol pathway, energy metabolism (eg TCA cycle) and amino acid metabolism. Fifteen top (VIP>1) differential metabolites specific for oxidative *vs* control **(Figure 16B)** were inosine, creatine, cAMP, lactic acid, xylitol, aspartic acid, homoserine, lactose, guanine,threonine, acetylglycine, thymine, 7-Methylguanosine, guanosine, leucine etc., thereby suggesting metabolites responsible for discrimination for this group fall under energy metabolism (eg TCA cycle), amino acid metabolism, nucleotide metabolism. Top fifteen differential metabolites (VIP>1) specific for iron stress *vs* control **(Figure 16C)** were inosine, isocitric acid, citric acid, GABA, aspartic acid, guanosine, AMP, guanine, cytosine, glutamic acid, thymidine, monoethylmalonate, glutamine, leucine, malic acid etc, indicating discriminating metabolites for this group fall under energy metabolism (eg TCA cycle), amino acid metabolism and nucleotide metabolism. While differential metabolites specific for nutrient starvation vs control **(Figure 16D)** were isocitric acid, citric acid, aspartic acid, glutamic acid, monoethyl malonic acid, oxoglutaric acid, GABA, Inosine, GMP, arginine, guanosine, N-acetylglutamic acid, aminoadipic acid, leucine, UDP etc suggesting that differential metabolites are contributed by energy metabolism (eg TCA cycle), amino acid metabolism and nucleotide metabolism. While the above described metabolites that are responsible in discriminating the various stress conditions, the biological significance of the same has been deliberated in the discussion section below.

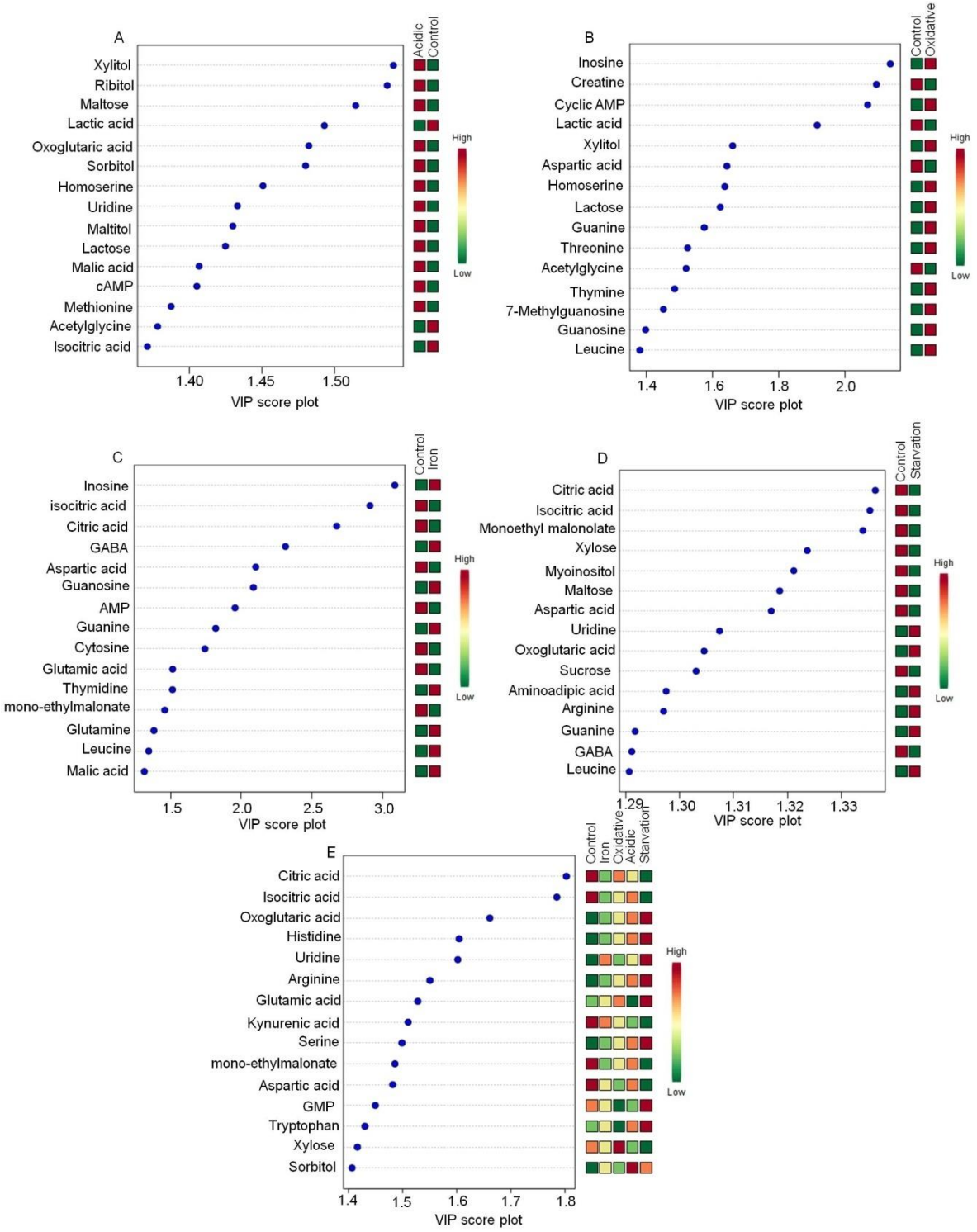

Figure 16: Variable Importance in Projection (VIP) plots indicating the most discriminating metabolite, identified through PLS-DA analyses, in descending order of importance for A) Acidic *vs* Control. B) Oxidative *vs* Control. C) Iron *vs* Control D) Nutrient starvation *vs* Control in *M. tb*. These important variables are responsible for the segregation among the groups as indicated.

3.3.1.2 Pathway Analysis

List of pathways perturbed in each condition was made based on p-value, exclusion of false discovery rate and impact factor (Table 4). It could be observed that across the different stress conditions, certain pathways appear in common. However, the significance of each pathway in the given stress condition seems to be different as reflected by the p-value. There appear to be common nodes to different stress response as well as differences reflected in terms of metabolite levels. For instance, cysteine and methionine metabolism appear as common pathways in both oxidative and acidic stresses. However, the significance score of this pathway in oxidative stress is higher than in acidic stress that is reflected in **Table 12**. For acidic stress, significance of TCA cycle was higher than in acidic stress as compared to others that is reflected in Table 4. Whereas for nutrient starvation condition, the major pathway perturbed was arginine and proline metabolism along with alanine, aspartate and glutamate metabolism (Table 4). For iron stress, the analysis points to perturbations in purine metabolism **(Table 12)**. The tables referred in this section are annexed at the end of this section. Some of the interpretations that could be made from the above metabolomics analysis are discussed below. These may serve as potential leads towards understanding early adaptive changes in *M. tb* to establish infection in host macrophages.

3.3.2 Discussion

3.3.2.1Adenylate energy charge (AEC) were maintained close to normal during stresses despite differences in absolute levels of adenine nucleotides

Energy being the fundamental requirement for viability of the cellular metabolic network, cellular homeostasis under stress is maintained by producing variations in adenosine nucleotide and their molecular turnover. The energy homeostasis is maintained not in terms of absolute levels of ATP, ADP and AMP, but the ratio of ATP/ADP or ATP/AMP. To evaluate the energy homeostasis in the energy metabolic pathways, we profiled the levels of ATP, ADP and AMP. We measured adenylate energy charge (AEC) for *M. tb* in each stress condition. The AEC values were about 0.87 when grown in Sauton's media, 0.91, 0.90, 0.91 and 0.96 for acidic, oxidative, iron and nutrient starvation stresses respectively. It was clear that though large fluctuations in the adenosine nucleotide concentrations were observed during microbicidal stresses in *M .tb* (**Figure**

17), the AEC values were maintained close to that of normal, indicating that *M. tb* could adapt and uphold homeostasis despite stress by shifting energy dynamics. With these observations, it could be inferred that H37Rv has the ability to change its energy dynamics to sustain functional metabolic viability for early adaptation to similar microbicidal stresses inside host macrophages.

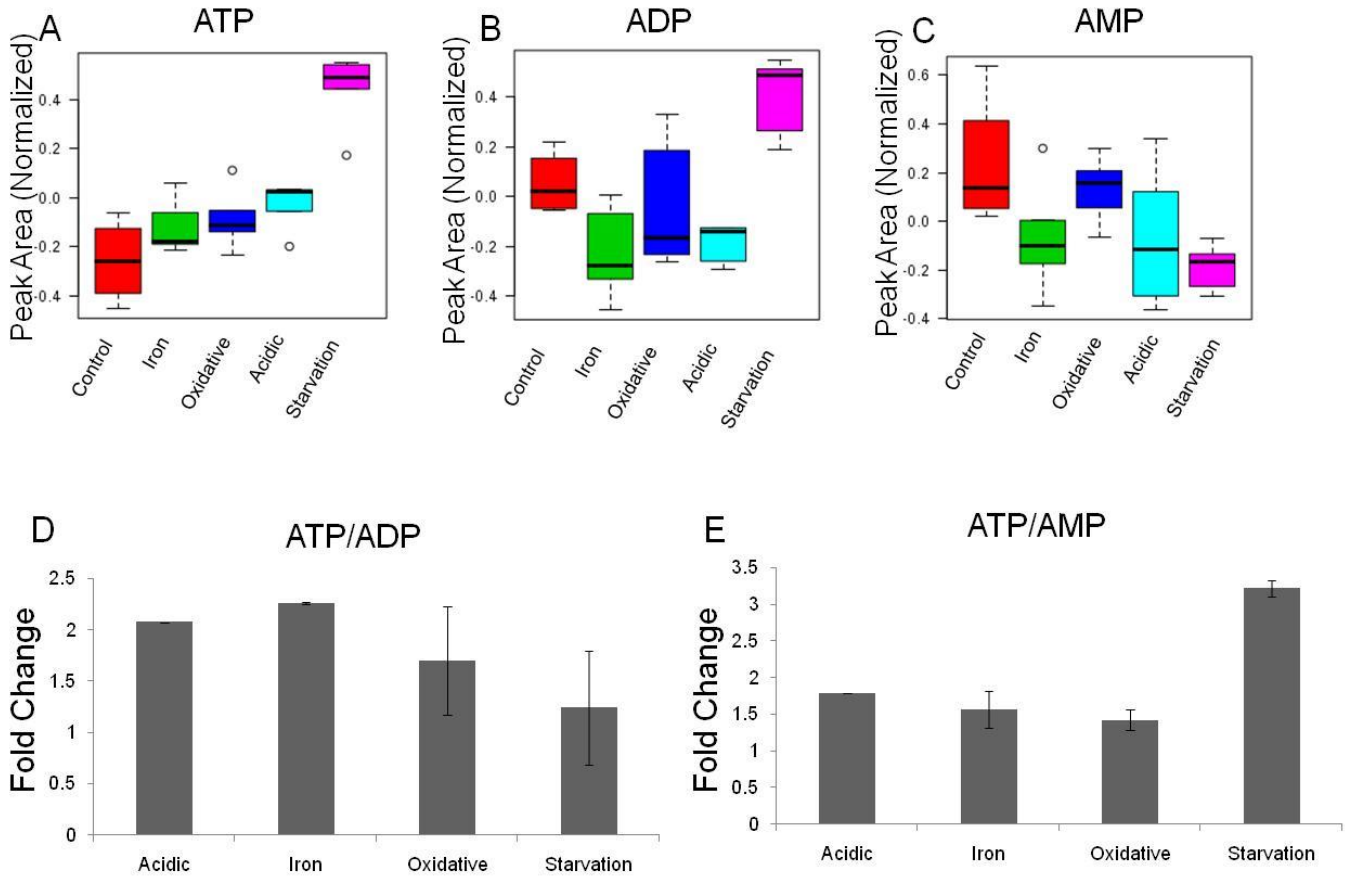

Figure 17: Adenosine nucleotide concentrations in *M. tb* during microbicidal stresses

The levels of (A) ATP, (B) ADP and (C) AMP in *M. tb* exposed to microbicidal conditions as indicated are shown. The ratios of ATP/ADP (D) and ATP/AMP (E) represented as fold change with respect to control.

3.3.2.2 Adjustments in TCA cycle, amino acid metabolism and possible significance of anaplerotic reactions stress adaptation

TCA cycle plays an essential role in cellular metabolism; it favours energy generation as well as provides biosynthetic precursors (Tian *et al*. 2005). TCA cycle is an important metabolic network and its metabolites apart from its usual function are also important to overcome stress (Mailloux *et al*. 2007). Our metabolomics data suggested TCA cycle to be perturbed in all stress conditions **(Figure 18)**. There was a drop in the levels of citrate and isocitrate in all the conditions, being lowest in nutrient starvation. However, a notable increase in α-ketoglutarate levels were observed with 3 fold increase during starvation, 1.8 fold increase in acidic and oxidative stresses while only 1.2 fold increase was observed in iron stress. One can expect that α-ketoglutarate may be replenished *via* glutamate using enzyme glutamate dehydrogenase (GDH). This anaplerosis node may serve as universally important step for early metabolic adaptation during these stresses **(Figure 19)**. GDH has been previously suggested to have importance in overcoming acidic, oxidative and nitrosative stresses (Gallant *et al*. 2016). The conversion of glutamate to α-ketoglutarate by GDH can replenish TCA cycle intermediate, while it also releases ammonia and reducing equivalents of energy metabolites (NAD^+).

Ammonia helps in neutralizing acidic stress while NAD+ helps in coping with oxidative stress. Some studies show that α-ketoglutarate can also get metabolized to 2-hydroxy-3-oxoadipate by peroxidase system, a component of KDH complex, to counter nitrosative stress (Maksymiuk *et al*. 2015). *M. tb* employs the membrane transporter AnsP1/Rv2127 to capture aspartate and AnsP2 (Rv0346c) to capture asparagine and use these amino acids as a nitrogen source during infection by hydrolysis of asparagine to aspartate (by asperginase) that also releases ammonia and further conversion of aspartate to glutamate (Gouzy *et al*. 2013; Gouzy *et al*. 2014). Our metabolic profiles show that aspartate is very low in all stress conditions, suggesting its possible utilization. At the same time, we observed low levels of glutamate in acidic, oxidative and iron stresses, which possibly is getting metabolized to α-ketoglutarate by GDH. Glutamate can be generated from glutamine **(Figure 19A)**.

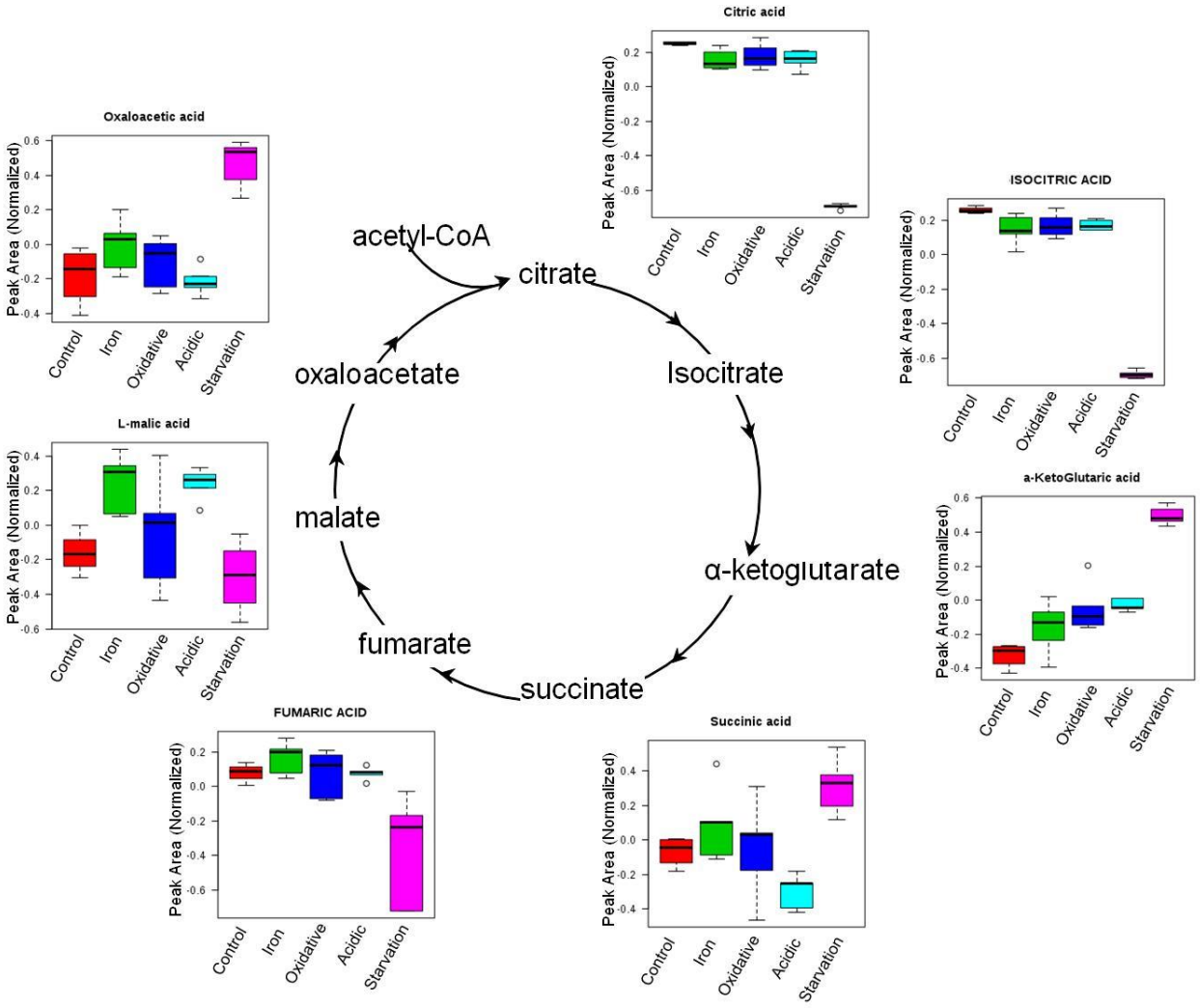

Figure 18: TCA cycle in *M. tb* is altered during all microbicidal stress conditions: Schematic representation of TCA cycle, along with the levels of intermediate metabolites of the pathway identified in each stress. It could be clearly observed that TCA cycle was perturbed under all microbicidal stress conditions. High levels of α-ketoglutarate, more predominant during starvation conditions, suggests occurrence of anapleurotic reactions indicating breakdown of amino acids to α- ketoglutarate.

With our metabolomics observations together with earlier reports, we propose glutamate dehydrogenase (GDH), glutamate synthetase (GS) and glutamine oxoglutarate aminotransferase (GOGAT) as important nodes for early adaptation to microbicidal stresses, especially acidic and oxidative stresses. In order to check if the differential levels of glutamate, glutamine, aspartate etc were due to differences in the expression levels of these enzymes, we evaluated the transcript levels of GS, GDH and GOGAT by RT-PCR during these stress conditions **(Figure 19B)**. It was

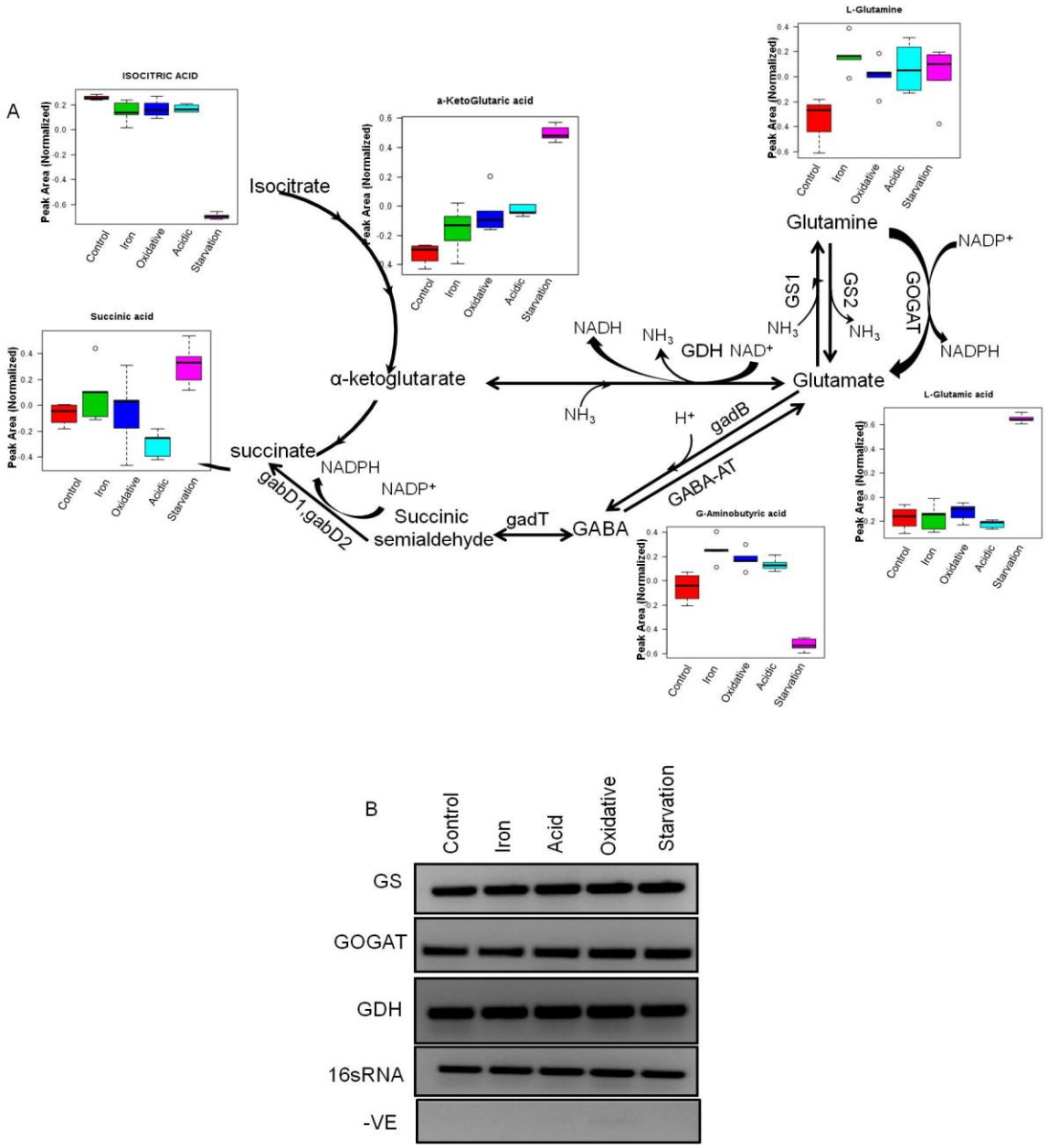

Figure 19: Anaplerotic reactions in TCA cycle and GABA shunt. A) TCA cycle showing anaplerosis node where glutamate is converted to α-keto glutarate resulting in the release of NH_3, which is essential for neutralizing acidic pH, while NADH helps to neutralize ROS and therefore contributes towards adaptation to acidic and oxidative stress. B) Expression profiling of GS, GDH and GOGAT by RT-PCR during the microbicidal stress conditions with 16sRNA as loading control. Lane corresponding to –ve control confirms that there was no genomic DNA contamination. GS- glutamate synthetase; GDH-glutamate dehydrogenase; GOGAT- Glutamine oxoglutarate aminotransferase; gadB -glutamate decarboxylase; GABA-aminotransferase- GABA-AT; gabD1 AND gabD2 - succinate-semialdehyde dehydrogenase.

clear that the regulation of these enzymes is not at transcription level, but possibly the enzymatic activities are altered during stresses.

Yet another set of the metabolites identified during the different microbicidal conditions were that of GABA shunt pathway as presented schematically along with their differential levels in figure 6A, which is described in the next section.

3.3.2.3 γ-Aminobutyric acid (GABA) and its possible role in adaptation to acidic and oxidative stresses in *Mycobacterium tuberculosis*

Another interesting observation was the accumulation of GABA during all stresses, except during starvation **(Figure 19A)**. This pathway is a shunt between TCA cycle and glutamate and glutamine metabolism. Previous studies performed on various organisms suggested that accumulation of GABA is beneficial for a cell under acidic stress (Tramonti *et al.* 2002; Lei *et al*. 2011). These studies convey that GABA is more beneficial to a cell than glutamate as GABA is less acidic (pI 7.0) than glutamate (pI 3.0) (Tramonti *et al*. 2002; Lei *et al*. 2011), though pKa of side chain carboxyl group and the α-amino group of glutamate and GABA have same pKa. The possible differences lie in their ability to release a proton at low pH (Foster *et al* 2004). In GABA shunt, glutamate gets metabolized to GABA, which then gets metabolized to succinate semialdehyde then to succinate (Tian *et al.* 2005). When glutamate gets metabolized to GABA by glutamate decarboxylase, a proton is utilized, which proves beneficial for cell **(Figure 19A)**. Apart from removing proton at low pH, GABA shunt can also reduce NAD^+ to form succinate by metabolizing succinate semialdehyde by action of succinate-semialdehyde dehydrogenase (Tian *et al.* 2005). In *M. tb*, GABA shunt is not well studied and only recently succinate-semialdehyde dehydrogenase enzyme activity was detected in *M. tb*, *M. smegmatis* and *M. bovis* BCG (Tian *et al*. 2005). GABA shunts can also regenerate glutamate by the action of GABA-aminotransferase (Feehily and Karatzas 2013). Thus, GABA synthesis can quench protons and at the same time participate in NAD^+/NADH balance, therefore contributing to early adaptations during both acidic and oxidative stresses.

3.3.2.4 Urease activity is not important for early adaptation to acidic stress in *M. tb*

From the above discussions, it is clear that proton quenching can be one of the ways to counter acidic stress. However, a very well established phenomenon that is also reported in *M. tb* is release of ammonia to neutralize low pH. The ammonia released in culture media during different stresses **(Figure 20A)** and the changes in the pH of media post-stress as compared to pre-stress were measured **(Figure 20B)**. It was observed that a basal level of NH_3 was released during iron stress, but it was nearly 2 fold more during acidic stress. NH_3 levels were below detectable levels in oxidative and starvation conditions, which points to the possibility of its utilization as a nitrogen source. When the pH of media was tested post-stress, as expected, we observed a rise in the pH of the media only during acidic stress and marginally during iron stress where a small quantum of ammonia release was observed **(Figure 20)**. Thereby, we could confirm that the early adaptation of *M. tb* to acidic stress is by neutralizing the environment by releasing molecules like NH_3.

