Characterising tuberculosis and treatment failure thereof using metabolomics

L Luies

21637156

Thesis submitted for the degree *Philosophiae Doctor* in *Biochemistry* at the Potchefstroom Campus of the North-West University

Promoter: Prof Du Toit Loots

April 2017

"In science one tries to tell people, in such a way as to be understood by everyone, something that no one ever knew before."

Paul Dirac

ACKNOWLEDGEMENTS

The contributions of the following institutions and individuals made towards the successful completion of this study are hereby acknowledged:

- NRF/DST Centre of Excellence in Biomedical Tuberculosis Research, Faculty of Health Sciences, University of Stellenbosch, South Africa, for providing the urine samples used in this investigation.

- The National Research Foundation (NRF) of South Africa, Technology Innovation Agency (TIA) and North-West University (NWU) for the research grants provided.

- Prof. Du Toit Loots, my study supervisor, and other members of the NWU Biochemistry Department (especially Prof. Japie Mienie, Dr. Ilse du Preez, Mrs. Mari van Reenen, Mrs. Derylize Beukes-Maasdorp and Dr. Zander Lindeque), for their guidance, expertise, insights, advice and patience throughout the years.

- My friends, for the moments of laughter and stress relief, which has kept me sane.

- My dearest parents, Leonard and Sandra, for their unending love, prayers, mental encouragement and financial support during my studies.

- My husband, Juan, for all your love, support and encouragement, every step of the way.

"I can do everything though Christ who gives me strength."
Philippians 4:13

TABLE OF CONTENTS

SUMMARY

Tuberculosis (TB), a highly contagious bacterial disease caused by *Mycobacterium tuberculosis*, is considered the leading cause of death globally from a single bacterial pathogen. The latest reports indicate 10.4 million new TB cases are diagnosed globally per annum, of which approximately only 5.7 million actually receive treatment, resulting in an estimated 1.8 million deaths. This high TB prevalence may be ascribed to a number of factors, including, amongst others, untimely and inaccurate diagnostics, inadequate treatment regimens, drug-resistant *M. tuberculosis* strains, human immunodeficiency virus (HIV) co-infection, and inadequate knowledge of the TB disease in general.

This study is novel in the sense that it used a metabolomics research approach to identify new metabolite markers from patient-collected urine, for better characterising/understanding the TB disease state and treatment failure thereof.

Using a validated urinary organic acid extraction and a two-dimensional gas chromatography time-of-flight mass spectrometry (GCxGC-TOFMS) metabolomics approach, we were able to differentiate a culture-confirmed active TB-positive group and TB-negative healthy control group, based on their detected metabolite differences, utilising a variety of multi- and univariate statistical methods. We identified the most significant urinary TB metabolite markers contributing to these differences, which shed light on previously unknown mechanisms/adaptations of the host in response to *M. tuberculosis* and other host–pathogen interactions. The most significant of these were the TB-induced changes resulting in an abnormal host fatty acid and amino acid metabolism, mediated through changes in interferon gamma and possibly insulin. This also explains some of the symptoms associated with TB and provides clues to better treatment approaches.

Thereafter, the same approach was used to compare TB-positive patients with an unsuccessful treatment outcome to those successfully treated. Differentiation of the groups was achieved using the urine samples collected from these patients at time of diagnosis, i.e. before any treatment was administered. The identified urinary biomarkers were then used to better understand the underlying biology related to TB treatment failure. The most significant observations were the elevated levels of those metabolites associated with a gut microbiome imbalance, which has been shown to alter an individual's response to anti-TB drugs and also negatively influence their immune function, contributing to an unsuccessful treatment outcome. Another interesting observation was those metabolites traditionally used for diagnosing inborn abnormalities in any of the three enzymes of the mitochondrial trifunctional protein complex in the treatment failure group. Since L-carnitine and various short-chain

fatty acids are also reduced in these individuals, and are well-known for their anti-mycobacterial properties, this metabolic profile may explain an additional mechanism responsible for these individuals having an increased disease severity and/or a poor response to TB treatment.

Considering the possible significance of these findings from a diagnostic perspective, the GCxGC-TOFMS data generated from the aforementioned treatment outcome experiment was reanalysed using a univariate statistical approach, in order to find possible diagnostic markers for predicting treatment outcome, utilising urine collected at time of diagnosis. Using a logistic regression model, two predictors, i.e. 3,5-dihydroxybenzoic acid and 3-(4-hydroxy-3-methoxyphenyl)propionic acid, displayed the capacity to predict an unsuccessful treatment with an area under the receiver operating characteristic curve value of 0.94, and a leave-one-out cross-validation value of 0.89, indicating high sensitivity and specificity. Furthermore, these two identified predictors are also associated with an imbalance in gut microbiota, confirming the previously proposed mechanisms related to treatment failure in these individuals.

Considering the results, this study not only proved the capability of a metabolomics research approach to identify new metabolite markers which could be used towards better understanding TB, and treatment failure thereof, but also possibly diagnostically for predicting treatment outcome of first-line anti-TB drugs at time of diagnosis. Furthermore, the fact that these markers can be detected from patient-collected urine, as opposed to sputum, has additional benefits for both research and diagnostic applications, considering the ease by which such samples can be obtained, with very little discomfort to the individual.

Key words: host–pathogen interactions; metabolomics; prediction; treatment failure; tuberculosis; urine

LIST OF TABLES

CHAPTER 1:

CHAPTER 2:

CHAPTER 4:

CHAPTER 5:

CHAPTER 6:

APPENDIX A:

LIST OF FIGURES

CHAPTER 2:

CHAPTER 3:

CHAPTER 4:

CHAPTER 5:

CHAPTER 6:

APPENDIX C:

LIST OF SYMBOLS AND ABBREVIATIONS

Abbreviation	Meaning	Abbreviation	Meaning
2MG	2-C-Methylglycerol	5-OH-PA	5-Hydroxypyrazinoic acid
ABC	ATP-binding cassette	ACAA1	3-Ketoacyl-CoA thiolase
AcHZ	Monoacetylhydrazine	AcINH	N1-Acetyl-N2-isonicotinylhydrazide
ACOX 1	Acyl-CoA oxidase 1	ACOX 2	Acyl-CoA oxidase 2
ACP	Acyl carrier protein	ADR	Adverse drug reaction
ARV	Antiretroviral	ATD	Automated thermal desorption
ATDH	Anti-tuberculosis drug-induced hepatotoxicity	ATP	Adenosine triphosphate
AUC	Area under the curve	BCAA	Branched chain amino acid
BCG	Bacille Calmette-Guerin	BH_4	Tetrahydrobiopterin
BMI	Body mass index	bp	Base pair
BSTFA	Bis(trimethylsilyl)-trifluoroacetamide	CACT	Carnitine-acylcarnitine translocase
CDC	Centres for Disease Control and Prevention	CE	Capillary electrophoresis
CFP	Culture filtrate protein	CNS	Central nervous system
CoA	Coenzyme A	COPD	Chronic obstructive pulmonary disease
CPT-1	Carnitine palmitoyltransferase 1	CPT-2	Carnitine palmitoyltransferase 2
CSF	Cerebrospinal fluid	CV	Coefficient of variation
CYP450	Cytochrome P450	DFA	Discriminant function analysis
DHPS	Dihydropteroate synthase	DiAcHZ	Diacetylhydrazine
DNA	Deoxyribonucleic acid	DOTS	Directly Observed Therapy Short-course
D-PBE	Peroxisomal bifunctional enzyme	DPG	Diphosphatidylglycerol
EDA	2,2'-(Ethylenediimino)-di-butyric acid	ELISA	Enzyme-linked immunosorbent assay
EMB	Ethambutol	EN	Electronic nose
ESAT	Early secreted antigenic target	ESI	Electrospray ionisation
etc.	Et cetera	FDA	Food and Drug Administration
GABA	Gamma-aminobutyric acid	GATB	Global Alliance for Tuberculosis Drug Development

(Continues on next page)

(Continues from previous page)

Abbreviation	Meaning	Abbreviation	Meaning
GC	Gas chromatography	GDH	Glutamate dehydrogenase
GSTM1	Glutathione S-transferase Mu 1	H_2O_2	Hydrogen peroxide
HCl	Hydrochloric acid	HIV	Human immunodeficiency virus
HPLC	High-performance liquid chromatography	HZ	Hydrazine
IDO1	Indoleamine 2,3-dioxygenase 1	IFN-γ	Interferon gamma
IGRA	Interferon-gamma release assay	INA	Isonicotinic acid
INH	Isoniazid/isonicotinylhydrazide	INH-KA	Isoniazid-ketoglutaric acid
INH-PA	Isoniazid hydrazones with pyruvic acid	KAT	Kynurenine aminotransferase
LAM	Lipoarabinomannan	LC	Liquid chromatography
LCFA	Long-chain fatty acid	LM	Lipomannan
M. tuberculosis	*Mycobacterium tuberculosis*	M/SCHAD	Medium/short chain hydroxyacyl-CoA dehydrogenase
MADD	Multiple acyl-CoA dehydrogenase deficiency	MAPc	Mycolyl arabinogalactan-peptidogalactan complex
MCAD	Medium chain acyl-CoA dehydrogenase	MCT	Medium chain 3-ketoacyl-CoA thiolase
MDR-TB	Multidrug-resistant tuberculosis	MEP	2-C-Methyl-D-erythritol-4-phosphate
MIC	Minimum inhibitory concentration	MS	Mass spectrometry
MS/MS	Tandem mass spectrometry	MTP	Mitochondrial trifunctional protein
Na_2SO_4	Anhydrous sodium sulphate	NAA	Nucleic acid amplification
NAD	Nicotinamide adenine dinucleotide	NADH	Nicotinamide adenine dinucleotide (NAD) + hydrogen (H)
NAT 2	N-acetyltransferase 2	NMR	Nuclear magnetic resonance
NRF	National Research Foundation	NWU	North-West University
OPLS-DA	Orthogonal partial least-squares discriminant analysis	PA	Pyrazinoic acid
PABA	*p*-Aminobenzoate	PAH	Phenylalanine hydroxylase
PAS	Para-aminosalicylic acid	PC	Principal component
PCA	Principle component analysis	PCR	Polymerase chain reaction
PDIM	Phthiocerol dimycocerosate	PE	Phosphatidylethanolamine
PHYH	Phytanoyl-CoA hydroxylase	PI	Phosphatidylinositol
PIM	Phosphatidylinositol mannoside	pistanalDH	Pristanaldehyde dehydrogenase

(Continues on next page)

(Continues from previous page)

Abbreviation	Meaning	Abbreviation	Meaning
PKU	Phenylketonuria	PLIEM	Potchefstroom Laboratory for Inborn Errors of Metabolism
PLS-DA	Partial least squares discriminant analysis	PMN	Polymorphonuclear
PPD	Purified protein derivative	PZA	Pyrazinamide
QC	Quality control	QC–CV	Quality control-coefficient of variation
Q-TOFMS	Quadruple time-of-flight mass spectrometry	RIF	Rifampicin
RNA	Ribonucleic acid	ROC	Receiver operating characteristic
RT	Retention time	S/N	Signal to noise ratio
SCAD	Short chain acyl-CoA dehydrogenase	SCFA	Short-chain fatty acid
SCT	Short chain 3-ketoacyl-CoA thiolase	SCYD	Short chain enoyl-CoA hydratase
SPME	Solid phase micro-extraction	STARD	Standards for the Reporting of Diagnostic Accuracy Studies
TAT	Tyrosine aminotransferase	TB	Tuberculosis
TBSA	Tuberculostearic acid	TCA	Tricarboxylic acid
TDM	Trehalose 6,6' dimycolate	TDO	Tryptophan dioxygenase
TDR-TB	Totally drug-resistant tuberculosis	TIA	Technology Innovation Agency
TL	Translocase	TMCS	Trimethylchlorosilane
TNF-α	Tumour necrosis factor alpha	TOFMS	Time-of-flight mass spectrometry
TST	Tuberculin skin test	U(H)PLC	Ultra-(high) performance liquid chromatography
UV	Ultra-violet	VIP	Variables' influence on the projection
VLCAD	Very-long-chain acyl-coenzyme A dehydrogenase	VOC	Volatile organic compound
WHO	World Health Organisation	XDR-TB	Extensively-drug-resistant tuberculosis
XME	Xenobiotic-metabolising enzyme	Xpert	GeneXpert MTB/RIF

Symbols:

Symbol	Meaning	Symbol	Meaning
<	Smaller than	↑	Increased
>	Greater than	↓	Decreased
α	Alpha	β	Beta
°C	Degrees Celsius	g	Gram
m	Metre	m/z	Mass to charge ratio
mg/g	Milligram per gram	mL	Millilitre
mL/min	Millilitre per minute	mmol/L	Millimole per litre
µg/mL	Microgram per millilitre	µL	Microliter
µL/µmol	Microliter per micromole	µm	Micrometre

CHAPTER 1: PREFACE

1.1 BACKGROUND AND MOTIVATION

Tuberculosis (TB) remains the world's foremost cause of death from a single bacterial agent, which is alarming as it is considered curable. Nearly 10.4 million new TB cases are reported per annum, resulting in approximately 1.8 million deaths globally. Additionally, TB is considered the leading cause of death in patients co-infected with the human immunodeficiency virus (HIV), contributing to approximately 0.4 million of the reported global TB deaths in 2015, with Africa accounting for the majority of these (± 84%). In South Africa, approximately 454 000 people were infected with TB in 2015, resulting in 73 000 deaths. In the same year, 910 124 people who were newly enrolled in HIV-care globally were also started on TB-preventive therapy, with South Africa being the country with the highest global prevalence of this (45% of the total cases) (World Health Organization, 2016).

Accessibility of treatment is another concern. In 2015, of the approximately 580 000 cases of multi-drug-resistant (MDR)-TB reported, only 125 000 (20%) actually received treatment, contributing to an estimated 250 000 deaths. Furthermore, an estimated 9.5% of these MDR cases were diagnosed with extensively-drug-resistant TB (XDR-TB) (World Health Organization, 2016). In 2007, the first cases of totally-drug-resistant TB (TDR-TB) were reported in India, Iran and Italy, with South Africa becoming the fourth country reporting the emergence of this virtually untreatable strain (World Health Organization, 2012). The latest treatment outcome data indicate a treatment success rate of 83% for fully drug-susceptible TB, 52% for MDR-TB and 28% for XDR-TB (World Health Organization, 2016). Treatment failure or an unsuccessful treatment outcome is an additional concern, and although it has been strongly associated with a number of factors, including (a) irregular or inadequate anti-TB drug supplies to rural areas in third-world countries, (b) poor patient TB-education, (c) poor socio-economic circumstances (such as poverty, malnutrition and overcrowding), (d) the prolonged treatment duration, (e) treatment non-adherence, (f) drug-resistance by the infectious organism, and (g) HIV co-infection, there are various biological/biochemical factors associated with this which are not yet well described or unknown (De Villiers & Loots, 2013).

Despite the major discovery in 1882 by Robert Koch, that the causal agent of TB is *M. tuberculosis* (Daniel, 2005), and all the genomics, transcriptomics and proteomics data collected on this organism to date, as well as the vaccination, diagnostic and treatment approaches developed since, TB is still considered a major health problem globally. Considering this, we still have a lot to learn about this infectious organism and the host–

microbe adaptations and interactions for the development of improved diagnostic and treatment strategies. An alternative research approach such as metabolomics, considering its capability for new biomarker identification and hypotheses generation, may serve well towards achieving these goals. To date, the North-West University's Research Focus Area: Human Metabolomics, has used this research strategy to better understand TB and to define the functionality of unknown mycobacterial genes (Loots *et al.*, 2013; Loots *et al.*, 2016), to better describe *M. tuberculosis* virulence (Meissner-Roloff *et al.*, 2012), drug-resistance (Du Preez & Loots, 2012; Loots, 2014; Loots, 2015), host and microbe interactions (Du Preez & Loots, 2013b; Luies & Loots, 2016), and also to identify or characterise various *Mycobacterium* species for possible diagnostic applications (Du Preez & Loots, 2013a; Olivier & Loots, 2012a; Olivier & Loots, 2012b). Despite these recent applications of metabolomics with a view towards a better understanding of this disease and related topics, little research has been done using metabolomics for identifying urinary biomarkers of TB, which would serve to possibly better describe host adaptations, or host–microbe interactions. Furthermore, very little is known about the biology behind treatment failure, and the possible prediction of this phenomenon.