It was discussed above that NH_3 can be released from several pathways including inter-conversions of aspartate, glutamate, glutamine and α-ketoglutarate (**Figure 19A**). A general pathway adopted by many bacteria is the generation of NH_3 by urea cycle where urea is metabolized by urease to release NH_3 (Mobley *et al.* 1995; Mendz and Hazell 1996). Conventionally, urea cycle is believed to be predominantly responsible for adaptation to acidic stress in bacteria, including in *M. tb*. In our study, we observed an accumulation of both arginine and citrulline under acidic and starvation stresses **(Figure 21A)**. Arginine can be acted upon by arginase to yield urea, which can be metabolized by urease to yield ammonia (Mendz and Hazell 1996). Arginine is also a precursor of polyamines to counter oxidative stress by arginine decarboxylase (Cunin, *et al.* 1986). Arginine can also be converted to citrulline by ADI pathway (Cunin, *et al.* 1986). In order to assess the role of urease in this context, the transcript levels and activity of urease was measured during the different stresses. While we observed no change at transcript level by RT-PCR **(Figure 21C)**, to our surprise, we observed a decrease in urease activity in all stresses as compared to control **(Figure 21B)**, suggesting that regulation of this pathway under acidic and other stresses is not predominant during early adaptation of *M. tb*. Our results are in agreement with a previous study, where, the intracellular growth of urease-deficient

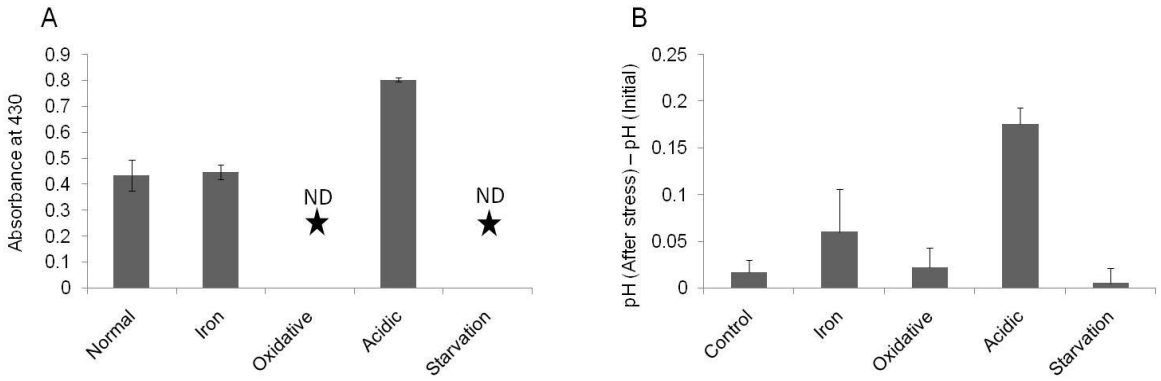

Figure 20: Estimation of ammonia release in *M. tb* during different microbicidal stress conditions. Ammonia release by *M. tb* in the different stress conditions was estimated by Nessler's assay as described (A) by absorbance at 430nm. The pH of the culture was measured pre and post stress (B). These experiments suggested that the ammonia release by M. tb contributed significantly to the rise in pH post stress.

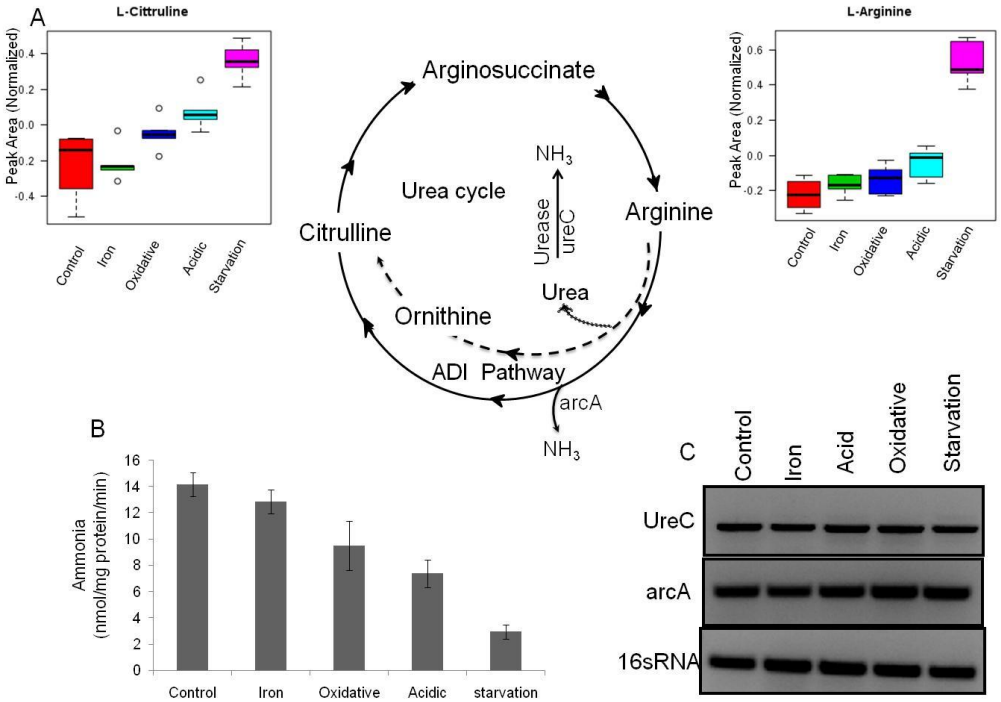

Figure 21: Urease activity in *M. tb* cultures exposed to microbicidal stress conditions
(A) Schematic representation of urea cycle with and the differential levels of its component metabolites, L-citrulline and L-arginine. B) Measurement of urease activity of mycobacteria under stress conditions reveals altered urease activity.The decrease in urease activity suggests that mycobacteria is utilizing Arginine deaminase pathway during stress conditions to yield NH_3 and citrulline. (C) Expression profiling of enzymes involved in urease cycle, ureC and in ADI pathway arcA with 16sRNA as loading control.

M. tb, wild type and complement strain in macrophages were studied, where it was shown that urease did not contribute any growth advantage to *M. tb* in macrophage (Lin *et al.* 2012).

3.3.2.5 Accumulation of sugar alcohol as possible stress adaptation

Another interesting observation from our study was the accumulation of polyols (**Figure 9**). Sugar alcohol (polyols) can be generated by pentose phosphate pathway (PPP) (Toivari *et al*. 2007; Lin *et al* 1996). Pentose phosphate pathway (PPP) is one of the fundamental pathways that provides reducing molecules for metabolism, overcoming oxidative stress, maintaining carbon homeostasis, and provides precursors for amino acid as well as nucleotide biosynthesis(Stincone *et al*. 2015). In our data, we observed a robust increase in polyols including xylitol (25 fold) and ribitol (21 fold) in acidic stress whereas a modest 1.9 fold increase in xylitol and 1.2 fold in ribitol in oxidative stress, and in iron stress, it was only 1.8 and 1.3 fold increase in xylitol and ribitol respectively **(Table 10A, 10B and 10C** for acidic, oxidative and iron stresses respectively**)**. However, during nutrient starvation, the levels of xylitol and ribitol were lower as compared to control, i.e showed 0.76 and 0.79 fold change respectively **(Table 10D)**. The differential levels of xylitol and ribitol in the different stress conditions suggests a greater impact of PPP pathway in acidic stress as compared to oxidative and iron stresses **(Figure 22)**. Literature support was sought to understand the significance of polyols in acidic stress. The different modes of adaptation to acidic stress of most bacteria have predominantly focused on ammonia release to neutralization of pH. The adaptation of mycobacteria to acid stress causes extensive changes in transcriptomic profile (Fisher *et al* 2002). Some of the genes induced during acid stress include isocitrate lyase, nonribosomal peptide synthetases and polyketide synthases. On the contrary, genes responsible for mycolic acid biosynthesis were repressed. The transcription factors activated during acidic stress includes a marP family transcription factor, Rv1404, which de-represses the expression of putative SAM-dependent methyl transferases (Healy *et al.*, 2016). Another TetR family transcription factor, Rv1685c induced by acid-nitrosative stress controls the expression of efflux pumps, carbon metabolism and virulence (Perrone *et al.*, 2017). The expression of efflux pumps in acid stress explains the mycobacterial tolerance to anti-biotics isoniazid and rifampin (Baker and Abramovitch, 2018). However, the role of polyols has not been explored much in *M. tb*. The drastic accumulation of polyols including xylitol and ribitol in acid stress suggest novel metabolic adaptive mechanisms. Xylitol

is a known bacteriostatic metabolite that is up taken by many bacteria but not metabolized. It also inhibits bacterial metabolism, including acid production (Trahan, *et al.* 1991; Roberts *et al.* 2002; Tanzer *et al*. 2006; Nayak *et al.* 2014). This possibly explains potential adaptation to acid stress, wherein the acid generating pathways are blocked or re-routed. In addition to these, we also observed high levels (~ 1.2 fold) of maltitol both in acidic and oxidative stress as compared to control. However, there was no change in iron stress but low levels were observed in nutrient starvation. The possible biosynthesis of maltitol could be from maltose which in turn is synthesized from trehalose. These biochemical conversions are known to occur in mycobacteria (De Smet *et al*. 2000).

In our study, in addition to polyols, we could also observe an accumulation of metabolites such as oxoglutaric acid, inosine which are the by-products of ammonia production. There is very little work that has focused on utilization of excess protons generated due to acidic stress. Thus we hypothesized that apart from ammonia formation; bacteria also activate pathways which can utilize protons **(Figure 22)**. In this context, the sudden rise in polyols supported our notion as these reactions utilize protons. As polyols are reduced form of aldose or ketose sugar and their conversion from sugar to sugar alcohol consumes excess protons from the milieu. A recent study had shown that xylitol can be formed in *M. smegmatis* by utilizing glucose or xylose as co-substrate (Izumori *et al* 1988). As xylose and glucose were not supplied in the media, but had glycerol, the intracellular xylitol and ribitol were possibly formed through pentose phosphate pathway by either metabolizing ribulose-5-phosphate or by glyceraldehyde-3-phosphate (Toivari *et al.* 2007). Sorbitol could possibly be formed by reduction of glucose by polyol pathway. This supports our hypothesis that *M. tb* resists acid stress by utilizing excess protons. However, low levels of xylitol and ribitol during nutrient starvation condition suggested that the biosynthesis of these metabolites does not occur as the carbon sources for xylitol and ribitol (glyceroldehyde-3-phosphate and ribulose-5-phosphate) are limiting **(Figure 22)**. Taken together, these observations suggest a novel role of sugar alcohols in overcoming acidic stress as an early adaptation, apart from overcoming oxidative stress by reducing energy metabolites.

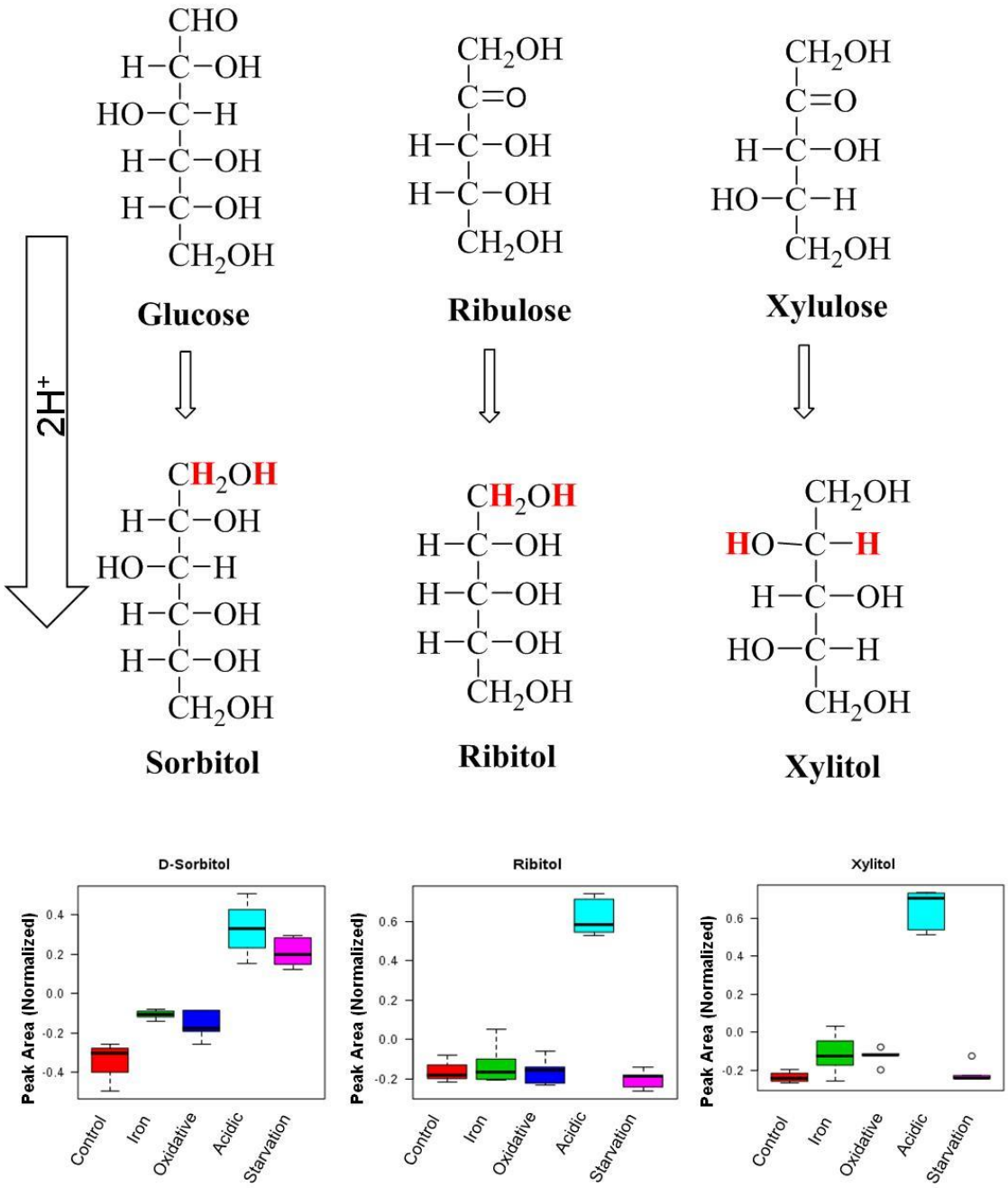

Figure 22: Chemical structures of sugar alcohols and their corresponding sugar moieties. Analogous sugar gets reduced by utilizing proton to yield sugar alcohol. The levels of these metabolites identified in the present study are also shown.

3.3.3 New observations and leads from this study

1. The study suggests quenching of protons as an important mechanism for early adaptation to acidic and oxidative stress and points to the role of GABA shunt and sugar alcohol (polyols) synthesis and accumulation, which if thoroughly investigated can lead to new anti-mycobacterial targets.

2. This study identified glutamate dehydrogenase (GDH), glutamate synthetase (GS) and glutamine oxoglutarate aminotransferase (GOGAT) as important nodes for early adaptation to microbicidal stresses, especially acidic and oxidative stresses and hence can be good anti-mycobacterial targets.

3. Urease, unlike the general belief, is not the major player in adaptation to acidic stress in *M. tb*.

Annexure I

(Tables)

Annexure I (PART 3A)

Tables for Comparative and quantitative metabolite profiling of non-pathogenic mycobacteria under microbicidal stresses

3.4.1 Tables for comparative and quantitative metabolite profiling of non-pathogenic mycobacteria under microbicidal stresses

Table 5: *Mycobacterium smegmatis* metabolites were assigned their respective ^{1}H chemical shifts (PPM)

S.No	Metabolites	PPM
1	2-Aminobutyrate	3.706(t), 1.887(m), 0.968(t)
2	2-Hydroxy-3-methylvalerate (HMVA)	3.88(d), 1.705(m), 1.351(m), 1.162(m), 0.932(d), 0.872(t)
3	3-Hydroxyisovalerate	1.233(s), 2.35(s)
4	3-Methyl-2-Oxovalerate	0.88(t), 1.086(d), 1.444(m), 1.687(m), 2.922(m)
5	Acetate	1.90(s)
6	Acetone	2.221(s)
7	ADP	8.534(s), 8.261(s), 6.122(d), 4.52(m), 4.35(m), 4.00(m)
8	Alanine	1.46(d), 3.805(q)
9	AMP	8.596(s), 8.25(s), 6.12(d), 4.49(t), 4.355(m), 4.02(m)
10	Aspargine	2.777(dd), 2.89(dd),3.98(dd)
11	Aspartate	2.66(dd), 2.79(dd), 3.916(dd)
12	ATP	8.52(s),8.25(s), 6.128(d), 4.56(t), 4.41(m), 4.23(m), 4.30(m)
13	Beta-alanine	3.16(t), 2.54(t)
14	Betaine	3.263(s), 3.885(s)
15	Caprate	0.839(t), 1.281(m), 1.528(m), 2.519(m)
16	Cholate	0.711(s), 0.905(s), 0.960(d), 1.00(m), 1.16(m), 1.130(m), 1.140(m), 1.491(m), 1.568(m), 1.623(m),
		1.75(m),1.89(m), 2.01(m), 2.05(m), 2.23(m), 3.501(m), 3.39(m),4.06(m)
17	Citrate	2.53(d), 2.666(d)
18	Citrulline	1.526(m), 1.59(m), 1.84(m), 1.889(m), 3.147(m), 3.126(m),3.74(m)
19	Dimethylamine	2.718(s)
20	DSS	0.0(s), 0.62(t),1.75(m), 2.91(t)
21	dTTP	7.68(s), 6.333(m), 4.618(m), 4.22(m), 4.17(m), 2.38(m), 1.918(s)
22	Ethanol	1.185(t), 3.664(q)
23	Formate	8.44(s)
24	Fumarate	6.50(s)
25	Glucose-1-Phosphate	5.45(dd), 3.908(m), 3.86(m), 3.76(m), 3.487(m), 3.398(t)

26	Glutamate	2.03(m),2.10(m), 2.34(m), 3.75(dd)
27	Glutamine	2.141(m), 2.459(m), 3.76(t)6
28	GTP	8.110(s), 6.12(s),5.96(s), 4.54(m), 4.35(m4.24(m)
29	Homoserine	2.01(m), 2.16(m), 3.77(m), 3.85(dd)
30	IMP	8.553(s), 8.231(s), 6.136(d), 4.51(m), 4.36(m), 4.01(m)
31	isoleucine	0.926(t), 0.992(d), 1.248(m), 1.457(m), 1.968(m), 3.661(d)
32	Lactate	4.096(q), 1.313(d)
33	Leucine	3.721(m), 1.701(m), 0.94(m)
34	Lysine	3.74(t), 3.02(t), 1.89(m), 1.71(m), 1.46(m)
35	Malonate	3.09(s)
36	Maltose	5.41(d), 5.39(d), 5.211(d), 3.96(m), 3.93(m), 3.84(m), 3.76(m), 3.70(m), 3.66(m),3.62(m),
		3.58(m), 3.421(m),3.27(m)
37	Methylamine	2.573(s)
38	N-Acetyl-glucosamine	
39	NAD	9.314(s), 9.121(d), 8.824(d), 8.406(s), 8.184(m), 8.154(s), 6.07(d), 6.021(d), 4.522(m),
40	NADP+	9.281(s), 9.08(d), 8.80(d), 8.41(s), 8.18(m), 8.13(s), 6.112(d), 6.022(d), 4.99(q), 4.60(t),
41	Phenyl alanine	7.42(m), 7.36(m), 7.32(m), 3.98(dd), 3.27(m), 3.13(m)
42	Succinate	2.385(s)
43	Threonine	4.244(m), 3.582(d), 1.313(d)
44	Trehalose	5.18(d), 3.84(m), 3.80(m), 3.76(m), 3.64(dd),3.43(t)
45	Tryptophan	7.72(d), 7.53(d), 7.32(s), 7.24(m), 7.19(m0, 4.04(dd), 3.47(dd), 3.29(dd)
46	Tyrosine	7.17(d), 6.87(d), 3.93(dd), 3.18(dd),3.04(dd)
47	UDP Galactose	7.93(d), 5.97(m), 5.63(dd), 4.36(m), 4.25(m), 4.15(m),4.02(d), 3.90(dd), 3.805(dt), 3.75(m)
48	UDP-Glucose	7.925(d), 5.977(m), 5.593(dd), 4.36(m), 4.27(m), 4.24(m), 4.189(m), 3.877(m), 3.77(m),
49	UDP-N-Acetylglucosamine	8.177(d), 8.077(m), 5.19(d), 4.70(d), 3.904(m), 3.841(m), 3.784(m), 3.74(m),
50	UMP	8.07(d), 5.98(m), 4.41(t), 4.34(t), 4.26(m), 3.97(m)
51	Valine	0.996(d), 1.047(d), 2.281(m),3.617(d)

Table 6A: Fold change and VIP score of Control vs Acidic stress in *M. smegmatis*

S.No	Metabolites	Fold Change	VIP score	p.value	FDR
1	D-Maltose	3.08	0.65	1.0E-09	1.3E-08
2	Capric acid	2.17	0.47	7.5E-05	2.0E-04
3	Taurine	1.77	0.62	3.6E-05	1.0E-04
4	Methylamine	1.59	0.22	1.3E-01	1.6E-01
5	2-Hydroxy-3-methylpentanoic acid	1.46	0.25	3.3E-03	6.4E-03
6	Dimethylamine	1.45	0.25	1.2E-01	1.4E-01
7	Malonic acid	1.34	0.73	1.5E-02	2.5E-02
8	3-Methyl-2-oxovaleric acid	1.34	0.09	1.3E-02	2.2E-02
9	Citric acid	1.29	0.63	2.3E-01	2.6E-01
10	L-Asparagine	1.28	1.78	1.8E-02	3.0E-02
11	Uridine diphosphate glucose	1.11	0.11	3.4E-01	3.7E-01
12	3-Hydroxyisovaleric acid	1.00	0.00	9.8E-01	1.0E+00
13	L-Lactic acid	1.00	0.00	1.0E+00	1.0E+00
14	Ethanol	1.00	0.01	9.8E-01	1.0E+00
15	Betaine	0.96	0.04	6.3E-01	6.9E-01
16	NADP	0.92	0.03	6.6E-01	7.0E-01
17	ADP	0.84	0.15	8.9E-02	1.2E-01
18	NAD	0.83	0.39	2.0E-02	3.2E-02
19	Nicotinamide N-oxide	0.83	0.13	8.5E-02	1.2E-01
20	Acetone	0.83	0.25	1.9E-05	6.3E-05
21	Trehalose	0.82	1.15	1.7E-01	2.0E-01
22	Citrulline	0.80	0.42	1.0E-01	1.3E-01
23	Uridine diphosphate-N-acetylglucosamine	0.80	0.13	3.0E-02	4.6E-02
24	Formic acid	0.78	0.28	1.0E-03	2.2E-03
25	Acetic acid	0.74	0.41	4.3E-02	6.1E-02
26	Adenosine monophosphate	0.72	0.55	6.0E-04	1.4E-03
27	Sucrose	0.70	0.17	1.4E-01	1.7E-01
28	Adenosine triphosphate	0.70	0.20	9.9E-02	1.3E-01
29	Inosinic acid	0.70	0.12	9.1E-03	1.6E-02
30	L-Lysine	0.67	0.29	1.3E-03	2.7E-03
31	L-Leucine	0.66	0.28	1.7E-03	3.5E-03
32	Uridine diphosphategalactose	0.66	0.20	2.5E-05	7.7E-05
33	L-Phenylalanine	0.65	0.21	9.1E-11	1.5E-09
34	Glucose 1-phosphate	0.62	0.25	4.0E-02	5.8E-02
35	Uridine 5'-monophosphate	0.61	0.30	1.1E-04	2.8E-04
36	Succinic acid	0.60	0.85	9.5E-06	3.4E-05

37	Cholic acid	0.60	0.27	3.9E-03	7.2E-03
38	Thymidine 5'-triphosphate	0.59	0.28	8.7E-06	3.4E-05
39	L-Glutamine	0.59	1.73	2.3E-06	1.0E-05
40	L-Alpha-aminobutyric acid	0.57	0.31	6.0E-04	1.4E-03
41	L-Aspartic acid	0.56	1.73	7.0E-05	2.0E-04
42	D-Glutamic acid	0.55	4.71	1.2E-08	9.0E-08
43	L-Homoserine	0.49	2.61	1.3E-08	9.0E-08
44	Isoleucine	0.37	0.32	3.9E-08	2.4E-07
45	1-Methylnicotinamide	0.33	0.42	8.2E-08	4.5E-07
46	Beta-Alanine	0.31	1.34	2.2E-06	1.0E-05
47	L-Threonine	0.31	1.25	4.5E-06	1.9E-05
48	L-Valine	0.30	0.77	4.4E-09	4.4E-08
49	L-Tyrosine	0.28	0.54	7.6E-11	1.5E-09
50	Alanine	0.26	1.19	5.7E-11	1.5E-09

FC: fold change; VIP score: variable of importance score obtained from PLS-DA plot, p value: p values obtained after performing t-test, FDR: value obtained after performing false discovery test.