1.2 AIMS AND OBJECTIVES OF THIS STUDY

1.2.1 Aims

The aims of this study are to use a GCxGC-TOFMS metabolomics approach to compare and differentiate the urine organic acid profiles of:

1. Culture-confirmed active TB-positive patients (n=38) and TB-negative healthy controls (n=30), to identify characteristic urinary metabolites, occurring as a result of the host–pathogen interactions and adaptations, for better characterising TB.

2. Culture-confirmed active TB-positive patients with a successful (n=27) and unsuccessful (n=11) treatment outcome as early as possible during the treatment regimen, and identify those metabolite markers which would (a) better characterise and explain the biological mechanisms related to TB treatment failure, and to (b) possibly be used diagnostically for predicting treatment failure.

1.2.2 Objectives

In view of the above mentioned aims, the objectives of this study are to:

1. Determine and validate the analytical repeatability of the methodology (analyst/ extraction repeatability) and that of the GCxGC-TOFMS analytical apparatus used (machine repeatability), for analysing samples in the context of a metabolomics research application.

2. Apply the validated methodology (mentioned in Objective 1) to extract all samples.

3. Utilise various multi- and univariate statistical approaches to identify metabolite markers for better characterising the TB disease state (Aim 1), treatment failure (Aim 2a), as well as to be used diagnostically for predicting treatment failure (Aim 2b).

1.3 STRUCTURE OF THESIS AND RESEARCH OUTPUTS

This thesis is written specifically to comply with the requirements of the North-West University, Potchefstroom Campus, South Africa, for the completion of the degree Philosophiae Doctor (Biochemistry) in article format. Thus, each chapter will have its own introduction, materials and methods, results, discussion, conclusion and reference sections, relevant to that chapter. Additionally, a comprehensive literature review (Chapter 2) and conclusion (Chapter 7) are also added in accordance to these guidelines.

Chapter 1 (the current chapter) gives a brief background and motivation for the study, as well as the aims and objectives. Also included in this chapter are the basic layout/structure of the thesis, and the research outputs/publications which emanated from this study.

Chapter 2 provides a literature review of TB in general and all other related aspects of this disease, relevant to this investigation. Parts of this chapter are published/submitted in three separate review papers (see Appendix D):

- De Villiers, L. & Loots, DT. (2013). Using metabolomics for elucidating the mechanisms related to tuberculosis treatment failure. *Current Metabolomics*, 1(4): 306-317.
- Du Preez, I., Luies, L. & Loots, DT. (2017). Metabolomics biomarkers for tuberculosis diagnostics: Current status and future objectives. *Biomarkers in Medicine,* 11 (2): 179-194.
- Luies, L., Du Preez, I. & Loots, DT. (2017). Improved tuberculosis treatment strategies using metabolomics. Submitted for publication to *Biomarkers in Medicine* (Manuscript number: BMM-2017-0141).

Chapter 3 describes the experimental design for the study, including the patient sample collection approach, as well as the metabolomics methodology used and validation thereof.

In Chapter 4, the methodology and data validated in Chapter 3 were used to compare and differentiate culture-confirmed active TB-positive (n=38) and TB-negative healthy control (n=30) groups from one another, based on the differences in their urinary metabolite profiles. Various multi- and univariate statistical analyses were used to identify which metabolite markers contributed most to the variation seen between the compared groups. These markers were then interpreted, considering their role in known metabolic pathways and biological mechanisms in both man and microbe, and subsequently new host–microbe interactions and adaptations better characterising TB were identified and discussed. This chapter was published in *Metabolomics* (see Appendix D):

- Luies, L. & Loots, DT. (2016). Tuberculosis metabolomics reveals adaptations of man and microbe in order to outcompete and survive. *Metabolomics*, 12(3): 1-9.

In Chapter 5, the methodology and data validated in Chapter 3 were used to differentiate TB-positive individuals with a successful (n=27) and unsuccessful (n=11) treatment outcome, on the basis of the metabolome differences detected between these two groups, from urine collected at various intervals throughout the course of treatment, i.e. from time of diagnosis, during treatment and two weeks after treatment completion. Multi- and univariate statistics were used to identify which metabolite markers differed most between the groups, which when described in the context of known metabolic pathways, better explains the underlying biological mechanisms of treatment failure. This chapter has been submitted for publication in *Metabolomics* (see Appendix D):

- Luies, L., Mienie, J., Motshwane, C., Ronacher, K., Walzl, G. & Loots, DT. (2017). Urinary metabolite markers characterising tuberculosis treatment failure. Submitted for publication in *Metabolomics* (Manuscript number: MEBO-D-17-00069).

In Chapter 6, we re-examined the data generated from the pre-treatment urine samples, which showed differentiation between the successfully cured and treatment failure groups (Chapter 5), from a purely univariate perspective, for the purpose of identifying markers, which when used in a logistic regression model can potentially serve as diagnostic predictors for TB treatment failure. This chapter has been submitted for publication in *Biomarkers in Medicine* (see Appendix D):

- Luies, L., Van Reenen, M., Ronacher, K., Walzl, G. & Loots, DT. (2017). Predicting tuberculosis treatment outcome using metabolomics. Submitted for publication to *Biomarkers in Medicine* (Manuscript number: BMM-2017-0133).

Chapter 7 is a comprehensive summary of all the results and conclusions, in the context of the aims originally set out in this investigation, in addition to future recommendations, which could be considered in light of the current findings.

1.4 AUTHOR CONTRIBUTIONS

The primary author/investigator of this thesis is Laneke Luies (neé de Villiers). The contributions of the co-authors, co-workers, and collaborators made towards this work, are summarised in Table 1-1.

The following is a statement from the primary investigator and supervisor, confirming their individual roles in the study and giving their permission that the data generated and conclusions made may form part of this thesis: I declare that my role in this study, as indicated in Table 1-1, is a representation of my actual contribution, and I hereby give my consent that this work may be published as part of the Ph.D. thesis of Laneke Luies.

Prof. Du Toit Loots

Laneke Luies

Table 1-1: The research team.

Co-author	Co-worker	Collaborator	Contribution
Laneke Luies (B.Sc. Hons. Biochemistry)			Responsible, together with the study leader, for the conceptualising, planning, execution, data analyses, and writing of the thesis, publications, and all other documentation associated with this study
Prof. Du Toit Loots (Ph.D. Biochemistry)			Study leader: Conceptualised, co-ordinated and supervised all aspects of the study, including the study design, planning, execution, writing of the thesis, publications, and all other documentation associated with this study

(Continues on next page)

(Continues from previous page)

Co-author	Co-worker	Collaborator	Contribution
	Dr. Ilse du Preez (Ph.D. Biochemistry)		Co-author on two review papers, and responsible, together with the other co-authors, for developing and conceptualising the review topic, working on data acquisition and drafting the article
	Prof. Japie Mienie (Ph.D. Biochemistry)		Co-author on a paper, where he assisted with data interpretation and drafting the article
	Mrs. Mari van Reenen (M.Sc. Statistics)		Co-author on a paper, where she assisted with statistical data analyses and interpretation and drafting the article
	Mrs. Derylize Beukes-Maasdorp (B.Sc. Biochemistry)		Assisted with sample analyses as officially appointed laboratory manager
	Christinah Motshwane (B.Sc. Hons. Biochemistry)		Performed the LC-MS sample analyses, used in Chapter 5
		AMPATH Laboratories	Determined the creatinine values of all collected urine samples
		Profs. Gerhard Walzl and Katharina Ronacher at the DST/NRF Centre of Excellence for Biomedical Tuberculosis Research/MRC Centre for Molecular and Cellular Biology, Division of Molecular Biology and Human Genetics, Faculty of Medicine and Health Sciences, Stellenbosch University (Tygerberg)	Provided the patient urine samples used in this metabolomics study

1.5 REFERENCES

Daniel, T. 2005. Robert Koch and the pathogenesis of tuberculosis [Founders of Our Knowledge]. *The International Journal of Tuberculosis and Lung Disease*, 9(11):1181-1182.

De Villiers, L. & Loots, D.T. 2013. Using metabolomics for elucidating the mechanisms related to tuberculosis treatment failure. *Current Metabolomics*, 1(4):306-317.