Table 6B: Fold change and VIP score of Control vs Oxidative stress in *M. smegmatis*

S.no	Metabolites	Fold Change	VIP score	p.value	FDR
1	L-Lysine	3.33	1.96	1.8E-07	1.1E-06
2	Methylamine	2.85	1.57	1.9E-05	6.6E-05
3	Uridine diphosphate-N-acetylglucosamine	2.36	1.40	3.5E-07	1.8E-06
4	Betaine	2.36	1.70	3.2E-09	8.1E-08
5	L-Aspartic acid	2.20	2.69	1.3E-07	9.0E-07
6	Capric acid	2.10	1.40	8.0E-06	3.1E-05
7	D-Maltose	2.08	1.37	2.5E-08	4.2E-07
8	2-Hydroxy-3-methylpentanoic acid	1.90	1.22	4.5E-05	1.5E-04
9	3-Methyl-2-oxovaleric acid	1.83	0.94	2.3E-07	1.3E-06
10	Isoleucine	1.50	0.83	1.2E-03	3.6E-03
11	Acetic acid	1.43	0.80	7.3E-02	1.4E-01
12	Dimethylamine	1.43	0.89	2.4E-03	6.6E-03
13	Taurine	1.36	0.71	3.5E-02	8.1E-02
14	L-Lactic acid	1.22	0.59	9.7E-02	1.6E-01
15	L-Glutamine	1.17	0.94	3.4E-03	9.0E-03
16	3-Hydroxyisovaleric acid	1.13	0.34	1.4E-01	2.2E-01
17	Uridine diphosphate glucose	1.13	0.41	8.5E-02	1.5E-01
18	Acetone	1.11	0.30	2.0E-01	2.9E-01
19	L-Leucine	1.10	0.25	3.6E-01	4.7E-01
20	L-Asparagine	1.10	0.61	1.7E-01	2.6E-01
21	Formic acid	1.10	0.30	2.2E-01	3.2E-01
22	L-Homoserine	1.09	0.48	2.9E-01	4.0E-01
23	Thymidine 5'-triphosphate	1.09	0.12	6.7E-01	7.0E-01
24	Citric acid	1.08	0.30	5.4E-01	6.2E-01
25	Ethanol	1.06	0.21	5.4E-01	6.2E-01
26	Citrulline	1.04	0.16	6.5E-01	6.9E-01
27	L-Phenylalanine	1.03	0.10	5.7E-01	6.4E-01
28	L-Threonine	0.99	0.01	9.8E-01	9.8E-01
29	Beta-Alanine	0.98	0.09	8.1E-01	8.3E-01
30	Succinic acid	0.95	0.24	4.9E-01	5.8E-01
31	Trehalose	0.93	0.28	6.3E-01	6.9E-01
32	Cholic acid	0.92	0.21	4.0E-01	4.9E-01
33	Alanine	0.91	0.42	9.4E-02	1.6E-01
34	L-Valine	0.91	0.30	3.4E-01	4.4E-01
35	Uridine diphosphategalactose	0.90	0.30	1.2E-01	1.9E-01

36	L-Alpha-aminobutyric acid	0.89	0.24	4.1E-01	5.1E-01
37	L-Tyrosine	0.89	0.41	6.3E-02	1.3E-01
38	ADP	0.83	0.52	4.1E-02	8.9E-02
39	Uridine 5'-monophosphate	0.82	0.51	5.4E-02	1.1E-01
40	Sucrose	0.80	0.34	3.0E-01	4.0E-01
41	Glucose 1-phosphate	0.76	0.56	7.4E-02	1.4E-01
42	D-Glutamic acid	0.71	2.16	8.9E-08	7.4E-07
43	Inosinic acid	0.70	0.60	1.6E-02	3.9E-02
44	Malonic acid	0.69	1.08	1.3E-02	3.1E-02
45	Nicotinamide N-oxide	0.67	0.90	7.8E-05	2.4E-04
46	NAD	0.59	1.47	1.7E-06	7.5E-06
47	Adenosine monophosphate	0.52	1.62	6.4E-08	6.6E-07
48	Adenosine triphosphate	0.44	1.35	4.6E-06	1.9E-05
49	1-Methylnicotinamide	0.42	1.40	6.6E-08	6.6E-07
50	NADP	0.13	1.70	2.1E-12	1.1E-10

FC: fold change; VIP score: variable of importance score obtained from PLS-DA plot, p value: p values obtained after performing t-test, FDR: value obtained after performing false discovery test.

Table 6C: **Fold change and VIP score of Control vs Nutrient starvation stress in *M. smegmatis***

S.no	Metabolites	Fold Change	VIP score	p.value	FDR
1	2-Hydroxy-3-methylpentanoic acid	10.45	1.05	2.80E-07	5.30E-07
2	L-Leucine	4.91	0.84	1.80E-05	2.70E-05
3	Capric acid	3.88	0.62	8.10E-08	2.00E-07
4	Taurine	3.88	1.04	1.90E-10	8.80E-10
5	Methylamine	3.45	0.48	2.10E-03	2.60E-03
6	L-Lysine	2.88	0.62	1.20E-07	2.70E-07
7	L-Tyrosine	2.83	0.66	1.90E-08	5.30E-08
8	Malonic acid	2.68	1.65	7.70E-08	2.00E-07
9	Acetone	1.93	0.5	7.50E-12	4.20E-11
10	3-Hydroxyisovaleric acid	1.87	0.4	4.30E-16	2.10E-14
11	3-Methyl-2-oxovaleric acid	1.83	0.13	3.50E-04	4.80E-04
12	Ethanol	1.8	1.14	3.10E-13	2.60E-12
13	Betaine	1.75	0.4	7.40E-06	1.20E-05
14	Acetic acid	1.43	0.48	5.50E-03	6.30E-03
15	Formic acid	1.38	0.29	4.80E-04	6.30E-04
16	Dimethylamine	1.31	0.12	4.40E-01	4.50E-01
17	L-Lactic acid	1.3	0.32	4.60E-02	4.90E-02
18	Citrulline	1.28	0.43	4.40E-02	4.70E-02
19	D-Glutamic acid	1.22	2.27	5.00E-04	6.40E-04
20	L-Alpha-aminobutyric acid	1	0.0001	1.00E+00	1.00E+00
21	L-Homoserine	0.79	1.06	4.60E-03	5.40E-03
22	Cholic acid	0.73	0.13	1.50E-01	1.50E-01
23	L-Valine	0.7	0.29	2.60E-02	2.90E-02
24	L-Threonine	0.63	0.61	4.90E-03	5.70E-03
25	Nicotinamide N-oxide	0.57	0.23	2.90E-04	4.10E-04
26	Trehalose	0.56	2.02	2.00E-03	2.40E-03
27	NAD	0.5	0.7	2.60E-07	5.30E-07
28	ADP	0.46	0.31	4.20E-06	7.20E-06
29	Uridine diphosphate glucose	0.45	0.4	7.00E-07	1.30E-06
30	Inosinic acid	0.43	0.15	7.70E-06	1.20E-05
31	Isoleucine	0.41	0.23	1.50E-05	2.30E-05
32	Alanine	0.4	0.83	7.80E-09	2.30E-08
33	Succinic acid	0.35	0.88	2.40E-07	5.10E-07
34	Adenosine triphosphate	0.33	0.35	1.50E-05	2.30E-05

35	L-Phenylalanine	0.3	0.24	4.10E-15	1.00E-13
36	L-Glutamine	0.28	1.87	5.30E-09	1.70E-08
37	Uridine 5'-monophosphate	0.21	0.38	1.50E-09	5.80E-09
38	L-Aspartic acid	0.18	2.06	1.90E-09	6.90E-09
39	NADP	0.11	0.23	1.00E-07	2.40E-07
40	Uridine diphosphategalactose	0.08	0.29	8.90E-14	8.90E-13
41	Adenosine monophosphate	0.08	0.93	6.40E-13	4.60E-12
42	Uridine diphosphate-N-acetylglucosamine	0.08	0.29	4.30E-10	1.80E-09
43	D-Maltose	0.06	0.35	2.40E-14	4.00E-13
44	Thymidine 5'-triphosphate	0.05	0.37	7.80E-13	4.80E-12
45	Sucrose	0.04	0.36	7.20E-05	1.00E-04
46	1-Methylnicotinamide	0.04	0.4	1.10E-10	5.60E-10
47	Glucose 1-phosphate	0.04	0.41	3.80E-06	6.70E-06
48	Beta-Alanine	0.004	1.32	3.50E-09	1.20E-08
49	Citric acid	0.0024	1.65	1.40E-07	3.10E-07
50	L-Asparagine	0.0006	3.6	4.00E-14	5.00E-13

FC: fold change; VIP score: variable of importance score obtained from PLS-DA plot, p value: p values obtained after performing t-test, FDR: value obtained after performing false discovery test.

Table 8: Metabolic pathway impact analysis revealing the significantly impacted metabolic pathways under acidic, oxidative and nutrient starvation in *M. smegmatis*

Pathway impact analysis					
Acidic Stress	Total Cmpd[a]	Hits[b]	-log(p)[c]	FDR[d]	Impact[e]
Glycine, serine and threonine metabolism	28	4	20.31	1.09E-08	0.21
D-Glutamine and D-glutamate metabolism	7	2	18.68	3.70E-08	0.38
beta-Alanine metabolism	9	2	15.17	7.28E-07	1
Alanine, aspartate and glutamate metabolism	20	5	15.12	7.28E-07	0.69
Pyrimidine metabolism	37	2	13.04	4.23E-06	0.11
Purine metabolism	66	5	12.75	5.41E-06	0.14
Nicotinate and nicotinamide metabolism	13	3	9.89	8.09E-05	0.15
Methane metabolism	13	1	6.9	1.44E-03	0.1
Amino sugar and nucleotide sugar metabolism	38	4	4	2.47E-02	0.21
Streptomycin biosynthesis	10	1	3.23	4.86E-02	0.22
Citrate cycle (TCA cycle)	20	2	3.15	5.10E-02	0.12
Oxidative stress					
Nicotinate and nicotinamide metabolism	13	3	21.63	8.69E-09	0.15
Glycine, serine and threonine metabolism	28	4	19.08	4.63E-08	0.21
D-Glutamine and D-glutamate metabolism	7	2	18.39	7.41E-08	0.38
Amino sugar and nucleotide sugar metabolism	38	4	18.13	8.22E-08	0.21
Purine metabolism	66	5	16.23	3.50E-07	0.14
Alanine, aspartate and glutamate metabolism	20	5	16.23	3.50E-07	0.69
beta-Alanine metabolism	9	2	15.64	5.06E-07	1
Pyrimidine metabolism	37	2	7.13	0.002	0.11
Starch and sucrose metabolism	30	5	4.77	0.019	0.39
Nutrient starvation					
Alanine, aspartate and glutamate metabolism	20	5	34.86	7.84E-15	0.69
beta-Alanine metabolism	9	2	26.77	1.13E-11	1
Nicotinate and nicotinamide metabolism	13	3	21.8	1.22E-09	0.15
Purine metabolism	66	5	21.23	1.99E-09	0.14
Pyrimidine metabolism	37	2	19.33	1.02E-08	0.11
Amino sugar and nucleotide sugar metabolism	38	4	19.11	1.19E-08	0.21
Glycine, serine and threonine metabolism	28	4	18.28	2.37E-08	0.21
Citrate cycle (TCA cycle)	20	2	18.08	2.74E-08	0.12
Glyoxylate and dicarboxylate metabolism	22	2	16.2	1.46E-07	0.43

D-Glutamine and D-glutamate metabolism	7	2	13.3	2.05E-06	0.38
Streptomycin biosynthesis	10	1	12.49	4.36E-06	0.22
Methane metabolism	13	1	7.64	5.41E-04	0.1
Starch and sucrose metabolism	30	5	6.86	1.16E-03	0.39

Note: [a] Total number of metabolites in the pathway,[b] Number of matched metabolites, explained in Data Analysis section, [c]-log(P) is the negative natural log of the P value for each pathway, [d] False Discovery Rate (Benjamini-Hochberg),[e] Impact is the pathway impact value on each antibiotic treatmentcalculated from pathway topology analysis

Annexure I (PART 3B)

Tables for Comparative and quantitative metabolite profiling of pathogenic mycobacteria under microbicidal stresses

3.4.2 Tables for comparative and quantitative metabolite profiling of pathogenic mycobacteria under microbicidal stresses

Table 9: Identified metabolites with their respective retention time

S.no	Analyte Peak Name	Analyte Retention Time (min)	S.no	Analyte Peak Name	Analyte Retention Time (min)
1	7-Methylguanosine	9.28	45	L-Histidine	4.87
2	Adenine	19	46	L-Homoserine	12
3	Adenosine	8.78	47	L-Isoleucine	2.97
4	Adipic acid	2.98	48	L-Leucine	11.5
5	ADP	5.14	49	L-malic acid	3.6
6	a-KetoGlutaric acid	4.77	50	L-Methionine	12.1
7	Allantoin	12.8	51	L-phenylalanine-3,3-D2	19.6
8	AMP	9.75	52	L-Serine,	12
9	Arabinose	3.98	53	L-Threonine	12
10	ATP	5.75	54	L-Tryptophan	21.8
11	cAMP	20.5	55	L-Tyrosine	10.6
12	CDP	19.2	56	L-Valine	11.7
13	Citric acid	3.42	57	Maltitol	4.5
14	CMP	9.83	58	Maltose	4.53
15	Creatine	5.2	59	Mannose	4.18
16	CTP	3.09	60	Methyl malonate	3.66
17	Cytidine	9.86	61	mono-ethylmalonate	3.32
18	Cytosine	4.44	62	myo-inositol	4.04
19	isocitric acid	3.41	63	N-Acetylgalactosamine	9.04
20	Phenylalanine	19.6	64	N-Acetylglucosamine	9.04
21	Sorbitol	9.17	65	N-Acetylglycine	3.95
22	Ectoine	12.9	66	N-Acetyl-L-Glutamic acid	5.83
23	Fructose	4.18	67	NAD+	17.5
24	fumaric acid	3.96	68	Oxaloacetic acid	3.09
25	Galactose	4.18	69	Pyruvate	2.97
26	GABA	11.7	70	Ribitol	4.27
27	Glucose	4.21	71	Riboflavin	34.8
28	Glutathione Reduced	3.03	72	Sarcosine	12.1
29	GMP	12.8	73	Succinic acid	3.67
30	Guanidineacetic acid	11.7	74	Sucrose	4.52
31	Guanine	9.62	75	Taurine	4.97

32	Guanosine	9.31	76	Thymidine	4.54
33	Hippuric acid	28.6	77	Thymine	4.52
34	Inosine	8.03	78	tyramine	5.32
35	Kynurenic acid	3.81	79	UDP	3.33
36	L-2,Aminoadipic acid	3.82	80	UMP	32.6
37	Lactic acid	3.98	81	Uracil	12.7
38	Lactose	4.5	82	Uric acid	8.88
39	L-Arginine	13.3	83	Uridine	16.4
40	L-Ascorbic acid	23.5	84	UTP	3.15
41	L-Aspartic acid	12.3	85	Xanthine	8.72
42	L-Cittruline	12.8	86	Xanthosine dihydrate	3.23
43	L-Glutamic acid	12	87	Xylitol	4.27
44	L-Glutamine	12.4	88	Xylose	4.61

Tables 10 A, B and C

Table 10 A: Fold change and VIP score of control vs Acidic stress (*M. tb*)

Metabolite	FC	p.value	FDR	VIP
D-Xylitol	25.6	1.1E-06	9.4E-05	1.54
Ribitol	21.5	2.1E-06	9.4E-05	1.54
Inosine	7.1	9.4E-03	3.1E-02	1.25
Sorbitol	5.5	1.1E-04	1.6E-03	1.48
Adenosine	4.5	3.7E-02	7.9E-02	1.09
D-Maltose	3.5	1.7E-05	5.0E-04	1.51
L-Methionine	2.7	1.4E-03	9.7E-03	1.39
D-Mannose	2.0	5.2E-03	2.5E-02	1.30
Oxoglutaric acid	1.9	9.8E-05	1.6E-03	1.48
L-Homoserine	1.9	3.1E-04	3.9E-03	1.45
7-Methylguanosine	1.9	8.3E-03	2.9E-02	1.26
L-Threonine	1.8	7.2E-03	2.9E-02	1.28
D-Glucose.1	1.8	7.1E-03	2.9E-02	1.28
Thymine	1.8	2.9E-02	6.4E-02	1.12
GABA	1.7	1.7E-02	4.5E-02	1.19
D-Fructose	1.7	1.4E-02	4.1E-02	1.22
L-Arginine	1.6	2.4E-02	5.9E-02	1.15
L-Malic acid	1.6	9.5E-04	7.4E-03	1.41
Cyclic AMP	1.6	9.8E-04	7.4E-03	1.41
Alpha-Lactose	1.6	6.2E-04	5.6E-03	1.42
L-Leucine	1.6	1.5E-02	4.3E-02	1.21
Allantoin	1.6	1.6E-02	4.3E-02	1.20
Citrulline	1.6	2.7E-02	6.4E-02	1.13
N-Acetylglutamic acid	1.6	2.9E-02	6.4E-02	1.12
Uracil	1.5	2.8E-02	6.4E-02	1.13
Adenine	1.5	4.4E-02	8.8E-02	1.06
Uridine	1.5	5.0E-04	5.5E-03	1.43
Thymidine	1.5	7.5E-02	1.4E-01	0.97
Guanidoacetic acid	1.4	5.4E-02	1.0E-01	1.03
D-Galactose	1.4	5.5E-03	2.5E-02	1.30
L-Valine	1.4	7.4E-02	1.4E-01	0.97
Maltitol	1.4	5.5E-04	5.5E-03	1.43
Xanthosine	1.4	4.7E-01	6.0E-01	0.43
L-Serine	1.4	2.6E-02	6.4E-02	1.14
Adenosine triphosphate	1.3	4.1E-02	8.6E-02	1.07

Sucrose	1.3	3.3E-03	1.9E-02	1.34
D-Phenylalanine	1.3	1.5E-02	4.3E-02	1.20
L-Glutamine	1.3	1.8E-02	4.6E-02	1.19
Myoinositol	1.3	1.2E-01	2.2E-01	0.86
Creatine	1.3	3.0E-01	4.4E-01	0.61
L-Tryptophan	1.2	4.2E-02	8.7E-02	1.07
Cytidine	1.2	6.4E-01	7.2E-01	0.28
Arabinose	1.2	1.6E-01	2.7E-01	0.80
Guanine	1.2	2.0E-01	3.3E-01	0.74
N-Acetylgalactosamine	1.2	2.6E-01	3.9E-01	0.66
Guanosine	1.2	2.8E-01	4.2E-01	0.63
N-Acetyl-D-glucosamine	1.2	2.5E-01	3.8E-01	0.67
Uridine 5'-monophosphate	1.2	5.1E-01	6.2E-01	0.40
L-Isoleucine	1.2	3.2E-01	4.5E-01	0.59
D-Glucose	1.2	6.8E-01	7.5E-01	0.25
Hippuric acid	1.2	4.5E-01	5.9E-01	0.45
Creatine.1	1.1	4.2E-01	5.6E-01	0.48
Uridine 5'-diphosphate	1.1	4.1E-01	5.5E-01	0.49
L-Histidine	1.1	7.9E-03	2.9E-02	1.27
L-Tyrosine	1.1	2.3E-01	3.6E-01	0.70
Adipic acid	1.1	6.2E-01	7.1E-01	0.30
Riboflavin	1.1	5.8E-01	6.9E-01	0.33
Sarcosine	1.1	6.7E-01	7.5E-01	0.26
Taurine	1.1	7.4E-01	7.9E-01	0.20
Fumaric acid	1.0	9.0E-01	9.1E-01	0.08
Pyruvic acid.1	1.0	8.9E-01	9.1E-01	0.09
Nicotinamide adenine dinucleotide	1.0	8.2E-01	8.5E-01	0.14
D-Xylose	1.0	5.6E-01	6.7E-01	0.36
Monoethyl malonic acid	0.9	2.0E-01	3.3E-01	0.73
Oxalacetic acid	0.9	7.0E-01	7.6E-01	0.23
L-Glutamic acid	0.9	3.1E-01	4.4E-01	0.60
CDP	0.9	6.1E-01	7.1E-01	0.31
Ascorbic acid	0.9	9.6E-01	9.6E-01	0.03
Tyramine	0.9	4.8E-01	6.0E-01	0.43
Aminoadipic acid	0.9	6.1E-01	7.1E-01	0.31
Xanthine	0.8	2.1E-01	3.4E-01	0.72
Uridine triphosphate	0.8	4.8E-01	6.0E-01	0.42
Cytidine monophosphate	0.8	2.6E-01	3.9E-01	0.66
Adenosine monophosphate	0.8	1.7E-01	2.8E-01	0.79

Isocitric acid	0.8	1.9E-03	1.1E-02	1.37
Uric acid	0.8	7.9E-01	8.2E-01	0.16
Guanosine monophosphate	0.8	7.9E-01	8.2E-01	0.16
Citric acid	0.8	1.2E-02	3.8E-02	1.23
Ectoine	0.8	3.8E-01	5.1E-01	0.52
Glutathione	0.7	3.7E-01	5.1E-01	0.53
Cytidine triphosphate	0.7	4.9E-01	6.0E-01	0.41
L-Aspartic acid	0.7	5.3E-02	1.0E-01	1.03
Methylmalonic acid	0.7	8.1E-03	2.9E-02	1.27
Succinic acid	0.7	7.9E-03	2.9E-02	1.27
Acetylglycine	0.6	1.7E-03	1.1E-02	1.38
Pyruvic acid	0.6	1.4E-01	2.5E-01	0.83
ADP	0.6	1.0E-02	3.3E-02	1.25
Cytosine	0.6	4.8E-03	2.5E-02	1.31
Kynurenic acid	0.5	5.3E-03	2.5E-02	1.30
L-Lactic acid	0.3	6.0E-05	1.3E-03	1.49

FC: fold change; VIP score: variable of importance score obtained from PLS-DA plot, p value: p values obtained after performing t-test, FDR: value obtained after performing false discovery test.

Table 10 B: Fold change and VIP score of Control vs Oxidative stress (*M. tb*)

Metabolite	FC	p.value	FDR	VIP
Cyclic AMP	4.6	3.4E-05	0.001	2.1
Inosine	3.0	3.1E-05	0.001	2.1
Cytidine	2.2	4.6E-02	0.16	1.1
Guanine	2.1	3.1E-03	0.05	1.6
Thymine	2.1	1.1E-02	0.07	1.5
GABA	2.0	1.2E-02	0.07	1.3
D-Xylitol	1.9	5.0E-03	0.06	1.7
Guanosine	1.9	8.8E-03	0.07	1.4
Oxoglutaric acid	1.9	2.7E-02	0.11	1.2
Sorbitol	1.7	1.7E-02	0.09	1.3
Aminoadipic acid	1.7	4.9E-02	0.16	1.1
Thymidine	1.7	2.3E-02	0.11	1.3
L-Threonine	1.7	1.2E-02	0.07	1.5
7-Methylguanosine	1.6	2.0E-02	0.10	1.5
L-Leucine	1.5	8.2E-03	0.07	1.4
Uridine 5'-monophosphate	1.5	4.7E-01	0.68	0.5
Taurine	1.5	3.1E-01	0.52	0.7
Uric acid	1.4	3.5E-01	0.58	0.5
Guanidoacetic acid	1.4	4.3E-02	0.16	1.1
N-Acetylglutamic acid	1.4	3.1E-01	0.52	0.8
L-Homoserine	1.4	1.0E-02	0.07	1.6
Uridine	1.4	4.9E-02	0.16	1.2
Allantoin	1.4	1.3E-01	0.30	0.8
L-Valine	1.3	8.1E-02	0.23	1.0
L-Serine	1.3	2.3E-02	0.11	1.2
Myoinositol	1.3	1.1E-01	0.26	1.2
Adenine	1.3	4.0E-01	0.61	0.6
D-Maltose	1.3	7.8E-02	0.23	1.0
Citrulline	1.3	1.3E-01	0.31	0.8
Hippuric acid	1.3	4.8E-01	0.68	0.5
Alpha-Lactose	1.3	2.7E-02	0.11	1.6
Ribitol	1.3	2.8E-01	0.50	0.8
L-Isoleucine	1.3	9.3E-02	0.24	1.0
Adenosine	1.3	3.7E-02	0.14	1.3
D-Fructose	1.3	1.7E-01	0.37	0.9
L-Arginine	1.3	2.1E-01	0.45	0.8
L-Methionine	1.3	2.5E-01	0.48	0.8

Adenosine triphosphate	1.2	1.3E-01	0.31	0.9
L-Glutamine	1.2	1.4E-02	0.08	1.3
D-Mannose	1.2	3.1E-01	0.52	0.7
Uracil	1.2	2.2E-01	0.46	0.7
Glutathione	1.2	7.8E-01	0.89	0.2
Maltitol	1.2	7.9E-02	0.23	1.3
L-Malic acid	1.2	4.8E-01	0.68	0.4
Nicotinamide adenine dinucleotide	1.2	3.9E-01	0.61	0.6
Adipic acid	1.2	2.5E-01	0.48	0.8
Sucrose	1.2	9.0E-02	0.24	1.2
Oxalacetic acid	1.1	5.1E-01	0.69	0.5
L-Histidine	1.1	1.7E-01	0.37	0.9
Uridine 5'-diphosphate	1.1	5.1E-01	0.69	0.5
D-Xylose	1.1	2.3E-01	0.46	0.8
Methylmalonic acid	1.1	7.1E-01	0.87	0.2
L-Glutamic acid	1.1	4.6E-01	0.68	0.5
Sarcosine	1.1	6.0E-01	0.76	0.3
Uridine triphosphate	1.1	7.3E-01	0.87	0.3
Succinic acid	1.1	7.7E-01	0.89	0.2
L-Tyrosine	1.1	6.7E-01	0.83	0.3
Cytidine triphosphate	1.0	9.4E-01	0.96	0.1
Fumaric acid	1.0	9.1E-01	0.96	0.1
Riboflavin	1.0	8.5E-01	0.94	0.1
Cytidine monophosphate	1.0	9.2E-01	0.96	0.1
Cytosine	1.0	9.8E-01	0.98	0.0
D-Galactose	1.0	9.6E-01	0.98	0.0
L-Tryptophan	1.0	8.9E-01	0.96	0.1
Kynurenic acid	1.0	8.2E-01	0.93	0.1
Xanthosine	1.0	9.1E-01	0.96	0.1
Tyramine	0.9	8.4E-01	0.93	0.1
N-Acetylgalactosamine	0.9	7.5E-01	0.89	0.2
Ascorbic acid	0.9	9.3E-01	0.96	0.1
Adenosine monophosphate	0.9	4.7E-01	0.68	0.4
Monoethyl malonic acid	0.9	5.7E-01	0.75	0.4
Adenosine diphosphate	0.9	7.3E-01	0.87	0.3
D-Phenylalanine	0.9	3.6E-01	0.58	0.5
N-Acetyl-D-glucosamine	0.9	5.8E-01	0.76	0.4
Citric acid	0.8	1.0E-01	0.26	1.1
Xanthine	0.8	4.0E-01	0.61	0.7

Isocitric acid	0.8	5.3E-02	0.17	1.3
Pyruvic acid	0.8	2.6E-01	0.49	0.8
Arabinose	0.8	2.4E-01	0.47	0.7
D-Glucose	0.8	2.8E-01	0.50	0.8
Guanosine monophosphate	0.8	4.9E-01	0.68	0.4
CDP	0.7	6.1E-01	0.77	0.4
Acetylglycine	0.7	6.4E-03	0.07	1.5
L-Aspartic acid	0.5	1.3E-03	0.02	1.6
Ectoine	0.4	2.6E-02	0.11	1.2
L-Lactic acid	0.3	2.3E-04	0.01	1.9
Creatine	0.2	7.3E-05	0.002	2.1

FC: fold change; VIP score: variable of importance score obtained from PLS-DA plot, p value: p values obtained after performing t-test, FDR: value obtained after performing false discovery test.

Table 10 C: Fold change and VIP score of Control vs Iron stress (*M. tb*)

Iron	FC	p.value	FDR	VIP
Inosine	5.1	1E-06	1.3E-04	3.1
Guanosine	2.6	7E-04	1.6E-02	2.1
Guanine	2.5	3E-05	1.4E-03	1.8
Thymine	2.1	2E-03	0.03	1.2
Cytidine	2.1	2E-02	0.1	0.8
GABA	2.0	2E-03	0.03	2.3
D-Xylitol	1.9	1E-01	0.3	0.6
Thymidine	1.7	1E-02	0.1	1.5
Sorbitol	1.7	1E-02	0.1	0.9
Aminoadipic acid	1.7	5E-02	0.2	1.1
7-Methylguanosine	1.5	2E-02	0.1	0.8
L-Threonine	1.5	4E-02	0.1	0.7
Cytidine triphosphate	1.4	4E-01	0.6	0.4
L-Leucine	1.4	3E-02	0.1	1.3
L-malic acid	1.4	6E-03	0.05	1.3
Uridine	1.3	2E-01	0.4	0.7
Uric acid	1.3	1E+00	1.0	0.0
Ribitol	1.3	5E-01	0.7	0.3
L-Homoserine	1.3	1E-01	0.3	0.6
Uridine triphosphate	1.2	5E-01	0.7	0.7
Adenosine	1.2	9E-02	0.2	0.9
Oxoglutaric acid	1.2	3E-01	0.5	0.9
L-Tyrosine	1.2	3E-02	0.1	0.9
Myoinositol	1.2	8E-02	0.2	0.5
Oxalacetic acid	1.2	6E-01	0.7	0.3
Taurine	1.2	6E-01	0.7	0.3
L-Glutamine	1.2	8E-02	0.2	1.4
Guanidoacetic acid	1.1	4E-01	0.6	0.6
Fumaric acid	1.1	6E-01	0.7	0.3
L-Valine	1.1	5E-01	0.7	0.5
Uridine 5'-diphosphate	1.1	2E-01	0.5	0.5
Methylmalonic acid	1.1	5E-01	0.7	0.5
Xanthosine	1.1	5E-01	0.7	0.4
Succinic acid	1.1	5E-01	0.7	0.4
tyramine	1.1	9E-01	1.0	0.1
Riboflavin	1.1	5E-01	0.7	0.3
Cyclic AMP	1.1	7E-01	0.9	0.2

Alpha-Lactose	1.1	4E-01	0.7	0.4
D-Maltose	1.0	4E-01	0.6	0.3
Sucrose	1.0	8E-01	0.9	0.2
Maltitol	1.0	9E-01	1.0	0.03
D-Mannose	1.0	9E-01	1.0	0.03
Adenosine triphosphate	1.0	9E-01	1.0	0.04
D-Fructose	1.0	1E+00	1.0	0.01
Adenine	1.0	1E+00	1.0	0.0
Uridine 5'-monophosphate	1.0	3E-02	0.1	0.8
L-Serine	1.0	8E-01	0.9	0.1
L-Arginine	1.0	9E-01	1.0	0.1
D-Glucose	0.9	5E-01	0.7	0.2
Ascorbic acid	0.9	9E-01	1.0	0.3
L-Isoleucine	0.9	6E-01	0.8	0.3
Sarcosine	0.9	7E-01	0.8	0.2
D-Phenylalanine	0.9	4E-01	0.7	0.4
Cytidine monophosphate	0.9	5E-01	0.7	0.3
Nicotinamide adenine dinucleotide	0.9	5E-01	0.7	0.7
Xanthine	0.9	3E-01	0.5	0.4
L-Histidine	0.9	9E-02	0.2	0.5
L-Methionine	0.9	4E-01	0.7	0.5
N-Acetylglutamic acid	0.9	5E-01	0.7	0.7
Hippuric acid	0.9	8E-01	1.0	0.1
Adipic acid	0.9	3E-01	0.5	0.6
L-Tryptophan	0.9	1E-01	0.3	0.6
L-Glutamic acid	0.8	5E-02	0.2	1.5
Allantoin	0.8	2E-01	0.4	0.4
Citrulline	0.8	3E-01	0.5	0.7
D-Xylose	0.8	4E-02	0.1	0.7
Guanosine monophosphate	0.8	7E-01	0.8	0.2
Acetylglycine	0.8	1E-01	0.3	0.5
Kynurenic acid	0.8	4E-03	0.0	0.8
Uracil	0.8	6E-01	0.7	0.7
Creatine	0.8	2E-01	0.4	0.4
D-Galactose	0.8	2E-02	0.1	0.9
N-Acetyl-D-glucosamine	0.7	2E-01	0.4	0.6
Arabinose	0.7	1E-02	0.1	0.9
N-Acetylgalactosamine	0.7	1E-01	0.3	0.7
Adenosine monophosphate	0.7	5E-03	0.0	2.0

Citric acid	0.7	4E-02	0.1	2.7
Isocitric acid	0.7	6E-02	0.2	2.9
Monoethyl malonic acid	0.7	8E-02	0.2	1.5
Glutathione	0.6	2E-01	0.4	0.7
Pyruvic acid	0.6	1E-01	0.3	0.8
L-Lactic acid	0.6	7E-03	0.1	1.0
L-Aspartic acid	0.6	8E-03	0.1	2.1
Adenosine diphosphate	0.5	4E-02	0.1	1.2
Ectoine	0.3	5E-03	0.0	1.2
Cytosine	0.3	1E-04	0.0	1.7
CDP	0.1	2E-01	0.4	1.1

FC: fold change; VIP score: variable of importance score obtained from PLS-DA plot, p value: p values obtained after performing t-test, FDR: value obtained after performing false discovery test.

Table 10 D: Fold change and VIP score of control vs Nutrient starvation stress (*M. tb*)

	FC	p.value	FDR	VIP
Guanosine monophosphate	52.8	1.0E-04	4.2E-04	1.27
Aminoadipic acid	8.3	1.5E-05	1.2E-04	1.30
Inosine	7.9	5.2E-05	2.3E-04	1.28
Cytidine	7.4	1.4E-04	5.4E-04	1.26
Guanosine	7.3	2.8E-04	8.9E-04	1.24
7-Methylguanosine	5.4	1.5E-04	5.4E-04	1.26
Guanine	4.4	2.4E-05	1.6E-04	1.29
Uridine	3.7	5.8E-06	6.52E-05	1.31
L-Arginine	3.5	1.6E-05	1.2E-04	1.30
Oxoglutaric acid	3.0	7.9E-06	7.91E-05	1.30
L-Leucine	2.9	2.6E-05	1.6E-04	1.29
Sorbitol	2.9	5.5E-04	1.6E-03	1.22
Cytidine triphosphate	2.9	3.2E-02	5.6E-02	0.95
L-Serine	2.8	4.4E-05	2.2E-04	1.28
L-Threonine	2.3	8.5E-04	2.3E-03	1.21
L-Glutamic acid	2.2	2.9E-05	1.6E-04	1.29
L-Histidine	2.2	4.4E-05	2.2E-04	1.28
Tyramine	2.1	2.8E-02	5.1E-02	0.97
Uridine triphosphate	2.1	6.5E-02	9.7E-02	0.85
Pyruvic acid	2.0	4.1E-02	6.7E-02	0.92
Adenine	2.0	2.9E-02	5.1E-02	0.96
L-Homoserine	1.9	2.1E-03	5.0E-03	1.17
Riboflavin	1.9	1.5E-04	5.4E-04	1.26
Oxalacetic acid	1.9	1.7E-02	3.1E-02	1.02
CDP	1.8	0.4	4.9E-01	0.41
Uridine 5'-monophosphate	1.8	0.1	1.9E-01	0.72
Guanidoacetic acid	1.8	2.1E-03	5.0E-03	1.17
L-Valine	1.7	2.2E-03	5.0E-03	1.17
Uridine 5'-diphosphate	1.6	3.4E-02	5.8E-02	0.94
Citrulline	1.6	2.9E-02	5.1E-02	0.96
L-Isoleucine	1.6	2.1E-03	5.0E-03	1.17
Allantoin	1.6	4.2E-02	6.7E-02	0.92
Xanthosine	1.6	1.7E-01	2.3E-01	0.67
Adenosine triphosphate	1.5	1.2E-02	2.5E-02	1.05
Adenosine	1.5	1.2E-02	2.4E-02	1.05
Uracil	1.5	4.7E-02	7.2E-02	0.90
D-Glucose	1.5	1.9E-01	2.5E-01	0.65

Ascorbic acid	1.5	6.2E-01	7.0E-01	0.26
Adipic acid	1.4	5.9E-03	0.01	1.11
Thymine	1.3	2.4E-01	0.31	0.59
Methylmalonic acid	1.3	2.0E-01	0.26	0.63
L-Tryptophan	1.3	9.7E-03	0.02	1.07
Succinic acid	1.2	3.0E-01	0.38	0.52
Adenosine diphosphate	1.1	0.4	0.44	0.46
Acetylglycine	1.1	0.7	0.75	0.22
L-Methionine	1.0	0.9	0.96	0.04
N-Acetyl-D-glucosamine	1.0	0.9	0.95	0.05
Ectoine	1.0	1.0	0.99	0.02
Taurine	1.0	0.9	0.95	0.05
N-Acetylgalactosamine	1.0	1.0	0.99	0.003
Nicotinamide adenine dinucleotide	1.0	0.9	0.91	0.09
D-Mannose	1.0	0.8	0.84	0.15
Thymidine	0.9	0.8	0.91	0.10
Creatine	0.9	0.7	0.80	0.18
Cyclic AMP	0.9	0.3	0.41	0.48
Sarcosine	0.9	0.5	0.61	0.32
Cytidine monophosphate	0.9	0.5	0.58	0.34
L-Glutamine	0.8	0.1	0.12	0.81
Ribitol	0.8	0.3	0.38	0.51
D-Fructose	0.8	0.1	0.16	0.75
Maltitol	0.8	1.8E-03	4.6E-03	1.18
D-Xylitol	0.8	0.1	0.12	0.81
D-Galactose	0.8	0.2	0.26	0.63
Arabinose	0.7	0.1	0.09	0.87
L-Lactic acid	0.7	4.6E-02	0.07	0.90
Alpha-Lactose	0.7	4.9E-05	2.3E-04	1.28
D-Phenylalanine	0.7	1.4E-02	0.03	1.04
Sucrose	0.7	9.2E-06	8.3E-05	1.30
Xanthine	0.6	0.1	0.14	0.78
L-Tyrosine	0.6	6.1E-05	2.6E-04	1.28
L-Malic acid	0.6	1.4E-02	0.03	1.04
Uric acid	0.5	0.3	0.33	0.56
Adenosine monophosphate	0.5	1.5E-04	5.4E-04	1.26
Glutathione	0.5	4.2E-02	0.07	0.92
Cytosine	0.5	7.4E-04	2.1E-03	1.21
D-Maltose	0.4	1.2E-06	1.8E-05	1.32

N-Acetylglutamic acid	0.3	3.3E-04	1.0E-03	1.24
Kynurenic acid	0.3	1.8E-04	6.0E-04	1.26
Hippuric acid	0.2	8.7E-03	1.9E-02	1.08
Myoinositol	0.2	7.2E-07	1.3E-05	1.32
Fumaric acid	0.2	8.5E-03	1.9E-02	1.08
D-Xylose	0.2	4.1E-07	9.2E-06	1.32
Gamma-Aminobutyric acid	0.1	2.6E-05	1.6E-04	1.29
Monoethyl malonic acid	0.1	4.7E-09	1.40E-07	1.33
L-Aspartic acid	0.05	1.6E-06	2.01E-05	1.32
Isocitric acid	0.02	1.2E-09	5.25E-08	1.34
Citric acid	0.01	2.2E-10	2.00E-08	1.34

FC: fold change; VIP score: variable of importance score obtained from PLS-DA plot, p value: p values obtained after performing t-test, FDR: value obtained after performing false discovery test.

Table 12: Metabolic pathway impact analysis revealing the significantly impacted metabolic pathways under acidic,oxidative,iron and nutrient starvation stresses (*M. tb*)

Pathway impact analysis					
Acidic stress	Total Cmpd[a]	Hits[b]	-log(p)[c]	FDR[d]	Impact[e]
Citrate cycle (TCA cycle)	20	8	8.0	7.0E-03	0.43
Pyruvate metabolism	21	4	7.6	7.0E-03	0.23
Galactose metabolism	26	6	7.4	7.0E-03	0.11
Alanine, aspartate and glutamate metabolism	20	9	7.1	7.0E-03	0.98
Starch and sucrose metabolism	30	5	6.9	7.0E-03	0.25
Cysteine and methionine metabolism	37	5	6.9	7.0E-03	0.18
Pyrimidine metabolism	37	12	6.4	7.0E-03	0.44
Glyoxylate and dicarboxylate metabolism	22	5	6.2	9.0E-03	0.26
Glycine, serine and threonine metabolism	28	6	5.8	1.1E-02	0.36
Purine metabolism	66	13	5.5	1.4E-02	0.23
Aminoacyl-tRNA biosynthesis	66	12	5.1	1.9E-02	0.18
Butanoate metabolism	22	4	4.5	2.8E-02	0.15
Histidine metabolism	19	2	4.3	3.0E-02	0.10
Arginine and proline metabolism	40	9	4.3	3.0E-02	0.35
D-Glutamine and D-glutamate metabolism	7	2	3.8	4.0E-02	0.17
Methane metabolism	13	1	3.6	4.5E-02	0.10
Oxidative stress					
Cysteine and methionine metabolism	37	5	6.7	2.0E-02	0.18
Glycine, serine and threonine metabolism	28	6	5.7	3.0E-02	0.36
Purine metabolism	66	13	5.1	4.0E-02	0.23
Iron stress					
Purine metabolism	66	13	11.4	5.2E-04	0.23
Arginine and proline metabolism	40	9	7.4	9.4E-03	0.35
Alanine, aspartate and glutamate metabolism	20	9	6.8	1.3E-02	0.98
Pyruvate metabolism	21	4	5.9	2.7E-02	0.23
Aminoacyl-tRNA biosynthesis	66	12	5.2	3.2E-02	0.18
Glycine, serine and threonine metabolism	28	6	5.1	3.2E-02	0.36
Cysteine and methionine metabolism	37	5	4.8	3.3E-02	0.18
D-Glutamine and D-glutamate metabolism	7	2	4.6	3.3E-02	0.17
Butanoate metabolism	22	4	4.4	3.8E-02	0.15
Nicotinate and nicotinamide metabolism	13	2	4.2	4.2E-02	0.15

Niutrient starvation stress					
Arginine and proline metabolism	40	9	19.7	1.00E-07	0.3
Alanine, aspartate and glutamate metabolism	20	9	15.1	3.00E-06	1
Aminoacyl-tRNA biosynthesis	66	12	14.1	6.00E-06	0.2
Inositol phosphate metabolism	15	1	14.1	6.00E-06	0.1
Citrate cycle (TCA cycle)	20	8	13.9	6.00E-06	0.4
Glycine, serine and threonine metabolism	28	6	12.7	2.00E-05	0.4
Starch and sucrose metabolism	30	5	12.3	2.00E-05	0.3
Galactose metabolism	26	6	11.8	3.00E-05	0.1
Glyoxylate and dicarboxylate metabolism	22	5	11.5	4.00E-05	0.3
Histidine metabolism	19	2	11.4	4.00E-05	0.1
Purine metabolism	66	13	10.9	5.00E-05	0.2
Nicotinate and nicotinamide metabolism	13	2	10.3	8.00E-05	0.2
Methane metabolism	13	1	10	1.00E-04	0.1
Pyrimidine metabolism	37	12	9.7	1.00E-04	0.4
D-Glutamine and D-glutamate metabolism	7	2	9.3	2.00E-04	0.2
Butanoate metabolism	22	4	9.3	2.00E-04	0.1
Riboflavin metabolism	11	1	8.8	3.00E-04	0.2
Cysteine and methionine metabolism	37	5	8.6	3.00E-04	0.2
Glycolysis or Gluconeogenesis	30	4	6.4	2.00E-03	0.1
Glutathione metabolism	17	2	6.1	3.00E-03	0.3
Pyruvate metabolism	21	4	5.8	4.00E-03	0.2
Fructose and mannose metabolism	19	2	5.3	6.00E-03	0.1

Note: [a] Total number of metabolites in the pathway,[b] Number of matched metabolites, explained in Data Analysis section, [c]-log(P) is the negative natural log of the P value for each pathway, [d] False Discovery Rate (Benjamini-Hochberg),[e] Impact is the pathway impact value on each antibiotic treatmentcalculated from pathway topology analysis.

Chapter 4

Comparative metabolic profiling of sera from TB patients, their household clinically healthy contacts and unrelated clinically healthy volunteers

4.1 Overview of the chapter

Chapter 1 introduced the pathogenesis of TB disease and various challenges associated with its control, new disease indicators/markers for early detection of latent and active TB being one of the fundamental challenges. This chapter pertains to my second objective where attempt to identify a possible TB disease indicator by comparing metabolites from sera of patients, their respective household contacts and healthy volunteers were made, which may serve as a lead disease indicator for further studies in lab. I hypothesized that a disease state brings about a plethora of changes in metabolism of human host that gets reflected in the alterations in the sera composition including levels of these intermediate products of metabolic reactions. Comparing sera from diseased (here TB patients) vs healthy (here comprising of two groups; household healthy contacts of patients and unrelated healthy volunteers) will lead to identification of metabolite profiles that can distinguish diseased from healthy state, serving as biological indicators for TB disease.

Under this objective, sera from well-defined cohort was collected, compared for metabolite profiles using NMR-based metabolomics, subjecting differential metabolites to "Biomarker Analysis tool" to explore their potential as TB disease indicators. The study encompassed population from two different geographical regions (i) India, region with high TB burden (ii) Portugal, region with low TB burden. The collection, processing and analysis of Indian samples were done by me and further resultant important variables were identified in my study was compared with another study done independently by our collaborators in Portugal. The study (carried under the DBT New-Indigo scheme) was designed such that it permitted comparative cross-sectional analyses of metabolite profiles of Indian and Portugal populations. The shortlisted differential metabolites from PLS-DA were compared to similar exercise performed on Portuguese population (Portugal representing low TB burden country) to finally fix on three metabolites that can have the potential to distinguish TB from healthy sera samples over a geographical range of population. From analysis of Indian samples alone, a set of four metabolites could be identified that may have more distinguishing potential for Indian population exclusively. The methodology used and the study cohort have been detailed in Chapter 2. To our

knowledge, this is the first unbiased NMR based comparative metabolomic study on TB patients, house hold contacts and healthy cohort from India.

4.1.1 Results

Identification of metabolites and chemical shift assignment

A total of 36 NMR spectra were considered for analyses, which included 12 healthy, 12 patients and 12 respective household contacts of the patients. All the ^{1}H NMR spectra were manually phased and baseline-corrected using Topspin (v3.5) software (www.bruker.com/bruker/topspin). ^{1}H chemical shift dimension was directly referenced to the TSP resonance. For total correlation spectroscopy (TOCSY), prior to Fourier transform, the FIDs were weighed in both dimensions by a sine-bell function and zero-filled to 2048 and 1024 data points in F1 and F2 dimensions respectively. The distinctive chemical shifts in the frequency domain originating from sets of spectral measurements were assigned to respective metabolites. Metabolite present in ^{1}H NMR spectrumwere identified using Chenomx NMR Suite 8.1 software and confirmed with biological magnetic resonance data bank (BMRB) database (Ulrich *et al*. 2008) and human metabolome database (HMDB) (Wishart *et al*. 2007). Around 33 metabolites were assigned using the above strategy and marked on the ^{1}H NMR spectra, which are tabulated in **Table 13** (Annexed at the end of this chapter) along with their respective ^{1}H chemical shifts (with reference to TSP). **Figure 23** provides snapshots comparing representative spectra from each group. It can be clearly seen that the spectra from same group were similar, while being distinct from other groups. The arrows indicated in the figure point to some of such distinguishing spectral peaks (**Figure 23**).

Following inferences could be drawn from these observations :(i) The perturbations in metabolism during TB disease is reflected in the peripheral sera of TB patients. (ii) There indeed are indeed distinctive patterns of NMR metabolite spectra for each category, healthy, patients and their respective household contactsthat can be used for distinguishing these categories.

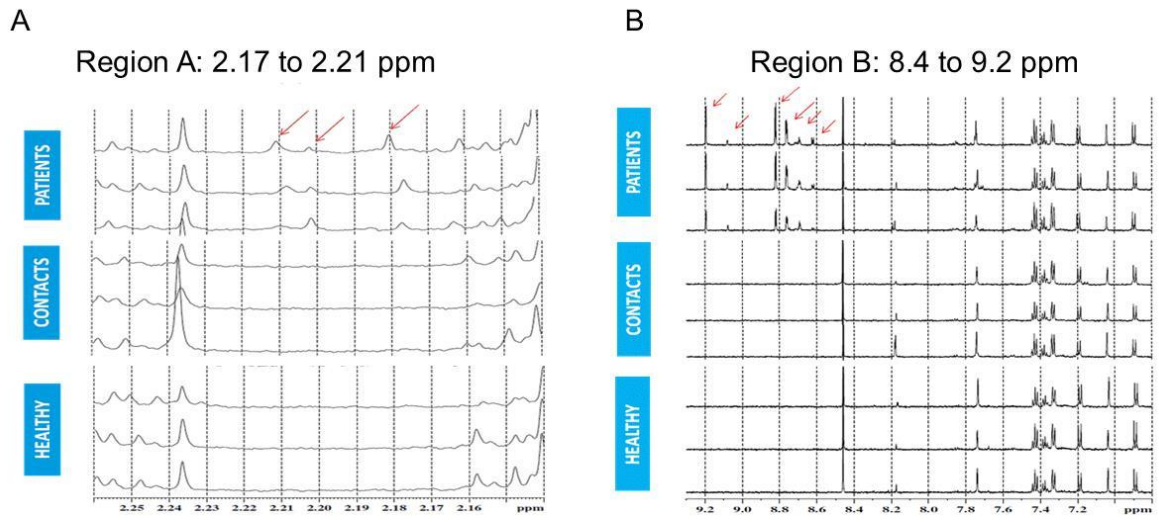

Figure 23: Representative NMR metabolite spectra from sera samples of Patients, Contacts (respective household contacts of patients) and Healthy categories of two regions (A) 2.17 to 2.21 ppm (B) 8.4 to 9.2 ppm. Arrows indicate some of the differential peaks in the spectra.

4.1.1.2 Quantitative Enrichment Analysis of the identified metabolites:

Analogous to Gene Set Enrichment Analysis (GSEA) software commonly used in transcriptomic data analysis is Metabolite Set Enrichment Analysis (MSEA) which supports identification and interpretation of human metabolite pattern and concentration changes in a biological perspective. MSEA indicated pathways such as phosphatidylcholine metabolism, phospholipid metabolism, phophoatidylethanolamine metabolism, aspartate metabolism, fatty acid metabolism, ammonia recycling, keto-bodies metabolism, phenylanine and tyrosine metabolism to be significantly associated with healthy as compared to 'patients' as well as 'contacts' categories **(Figure 24A and 24B)**. While for contacts, as compared to patients, the pathways perturbed **(Figure 24C)** include sperimidine and spermine biosynthesis, fructose and mannose degradation, propanoate metabolism, urea cycle, glycine, serine metabolism and homocysteine degradation. The biological significance of such differences has been briefly discussed in later under 'discussion' section.