Du Preez, I. & Loots, D.T. 2012. Altered fatty acid metabolism due to rifampicin-resistance conferring mutations in the rpoB gene of *Mycobacterium tuberculosis*: Mapping the potential of pharmaco-metabolomics for global health and personalized medicine. *OMICS: A Journal of Integrative Biology*, 16(11):596-603.

Du Preez, I. & Loots, D.T. 2013a. Detection limit for differentiating between various *Mycobacterium* species and *Pseudomonas aeruginosa* using gas chromatography-mass spectrometry (GC-MS) metabolomics: A comparison of two extraction methods. *African Journal of Microbiology Research*, 7(9):797-801.

Du Preez, I. & Loots, D.T. 2013b. New sputum metabolite markers implicating adaptations of the host to *Mycobacterium tuberculosis*, and vice versa. *Tuberculosis*, 93(3):330-337.

Loots, D.T. 2014. An altered *Mycobacterium tuberculosis* metabolome induced by katG mutations resulting in isoniazid resistance. *Antimicrobial Agents and Chemotherapy*, 58(4):2144-2149.

Loots, D.T. 2015. New insights into the survival mechanisms of rifampicin-resistant *Mycobacterium tuberculosis*. *The Journal of Antimicrobial Chemotherapy*, 71(3):655-660.

Loots, D.T., Meissner-Roloff, R.J., Newton-Foot, M. & Van Pittius, N.C.G. 2013. A metabolomics approach exploring the function of the ESX-3 type VII secretion system of *M. smegmatis*. *Metabolomics*, 9(3):631-641.

Loots, D.T., Swanepoel, C.C., Newton-Foot, M. & van Pittius, N.C.G. 2016. A metabolomics investigation of the function of the ESX-1 gene cluster in mycobacteria. *Microbial Pathogenesis*, 100:268-275.

Luies, L. & Loots, D. 2016. Tuberculosis metabolomics reveals adaptations of man and microbe in order to outcompete and survive. *Metabolomics*, 12(3):1-9.

Meissner-Roloff, R.J., Koekemoer, G., Warren, R.M. & Loots, D.T. 2012. A metabolomics investigation of a hyper-and hypo-virulent phenotype of Beijing lineage *M. tuberculosis*. *Metabolomics*, 8(6):1194-1203.

Olivier, I. & Loots, D.T. 2012a. A comparison of two extraction methods for differentiating and characterising various *Mycobacterium* species and *Pseudomonas aeruginosa* using GC-MS metabolomics. *African Journal of Microbiology Research*, 6(13):3159-3172.

Olivier, I. & Loots, D.T. 2012b. A metabolomics approach to characterise and identify various *Mycobacterium* species. *Journal of Microbiological Methods*, 88(3):419-426.

World Health Organization. 2012. Global tuberculosis control 2012. Geneva, Switzerland (WHO Press).

World Health Organization. 2016. Global tuberculosis report 2016. Geneva, Switzerland (WHO Press).

CHAPTER 2: LITERATURE REVIEW

Parts of this chapter are in various stages of publication:

- De Villiers, L. & Loots, DT. (2013). Using metabolomics for elucidating the mechanisms related to tuberculosis treatment failure. *Current Metabolomics*, 1(4): 306-317.
- Du Preez, I., Luies, L. & Loots, DT. (2017). Metabolomics biomarkers for tuberculosis diagnostics: Current status and future objectives. *Biomarkers in Medicine*, 11 (2): 179-194.
- Luies, L., Du Preez, I. & Loots, DT. (2017). The role of metabolomics in tuberculosis treatment research. Submitted for publication to *Biomarkers in Medicine* (Manuscript number: BMM-2017-0141).

Since this thesis focuses on identifying new metabolite markers better describing the active TB disease state from a metabolomics perspective, in addition to using the same approach for better describing and predicting treatment failure, the majority of the literature described in Chapter 2 will be in accordance with this. However, for the sake of presenting a holistic perspective of TB, a short description of the bacteriology and pathophysiology of the disease will also be given in section 2.2.

2.1 METABOLOMICS

One of the newcomers to the "omics" revolution, metabolomics, can be defined as the unbiased identification and quantification of all the intra- and extra-cellular metabolites (small molecule intermediates and products of metabolism) present in a biological system, using highly selective and sensitive analytical techniques (Dunn *et al.*, 2005; Orešič, 2009; Van der Werf *et al.*, 2007), in conjunction with biostatistical and mathematical analyses for identifying new metabolite markers (Olivier & Loots, 2011).

Metabolomics is the holistic study of an organisms' metabolism, and is considered an important addition to systems biology when the data generated are interpreted together with that generated using genomics, transcriptomics and proteomics (Chen *et al.*, 2007). The application of metabolomics for biomarker discovery, is based on the principle that an external stimulus, such as TB disease or infection, an anti-TB drug, or a mutation resulting in drug-resistance, may disrupt normal metabolism, altering the overall physiological status of

an organism or host, and these metabolic changes are specific to the perturbation investigated and not due to an overall state of inflammation or unrelated disease processes (Chen *et al.*, 2007; Kaddurah-Daouk *et al.*, 2008). Thus, an individual's metabolic state is a representation of their overall physiological status. Considering this, the new metabolite markers identified due to the perturbation not only allow us to better understand the underlying disease mechanisms, but can also be used diagnostically and/or to predict individual/organism drug responses (Kaddurah-Daouk *et al.*, 2008). In the context of TB, *M. tuberculosis* disrupts the host's normal metabolism, initiating metabolite changes that can be visualised as metabolic biosignatures or metabolite patterns (Parida & Kaufmann, 2010; Weckwerth & Morgenthal, 2005). Metabolomics analyses the end-products of these perturbations present in blood, sputum and/or urine (or any other fluid or tissue for that matter), and identifies metabolites that are both endogenous and exogenous to the perturbation, capturing information with regards to the mechanisms of disease or drug action (Schoeman & Loots, 2011).

Various chromatographic instruments are used as tools in metabolomics, e.g. gas chromatography (GC), liquid chromatography (LC), capillary electrophoresis (CE) and nuclear magnetic resonance (NMR) (Kaddurah-Daouk *et al.*, 2008), for the detection of these alterations. Usually, these instruments are combined with mass spectrometry (MS) for compound detection, of which a variety are available, including the quadrupole MS, ion trap MS and the time-of-flight MS (TOFMS), each of which have their own advantages and disadvantages pertaining to sensitivity, specificity and molecular preferences (Weckwerth & Morgenthal, 2005). The type of analytical instrumentation selected for a metabolomics experiment would dependent on the goal(s) of the study, and the type of molecule or metabolite pathways one would expect to have changed, or would be interested in investigating (Weckwerth & Morgenthal, 2005), in addition to the various advantages and disadvantages of the respective instrumentation (see Table 2-1). However, as per definition, since metabolomics is the "study of all the metabolites present in a biological system/sample", many approaches have been developed in order to analyse the entire metabolome, or as much of it as possible, using one or more analytical methods (Du Preez & Loots, 2012; Du Preez & Loots, 2013b; Loots *et al.*, 2005; Schoeman & Loots, 2011). In both of the aforementioned instances, as part and parcel of true metabolomics research, multivariate statistical analysis, such as principle component analysis (PCA), partial least squares discriminant analysis (PLS-DA) or orthogonal partial least-squares discriminant analysis (OPLS-DA), are employed to extract information from the large metabolite datasets generated, in order to identify possible biomarkers (Chen *et al.*, 2007; Halouska *et al.*, 2007; Powers, 2009).

Table 2-1: A summary of the advantages and disadvantages of the common techniques and tools used in metabolomics (Aligent Technologies, 2007; Lei et al., 2011; Shulaev, 2006; Wallace et al., 2010; Want et al., 2007).