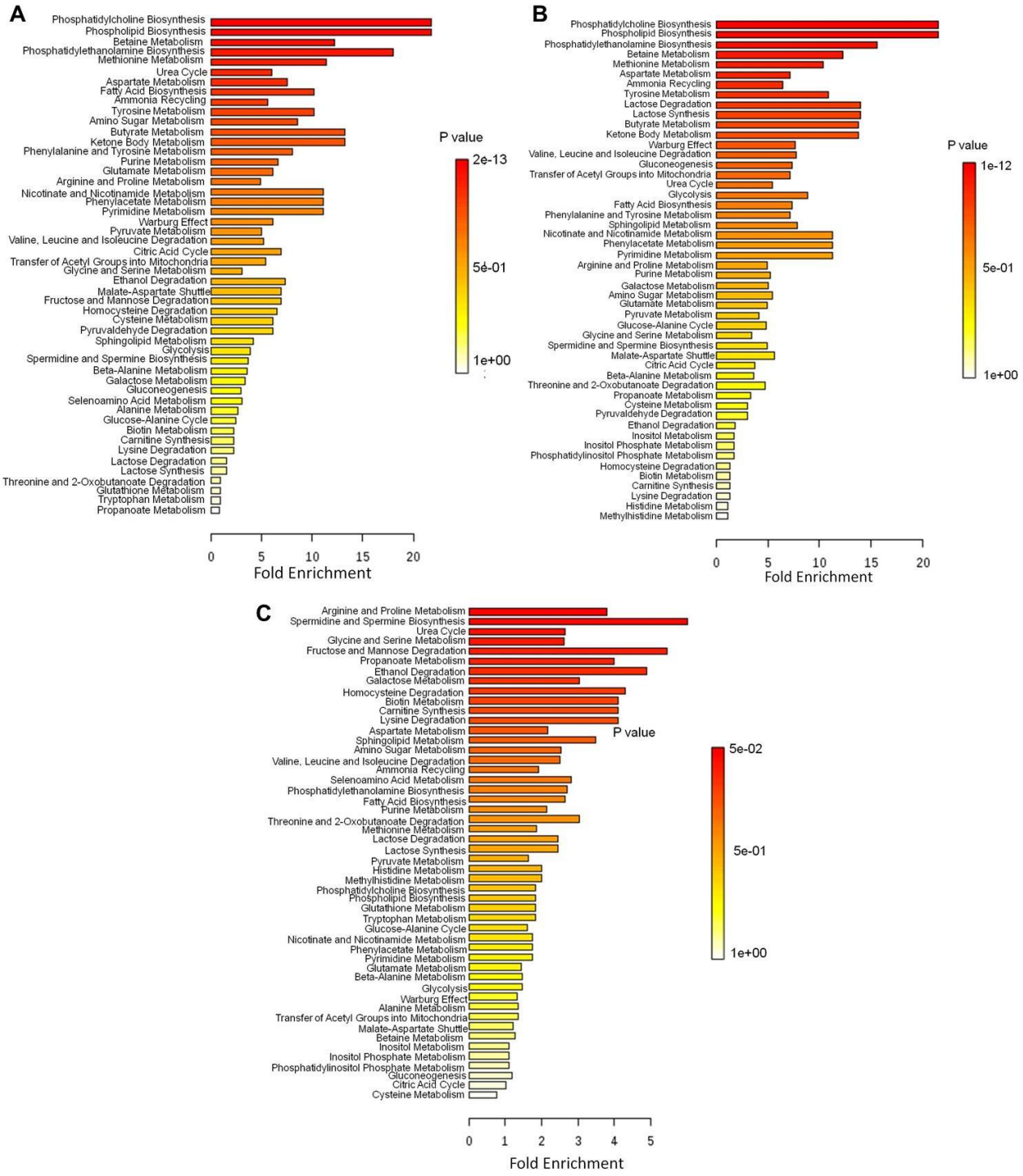

Figure 24: Comparative pathway enrichment analyses: Quantitative pathway enrichment with 33 metabolites were evaluated using the MSEA amongst (A) Patients vs Healthy, (B) Contacts vs. Healthy (C) Contacts vs Patients showing the pathways enriched in each category.Colourdifferences demonstrate the relative significance (P value) of pathway acrossthe different groups, while length of horizontal bars indicates fold enrichment with respect to pathway.

4.1.1.3 Identification of differential metabolic signatures using statistical analysis

To identify differential metabolomic signatures associated with the given groups (patients, contacts and healthy), uni-variate and multi-variate statistical data analysis were carried out using MetaboAnalyst 3.1. To begin with, uni-variate analysis, one-way Analysis of Variance (ANOVA) was performed, to identify differences between all the three groups. The F (variation between sample means / variation within the samples) and P (statistical significance) values of each metabolite are tabulated in **Table 14**. Principal component analysis (PCA) was then performed to examine the intrinsic variation in groups. Principal component analysis (PCA) was then performed to examine the metabolic signatures associated with different groups. The total variance explained by five components PCA analyses were 70.94% for healthy vs patients, 68.9% for contact vs healthy, 61.4% for contact vs patients. **Figure 25** represents divergent separation on the score plot of the first two principal components PC1 and PC2. Percentage variance explained by PC1 and PC2 for respective groups are mentioned in their respective plots. The 2D-PCA score plots for healthy vs patients (**Figure 25A**) and contact *vs* healthy (**Figure 25B**), though segregated the groups from each other, do not show significant differences from each other. Interestingly, we observed a clustering of 'contacts' and 'patients' in similar quadrant of PCA plot in these analyses (**Figure 25C**), which suggests that the contacts, which are supposedly exposed to *M. tb* through patients, showed almost similar sera metabolite profile to that of drug naïve patients at the point of diagnosis. Comparing all the groups showed partial discrimination (**Figure 25D**).

The data was then subjected to partial least square discrimination (PLS-DA) to maximize the separation between healthy, contacts and patients (**Figure 26**)**.** As detailed in previous chapters, PLS-DA considers the categories and tries to reduce the dimension while maximizing the separation of categories, hence has the discriminatory potential ideal for classification of categories.Clear discrimination between patients, contacts and healthy, could be observed (**Figure 26D**).The quality (goodness of fit) of the PLS-DA models was described by R^2 and Q^2values whichare given in table (**Table 15).** Values when closer to 1 indicate good discriminatory abilities of the model. The table clearly shows that while Patient from Healthy and Contact from Healthy can be discriminated, Contacts showed considerable overlap with

Patients in terms of metabolite signatures. This may be because of their exposure to *M. tb*, though they remained asymptomtic.

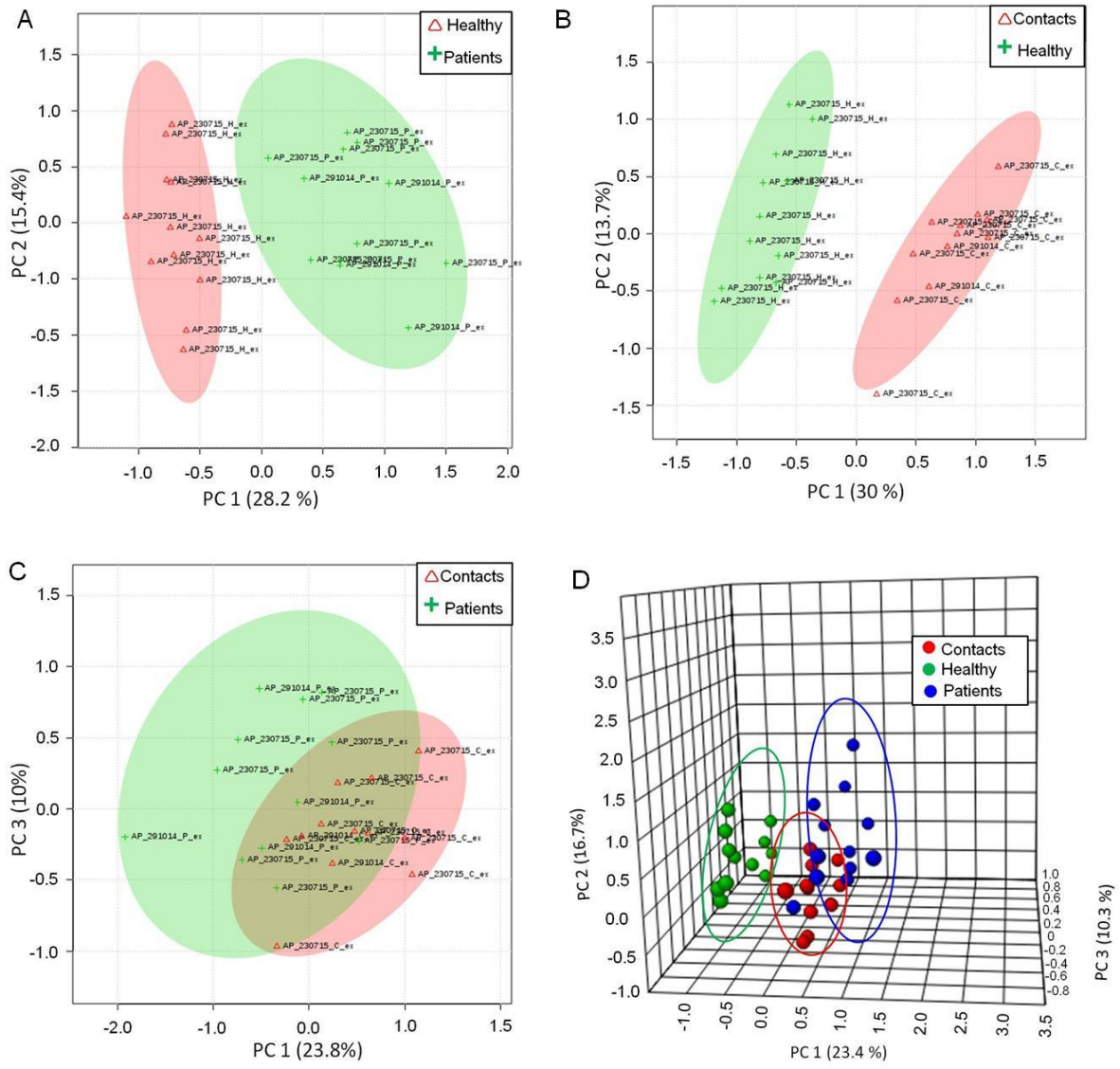

Figure25: Principal Component Analysis differentially segregates metabolites from patients, contacts and healthy: Principal Component Analysis (PCA) 2D score plot of A) Healthy *vs* Patients. B) Contact *vs* Healthy. C) Contact *vs* Patients. D) 3D score plot of Healthy *vs* Patients and contacts. The dots inside the all the plots correspond to samples numbers under each category.

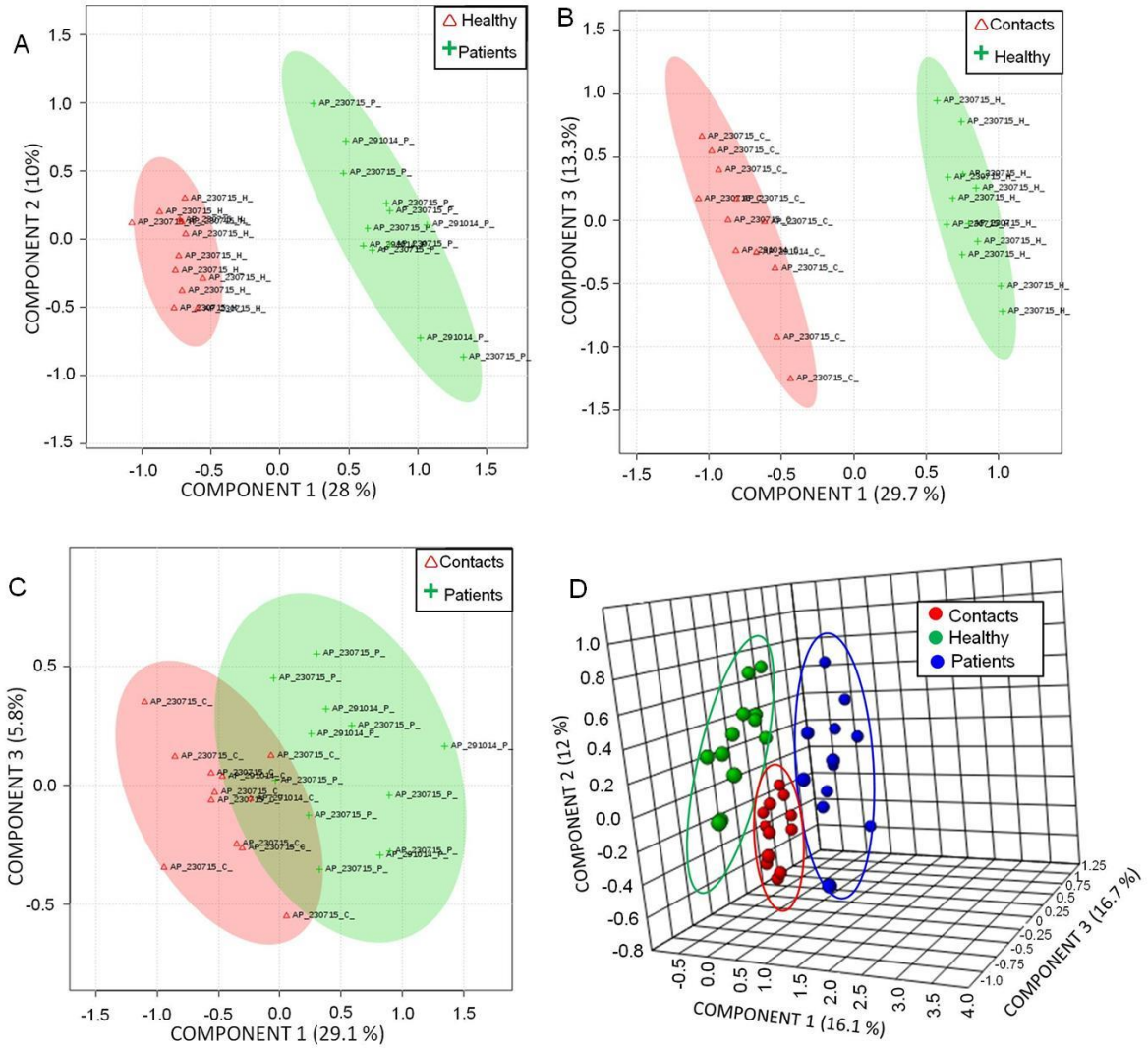

Figure 26: Partial Least Squares - Discriminant Analysis segregates healthy from patients and contacts

Partial Least Squares - Discriminant Analysis (PLS-DA) 2D score plot of A) Healthy vs Patients. B) Contacts vs Healthy. C) Contacts vs Patients. D) 3D score plot of Healthy vs Patients and Contacts. PLS-DA shows better segregation between the groups analysed under A, B and C as compared to PCA analysis. When all the groups were analysed together (D) showed distinct segregation of contacts from healthy and patients. The dots inside the all the plots correspond to samples numbers under each category.

Groups	**Measure**	**1 comps**	**2 comps**	**3 comps**	**4 comps**	**5 comps**
	Accuracy	1	1	1	1	1
Healty vs Patients	R2	0.92	0.97	0.98	0.99	1.00
	Q2	0.88	0.89	0.90	0.88	0.87
Healthy vs Contacts	R2	0.95	0.99	0.99	1.00	1.00
	Q2	0.90	0.95	0.96	0.96	0.96
Contacts vs Patients	R2	0.64	0.85	0.93	0.97	0.98
	Q2	0.41	0.51	0.40	0.28	0.25
Healty vs Contacts vs Patients	R2	0.59	0.78	0.83	0.88	0.92
	Q2	0.38	0.44	0.44	0.27	0.12

Table 15: R^2 and Q^2 values of Partial Least Squares - Discriminant Analysis (PLS-DA) Comps: component

The variable importance of projection (VIP) scores obtained from the PLS-DA model were used to identify key metabolic features significantly altered in the sera of healthy from patients (**Figure 27A**), healthy from contacts (**Figure 27B**), contacts from patients (**Figure 27C**) and comparing all the three groups (**Figure 27D**). Metabolites differentiating various categories with their VIP scores are tabulated as **Tables 16A, B and C.** The analyses showed distinct differences in healthy metabolite profile as comparedwith that of house hold contacts and patients. Based on VIP score of more than 1, common metabolites distinguishing healthy *vs* patients and contacts areornithine, creatinine, proline, lysine, hypoxanthine, 2-hydroxybutyric acid, S-3-hydroxyisobutyric acid (3HIBA), valine, mannose, acetic acid, glutamic acid serine and threonine **(Figure 27D)**. It was interesting to note that metabolites ornithine and creatinine showed unusual trend in levels with the serum levels highest in contacts, intermediate in healthy and least in patients, while metabolites proline, lysine, hypoxanthine, valine and hydroxybutyrate showed the reverse with contacts showing the least levels and patients showing the highest. These metabolites remained at intermediate levels in healthy. If the sera levels of healthy for

these metabolites are considered as normal, then it is clear that both infection as well as exposure to *M. tb* disturbs these metabolite levels which gets reflected in peripheral blood. Mannose, acetic acid and glutamic acid levels were in ascending pattern with healthy least and patient highest (healthy < contacts < patients). Serine and threonine levels were high in healthy while in patients and contacts their levels are lowered. Distinguishing metabolites specific to healthy vs patients were citric acid and aspartic acid. The citric acid levels were low in patients and high in healthy while for aspartic acid, it was opposite. Similarly, distinguishing metabolites specific to healthy vs contacts were asparagine and glucose, with their levels higher in contacts than healthy. 2-hydroxybutyric acid, S-3-hydroxyisobutyric acid, lysine, proline, valine, hypoxanthine could distinguish between patients and their close contacts where all these metabolites were high in patients compare to contacts. Creatinine however was lower in patients than contacts.
From this set of analysis,the first three metabolites that had high scores in discriminating all the three groups in Indian cohort **(Figure 27D)**, **ornithine, creatinine and proline** were selected for further validation through Biomarker Analysis Tool.

In addition to that, to search for a more universal disease indicator, the data generated independently on Portugal cohort was used. As mentioned in Chapter 2 (Materials and Methods) under 'Study Cohort' the study included parallel analyses on population from two different geographical regions (i) India, region with high TB burden (ii) Portugal, region with low TB burden.The independent analyses by Portugal collaborators shortlisted three metabolites showing the highest VIP scores, mannose (VIP score: 2.48), hypoxanthine (VIP score: 2.54) and glutamic acid (VIP score: 1.93), the levels of which were low in Healthy and high in Patients (unpublished data shared). We observed that these metabolites also appeared in our PLS-DA analysis (VIP scores > 1) where all the three groups were compared **(Figure 27D)**. The levels of metabolites were lower in Healthy sera than in Patients from India, similar to that seen in Portuguese cohort. Hence, **glutamic acid, mannose and hypoxanthine** were selected for further validation through Biomarker Analysis Tool.

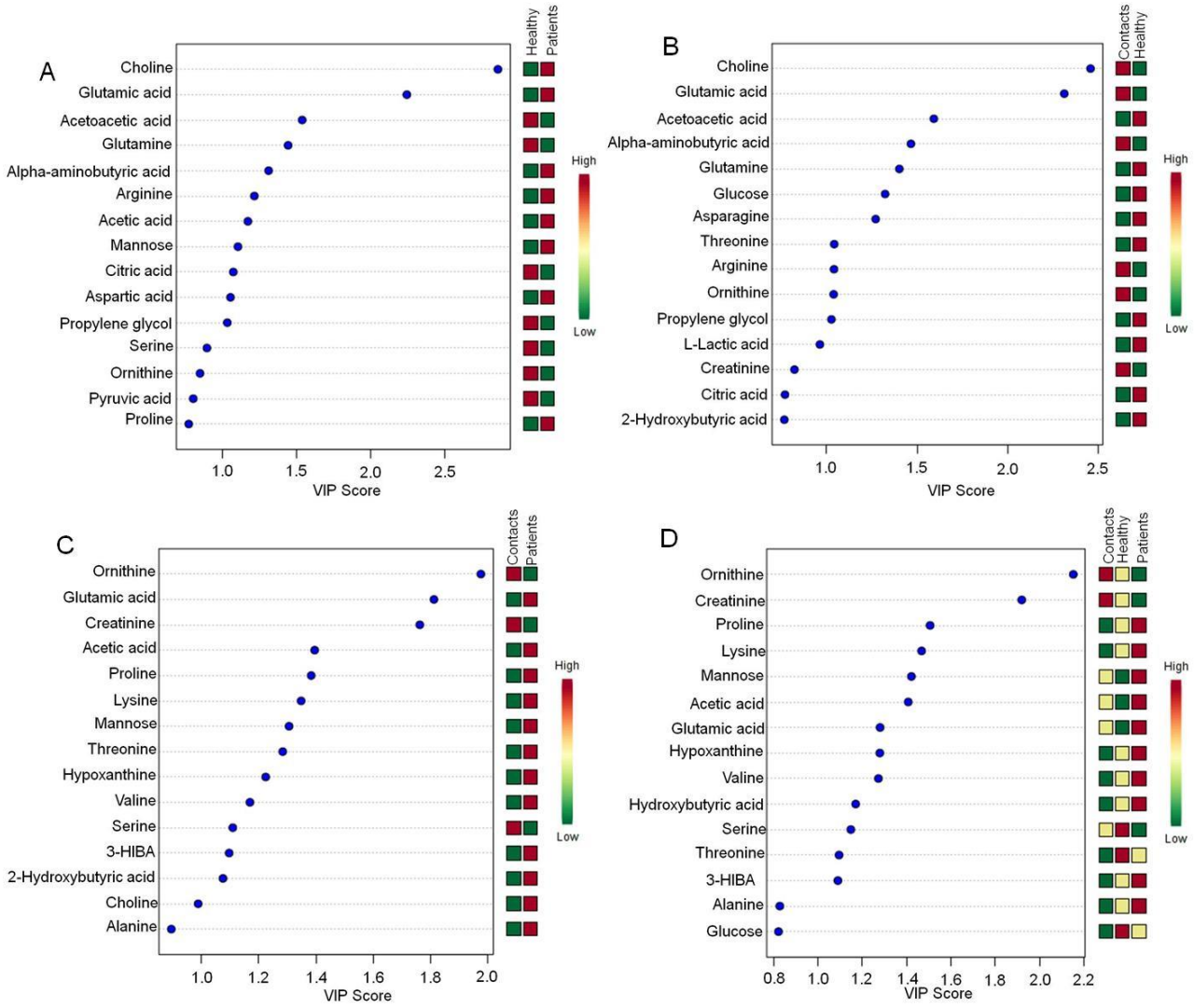

Figure 27: Variable Importance in Projection (VIP) plots indicating the most discriminating metabolite, identified through PLS-DA analyses, in descending order of importance for A) Healthy *vs.* Patients; B) Contacts *vs.* Healthy; C) Contacts *vs.* Patients; D) Healthy vs. Patients and contacts. These important variables are responsible for the segregation among the groups as indicated.

4.1.1.4 Evaluation and validation of shortlisted metabolites for their potency disease indicators (biomarkers) using receiver operating characteristic (ROC) curve (Biomarker analysis services from MetaboAnalyst)

From the above analyses, the following two shortlisted set of metabolites were subjected to further in-silico evaluations and validations:

SET-A: Ornithine, Creatinine and Proline

SET-B: Glutamic acid, Mannose and Hypoxanthine

The metabolite sets were evaluated for their potency as (a) active disease marker using Healthy and Patients data and (b) latency marker using Contacts and Patients data

4.1.1.4.1 Potency as active disease indicators (Healthy vs Patients)

ROC curve based model evaluation was performed selecting either SET-A or SET-B to create biomarker models using PLS-DA for Healthy and Patients. Figure 6 shows the images of the ROC generated using SET-A (Figure 28A) and SET-B (Figure 28B). The algorithm in MetaboAanalyst performs 100 cross validations (CV) and results are averaged to generate a smooth ROC curve.

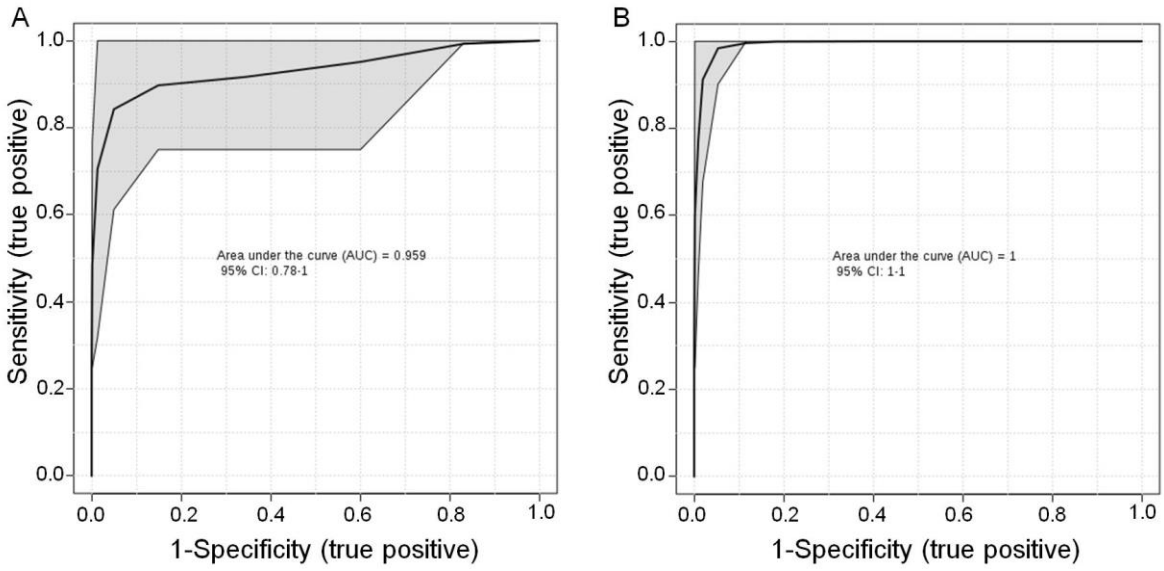

Figure 28: ROC curve analysis for the trained model for Healthy vs Patients (A) ROC analysis of SET-A shows AUC 0.959 and CI 0.78-1(B) ROC analysis of SET-B shows AUC 1.0 and CI 1.0-1.0. ROC analysis were calculated for the created biomarker model and expressed as areas under the curve (AUC). A confidence level of 95% was achieved through this model which defines the variables as promising potential biomarkers.

The predicted class probabilities (average of the cross-validation) for each sample using SET-A and SET-B are illustrated in **Figure 29**. For SET-A (**Figure 29A**), the verification results showed that in a total of 24 samples comprising of healthy (n=12) and Patients (n=12), 23 were predicted correctly. One can see that while all Healthy samples (open circles) could be predicted correctly (clustered in distinct quadrants), out of 12, 11 Patient samples (closed circles) could be predicted correctly. Therefore, using SET-A, the PLS-DA prediction model exhibited a sensitivity of 91.6% but an absolute specificity (100%).

For SET-B **(Figure 29B)**, the verification results showed that all 24 samples could be predicted correctly. One can see that all Healthy samples (open circles) and Patient samples (closed circles) could be predicted correctly and clustered in distinct quadrants. Therefore, using SET-B the PLS-DA prediction model exhibited anabsolute sensitivity and specificity.

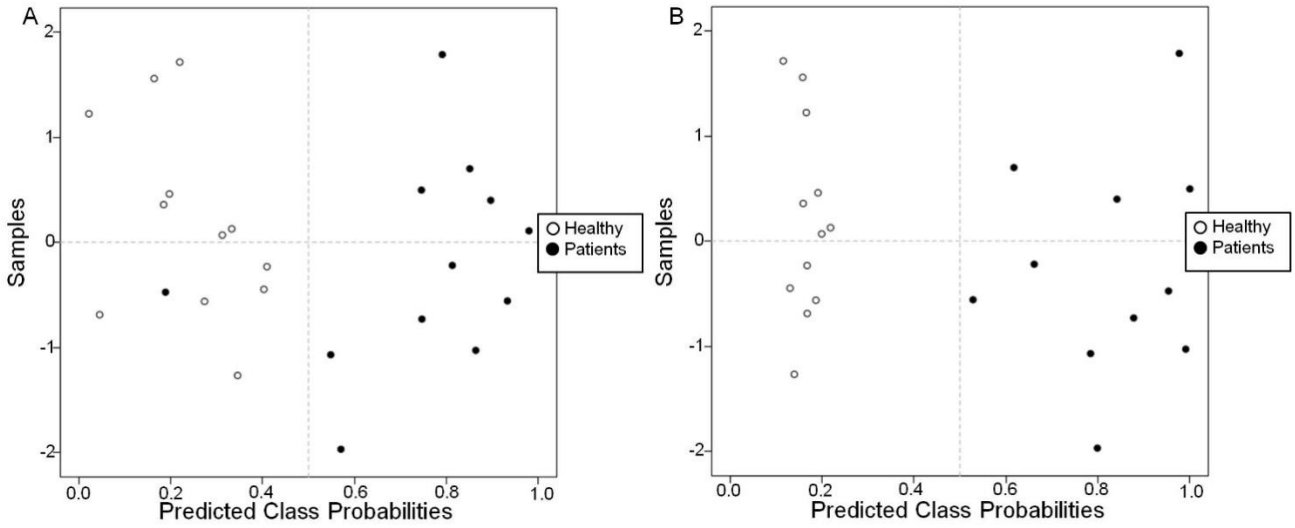

Figure29: ROC curve based model evaluation based on SET-A and SET-B for Healthy vs Patients A) Image showing the predicted class probabilities of all samples using SET-A. B) Image showing the predicted class probabilities of all samples using SET-B. The classification boundary is at the center (x=0.5, dotted line).

4.1.1.4.2 Potency as latency indicators (Contacts vs Patients)

Similarly, ROC curve based model evaluation was performed with either SET-A or SET-B to create biomarker models using PLS-DA for Contacts and Patients. Figure 30 shows the images of the ROC generated using SET-A (Figure 30A) and SET-B (Figure 30B).

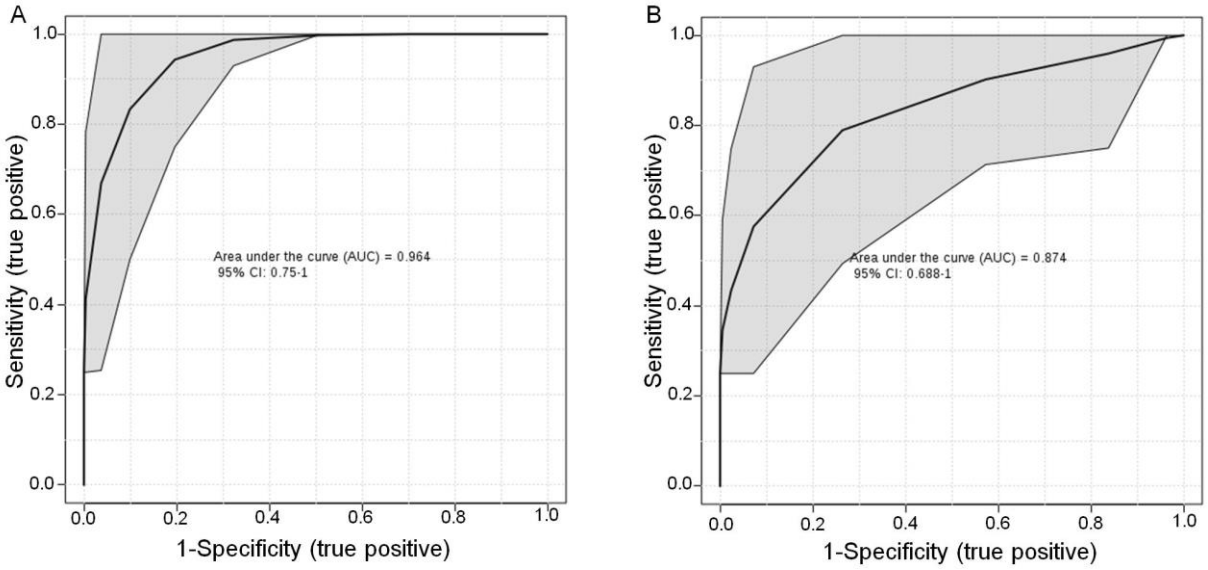

Figure 30: ROC curve analysis for the trained model for Contacts vs Patients: (A) ROC analysis of SET-A shows AUC 0.964 and CI 0.75-1(B) ROC analysis of SET-B shows AUC 0.87 and CI 0.688-1.0. ROC analysis were calculated for the created biomarker model and expressed as areas under the curve (AUC). A confidence level of 95% was achieved through this model which defines the variables as promising potential biomarkers.

Figure 31 shows predicted class probabilities (average of the cross-validation) for each sample (Contacts and Patients) using SET-Aand SET-B. For SET A (**Figure 31A**), the verification results showed that in a total of 24 samples comprising of Contacts (n=12) and Patients (n=12), 21 were predicted correctly. One sample of Contacts (open circles) was incorrectly placed, while one was place on borderline, hence considered incorrect.Out of 12, 11 Patient samples (closed circles) could be predicted correctly.Therefore, using SET-A, the PLS-DA prediction model exhibited 91.6% of sensitivity and 83% specificity for SET-A as latency indicators.For SET-B **(Figure 31B)**, the verification results showed that in a total of 24 samples (Contacts+ Patients), 19 were predicted correctly. It was observed that two samples each from Contacts and Patients

were predicted wrong and one sample of Contacts was at borderline and was considered incorrect prediction (**Figure 31B**). Therefore, the PLS-DA prediction model exhibited 83% sensitivityand 75% specificityusing SET-B for latency.

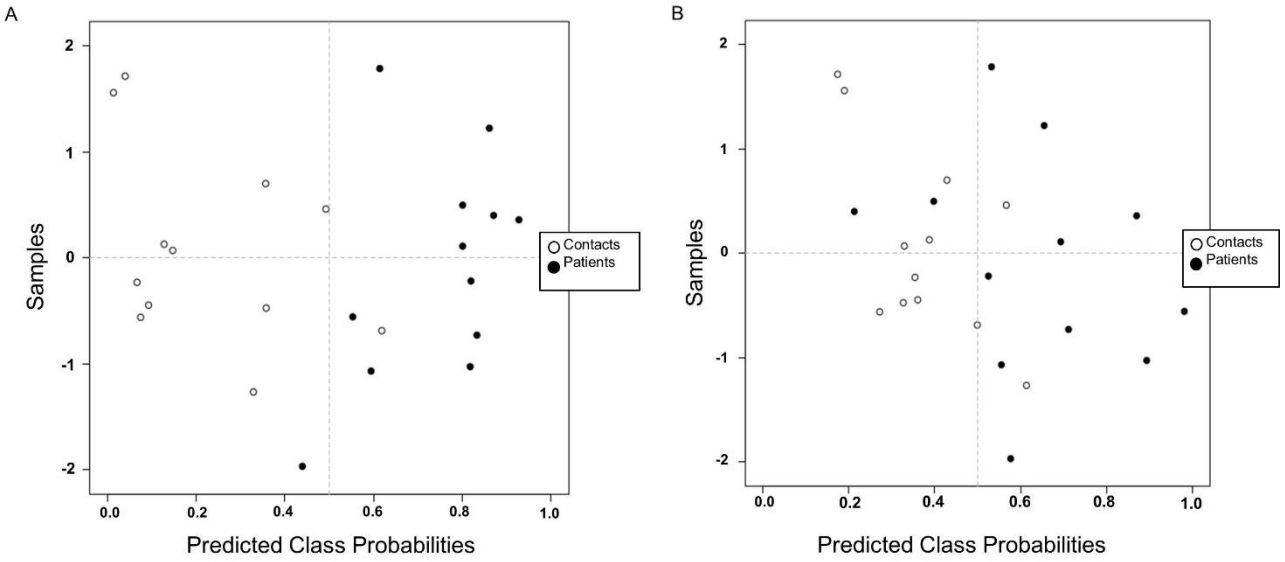

Figure 31: ROC curve based model evaluation based on SET-A and SET-Bfor Contacts vs Patients A) Image showing the predicted class probabilities of all samples using SET-A. B) Image showing the predicted class probabilities of all samples using SET-B. The classification boundary is at the center (x=0.5, dotted line).

Table 17 below summarizes the results regarding the prediction of the potencies of these sets of metabolites for active and latent disease indicators.

Metabolite Set	Metabolites	Sensitivity as active disease indicator	Specificity as active disease indicator	Sensitivity as latency indicator	Specificity as latency indicator
SET-A	Ornithine, Creatinine and Proline	91.6%	100%	91.6%	83%
SET-B	Glutamic acid, Mannose and Hypoxanthine	100%	100%	83%	75%

`**Table 17:** Summary of predictions for SET-A and SET-B metabolites for their potency as disease indicators based on analyses through Biomarker Analysis Tool from MetaboAnalyst

Considering the prevalence of *M. tb* amongst Indian population, SET-A may be a better early TB disease indicator as it can discriminate *M. tb* exposed close household asymptomatic healthy contacts (may be equivalent to latent TB cases) more significantly than SET-B.

4.1.2 Discussion

The present study is the first untargeted NMR based comparative metabolomic study on TB patients, their asymptomatic house hold contacts and healthy cohort from India. In one of the earlier reports, metabolomics study using GC-MS had shown 20 differential metabolites to differentiate between patients and healthy (Weiner *et al.* 2012), where they also arrived at pathways such as amino acid, lipid metabolism etc, similar to our analysis. This study identified 33 metabolites as against 20 from the previous report. Inclusion of household contacts was unique to this study. The household contacts used in the study include the individuals in the household with whom the patients were in close proximity for 3 months and at least seven consecutive days prior to the diagnosis of tuberculosis in patients. Though no tests were performed, it was assumed that household contacts were exposed to TB infection while they remained clinically asymptomatic, and therefore may represent some latent cases. Comparing them with TB patients and unrelated healthy volunteers was an attempt to find a serum indicator for latent TB as well. The TB patients, as per the historical information provided by the patients, were drug naïve that is they had not taken any TB drug before. However, it was observed that 18 out of 40 TB patient samples had rifampicin or pyrazinamide signatures in their NMR profiles. Besides that, for 10 samplesacquired spectra were not decipherable, this left us with 12 samples in this category. Since the study required drug naïve TB population, these 28 samples along with their respective household contact samples, were excluded from analyses. Consequently, the data from their respective house hold contacts were also not included in the analyses. Yet another approach unique to this study was comparison of data from high TB burden (India) and low TB burden (Portugal) countries to arrive at more universal metabolite markers.

All 33 metabolites identified from sera samples were subjected to pathway analysis using MSEA (**Figure 24**). MSEA showed an enrichment of certain pathways in patients as compared to healthy in (i) lipid metabolism which is well reported to be essential for TB disease progression; (ii) keto-bodies metabolism, yet again related to lipid metabolism (iii) methane and betaine metabolism

which contribute to stress adaptation; (iv) urea cycle that contributes to adaptation of *M. tb* to host-induced acid stress. Interestingly, both TB patients and their household asymptomatic contacts showed similar enrichment pattern in their amino acid metabolism such as that glutamic acid, glutamine, arginine, aspartic acid, serine, proline, alanine, lysine, valine and threonine as compared to healthy. Amino acid balance is maintained by contribution of gluconeogenesis, oxidative catabolism and proteolysis. Previous studies have shown that amino acid proteolysis are common features of TB infection and malnutrition (Macallan 1998).This data is in agreement with earlier reports which says that malnutrition and wasting are hallmark of TB (Eddleston *et al*., 2008; Fauci *et al*., 2009). In the case of contacts, which are placed in an intermediate position with respect to M. tb exposure, also showed similar enrichment in pathways associated with lipid metabolism, urea cycle etc suggesting that this may be a reflection of *M. tb* infection without developing the disease (asymptomatic). Such observation may come useful in narrowing down on metabolic markers for latent TB. The enrichment of spermine and spermidine biosynthesis pathways in household contacts potentially contributes to anti-oxidant state. This change in metabolism probably helps the close contacts of TB patients in containing mycobacterial infection.

Statistical analysis such as PCA 2D plots shows segregation between healthy and patients suggesting difference among these groups. One of the interesting observations from this study was the clustering of contact and patients in similar quadrant in these analyses (**Figure 25C**), suggesting that asymptomatically healthy contacts may have latent TB. The similarity in the metabolites of patients and their contacts can also be because they shared exposure to similar external cues like food, water etc. Important variables responsible for segregation were identified by PLS-DA analysis. PLS-DA analysis of healthy as compared to patients showed higher accumulation of metabolites in patients such as glutamic acid, choline, mannose, proline, arginine and aspartic acid. Previous studies (Zhou *et al*. 2013) had shown increased level of glutamic acid and aspartic acid (Weiner *et al.* 2012) in TB patients. Cellular processes such as proteolysis, gluconeogenesis and oxidative catabolism are known to contribute to amino acid balance. Earlier studies also suggest during both TB and malnutrition, amino acid proteolysis occurs, suggesting that TB is a wasting disease. Metabolites such as serine, glutamine, ornithine were found in low concentration in patients. Weiner and group (Weiner *et al.* 2012) showed that under stress

conditions, glutamine can become limiting leading to impaired immune function. PLS-DA features for healthy as compared to patients and contacts showed important variables including mannose, acetic acid, glutamic acid showing an increase in their levels as a function of relative exposure to *M. tb* with healthy showing the least and patients highest (health < contacts < patients); while hypoxanthine was high in patients as compared to healthy and contacts. Previous study had shown hypoxia prevails and may contribute to the containment of *M. tb* leads to tuberculous granulomas and hypoxanthine is a hypoxia biomarker (Weiner *et al.* 2012). Further, serine and threonine were showing decrease in their levels in the patients with the trend being high in healthy (**Figure 27**).

As elaborated in results, I arrived at two sets of metabolites that were further evaluated in-silico through Biomarker Analysis Tool. SET-A, comprising Ornithine, Creatinine and Proline was an outcome of metabolite signatures from sera of Indian cohort only, while SET-B metabolites (Glutamic acid, Mannose and Hypoxanthine) were deduced from the metabolite signatures of Portuguese sera samples, but were also found in PLS-DA of Indian data with a VIP score of >1. Therefore, beside SET-A, SET-B was also tested on Indian samples for its potency as disease indicator on Indian sample. The properties of SET-A and SET-B metabolites as potential active disease and latency indicators were arrived at using ROC curve based model evaluation for Healthy vs Patients (**Figures 28 and 29**) and Contacts vs Patients (**Figures 30 and 31**) respectively (**Table 17**). While SET-B (Hypoxanthine, Glutamic acid and Mannose), has the potential to differentiate TB patients from healthy across two geographically different populations better than SET-B, its discriminatory ability of Patients from *M. tb* exposed asymptomatic subjects (here, household contacts) was less than SET-A. Hence, in my judgement, SET-A is a better metabolite signature that can distinguish early stages of active TB from asymptomatic TB (latent or no TB) in Indian population.

This study provides preliminary information to further study and answer the question if the impact on metabolism during TB disease is reflected in the peripheral sera of TB patients, because collecting peripheral blood samples is one of the easiest and preferred methods for diagnostics purpose. The study has suggested some differential metabolites, but to be sure if these are TB specific disease indicators, a control group of other infections or inflammatory disorders

requires to be included in subsequent studies. A follow-up study where patients can be studied post treatment will help understanding if these metabolites can also indicate treatment outcome. Similar studies, as a continuation of observations from my PhD thesis is being designed in our laboratory.

Annexure II

(Tables)

Annexure II (Chapter 4)

Tables for comparative metabolic profiling of sera from TB patients, their household clinically healthy contacts and unrelated clinically healthy volunteers

4.2.1 Tables for comparative metabolic profiling of sera from TB patients, their household clinically healthy contacts and unrelated clinically healthy volunteers.

Table 13: Metabolites identified and their chemical shift

S.no	Metabolite name	chemical shift (ppm)
1	Choline	3.2(s), 4.05(t), 3.51(t)
2	D-Glutamic acid	3.76(m) 2.05(m) 2.35(dd)
3	Acetoacetic acid	2.28 (s)
4	L-Glutamine	2.15(m), 2.44(m), 3.77(m)
5	L-Arginine	1.72(m), 1.93(m), 3.77(m)
6	L-Serine	3.94 (m) 3.83 (dd)
7	L-Aspartic acid	3.88 (dd) 2.80 (dd)
8	Citric acid	2.53(d), 2.666(d)
9	Acetic acid	1.90(s)
10	L-Alpha-aminobutyric acid	0.9(t)
11	Inosine	8.34(s), 6.09(d), 8.22(s), 4.76(t),4.47(m)
12	Pyruvic acid	2.38(s)
13	D-Mannose	5.2(d)
14	L-Proline	4.12(dd) 2.01(m) 3.34(m) 3.42(m)
15	Propylene glycol	1.1(d)
16	L-Methionine	2.63 (t) 2.12 (s)
17	Ornithine	3.8(t),1.84(m),1.94(m),3.05(t)
18	D-Glucose	5.23(d)
19	Hypoxanthine	8.20(s) 8.21(s)
20	L-Lactic acid	4.096(q), 1.313(d)
21	Creatine	3.03(s), 3.92(s)
22	L-Lysine	3.74(t), 3.02(t), 1.89(m), 1.71(m), 1.46(m)
23	L-Valine	0.996(d), 1.047(d), 2.281(m),3.617(d)
24	L-Phenylalanine	7.40(m), 7.33(m), 7.35(m)
25	L-Asparagine	2.84 (dd) 2.94 (dd) 4.00 (dd)
26	Myoinositol	3.5
27	L-Alanine	3.77(q), 1.48(d)
28	L-Histidine	7.75(t), 7.08(d), 6.05(d)
29	L-Threonine	1.31 (d) 4.24 (m)
30	S-3-Hydroxyisobutyric acid	2.30(dd) 4.16(m) 1.19(d)
31	2-Hydroxybutyric acid	0.9(t)
32	Betaine	3.27(s), 3.93(s)
33	Creatinine	3.04 (s) 4.06 (s)

Table 14: Analysis of variance (ANOVA)

Metabolite	f.value	p.value	-LOG10(p)	FDR
Choline	194.4	5.5E-19	18.3	1.82E-17
D-Glutamic acid	70.5	1.2E-12	11.9	2.03E-11
Acetoacetic acid	26.2	1.6E-07	6.8	1.71E-06
L-Glutamine	12.5	8.9E-05	4.1	7.3E-04
L-Alpha-aminobutyric acid	8.3	1.2E-03	2.9	6.6E-03
Ornithine	8.2	1.3E-03	2.9	6.6E-03
L-Arginine	8.1	1.4E-03	2.9	6.6E-03
Acetic acid	7.2	2.5E-03	2.6	1.0E-02
L-Proline	6.8	3.3E-03	2.5	0.011
D-Mannose	6.8	3.3E-03	2.5	0.011
L-Serine	6.2	5.3E-03	2.3	0.02
D-Glucose	5.7	7.6E-03	2.1	0.02
Citric acid	5.5	8.8E-03	2.1	0.02
Creatinine	5.4	9.7E-03	2.0	0.02
L-Aspartic acid	5.2	1.1E-02	1.9	0.02
Propylene glycol	4.8	1.5E-02	1.8	0.03
Pyruvic acid	3.6	3.8E-02	1.4	0.07
L-Threonine	3.5	4.3E-02	1.4	0.08
2-Hydroxybutyric acid	3.2	5.6E-02	1.3	0.10
S-3-Hydroxyisobutyric acid	3.1	5.9E-02	1.2	0.10
L-Lysine	3.0	6.5E-02	1.2	0.10
L-Methionine	2.9	6.7E-02	1.2	0.10
L-Valine	2.8	7.4E-02	1.1	0.11
Inosine	2.7	8.3E-02	1.1	0.11
Hypoxanthine	2.7	8.4E-02	1.1	0.11
L-Lactic acid	2.5	9.5E-02	1.0	0.12
L-Asparagine	2.5	9.7E-02	1.0	0.12
Creatine	1.2	3.0E-01	0.5	0.36
Betaine	0.9	4.0E-01	0.4	0.44
L-Histidine	0.9	4.2E-01	0.4	0.44
L-Alanine	0.9	4.2E-01	0.4	0.44
Myoinositol	0.9	4.2E-01	0.4	0.44
L-Phenylalanine	0.2	8.2E-01	0.1	0.82

F value indicates the ratio of the variance of the group means to that of the pooled within group variance. The larger the F value the greater the relative variance among the group means. The p value tells you the probability of obtaining an F value as extreme or more extreme as the one observed under the assumption that the null hypothesis is true.FDR: value obtained after performing false discovery test.

Tables 16A, B and C

Table 16A: Metabolites differentiating pateints from healthy subjects

Metabolite	FC	VIP	p.value	FDR
Choline	3.6	2.9	2.9E-15	9.5E-14
D-Glutamic acid	2.7	2.2	9.6E-10	1.6E-08
L-Aspartic acid	1.7	1.1	0.01	0.02
Acetic acid	1.5	1.2	3.9E-03	0.02
D-Mannose	1.4	1.1	0.01	0.02
L-Arginine	1.4	1.2	1.8E-03	0.01
L-Alpha-aminobutyric acid	1.3	1.3	5.0E-03	0.02
2-Hydroxybutyric acid	1.3	0.3	0.4	0.5
L-Proline	1.2	0.8	0.02	0.1
Hypoxanthine	1.2	0.6	0.2	0.3
L-Asparagine	1.1	0.1	0.8	0.9
L-Lysine	1.1	0.6	0.1	0.3
L-Phenylalanine	1.1	0.2	0.6	0.7
L-Alanine	1.1	0.4	0.3	0.4
L-Valine	1.1	0.3	0.3	0.4
Myoinositol	1.0	0.0	1.0	1.0
S-3-Hydroxyisobutyric acid	1.0	0.1	0.8	0.9
Betaine	1.0	0.2	0.7	0.8
L-Histidine	0.9	0.1	0.8	0.8
D-Glucose	0.9	0.4	0.2	0.3
L-Lactic acid	0.9	0.5	0.2	0.3
L-Threonine	0.9	0.5	0.3	0.4
Creatine	0.9	0.5	0.2	0.4
L-Methionine	0.9	0.8	0.05	0.1
Creatinine	0.8	0.6	0.2	0.3
Ornithine	0.8	0.8	0.1	0.1
L-Serine	0.8	0.9	0.01	0.02
Citric acid	0.8	1.1	0.004	0.02
Propylene glycol	0.7	1.0	0.02	0.1
Pyruvic acid	0.7	0.8	0.01	0.03
L-Glutamine	0.7	1.4	1.6E-04	0.001
Inosine	0.5	0.7	0.02	0.05
Acetoacetic acid	0.2	1.5	1.8E-05	2.0E-04

FC: fold changeVIP score: variable of importance score obtained from PLS-DA plot, p value: p values obtained after performing t-test, FDR: value obtained after performing false discovery test.

Table 16B: Metabolites differentiating contacts from healthy subjects

Metabolite	FC	VIP	p.value	FDR
Choline	3.30	2.46	2.3E-14	7.7E-13
D-Glutamic acid	1.94	2.31	8.4E-14	1.4E-12
L-Alpha-aminobutyric acid	1.37	1.47	2.4E-04	1.2E-03
L-Aspartic acid	1.35	0.76	0.01	0.04
L-Arginine	1.28	1.04	0.003	0.01
Creatinine	1.25	0.82	0.04	0.09
Ornithine	1.18	1.04	0.01	0.04
L-Histidine	1.12	0.31	0.29	0.33
Acetic acid	1.11	0.63	0.18	0.24
Myoinositol	1.10	0.54	0.20	0.25
D-Mannose	1.03	0.11	0.74	0.77
L-Phenylalanine	1.03	0.18	0.68	0.73
L-Alanine	0.98	0.08	0.86	0.86
L-Serine	0.95	0.40	0.25	0.30
L-Lysine	0.91	0.45	0.26	0.31
L-Valine	0.90	0.58	0.13	0.21
Hypoxanthine	0.88	0.30	0.42	0.47
L-Proline	0.86	0.56	0.10	0.16
Creatine	0.86	0.44	0.16	0.23
Betaine	0.85	0.55	0.18	0.24
Citric acid	0.84	0.77	0.04	0.08
L-Lactic acid	0.83	0.96	0.02	0.05
Pyruvic acid	0.81	0.59	0.08	0.14
L-Methionine	0.81	0.59	0.05	0.10
2-Hydroxybutyric acid	0.78	0.77	0.03	0.07
L-Threonine	0.77	1.04	0.03	0.06
D-Glucose	0.76	1.32	6.7E-06	7.4E-05
S-3-Hydroxyisobutyric acid	0.75	0.71	0.02	0.06
L-Glutamine	0.75	1.40	1.4E-04	9.2E-04
Propylene glycol	0.72	1.03	0.01	0.02
Inosine	0.71	0.56	0.16	0.23
L-Asparagine	0.65	1.27	2.4E-04	1.2E-03
Acetoacetic acid	0.23	1.59	9.6E-06	7.9E-05

FC: fold change, VIP score: variable of importance score obtained from PLS-DA plot, p value: p values obtained after performing t-test, FDR: value obtained after performing false discovery test.