	Advantages	Disadvantages/Limitations
Direct MS	• Good combination of sensitivity and selectivity • Highly specific chemical information (accurate mass, isotope distribution patterns, characteristic fragment ions)	• Susceptible to ion suppression or enhancement • Data interpretation can be challenging • Inability to differentiate isomers
GC-MS	• Suited for volatile and non-volatile (after derivatisation) compound analyses • Affordable, with relatively low running costs • High sensitivity, mass resolution and accuracy • Good dynamic range • Compound identification using mass spectral library matching • Provides additional and orthogonal data (i.e. retention time/factor/index) • Reproducible chromatographic separations • Identification of stereoisomers possible • Shorter run times • Lower bleed (thinner films)	• Limited to volatile, thermally stable, and energetically stable compounds • Requires additional sample preparation, such as derivatisation, as this approach depends on the analytes being volatile and thermally stable, and few metabolites meet this requirement in their natural state • Less amenable to large, highly polar metabolites (poor volatility) • Co-eluting analytes in single dimension GC (corrected for using GCxGC) • Careful attention required for splitless injections • Slower scan rates/speed, unless coupled with TOFMS • Lower mass accuracy, unless coupled with TOFMS detectors
LC-MS	• Applicable for targeted and non-targeted metabolomics • Relatively low reagent cost • High analytical sensitivity, specificity and coverage depth • Able to analyse a wide range of compound classes • Ideal for highly polar and ionic compounds/ metabolites • Requires minimum sample preparation • Low matrix effects and interferences • The application of both positive and negative ionisation, hence more comprehensive metabolome coverage	• Lower chromatographic resolution compared to GC-MS • Comparatively higher running costs • Electrospray ionisation (ESI) can suffer from ionisation suppression • Retention time shifts are known to occur • Higher signal to noise (S/N) ratios compared to GC-MS • Low sample throughput
CE-MS	• Applicable for targeted and non-targeted metabolomics • Ideal for highly polar and ionic compounds/ metabolites • Fast, relatively affordable, and highly efficient separation technique • Typically coupled with TOFMS for fast acquisition rates	• Lower sensitivity compared to other techniques • Poor reproducibility • Electrochemical reactions of metabolites • May lack the necessary robustness • Least suitable for analysing complex biological samples
NMR-MS	• Highly selective • Non-destructive to sample material • Metabolite structural elucidation	• Lower sensitivity compared to other techniques

A number of metabolomics studies have already been done to date, identifying TB biomarkers which can be used for various applications, including improved diagnostics (as will be described in detail in section 2.3.3), and expanding on the existing knowledge of the biology of the causative pathogen (Rhee *et al.*, 2011), as well as to better explain various underlying disease mechanisms, including those related to the aforementioned drug-resistant *M. tuberculosis* strains (Du Preez & Loots, 2012; Loots, 2014; Loots, 2015), identifying new virulence factors (Meissner-Roloff *et al.*, 2012), and contributing to a better understanding of anti-TB drug mechanisms (Halouska *et al.*, 2012; Halouska *et al.*, 2013) and side-effects in the host (Loots *et al.*, 2005). What is of even greater interest currently is that metabolomics has allowed for a better understanding of the adaptations of *M. tuberculosis* to the host defence and vice versa (Du Preez & Loots, 2013b). This has shed light on never before identified metabolic pathways in both man and *M. tuberculosis*, which in time will undoubtedly contribute to improved treatment strategies, and ultimately assist in curbing this pandemic.

2.2 *M. TUBERCULOSIS* BACTERIOLOGY AND PATHOPHYSIOLOGY

M. tuberculosis, the highly infectious causative pathogen of TB, was first identified in 1882 by Robert Koch. These non-motile, rod-shaped, non-spore-forming, aerobic bacteria are facultative intracellular parasites, characterised with a slow growth rate. Mycobacteria usually measure 0.5 x 3 µm and are classified as acid-fast bacilli, due to fact that their cell walls are impermeable to certain dyes and stains (Knechel, 2009; Todar, 2009).

2.2.1 The *M. tuberculosis* cell wall

The cell wall composition of *M. tuberculosis* is considered unique among prokaryotes and crucial to its survival, virulence and growth. The cell wall consists of three layers, namely the (a) capsule (an outer impermeable coating), (b) mycolyl arabinogalactan-peptidogalactan complex (MAPc), and (c) inner cell membrane (see Figure 2-1). It is mainly composed of a variety of complex lipids (over 60%), which are largely considered toxic to eukaryotic cells. These cell wall lipids provide an extraordinary lipid barrier, which gives these bacteria a number of unique properties, including impermeability to stains and dyes, as well as resistance to host defence mechanisms, anti-TB drugs, acidic and alkaline compounds, osmotic lysis and lethal oxidations, allowing for its survival inside the host macrophages (Brennan, 2003; Knechel, 2009; Todar, 2009).

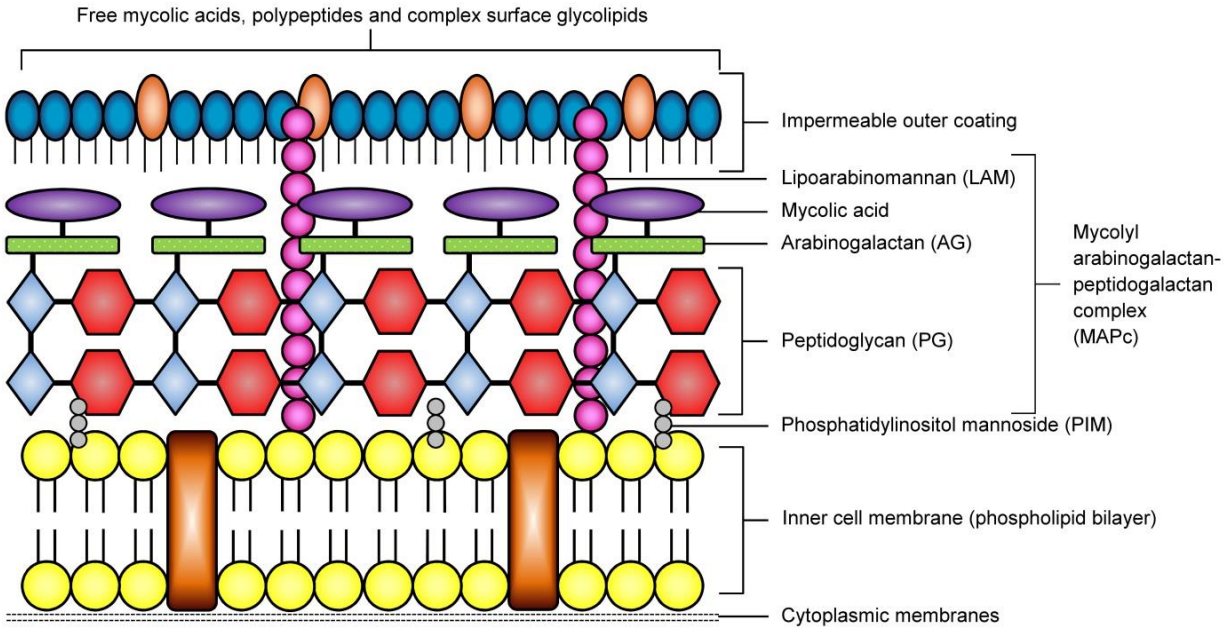


Figure 2-1: The cell wall composition of M. tuberculosis. The cell wall consists of multiple, complex layers, each contributing to M. tuberculosis virulence and its survival within the host. The well-developed cell wall consists of three layers, namely the capsule or outer impermeable coating, mycolyl arabinogalactan-peptidogalactan complex (MAPc), and the inner cell membrane, mostly composed of phospholipids.

The inner cell membrane contains primarily polar phospholipids, including phosphatidylethanolamine (PE), phosphatidylinositol (PI) and diphosphatidylglycerol (DPG), which form the basis of the membrane bilayer. *M. tuberculosis* has a family of four phosphatidylinositol mannosides (PIMs), which function to reinforce the cell membrane, adding additional coverage to ensure impermeability (Minnikin *et al.*, 2015). A peptidoglycan polymer, just outside the phospholipid inner cell membrane, is covalently attached to heteropolysaccharide arabinogalactan, which in turn is esterified to long-chain mycolic acids. These mycolic acids are complex β-hydroxylated α-alkyl-branched very long chain fatty acids (strong hydrophobes) with approximately 70–90 carbons and contain various diverse functional groups. These cell wall components form part and parcel of what is known as the MAPc, which functions by conferring rigidity to the cell wall as well as resistance and impermeability to various medicinal interventions (Brennan, 2003; Grzegorzewicz *et al.*, 2016; Knechel, 2009; Minnikin *et al.*, 2015; Todar, 2009). Cord factors, or trehalose dimycolates (TDMs), are composed of a trehalose sugar and disaccharide, esterified to two mycolic acid residues, and occur abundantly in the outer cell wall of virulent *M. tuberculosis* strains, allowing for these cells to grow in slender cords. This is also known to contribute to *M. tuberculosis* virulence, since it is toxic to mammalian cells, inhibits polymorphonuclear (PMN) leukocyte migration, induces granulomatous reactions, and furthermore protects the pathogen against the host defence mechanisms by preventing phagolysosomal fusion.

However, its main role is to act as an intermediate for the transfer of mycolic acids onto arabinosyl units in the cell envelope (Brennan, 2003; Minnikin *et al.*, 2015; Todar, 2009). Wax-D, on the other hand, is one of the major constituents found within the cell envelope, and may be characterised as an autolytic product of the cell wall's MAPc. It is thought to contribute to *M. tuberculosis* adjuvant activity and, hence, is necessary for cell wall growth regulation — since fractions of wax-D serve as building blocks from which the cell wall polymers are constructed (Kawabata *et al.*, 1998; Todar, 2009).

The outer layer consists of free mycolic acids, polypeptides and complex surface (glycol)lipids, interspersed with proteins, phthiocerol-containing lipids (e.g. phthiocerol dimycocerosate), lipomannan (LM), and lipoarabinomannan (LAM) (Brennan, 2003). LM and LAM are well-known mycobacterial glycolipids with long mannose polymer skeletons. LAM in particular contains arabinose-mannose disaccharide subunit repeats, and is considered a carbohydrate structural antigen, that is associated with pathogenic functionality, crucial to *M. tuberculosis* survival within host macrophages (Knechel, 2009), and functions as a potent inhibitor of interferon gamma-mediated activation of host murine macrophages. Furthermore, LAM scavenges oxygen radicals and inhibits host protein kinase C, thereby down-regulating the host immune response against this pathogen (Todar, 2009). Phthiocerol dimycocerosate (PDIM), on the other hand, is a highly apolar component of the *M. tuberculosis* cell wall, resulting in its wax-like characteristics, and consists of a long chain β-diol (phthiocerol moiety), esterified with two mycocerosic acids. Previous studies have also linked PDIM with *M. tuberculosis* virulence (Brennan, 2003).

2.2.2 Transmission and pathogenicity

TB is considered to be highly infectious, and the *M. tuberculosis* pathogen is transmitted from an individual with an active TB disease state, through small infectious aerosol droplets (typically 1–5 µm in diameter), by means of coughing, sneezing or talking/singing. Transmission is influenced by various factors, including the number of bacilli contained in these infectious droplets, its virulence, exposure to ultra-violet (UV) light, and ventilation of the environment where infection occurs. *M. tuberculosis* usually manifests in the lungs, and this is referred to as pulmonary TB (Knechel, 2009). However, extra-pulmonary TB may also occur, in which case *M. tuberculosis* disseminates and infects other parts of the human body, causing clinical manifestations (in order of prevalence) in the pleura (pleural TB) (Porcel, 2009), lymph nodes (TB lymphadenitis) (Golden & Vikram, 2005), urinary tract and reproductive system (genitourinary TB), bones and joints (osteal/skeletal TB) (Sharma & Bhatia, 2004), central nervous system (CNS/meningeal TB) (Golden and Vikram 2005), abdominal organs (including gastrointestinal tract, peritoneum, omentum, mesentery, liver, spleen and pancreas) (peritoneal TB) (Sharma & Bhatia, 2004), and various other organs.

The pathogenicity of *M. tuberculosis* depends its ability to secrete virulence factors that are displayed on the bacterial cell surface (as previously discussed). As illustrated in Figure 2-2, the *M. tuberculosis*-containing droplets travel through the respiratory tract, where the majority of the bacilli become trapped by mucus-secreting goblet cells, that latter of which are tasked with blocking entry and/or removing foreign entities. However, this mucociliary system first-line defence can sometimes be bypassed by these droplets, allowing them to reach the upper, aerated parts of the lungs (pulmonary TB), where the bacilli may undergo rapid replication. At this point during infection, the host's second-line/innate immune mechanisms may be recruited, by which the host macrophages engulf the infecting bacilli and attempt to destroy these using various proteolytic enzymes and cytokines, including tumour necrosis factor alpha (TNF-α) and interferon gamma (IFN-γ). This signals T-lymphocyte transfer to the site of infection, initiating a cell-mediated immune response, which may either eliminate the infecting organism or result in a granuloma formation, the latter of which is a well-known pathological occurrence characterising TB (Knechel, 2009; Philips & Ernst, 2012). A granuloma contains the bacteria (known as Ghon's focus) (Behr & Waters, 2014) and is defined as an amorphous mass of macrophages, monocytes and neutrophils, located in the lung, and functions to restrict the replication and spread of the infecting mycobacteria. However, after macrophage internalisation and granuloma formation, *M. tuberculosis* may avoid death by modulating the host immune system and blocking phagolysosomal fusion, creating a hospitable niche within these phagosomes for the bacilli to persist in a non- or slowly-replicating state, where they may survive for decades (Knechel, 2009; Nunes-Alves *et al.*, 2014; Pai *et al.*, 2016; Philips & Ernst, 2012; Warner, 2014). Considering this, although the host immune response may be unable to eradicate *M. tuberculosis*, a fully immune-competent host can suppress the infection indefinitely. This asymptomatic and non-infectious state is referred to as latent TB. Current reports indicate that one-third of the global population is infected with latent TB. The majority of these individuals (±90%) never manifest any signs of disease, however, when the immune system becomes compromised, for instance during HIV co-infection, the granuloma becomes caseous, loses its rigid integrity and ruptures, releasing the bacteria, which develops into active TB (the symptomatic and highly infectious state of the disease). Although HIV co-infection is considered the primary cause of active TB, other conditions, such as malnutrition, renal failure, uncontrolled diabetes mellitus, sepsis, chemotherapy, uncontrolled alcohol use, smoking and organ transplants, may also trigger the disease conversion from latent to active TB (Behr & Waters, 2014; Fennelly & Jones-López, 2015; Knechel, 2009; Pai *et al.*, 2016).

The clinical symptoms associated with pulmonary TB are often non-specific, but may include constant coughing with mucus (lasting three weeks or longer), pleurisy, haemoptysis, dyspnoea, wheezing, weakness and/or progressive fatigue, weight-loss, no or loss of appetite, chills/fever, and night sweats (Bakhsi, 2006; Leung, 1999; Pai *et al.*, 2016). In

approximately 5% of all adult active TB cases, clinical symptoms are absent; while up to 60% of all paediatric pulmonary TB cases are asymptomatic (Pineda *et al.*, 1993).