Table 16C: Metabolites differentiating contacts from patients

Metabolites	FC	VIP	p.value	FDR
L-Asparagine	1.74	0.71	0.14	0.31
2-Hydroxybutyric acid	1.63	1.08	0.05	0.14
D-Mannose	1.39	1.31	0.02	0.10
D-Glutamic acid	1.38	1.81	0.002	0.03
L-Proline	1.37	1.38	0.003	0.03
Acetic acid	1.36	1.40	0.02	0.11
S-3-Hydroxyisobutyric acid	1.31	1.10	0.07	0.18
Hypoxanthine	1.31	1.22	0.02	0.10
L-Aspartic acid	1.24	0.71	0.28	0.44
L-Lysine	1.23	1.35	0.04	0.13
D-Glucose	1.21	0.76	0.12	0.28
L-Valine	1.20	1.17	0.04	0.13
L-Threonine	1.16	1.28	0.08	0.21
Betaine	1.13	0.60	0.37	0.51
L-Alanine	1.11	0.89	0.18	0.33
L-Lactic acid	1.10	0.52	0.40	0.51
Choline	1.10	0.99	0.18	0.33
L-Arginine	1.09	0.59	0.36	0.51
Acetoacetic acid	1.07	0.21	0.74	0.78
L-Phenylalanine	1.06	0.16	0.77	0.79
L-Methionine	1.05	0.37	0.44	0.53
Creatine	1.05	0.29	0.58	0.64
Propylene glycol	1.02	0.01	0.99	0.99
L-Alpha-aminobutyric acid	0.94	0.49	0.48	0.56
Myoinositol	0.92	0.61	0.30	0.45
Citric acid	0.90	0.73	0.28	0.44
L-Glutamine	0.88	0.86	0.19	0.33
Pyruvic acid	0.87	0.51	0.39	0.51
L-Histidine	0.84	0.64	0.16	0.33
L-Serine	0.82	1.11	0.04	0.13
Ornithine	0.71	1.98	0.002	0.03
Inosine	0.67	0.25	0.58	0.64
Creatinine	0.67	1.76	0.00	0.03

FC: fold changeVIP score: variable of importance score obtained from PLS-DA plot, p value: p values obtained after performing t-test, FDR: value obtained after performing false discovery test.

Chapter 5

Summary and future prospects

5 Summary and future prospects

The study is a part of the on-going quest in the laboratory where multiple approaches are being employed to address two fundamental problems in tuberculosis, i.e, requirement of biomarkers for early and accurate TB diagnosis and requirement of new drug target to design anti-mycobacterial molecule to overcome the long chemotherapy and drug resistance.

I used a comparative metabolomics / metabonomics approach towards these two goals and two main objectives were designed for my thesis work, that is, to understand immediate adaptation of *M. tb* to microbicidal stress to find important pathways as possible drug targets and identify differential metabolites from sera of host (TB patient) as compared to healthy individuals to identify disease indicators.

5.1 Highlights and future prospects

5.1.1 Objective 1: Comparative and quantitative metabolite profiling of pathogenic mycobacteria H37Rv and non-pathogenic soil mycobacteria, *M. smegmatis* MC2155, under microbicidal stresses

The objective attempted to address what are the early adaptive metabolic changes in mycobacteria in response to micobicidal stresses like low pH, oxidative stress, iron deficiency and nutrient deprivation. The study was performed by using well-established *in-vitro* models mimicking microbicidal stress conditions. Two strains of mycobacteria were used, the non-pathogenic *Mycobacterium smegmatis* and the pathogenic *M. tb* strain H37Rv.

5.1.1.1 *Metabolite profiling of M. smegmatis (MC^{2}155) subjected to nutrient starvation, acidic and oxidative stresses*

^{1}H NMR-based metabolic profiling was performed and a total of 40 ^{1}H NMR spectra were obtained corresponding to control and three stress i.e acidic, oxidative and nutrient starvation conditions with about 10 replicates for each condition. 31, 20 and 47 differential metabolites for acidic, oxidative and nutrient starvation stresses respectively were detected as compare to normal growth with a fold change cut-off of 1.2 and $p < 0.05$. Pathway analyses showed perturbation in common pathway such as glutamine and glutamate metabolism, purine metabolism, pyrimidine

metabolism, nicotinate and nicotinamide metabolism, alanine, aspartate and glutamate metabolism and beta-Alanine metabolism. Besides tabulating the differential metabolites and tracing them to pathways differentially perturbed during various stresses, some distinctive observations not reported earlier in *M. smegmatis* could be made that are enlisted below.

1. Identification of a previously unreported possible pathway of intracellular biosynthesis of betaine, methylamine and dimethylamine in *M. smegmatis*. With the genes encoding the enzymes of these pathways transcribed during stresses, the physiological significance of this new pathway that can convert choline to betaine or carnitine to TMA to methylamines in adaptation to environmental microbicidal stresses in *M. smegmatis* can be studied further.

2. The study reports for the first time the differential regulation of α-Glucan biosynthesis at transcript levels of selected ORFs in *M. smegmatis* during different stress conditions. With implication of α-Glucan biosynthesis in capsule formation, the regulation of this pathway can be studied to understand critical adaptive nodes for different stress conditions in *M. smegmatis.*

3. The orthologues of these pathways can be traced in pathogenic mycobacteria and their significance may be studies to adapt to microbicidal stresses mimicking intracellular survival.

5.1.1.2 *Metabolite profiling of M. tb strain H37Rv subjected to nutrient starvation, iron deprivation, acidic and oxidative stresses*

LC-MRM-based targeted metabolomic approach was used to assess the influence of microbicidal effect such acidic, oxidative, iron and nutrient starvation stress on H37Rv.

This approach was used to overcome the limitations faced during untargeted metabolomics study using NMR for *M. smegmatis.*

It was observed that H37Rv successfully maintained adenylate energy charge (AEC) within narrow physiological values despite large fluctuations in the adenosine nucleotide, suggestive of its ability to maintain homeostasis during microbicidal stresses.

Common pathways that were perturbed during all stress conditions reflected adjustment predominantly in purine and amino acid metabolism. With respect to amino acid the pathways

that were observed to be primarly altered were those belonging to glycine, serine, threonine, cysteine and methionine metabolism.Analysis of data suggested glutamate dehydrogenase (GDH) as a metabolic node to be active in acid, oxidative and iron stress conditions in the present study. The resultant α-ketoglurate formed by GDH activity further replenished the TCA cycle during the stress conditions. Complimentary experiment was done to estimation released ammonia and extracellular pH, which suggested increased release of ammonia and pH difference under acidic condition compare to other stresses. This also supported the possibility of GDH as an important target. The study also proposes that pathways/ enzymes like GlgE, GDH, ADI etc are important nodes that can be explored as targets. The new observations and future leads that have emerged from this part of the study are:

1. The study suggests quenching of protons as an important mechanism for early adaptation to acidic and oxidative stress where GABA shunt pathway and sugar alcohol synthesis and accumulation may be playing important role. These pathways have not been studied much in pathogenic mycobacteria, which if thoroughly investigated can lead to new anti-mycobacterial targets.

2. Glutamate dehydrogenase (GDH), glutamate synthetase (GS) and glutamine oxoglutarate aminotransferase (GOGAT) have emerged as important nodes for early adaptation to microbicidal stresses, especially acidic and oxidative stresses and hence can be good anti-mycobacterial targets. This would require pursuing these observations using KO strains.

3. Urease, unlike the general belief, is not the major player in early adaptation to acidic stress in *M. tb.*

5.1.2 Objective 2: Comparative metabolic profiling of sera from TB patients, their clinically healthy household contacts and unrelated clinically healthy volunteers

Under this objective, sera from well-defined cohort was collected, compared for metabolite profiles using NMR-based metabolomics, subjecting differential metabolites to “Biomarker Analysis tool” to explore their potential as TB disease indicators. The study encompassed population from two different geographical regions (i) India, region with high TB burden (ii) Portugal, region with low TB burden. The collection, processing and analysis of Indian sample

were done by me and further resultant important variables were identified in my study was compared with another study done independently by our collaborators in Portugal. The study (carried under the DBT New-Indigo scheme) was designed such that it permitted comparative cross-sectional analyses of metabolite profiles of Indian and Portugal populations. The following set of metabolites could be shortlisted as potential disease indicators which were validated through Biomarker Analysis Tool.

SET A: Ornithine, Creatinine and Proline: Indian cohort specific

SET B: Glutamic acid, Mannose and Hypoxanthine: common for Indian and Portugal population.

These metabolites can now be further validated using a larger cohort.

5.2 Limitations of this study

While it was intended to compare the differences in the early adaptive mechanisms of non-pathogenic *vs* pathogenic mycobacteria, it could not be performed because of using two different technology platforms, that is, NMR and LC-MS. While we had reasons to use these two approaches, a systematic study on metabolite profiling to measure temporal changes in both *M. smegmatis* and *M. tb* using LC-MS is being pursued as a continuation of these studies.

The disease indicators shortlisted could be validated only *in-silico* and testing in a small cohort to support the *in-silico* validation would have helped. This could not be performed because of difficulties in getting samples and required permissions.

Finding rifampicin or pyrazinamide signatures in 18 out of 40 TB patient samples was disappointing. Besides that, for 10 samples acquired spectra were not decipherable. Since the study required drug naïve population, these 28 samples along with their respective household contact samples, were excluded from analyses, which reduced the effective sample numbers tremendously.

References

Akaki, T. *et al*. (2000) ‘Comparative roles of free fatty acids with reactive nitrogen intermediates and reactive oxygen intermediates in expression of the anti-microbial activity of macrophages against Mycobacterium tuberculosis’, ClinExpImmunol. doi: cei1298 [pii].

Andersen, M. R., M. L. Nielsen, *et al*. (2008). "Metabolic model integration of the bibliome, genome, metabolome and reactome of Aspergillus niger." MolSystBiol 4: 178.

Armstrong, J. A. and Hart, P. D. (1971) ‘Response of cultured macrophages to Mycobacterium tuberculosis, with observations on fusion of lysosomes with phagosomes.’, The Journal of experimental medicine. doi: 10.1084/jem.134.3.713.

Atkinson, D. E. and G. M. Walton (1967). "Adenosine triphosphate conservation in metabolic regulation. Rat liver citrate cleavage enzyme." J BiolChem 242(13): 3239-3241.

Axelrod, S. *et al*. (2008) ‘Delay of phagosome maturation by a mycobacterial lipid is reversed by nitric oxide’, Cellular Microbiology. doi: 10.1111/j.1462-5822.2008.01147.x.

Baek, S. H., Li, A. H. and Sassetti, C. M. (2011) ‘Metabolic regulation of mycobacterial growth and antibiotic sensitivity’, PLoS Biology. doi: 10.1371/journal.pbio.1001065.

Baker, J. J. and Abramovitch, R. B. (2018) ‘Genetic and metabolic regulation of Mycobacterium tuberculosis acid growth arrest’, Scientific Reports. doi: 10.1038/s41598-018-22343-4.

Baptista, R., D. M. Fazakerley, *et al*. (2018). "Untargeted metabolomics reveals a new mode of action of pretomanid (PA-824)." Sci Rep 8(1): 5084.

Barkan, D., Stallings, C. L. and Glickman, M. S. (2011) ‘An improved counterselectable marker system for mycobacterial recombination using galK and 2-deoxy-galactose’, Gene. doi: 10.1016/j.gene.2010.09.005.

Barksdale, L. and K. S. Kim (1977). "Mycobacterium." Bacteriol Rev 41(1): 217-372.

Barsch, A., Patschkowski, T. and Niehaus, K. (2004) 'Comprehensive metabolite profiling of Sinorhizobiummeliloti using gas chromatography-mass spectrometry', Functional & Integrative Genomics. doi: 10.1007/s10142-004-0117-y.

Beste, D. J. V and McFadden, J. (2010) 'System-level strategies for studying the metabolism of Mycobacterium tuberculosis', Molecular BioSystems. doi: 10.1039/c003757p.

Biasini, M., S. Bienert, *et al*. (2014). "SWISS-MODEL: modelling protein tertiary and quaternary structure using evolutionary information." Nucleic Acids Research 42(W1): W252-W258.

Boehme, C. C. *et al*. (2010) 'Rapid Molecular Detection of Tuberculosis and Rifampin Resistance', New England Journal of Medicine. doi: 10.1056/NEJMoa0907847.

Boender, L. G., M. J. Almering, *et al*. (2011). "Extreme calorie restriction and energy source starvation in Saccharomyces cerevisiae represent distinct physiological states." BiochimBiophysActa 1813(12): 2133-2144.

Borum, P. R. and H. P. Broquist (1977). "Purification of S-adenosylmethionine: epsilon-N-L-lysine methyltransferase. The first enzyme in carnitine biosynthesis." J Biol Chem 252(16): 5651-5655.

Buchmeier, N. *et al*. (2000) 'A parallel intraphagosomal survival strategy shared by Mycobacterium tuberculosis and Salmonella enterica', Molecular Microbiology. doi: 10.1046/j.1365-2958.2000.01797.x.

Bullen, J. J., Griffiths, E. and Edmiston, C. E. (1999) 'Iron and infection: molecular, physiological and clinical aspects, 2nd edition', Shock, 12(5).

Burg, M. B. and J. D. Ferraris (2008). "Intracellular organic osmolytes: function and regulation." J BiolChem 283(12): 7309-7313.

Caulfield, A. J. and Wengenack, N. L. (2016) 'Diagnosis of active tuberculosis disease: From microscopy to molecular techniques', Journal of Clinical Tuberculosis and Other Mycobacterial Diseases. doi: 10.1016/j.jctube.2016.05.005.

Chassagnole, C., F. Letisse, *et al*. (2002). "Carbon flux analysis in a pantothenate overproducing Corynebacterium glutamicum strain." MolBiol Rep 29(1-2): 129-134.

Chauhan, S. *et al*. (2011) 'Comprehensive insights into Mycobacterium tuberculosis DevR (DosR) regulon activation switch', Nucleic Acids Research. doi: 10.1093/nar/gkr375.

Che, N. *et al*. (2013) 'Decreased serum 5-oxoproline in TB patients is associated with pathological damage of the lung', ClinicaChimicaActa. doi: 10.1016/j.cca.2013.04.010.

Chen, J. (2010) 'Identifying Biosynthetic Pathways for Mycobacterial Cell Wall Components Using Transposon Mutagenesis degree of Doctor of Philosophy', Thesis.

Cheruvu, M., Plikaytis, B. B. and Shinnick, T. M. (2007) 'The acid-induced operon Rv3083-Rv3089 is required for growth of Mycobacterium tuberculosis in macrophages', Tuberculosis. doi: 10.1016/j.tube.2006.01.021.

Chiarelli, L. R., G. Mori, *et al*. (2016). "New and Old Hot Drug Targets in Tuberculosis." Curr Med Chem 23(33): 3813-3846.

Chionh, Y. H. *et al*. (2016) 'TRNA-mediated codon-biased translation in mycobacterial hypoxic persistence', Nature Communications. doi: 10.1038/ncomms13302.

Cole, S. T. *et al*. (1998) 'Deciphering the biology of mycobacterium tuberculosis from the complete genome sequence', Nature. doi: 10.1038/31159.

Combet, C., M. Jambon, *et al*. (2002). "Geno3D: automatic comparative molecular modelling of protein." Bioinformatics 18(1): 213-214.

Cunin, R., N. Glansdorff, *et al*. (1986). "Biosynthesis and metabolism of arginine in bacteria." Microbiol Rev 50(3): 314-352.

Darrah, P. A. *et al*. (2000) 'Cooperation between reactive oxygen and nitrogen intermediates in killing of Rhodococcusequi by activated macrophages', Infection and Immunity. doi: 10.1128/IAI.68.6.3587-3593.2000.

Das, M. K. *et al.* (2015) ‘Deregulated tyrosine-phenylalanine metabolism in pulmonary tuberculosis patientS’, Journal of Proteome Research. doi: 10.1021/acs.jproteome.5b00016.

De Voss, J. J. *et al.* (1999) ‘Iron acquisition and metabolism by mycobacteria’, Journal of Bacteriology. doi: 10.1128/JB.186.2.374-382.2004.

Dhariwal, K. R., A. Chander, *et al.* (1978). "Turnover of lipids in Mycobacterium smegmatis CDC 46 and Mycobacterium phlei ATCC 354." Archives of Microbiology 116(1): 69-75.

Dobson, P. D., K. Smallbone, *et al.* (2010). "Further developments towards a genome-scale metabolic model of yeast." BMC SystBiol 4: 145.

Dookie, N. *et al.* (2018) ‘Evolution of drug resistance in Mycobacterium tuberculosis: A review on the molecular determinants of resistance and implications for personalized care’, Journal of Antimicrobial Chemotherapy. doi: 10.1093/jac/dkx506.

Drakesmith, H. and Prentice, A. M. (2012) ‘Hepcidin and the iron-infection axis’, Science. doi: 10.1126/science.1224577.

Drapal, M., L. Perez-Fons, *et al.* (2014). "The application of metabolite profiling to Mycobacterium spp.: determination of metabolite changes associated with growth." J Microbiol Methods 106: 23-32.

Du Preez, I. and Loots, D. T. (2013) ‘New sputum metabolite markers implicating adaptations of the host to Mycobacterium tuberculosis, and vice versa’, Tuberculosis. doi: 10.1016/j.tube.2013.02.008.

Dwivedi, P. *et al.* (2010) ‘Metabolic profiling of Escherichia coli by ion mobility-mass spectrometry with MALDI ion source’, Journal of Mass Spectrometry. doi: 10.1002/jms.1850.

Eddleston M, Davidson R, Brent A, Wilkinson R (2008) Oxford Handbook of Tropical Medicine, 3rd Edition. Oxford, UK: Oxford University Press. 856 p.

Edwards, J. M., T. H. Roberts, *et al.* (2012). "Quantifying ATP turnover in anoxic coleoptiles of rice (Oryza sativa) demonstrates preferential allocation of energy to protein synthesis." J Exp Bot 63(12): 4389-4402.

Ehrt, S. and Schnappinger, D. (2009) ‘Mycobacterial survival strategies in the phagosome: Defence against host stresses’, Cellular Microbiology. doi: 10.1111/j.1462-5822.2009.01335.x.

Fauci AS, Braunwald E, Kasper DL, Hauser SL (2009) Harrison’s Principles of Internal Medicine, 17th Edition. New York: McGraw-Hill Professional. 2958 p.

Feehily, C. and K. A. Karatzas (2013). "Role of glutamate metabolism in bacterial responses towards acid and other stresses." J ApplMicrobiol 114(1): 11-24.

Fiehn, O. (2002) ‘Metabolomics - The link between genotypes and phenotypes’, Plant Molecular Biology. doi: 10.1023/A:1013713905833.

Fisher, M. a, Plikaytis, B. B. and Shinnick, T. M. (2002) ‘Microarray analysis of the Mycobacterium tuberculosis transcriptional response to the acidic conditions found in phagosomes’, J Bacteriol. doi: 10.1128/JB.184.14.4025.

Foster, J. W. (2004). "Escherichia coli acid resistance: tales of an amateur acidophile." Nat Rev Microbiol 2(11): 898-907.

Galagan, J. E. *et al*. (2013) ‘The Mycobacterium tuberculosis regulatory network and hypoxia’, Nature. doi: 10.1038/nature12337.

Gallant, J. L., A. J. Viljoen, *et al*. (2016). "Glutamate Dehydrogenase Is Required by Mycobacterium bovis BCG for Resistance to Cellular Stress." PLoS One 11(1): e0147706.

Ganji, R., S. Dhali, et al. (2016a). "Understanding HIV-Mycobacteria synergism through comparative proteomics of intra-phagosomal mycobacteria during mono- and HIV co-infection." Sci Rep 6: 22060.

Ganji, R., S. Dhali, et al. (2016b). "Proteomics approach to understand reduced clearance of mycobacteria and high viral titers during HIV-mycobacteria co-infection." Cell Microbiol 18(3): 355-368.

Gehlenborg, N. *et al*. (2010) ‘Visualization of omics data for systems biology’, Nature Methods. doi: 10.1038/nmeth.1436.

Gerosa, L. and U. Sauer (2011). "Regulation and control of metabolic fluxes in microbes." CurrOpinBiotechnol 22(4): 566-575.

Gouzy, A., G. Larrouy-Maumus, *et al*. (2013). "Mycobacterium tuberculosis nitrogen assimilation and host colonization require aspartate." Nat Chem Biol 9(11): 674-676.

Gouzy, A., G. Larrouy-Maumus, *et al*. (2014). "Mycobacterium tuberculosis exploits asparagine to assimilate nitrogen and resist acid stress during infection." PLoS Pathog 10(2): e1003928.

Haas, A. (2007). "The phagosome: Compartment with a license to kill." Traffic 8(4): 311-330.

Harrigan, G. G., R. H. LaPlante, *et al*. (2004). "Application of high-throughput Fourier-transform infrared spectroscopy in toxicology studies: contribution to a study on the development of an animal model for idiosyncratic toxicity." Toxicology Letters 146(3): 197-205.

Hartmans S., de Bont J.A.M., Stackebrandt E. (2006) The Genus Mycobacterium--Nonmedical. In: Dworkin M., Falkow S., Rosenberg E., Schleifer KH., Stackebrandt E. (eds) The Prokaryotes. Springer, New York, NY.

Hawkins, R. D., Hon, G. C. and Ren, B. (2010) 'Next-generation genomics: An integrative approach', Nature Reviews Genetics. doi: 10.1038/nrg2795.

Hawn, T. R., T. A. Day, *et al*. (2014). "Tuberculosis vaccines and prevention of infection." MicrobiolMolBiol Rev 78(4): 650-671.

Hawn, T. R., T. A. Day, *et al*. (2014). "Tuberculosis vaccines and prevention of infection." Microbiol Mol Biol Rev 78(4): 650-671.

Healy, C. *et al*. (2016) 'The MarR family transcription factor Rv1404 coordinates adaptation of Mycobacterium tuberculosis to acid stress via controlled expression of Rv1405c, a virulence-associated methyltransferase', Tuberculosis. doi: 10.1016/j.tube.2015.10.003.

Hiller, J., E. Franco-Lara, *et al*. (2007). "Fast sampling and quenching procedures for microbial metabolic profiling." Biotechnol Lett 29(8): 1161-1167.

Hochachka, P. W. and G. B. McClelland (1997). "Cellular metabolic homeostasis during large-scale change in ATP turnover rates in muscles." J ExpBiol 200(Pt 2): 381-386.

Holms, H. (1996). "Flux analysis and control of the central metabolic pathways in Escherichia coli." FEMS Microbiol Rev 19(2): 85-116.

Holmstrom, K. O., B. Welin, *et al*. (1994). "Production of the Escherichia coli betaine-aldehyde dehydrogenase, an enzyme required for the synthesis of the osmoprotectant glycine betaine, in transgenic plants." Plant J 6(5): 749-758.

Huard, R. C. *et al*. (2006) 'Novel genetic polymorphisms that further delineate the phylogeny of the Mycobacterium tuberculosis complex', Journal of Bacteriology. doi: 10.1128/JB.01783-05.

Hubbard, P. A., X. Liang, *et al*. (2003). "The crystal structure and reaction mechanism of Escherichia coli 2,4-dienoyl-CoA reductase." J BiolChem 278(39): 37553-37560.

Izumori K, Tuzaki K (1988). Production of xylitol from D-xylulose by Mycobacterium smegmatis. J Ferment. Technol. 66:33–6.

Jackett, P. S., Aber, V. R. and Lowrie, D. B. (1978) 'Virulence and resistance to superoxide, low pH and hydrogen peroxide among strains of Mycobacterium tuberculosis.', Journal of general microbiology. doi: 10.1099/00221287-104-1-37.

Jeremy J. Ramsden (1). "Metabolomics and Metabonomics." Springer. Springer London, 01 Jan. 1970. Web. 25 Apr. 2017.

Johnson, H. E., D. Broadhurst, *et al*. (2004). "High-throughput metabolic fingerprinting of legume silage fermentations via Fourier transform infrared spectroscopy and chemometrics." Applied and Environmental Microbiology 70(3): 1583-1592.

Johnson, H. E., D. Broadhurst, *et al*. (2004). "High-throughput metabolic fingerprinting of legume silage fermentations via Fourier transform infrared spectroscopy and chemometrics." Applied and Environmental Microbiology 70(3): 1583-1592.

Julia Bespyatykh, EgorShitikov, Ivan Butenko, Ilya Altukhov, Dmitry Alexeev, Igor Mokrousov, Marine Dogonadze, ViacheslavZhuravlev, Peter Yablonsky, Elena Ilina, Vadim Govorun.

Proteome analysis of the Mycobacterium tuberculosis Beijing B0/W148 cluster. Sci Rep. 2016; 6: 28985.

Kanabus, M. *et al*. (2016) 'The pleiotropic effects of decanoic acid treatment on mitochondrial function in fibroblasts from patients with complex I deficient Leigh syndrome', Journal of Inherited Metabolic Disease. doi: 10.1007/s10545-016-9930-4.

Kapopoulou, A., J. M. Lew, *et al*. (2011). "The MycoBrowser portal: a comprehensive and manually annotated resource for mycobacterial genomes." Tuberculosis (Edinb) 91(1): 8-13.

Karupiah, G. *et al*. (2000) 'NADPH oxidase, Nramp1 and nitric oxide synthase 2 in the host antimicrobial response', Rev Immunogenet.

Katti, M. K. *et al*. (2008) 'The ??fbpA mutant derived from Mycobacterium tuberculosis H37Rv has an enhanced susceptibility to intracellular antimicrobial oxidative mechanisms, undergoes limited phagosome maturation and activates macrophages and dendritic cells', Cellular Microbiology. doi: 10.1111/j.1462-5822.2008.01126.x.

Kaufmann, S. H. E. (2001). "How can immunology contribute to the control of tuberculosis?" Nature Reviews Immunology 1(1): 20-30.

Keating, M. J. *et al*. (2005) 'Early results of a chemoimmunotherapy regimen of fludarabine, cyclophosphamide, and rituximab as initial therapy for chronic lymphocytic leukemia', Journal of Clinical Oncology. doi: 10.1200/JCO.2005.12.051.

Kim, S. G., Bae, H. S. and Lee, S. T. (2001) 'A novel denitrifying bacterial isolate that degrades trimethylamine both aerobically and anaerobically via two different pathways', Archives of Microbiology. doi: 10.1007/s002030100319.

Kochan, I., Kvach, J. T. and Wiles, T. I. (1977) 'Virulence-Associated Acquisition of Iron in Mammalian Serum by Escherichia coli', The Journal of Infectious Diseases.

Koen, N., S. V. van Breda, *et al*. (2018). "Elucidating the antimicrobial mechanisms of colistin sulfate on Mycobacterium tuberculosis using metabolomics." Tuberculosis (Edinb) 111: 14-19.