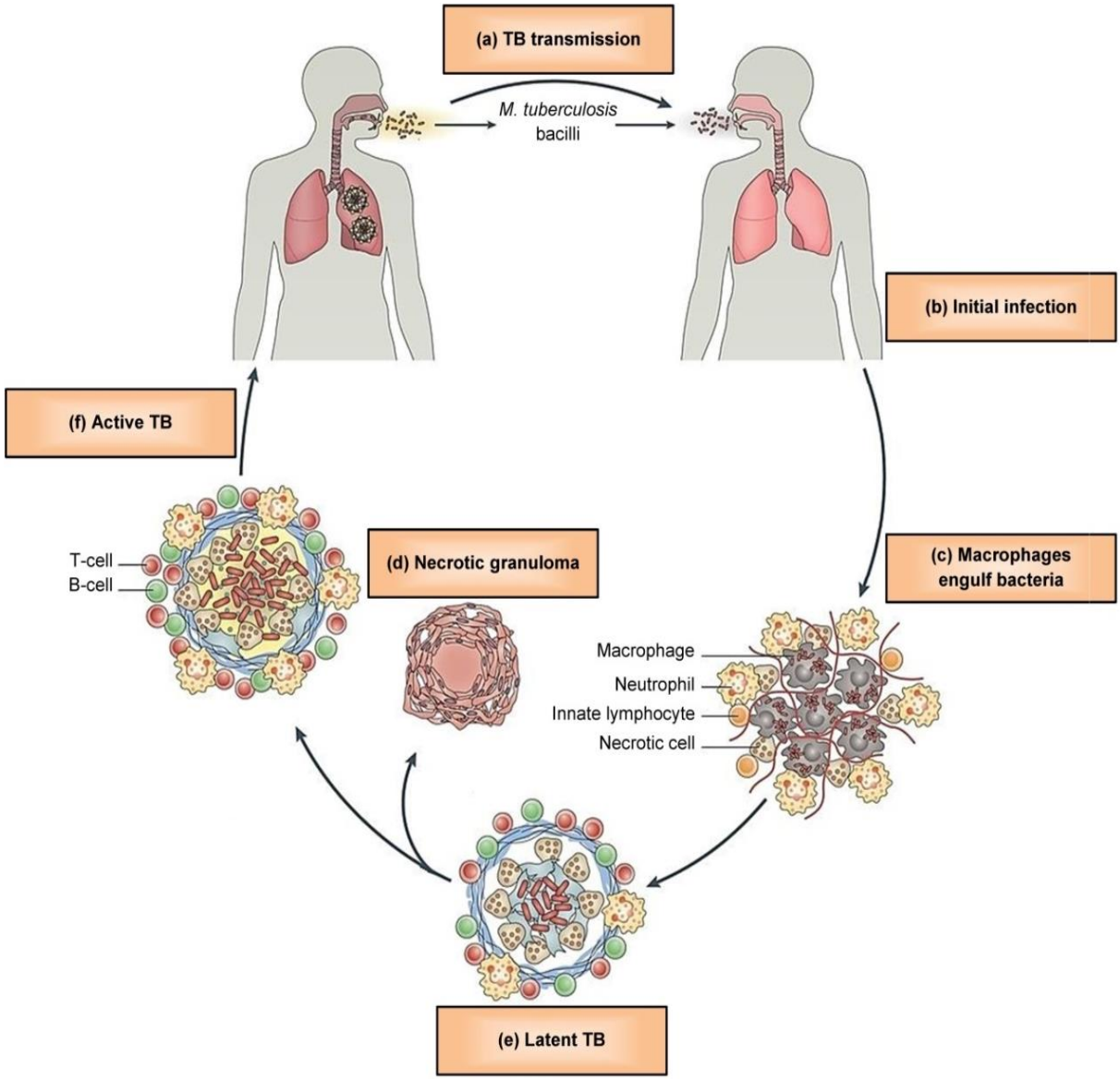


Figure 2-2: The transmission and pathogenicity of tuberculosis. (a) TB transmission occurs from host-to-host by means of coughing, sneezing or talking/singing. (b) The majority of M. tuberculosis-containing droplets are trapped and excreted by mucus-secreting goblet cells; however, some may still reach the lungs, initiating infection. (c) The invading bacteria are engulfed by host alveolar macrophages as part of the innate immune response, which results in the release of proteolytic enzymes and cytokines. (d) This in turn attracts T-lymphocytes to the site of infection, initiating cell-mediated immunity and necrotic granuloma formation, limiting the spread of the mycobacteria. (e) A fully immune-competent host can suppress the infection and keep it in its asymptomatic and non-infectious state (latent TB). (f) However, when the immune system becomes compromised, the granuloma becomes caseous, loses its rigid integrity and ruptures, releasing the bacteria, which develops into active disease. Adapted from Nunes-Alves et al. (2014).

2.3 TUBERCULOSIS DIAGNOSTICS

A definitive TB diagnostic test functions on the basis of directly detecting *M. tuberculosis*, or one or more specific biomarkers of the organism present in a diagnostic specimen (e.g. blood, urine or sputum) (Knechel, 2009). Most tests and techniques used for diagnosing TB are surprisingly inexpensive, however, they vary largely with regards to their speed, sensitivity and specificity (Frieden & Driver, 2003).

2.3.1 Diagnosing latent *M. tuberculosis* infection

The current diagnosis of latent TB relies primarily on the detection of the host immune response to the *M. tuberculosis* infection. Two methods based on this principle, the tuberculin skin test (TST) and the interferon-gamma release assay (IGRA), are currently the only methods recommended by the World Health Organisation (WHO) for the detection of *M. tuberculosis* infection (World Health Organization, 2006). Although easy to perform, these tests have a number of limitations, since false-positive results can occur in individuals who were previously vaccinated or previously infected with *M. tuberculosis*, and false-negative results are common in patients with a compromised immune system, such as that caused by HIV (Ferrara *et al.*, 2006; Pai *et al.*, 2006a; Pouchot *et al.*, 1997).

In the light of the current investigation, when used for predicting treatment outcome or monitoring treatment progression, IGRAs were shown to be the more accurate indicator of progression to active TB over time compared to the TST (Kik *et al.*, 2010). However, it is important to note that individuals who are able to eliminate the infection may still have a positive TST or IGRA result due to memory T-lymphocyte responses, hence these tests have a low predictive/prognostic value (Pai *et al.*, 2016).

2.3.1.1 Tuberculin skin test

The tuberculin skin test (TST) uses purified protein derivative (PPD), also known as tuberculin, a combination of proteins obtained from heat-killed *M. tuberculosis*, *M. bovis* and other environmental mycobacteria, which reduces the specificity of the TST. This test is based on an acquired, delayed-type hypersensitivity reaction in *M. tuberculosis*-infected individuals, when PPD is intra-cutaneously injected into the ventral forearm. T-cells that are produced due to a *M. tuberculosis* infection migrate to the site of injection where they release lymphokines, resulting in a thickening of the skin, which can be measured by a trained health-care professional, 48–72 hours after injection (González-Martín *et al.*, 2010; Olivier & Loots, 2011).

The TST, however, cannot distinguish between latent or active infection, and will result in a positive diagnostic outcome in both instances. Another disadvantage is that false-positive results are also prevalent in individuals who previously received the Bacille Calmette-Guerin (BCG) vaccination, those who are infected with different *Mycobacterium* species other than *M. tuberculosis*, and those individuals who have previously had TB. On the other hand, false-negative results may also occur in immune-compromised individuals (i.e. patients with HIV, those who recently had an organ transplant, or those treated with corticosteroids or chemotherapy), or due to other bacterial, fungal or viral infections, chronic renal failure, severe malnutrition, lymphoid organ disease, and in infants and the elderly. Other limitations include administration/technical mistakes, reading errors (which is subjective to interpretation), the need for a second visit to read the test results, and a lack of privacy. Thus, the reported sensitivity and specificity of this technique varies (González-Martín *et al.*, 2010; Mazurek *et al.*, 2001; Olivier & Loots, 2011).

2.3.1.2 Interferon gamma release assays

Another approach for detecting latent TB is the whole-blood cytokine detection assay. This test is based on the *in vitro* detection of a cell-mediated host immune reaction, resulting in the release of cytokines (IFN-γ in particular), in response to the antigens presented during *M. tuberculosis* infection (Mazurek *et al.*, 2001). Early IGRAs used PPD as the preferred antigen, however, these were replaced by those using alternative antigens more specific to *M. tuberculosis*, such as early secreted antigenic target (ESAT)-6, culture filtrate protein (CFP)-10 and TB7.7 (Rv2654). These tests are based on the principle of an enzyme-linked immunosorbent assay (ELISA) to assess an individual's acquired immunological response to ESAT-6 and CFP-10, based on the amount of IFN-γ released. Currently, there are two commercially available IGRAs named Quanti-FERON®-TB Gold In-Tube, using an ELISA-based method, and the T-SPOT-TB, based on the ELISPOT technique (González-Martín *et al.*, 2010; Mazurek *et al.*, 2001; Olivier & Loots, 2011).

A number of comparative studies show that IGRAs have a high specificity (>95%), but a reduced sensitivity (75–97%) when compared to the TST. However, IGRAs are significantly more specific in the vaccinated population when compared to that of TST. Furthermore, IGRAs have additional operational advantages, including (a) an ability to generate a diagnostic result immediately (point-of-care), (b) it being a single-step procedure, (c) ease of standardisation and implementation, as well as (d) various technical, logistic, and other cost advantages (González-Martín *et al.*, 2010; Mazurek *et al.*, 2001). Despite these however, this approach is not without its disadvantages, which includes follow-up testing in order to determine if the patient does in fact have active TB, which is also the case for TST (Mazurek *et al.*, 2001).