Koeth, R. A., B. S. Levison, *et al*. (2014). "gamma-Butyrobetaine is a proatherogenic intermediate in gut microbial metabolism of L-carnitine to TMAO." Cell Metab 20(5): 799-812.

Kurthkoti, K., H. Amin, *et al*. (2017). "The Capacity of Mycobacterium tuberculosis To Survive Iron Starvation Might Enable It To Persist in Iron-Deprived Microenvironments of Human Granulomas." MBio 8(4).

Landfald, B. and A. R. Strom (1986). "Choline-glycine betaine pathway confers a high level of osmotic tolerance in Escherichia coli." J Bacteriol 165(3): 849-855.

Laughon, B. E. and C. A. Nacy (2017). "Tuberculosis - drugs in the 2016 development pipeline." Nat Rev Dis Primers 3: 17015.

Lee, J. J., J. Lim, *et al*. (2018). "Glutamate mediated metabolic neutralization mitigates propionate toxicity in intracellular Mycobacterium tuberculosis." Sci Rep 8(1): 8506.

Leemans, J. C., N. P. Juffermans, *et al*. (2001). "Depletion of alveolar macrophages exerts protective effects in pulmonary tuberculosis in mice." Journal of Immunology 166(7): 4604-4611.

Lei, G. S., W. J. Syu, *et al*. (2011). "Repression of btuB gene transcription in Escherichia coli by the GadX protein." BMC Microbiol 11: 33.

Lew, J. M., A. Kapopoulou, *et al*. (2011). "TubercuList--10 years after." Tuberculosis (Edinb) 91(1): 1-7.

Li, F. *et al*. (2011) 'Metabolomic Analysis Reveals Novel Isoniazid Metabolites and Hydrazones in Human Urine', Drug Metabolism and Pharmacokinetics. doi: 10.2133/dmpk.DMPK-11-RG-055.

Lim, E. L., K. G. Hollingsworth, *et al*. (2010). "Measuring the acute effect of insulin infusion on ATP turnover rate in human skeletal muscle using phosphorus-31 magnetic resonance saturation transfer spectroscopy." NMR Biomed 23(8): 952-957.

Lin ECC. Escherichia Coli and Salmonella: Cellular and Molecular Biology. 2. ASM Press; Washington: 1996. Dissimilatory pathways for sugars, polyols, and carboxylates; pp. 307–342

Lin, W., V. Mathys, *et al*. (2012). "Urease activity represents an alternative pathway for Mycobacterium tuberculosis nitrogen metabolism." Infect Immun 80(8): 2771-2779.

Loebel, R. O., Shorr, E. & Richardson, H. B. (1933). The Influence of Adverse Conditionsupon the Respiratory Metabolism and Growth of Human Tubercle Bacilli. J Bacteriol 26,167-200.

Loots, D. T. *et al*. (2016) 'A metabolomics investigation of the function of the ESX-1 gene cluster in mycobacteria', Microbial Pathogenesis. doi: 10.1016/j.micpath.2016.10.008.

Luca, S. and T. Mihaescu (2013). "History of BCG Vaccine." Maedica (Buchar) 8(1): 53-58.

Luca, S. and T. Mihaescu (2013). "History of BCG Vaccine." Maedica (Buchar) 8(1): 53-58.

Luies L, du Preez I, Loots DT (2017) . The role of metabolomics in tuberculosis treatment research. Biomark Med.;11(11):1017-1029.

Luies, L., Du Preez, I. and Loots, D. T. (2017) 'The role of metabolomics in tuberculosis treatment research', Biomarkers in Medicine. doi: 10.2217/bmm-2017-0141.

Macallan, D. C. *et al*. (1998) 'Whole body protein metabolism in human pulmonary tuberculosis and undernutrition: evidence for anabolic block in tuberculosis.' Clinical science (London, England : 1979).

MacMicking, J. D., G. A. Taylor, *et al*. (2003). "Immune control of tuberculosis by IFN-gamma-inducible LRG-47." Science 302(5645): 654-659.

Magnus, J. B., D. Hollwedel, *et al*. (2006). "Monitoring and modeling of the reaction dynamics in the valine/leucine synthesis pathway in Corynebacterium glutamicum." Biotechnology Progress 22(4): 1071-1083.

Mailloux, R. J., R. Beriault, *et al*. (2007). "The tricarboxylic acid cycle, an ancient metabolic network with a novel twist." PLoS One 2(8): e690.

Maksymiuk, C., A. Balakrishnan, *et al*. (2015). "E1 of alpha-ketoglutarate dehydrogenase defends Mycobacterium tuberculosis against glutamate anaplerosis and nitroxidative stress." Proc Natl AcadSci U S A 112(43): E5834-5843.

Målen, H. *et al*. (2011) ‘Comparison of membrane proteins of Mycobacterium tuberculosis H37Rv and H37Ra strains’, BMC Microbiology. doi: 10.1186/1471-2180-11-18.

Man, D. K., T. Kanno, *et al*. (2018). "Rifampin- or Capreomycin-Induced Remodeling of the Mycobacterium smegmatis Mycolic Acid Layer Is Mitigated in Synergistic Combinations with Cationic Antimicrobial Peptides." mSphere 3(4).

Meissner-Roloff, R. J. *et al*. (2012) ‘A metabolomics investigation of a hyper- and hypo-virulent phenotype of Beijing lineage M. tuberculosis’, Metabolomics. doi: 10.1007/s11306-012-0424-6.

Mendz, G. L. and S. L. Hazell (1996). "The urea cycle of Helicobacter pylori." Microbiology 142 (Pt 10): 2959-2967.

Mobley, H. L., M. D. Island, *et al*. (1995). "Molecular biology of microbial ureases." Microbiol Rev 59(3): 451-480.

Mohr, S. E. *et al*. (2014) ‘RNAi screening comes of age : improved techniques and complementary approaches’, Nature Publishing Group. doi: 10.1038/nrm3860.

Molle, V. *et al*. (2006) ‘pH-dependent pore-forming activity of OmpATb from Mycobacterium tuberculosis and characterization of the channel by peptidic dissection’, Molecular Microbiology. doi: 10.1111/j.1365-2958.2006.05277.x.

Muñoz-Elías, E. J. and McKinney, J. D. (2006) ‘Carbon metabolism of intracellular bacteria’, Cellular Microbiology. doi: 10.1111/j.1462-5822.2005.00648.x.

Murray, J. F., Schraufnagel, D. E. and Hopewell, P. C. (2015) ‘Treatment of tuberculosis: A historical perspective’, Annals of the American Thoracic Society. doi: 10.1513/AnnalsATS.201509-632PS.

Nandakumar, M., Nathan, C. and Rhee, K. Y. (2014) ‘Isocitratelyase mediates broad antibiotic tolerance in Mycobacterium tuberculosis’, Nature Communications. doi: 10.1038/ncomms5306.

Nathan, C. and Shiloh, M. U. (2000) ‘Reactive oxygen and nitrogen intermediates in the relationship between mammalian hosts and microbial pathogens’, Proceedings of the National Academy of Sciences, 97(16), pp. 8841–8848. doi: 10.1073/pnas.97.16.8841.

Nayak, P. A., U. A. Nayak, *et al*. (2014). "The effect of xylitol on dental caries and oral flora." ClinCosmetInvestig Dent 6: 89-94.

Noble, K., J. Zhang, *et al*. (2006). "Lipid rafts, the sarcoplasmic reticulum and uterine calcium signalling: an integrated approach." Journal of Physiology-London 570(1): 29-35.

Oliveira, A. P., J. Nielsen, *et al*. (2005). "Modeling Lactococcuslactis using a genome-scale flux model." BMC Microbiol 5: 39.

Otchere, I. D. *et al*. (2017) 'Comparative genomics of Mycobacterium africanum Lineage 5 and Lineage 6 from Ghana suggests different ecological niches', bioRxiv. doi: 10.1101/202234.

Ozalp, V. C., T. R. Pedersen, *et al*. (2010). "Time-resolved measurements of intracellular ATP in the yeast Saccharomyces cerevisiae using a new type of nanobiosensor." J BiolChem 285(48): 37579-37588.

Pan, Y. T., V. Koroth Edavana, *et al*. (2004). "Trehalose synthase of Mycobacterium smegmatis: purification, cloning, expression, and properties of the enzyme." Eur J Biochem 271(21): 4259-4269.

Park, S., L. T. Smith, *et al*. (1995). "Role of Glycine Betaine and Related Osmolytes in Osmotic Stress Adaptation in Yersinia enterocolitica ATCC 9610." Appl Environ Microbiol 61(12): 4378-4381.

Paziak-Domanska B, Klink M, Jurkiewicz M, Rudnicka W. Production of reactive nitrogen and oxygen intermediates in human granulocytes and monocytes during internalization of live BCG bacilli. Med DoswMikrobiol. 2000;52:353–60.

Perrone, F. *et al*. (2017) 'A Novel TetR-like transcriptional regulator is induced in acid-nitrosative stress and controls expression of an efflux pump in mycobacteria', Frontiers in Microbiology. doi: 10.3389/fmicb.2017.02039.

Phelan, J. *et al*. (2016) 'Mycobacterium tuberculosis whole genome sequencing and protein structure modelling provides insights into anti-tuberculosis drug resistance', BMC Medicine. doi: 10.1186/s12916-016-0575-9.

Phillips, M. *et al*. (2010) ‘Breath biomarkers of active pulmonary tuberculosis’, Tuberculosis, 90(2), pp. 145–151. doi: 10.1016/j.tube.2010.01.003.

Piddington, D. L., A. Kashkouli, *et al*. (2000). "Growth of Mycobacterium tuberculosis in a defined medium is very restricted by acid pH and Mg(2+) levels." Infect Immun 68(8): 4518-4522.

Pluskal, T., M. Ueno, *et al*. (2014). "Genetic and metabolomic dissection of the ergothioneine and selenoneine biosynthetic pathway in the fission yeast, S. pombe, and construction of an overproduction system." PLoS One 9(5): e97774.

Pratima, K., N. Goud, *et al*. (2018). "An efficient and facile Green synthesis of Bisindole Methanes as potential Mtb FtsZ inhibitors." Chem Biol Drug Des.

Price, C. T., A. Bukka, *et al*. (2008). "Glycine betaine uptake by the ProXVWZ ABC transporter contributes to the ability of Mycobacterium tuberculosis to initiate growth in human macrophages." J Bacteriol 190(11): 3955-3961.

Rabinowitz, J. D. and Kimball, E. (2007) ‘Acidic acetonitrile for cellular metabolome extraction from Escherichia coli’, Analytical Chemistry. doi: 10.1021/ac070470c.

Ratledge, C. and Dover, L. G. (2000) ‘Iron Metabolism in Pathogenic Bacteria’, Annual Review of Microbiology. doi: 10.1146/annurev.micro.54.1.881.

Raynaud, C. *et al*. (2002) ‘The functions of OmpATb, a pore-forming protein of Mycobacterium tuberculosis’, Molecular Microbiology. doi: 10.1046/j.1365-2958.2002.03152.x.

Richard M. Hall Colin Ratledge (1982). A simple method for the production of mycobactin, the lipid-soluble siderophore, from mycobacteria. FEMS Microbiology Letter 15: 133-136

Roberts, M. C., C. A. Riedy, *et al*. (2002). "How xylitol-containing products affect cariogenic bacteria." J Am Dent Assoc 133(4): 435-441; quiz 492-433.

Robinson, N. *et al*. (2008) ‘Mycobacterial phenolic glycolipid inhibits phagosome maturation and subverts the pro-inflammatory cytokine response’, Traffic. doi: 10.1111/j.1600-0854.2008.00804.x.

Rohde, K. *et al*. (2007) ‘Mycobacterium tuberculosis and the environment within the phagosome’, Immunological Reviews. doi: 10.1111/j.1600-065X.2007.00547.x.

Rouse, D. A. *et al*. (1996) ‘Site-directed mutagenesis of the katC gene of Mycobacterium tuberculosis: Effects on catalase-peroxidase activities and isoniazid resistance’, Molecular Microbiology. doi: 10.1046/j.1365-2958.1996.00133.x.

Russell, D. G. (2001). "Mycobacterium tuberculosis: Here today, and here tomorrow." Nature Reviews Molecular Cell Biology 2(8): 569-577.

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

Saini, D. K. *et al*. (2004) ‘DevR-DevS is a bona fide two-component system of Mycobacterium tuberculosis that is hypoxia-responsive in the absence of the DNA-binding domain of DevR’, Microbiology. doi: 10.1099/mic.0.26218-0.

Sakamoto, K. (2012). "The Pathology of Mycobacterium tuberculosis Infection." Veterinary Pathology 49(3): 423-439.

Sambou, T., P. Dinadayala, *et al*. (2008). "Capsular glucan and intracellular glycogen of Mycobacterium tuberculosis: biosynthesis and impact on the persistence in mice." MolMicrobiol 70(3): 762-774.

Sassetti, C. M. and Rubin, E. J. (2003) ‘Genetic requirements for mycobacterial survival during infection’, Proceedings of the National Academy of Sciences. doi: 10.1073/pnas.2134250100.

Sauer, U. (2006). "Metabolic networks in motion: 13C-based flux analysis." Mol Syst Biol 2: 62.

Sauer, U., V. Hatzimanikatis, *et al*. (1997). "Metabolic fluxes in riboflavin-producing Bacillus subtilis." Nat Biotechnol 15(5): 448-452.

Schoeman, J. C., du Preez, I. and Loots, D. T. (2012) ‘A comparison of four sputum pre-extraction preparation methods for identifying and characterising Mycobacterium tuberculosis using GCxGC-TOFMS metabolomics’, Journal of Microbiological Methods. doi: 10.1016/j.mimet.2012.09.002.

Sharma, R., D. Keshari, *et al*. (2016). "MRA_1571 is required for isoleucine biosynthesis and improves Mycobacterium tuberculosis H37Ra survival under stress." Sci Rep 6: 27997.

Sherman, D. R. *et al*. (2001) 'Regulation of the Mycobacterium tuberculosis hypoxic response gene encoding alpha -crystallin.', Proceedings of the National Academy of Sciences of the United States of America. doi: 10.1073/pnas.121172498.

Shiloh, M. U. and Nathan, C. F. (2000) 'Reactive nitrogen intermediates and the pathogenesis of Salmonella and Mycobacteria', Current Opinion in Microbiology. doi: 10.1016/S1369-5274(99)00048-X.

Sibley, L. D. and Krahenbuhl, J. L. (1987) 'Mycobacterium leprae-burdened macrophages are refractory to activation by gamma interferon', Infection and Immunity.

Simão, E. *et al*. (2005) 'Qualitative modelling of regulated metabolic pathways: Application to the tryptophan biosynthesis in E. Coli', Bioinformatics. doi: 10.1093/bioinformatics/bti1130.

Solans, L. *et al*. (2014) 'The PhoP-Dependent ncRNA Mcr7 Modulates the TAT Secretion System in Mycobacterium tuberculosis', PLoS Pathogens. doi: 10.1371/journal.ppat.1004183.

Song, H., J. Huff, *et al*. (2011). "Expression of the ompATb operon accelerates ammonia secretion and adaptation of Mycobacterium tuberculosis to acidic environments." MolMicrobiol 80(4): 900-918.

Sreevatsan, S. *et al*. (1997) 'Restricted structural gene polymorphism in the Mycobacterium tuberculosis complex indicates evolutionarily recent global dissemination', Proceedings of the National Academy of Sciences. doi: 10.1073/pnas.94.18.9869.

Sritharan, M. (2000) 'Iron as a candidate in virulence and pathogenesis in mycobacteria and other microorganisms', World Journal of Microbiology and Biotechnology. doi: 10.1023/A:1008995313232.

Stincone, A., A. Prigione, *et al*. (2015). "The return of metabolism: biochemistry and physiology of the pentose phosphate pathway." Biol Rev CambPhilosSoc 90(3): 927-963.

Stuehr, D. J. and Nathan, C. F. (1989) ‘Nitric oxide. A macrophage product responsible for cytostasis and respiratory inhibition in tumor target cells.’, The Journal of experimental medicine. doi: 10.1084/jem.169.5.1543.

Sturgill-Koszycki, S. *et al*. (1994) ‘Lack of acidification in Mycobacterium phagosomes produced by exclusion of the vesicular proton-ATPase’, Science. doi: 10.1126/science.8303277.

Syhre, M. and Chambers, S. T. (2008) ‘The scent of Mycobacterium tuberculosis’, Tuberculosis. doi: 10.1016/j.tube.2008.01.002.

Szymanska, E., E. Saccenti, *et al*. (2012). "Double-check: validation of diagnostic statistics for PLS-DA models in metabolomics studies." Metabolomics 8(Suppl 1): 3-16.

Tanzer, J. M., A. Thompson, *et al*. (2006). "Streptococcus mutans: fructose transport, xylitol resistance, and virulence." J Dent Res 85(4): 369-373.

Tian, J., R. Bryk, *et al*. (2005). "Variant tricarboxylic acid cycle in Mycobacterium tuberculosis: identification of alpha-ketoglutarate decarboxylase." Proc Natl AcadSci U S A 102(30): 10670-10675.

Toivari, M. H., L. Ruohonen, *et al*. (2007). "Metabolic engineering of Saccharomyces cerevisiae for conversion of D-glucose to xylitol and other five-carbon sugars and sugar alcohols." Appl Environ Microbiol 73(17): 5471-5476.

Trahan, L., S. Neron, *et al*. (1991). "Intracellular xylitol-phosphate hydrolysis and efflux of xylitol in Streptococcus sobrinus." Oral MicrobiolImmunol 6(1): 41-50.

Tramonti, A., P. Visca, *et al*. (2002). "Functional characterization and regulation of gadX, a gene encoding an AraC/XylS-like transcriptional activator of the Escherichia coli glutamic acid decarboxylase system." J Bacteriol 184(10): 2603-2613.

Tsai, M. C., S. Chakravarty, *et al*. (2006). "Characterization of the tuberculous granuloma in murine and human lungs: cellular composition and relative tissue oxygen tension." Cellular Microbiology 8(2): 218-232.

Tufariello, J. M. *et al*. (2016) ‘Separable roles for Mycobacterium tuberculosis ESX-3 effectors in iron acquisition and virulence’, Proceedings of the National Academy of Sciences. doi: 10.1073/pnas.1523321113.

Turk, V., Turk, B. and Turk, D. (2001) ‘Lysosomal cysteine proteases: Facts and opportunities’, EMBO Journal. doi: 10.1093/emboj/20.17.4629.

Turner, O. C., Basaraba, R. J. and Orme, I. M. (2003) ‘Immunopathogenesis of pulmonary granulomas in the guinea pig after infection with Mycobacterium tuberculosis.’, Infection and immunity, 71(2), pp. 864–71. doi: 10.1128/IAI.71.2.864-871.2003.

Ulrich, E. L., H. Akutsu, *et al*. (2008). "BioMagResBank." Nucleic Acids Res 36(Database issue): D402-408.

Usuda, Y., Y. Nishio, *et al*. (2010). "Dynamic modeling of Escherichia coli metabolic and regulatory systems for amino-acid production." J Biotechnol 147(1): 17-30.

van Kessel, J. C. and Hatfull, G. F. (2007) ‘Recombineering in Mycobacterium tuberculosis’, Nature Methods. doi: 10.1038/nmeth996.

Vandal, O. H. *et al*. (2008) ‘A membrane protein preserves intrabacterial pH in intraphagosomal Mycobacterium tuberculosis’, Nature Medicine. doi: 10.1038/nm.1795.

Vergne, I. *et al*. (2005) ‘Mechanism of phagolysosome biogenesis block by viable Mycobacterium tuberculosis’, Proceedings of the National Academy of Sciences. doi: 10.1073/pnas.0409716102.

Voskuil, M. I., I. L. Bartek, *et al*. (2011). "The response of mycobacterium tuberculosis to reactive oxygen and nitrogen species." Front Microbiol 2: 105.

Waddell, S. J. and Butcher, P. D. (2007) ‘Microarray analysis of whole genome expression of intracellular Mycobacterium tuberculosis’, Current molecular medicine. doi: 10.2174/156652407780598548.

Walburger, A. *et al*. (2004) ‘Protein kinase G from pathogenic mycobacteria promotes survival within macrophages’, Science. doi: 10.1126/science.1099384.

Wang, R. and Marcotte, E. M. (2008) ‘The proteomic response of mycobacterium smegmatis to anti-tuberculosis drugs suggests targeted pathways’, Journal of Proteome Research. doi: 10.1021/pr0703066.

Weiner, J., 3rd, S. K. Parida, *et al*. (2012). "Biomarkers of inflammation, immunosuppression and stress with active disease are revealed by metabolomic profiling of tuberculosis patients." PLoS One 7(7): e40221.

Weiner, J., 3rd, S. K. Parida, *et al*. (2012). "Biomarkers of inflammation, immunosuppression and stress with active disease are revealed by metabolomic profiling of tuberculosis patients." PLoS One 7(7): e40221.

Whatmore, A. M., J. A. Chudek, *et al*. (1990). "The effects of osmotic upshock on the intracellular solute pools of Bacillus subtilis." J Gen Microbiol 136(12): 2527-2535.

Wishart, D. S., D. Tzur, *et al*. (2007). "HMDB: the Human Metabolome Database." Nucleic Acids Res 35(Database issue): D521-526.

Wolf, A. J., B. Linas, *et al*. (2007). "Mycobacterium tuberculosis infects dendritic cells with high frequency and impairs their function in vivo." Journal of Immunology 179(4): 2509-2519.

World Health Organization releases 2015 global report on tuberculosis. (2015). Breathe 11(4): 244-244.

Xia, J., R. Mandal, *et al*. (2012). "MetaboAnalyst 2.0--a comprehensive server for metabolomic data analysis." Nucleic Acids Res 40(Web Server issue): W127-133.

Xia, J., T. C. Bjorndahl, *et al*. (2008). "MetaboMiner--semi-automated identification of metabolites from 2D NMR spectra of complex biofluids." BMC Bioinformatics 9: 507.

Yancey, P. H., W. R. Blake, *et al*. (2002). "Unusual organic osmolytes in deep-sea animals: adaptations to hydrostatic pressure and other perturbants." Comp BiochemPhysiol A MolIntegrPhysiol 133(3): 667-676.

Yang, A. J., D. Chandswangbhuvana, *et al*. (1999). "Intracellular accumulation of insoluble, newly synthesized abetan-42 in amyloid precursor protein-transfected cells that have been treated with Abeta1-42." J BiolChem 274(29): 20650-20656.

Yimer, S. A. *et al*. (2015) 'Mycobacterium tuberculosis lineage 7 strains are associated with prolonged patient delay in seeking treatment for pulmonary tuberculosis in Amhara region, Ethiopia', Journal of Clinical Microbiology. doi: 10.1128/JCM.03566-14.

Ytting, C. K., A. T. Fuglsang, *et al*. (2012). "Measurements of intracellular ATP provide new insight into the regulation of glycolysis in the yeast Saccharomyces cerevisiae." IntegrBiol (Camb) 4(1): 99-107.

Yu, K. *et al*. (1999) 'Toxicity of nitrogen oxides and related oxidants on mycobacteria: M. tuberculosis is resistant to peroxynitrite anion', Tubercle and Lung Disease. doi: 10.1054/tuld.1998.0203.

Zamboni, N., S.-M. Fendt, *et al*. (2009). "13C-based metabolic flux analysis." Nature Protocols 4: 878.

Zhou, A., J. Ni, *et al*. (2013). "Application of (1)h NMR spectroscopy-based metabolomics to sera of tuberculosis patients." J Proteome Res 12(10): 4642-4649.

Zhu, Y., E. Jameson, *et al*. (2014). "Carnitine metabolism to trimethylamine by an unusual Rieske-type oxygenase from human microbiota." Proc Natl AcadSci U S A 111(11): 4268-4273.

Zimmermann, M., M. Kogadeeva, *et al*. (2017). "Integration of Metabolomics and Transcriptomics Reveals a Complex Diet of Mycobacterium tuberculosis during Early Macrophage Infection." mSystems 2(4).

Report of Plagiarism check

Metabolomics approach to understand mycobacterial adaptation to stress and identify disease markers

by Arshad Rizvi

Metabolomics approach to understand mycobacterial adaptation to stress and identify disease markers

ORIGINALITY REPORT

9%	3%	8%	3%
SIMILARITY INDEX	INTERNET SOURCES	PUBLICATIONS	STUDENT PAPERS

PRIMARY SOURCES

1	Submitted to Indian Institute of Science Education and Research Student Paper	1%
2	www.frontiersin.org Internet Source	1%
3	Sourav RoyChoudhury, Tushar H. More, Ratna Chattopadhyay, Indrani Lodh et al. "Polycystic ovary syndrome in Indian women: a mass spectrometry based serum metabolomics approach", Metabolomics, 2017 Publication	1%
4	Tushar H. More, Muralidhararao Bagadi, Sourav RoyChoudhury, Mainak Dutta, Annu Uppal, Anupama Mane, Manas K. Santra, Koel Chaudhury, Srikanth Rapole. "Comprehensive quantitative lipidomic approach to investigate serum phospholipid alterations in breast cancer", Metabolomics, 2016 Publication	<1%

Metabolomics approach to understand mycobacterial adaptation to stress and identify disease markers

by Arshad Rizvi

Metabolomics approach to understand mycobacterial adaptation to stress and identify disease markers

ORIGINALITY REPORT

9%	3%	8%	3%
SIMILARITY INDEX	INTERNET SOURCES	PUBLICATIONS	STUDENT PAPERS

PRIMARY SOURCES

1	Submitted to Indian Institute of Science Education and Research Student Paper	1%
2	www.frontiersin.org Internet Source	1%
3	Sourav RoyChoudhury, Tushar H. More, Ratna Chattopadhyay, Indrani Lodh et al. "Polycystic ovary syndrome in Indian women: a mass spectrometry based serum metabolomics approach", Metabolomics, 2017 Publication	1%
4	Tushar H. More, Muralidhararao Bagadi, Sourav RoyChoudhury, Mainak Dutta, Annu Uppal, Anupama Mane, Manas K. Santra, Koel Chaudhury, Srikanth Rapole. "Comprehensive quantitative lipidomic approach to investigate serum phospholipid alterations in breast cancer", Metabolomics, 2016 Publication	<1%