2.3.2 Diagnosing active *M. tuberculosis* infection

2.3.2.1 Microscopy smear techniques

Smear microscopy, as first demonstrated by Robert Koch in 1882, is currently the most commonly used method for diagnosing active TB. This technique is based on the acid-fast staining of mycobacteria after treatment with an acid-alcohol solution. *M. tuberculosis* is classified as a gram positive, acid-fast bacterium, a characteristic attributed to its high cell wall lipid content. The presence of mycolic acids in the cell wall prevents the binding of conventional dyes, and hence it is necessary to drive dyes into the cells using heat and phenol (known as the Ziehl-Neelsen method). During sputum smear microscopy, sputum is smeared onto a slide, stained with dye, dried using heat, and treated with acid-alcohol, after which all the acid-fast bacteria present will colour red and be visible under a microscope (Knechel, 2009; Konstantinos, 2010; Willey, 2008). Despite the fact that it is a low-cost, fast (<2 hours) and simple to use method, it requires a high number of bacilli (5 000–10 000 bacteria/mL sample) before a definitive diagnosis can be made, resulting in a sensitivity of only 62%. A further disadvantage is that it cannot distinguish between various *Mycobacterium* species, nor can it detect if they are drug-resistant (Dhingra *et al.*, 2003).

Apart from its role in diagnosing TB, this technique can be also used for monitoring treatment efficacy and progression, since the bacteria should visibly decrease with each smear done on successively collected follow-up samples, during the treatment duration, if treatment is successful. To date, several studies have reported that a positive sputum microscopy during the second month of treatment is an indicator of treatment failure; however, conversion may take some time, depending on the initial bacterial load (Bernabe-Ortiz *et al.*, 2011; González-Martín *et al.*, 2010).

2.3.2.2 Bacteriological cultures

Bacteriological culture is based on the observation of growth in culture media of *M. tuberculosis* harvested from patient sputum. It is considered the gold standard for TB diagnostics since it has a reported sensitivity and specificity of almost 100%, and only requires as little as 10–100 bacteria/mL sample. This technique can be used to detect drug-resistance when antibiotics are added to the culture media, however, this method is time-consuming, considering the slow growth rates of mycobacteria, thereby delaying treatment onset (Moore & Curry, 1995). Either solid or liquid media are used to grow cultures, each of which has its own advantages and disadvantages. Solid media can be purchased at lower costs, but requires 2–6 weeks of incubation before macroscopically visible growth can be seen. Although liquid culture media methods are more costly, they are considerably faster

(2–4 weeks before growth is visible) and more sensitive. A disadvantage of this method however, is the false-positive culture results which may occur due to bacilli transfer from TB-positive to TB-negative samples during laboratory handling, despite rigorous anti-contamination procedures (Olivier & Loots, 2011).

Currently, although the radiometric BACTEC 460TB system remains the preferred method for diagnosing TB (Cruciani *et al.*, 2004), the non-radiometric, fully automated BACTEC 960/MGIT system is also widely used (Trivedi, 2015). The BACTEC 460TB and BACTEC 960/MGIT systems have a reported sensitivity and specificity for detecting mycobacteria of 81.5% and 99.6% and 85.8% and 99.9%, respectively. Additionally, when combined with solid media (such as the traditional Löwenstein–Jensen medium), the sensitivities of these two systems increase to approximately 87.7% and 89.7%, respectively (Cruciani *et al.*, 2004; Lee *et al.*, 2010; Zhao *et al.*, 2014). Furthermore, the above-mentioned systems can not only distinguish between various *Mycobacterium* species (Olivier & Loots, 2011), but also detect drug resistance. Numerous studies investigated the accuracy of these two methods when testing susceptibility to first-line anti-TB drugs, and reported a level of agreement between results of approximately 97.2% with concordance values for isoniazid, rifampicin and ethambutol of 96.3%, 98.8% and 98.8%, respectively (Garrigó *et al.*, 2007; Piersimoni *et al.*, 2013).

Bacteriological cultures are also the preferred method for following treatment efficacy and progression. For this purpose, patient samples are collected bimonthly at every follow-up visit, at 15 days, and again at 30 days after treatment onset, provided that a sputum sample can be obtained from the lower respiratory tract. A favourable treatment outcome may be predicted by a negative *M. tuberculosis* culture within two months of commencing treatment (Antoine *et al.*, 2007; González-Martín *et al.*, 2010).

2.3.2.3 Nucleic acid amplification techniques

Molecular methods, such as nucleic acid amplification (NAA), also referred to as direct amplification, are based on the amplification of nucleic acid regions specific to the *M. tuberculosis* complex, and can hence distinguish between *M. tuberculosis* and various other *Mycobacterium* species. This approach is used as a follow-up to confirm sputum smear positive or culture results, and shows a 100% specificity and sensitivity compared to smear microscopy. Although this method has the capacity to outperform the current gold standard with regards to diagnostic turnover time and improve diagnostic certainty, it is far less sensitive (Knechel, 2009; Konstantinos, 2010; Pai *et al.*, 2003) when compared to bacteriological cultures.

The GeneXpert MTB/RIF (Xpert) assay is currently the fastest NAA TB diagnostic technique commercially available, reportedly able to obtain a diagnostic result within 90 minutes. It has been developed and packaged into a nearly fully automated cartridge-based diagnostic test, able to simultaneously identify *M. tuberculosis* and detect rifampicin-resistance. The Xpert assay purifies and concentrates *M. tuberculosis* bacilli from collected patient samples and utilises a hemi-nested real-time polymerase chain reaction (PCR) to detect and amplify clinically relevant *M. tuberculosis*-specific deoxyribonucleic acid (DNA) sequences, as well as mutations on the *rpoB* gene conferring rifampicin-resistance. This method has a reported sensitivity of 77.3% in smear-negative TB cases, and 98.2% in smear-positive TB cases (Boehme *et al.*, 2011; Hillemann *et al.*, 2011). Disadvantages, however, include various operational issues, the high cost and requirement for a stable flow of electricity, annual calibration, low limits for detecting rifampicin-resistance, and that the operating and storage temperatures may not exceed 30°C (Boehme *et al.*, 2011; Small & Pai, 2010).

Since this assay is able to quantify bacterial load, it can potentially predict treatment response. However, the Centres for Disease Control and Prevention (CDC) does not recommend the use of NAA tests such as the Xpert assay to monitor treatment response, since dead bacilli persist in the host for quite some time after, making it impossible for this assay to distinguish between the active disease and successfully cured disease (Blakemore *et al.*, 2011; Pai *et al.*, 2003; Shenai *et al.*, 2016).

2.3.2.4 Serological (immunological) methods

Serological (or immunological) methods function by detecting host antibodies and immune complexes produced in response to *M. tuberculosis* antigens. Despite a lack of accuracy and the inadequate sensitivity and specificity, this method remains popular in developing countries, mainly due to its speed, technological simplicity, low costs and modest training requirements (Steingart *et al.*, 2011). However, using serological methods for diagnosing TB is challenging due to the various stages associated with *M. tuberculosis* infection (exposure, latent/active infection and severe disease), each having their own antibody patterns. Additionally, cross-contamination from environmental mycobacteria can cause false-positive results and serological tests are unable to distinguish between latent and active TB or between different *Mycobacterium* species (Olivier & Loots, 2011; Pai *et al.*, 2006b).

Considering all the disadvantages and limitations of serological methods, the WHO strongly recommends not using these methods to diagnose TB. Also considering that there is no evidence that serological methods can be used to monitor response to treatment, these tests are not recommended for use to monitor treatment progression (World Health Organization, 2011).

2.3.2.5 Phage-based assay

The phage-based assay allows for the rapid detection of TB in individuals with no previous treatment history, giving results within 48 hours. During this technique, clinical samples are infected with mycobacteriophages (viruses infecting mycobacteria), after which a solution is added to destroy all phages that have not infected the bacteria. The remaining phages replicate, are released during cell lysis, and amplified. The phages can then be visualised as clear areas in a lawn of host cells, which indicate the presence of *M. tuberculosis*. The FASTPlaqueTB assay is currently the only commercially available phage assay for detecting TB, and has a reported sensitivity of 75% and specificity of 98% when compared to smears, and 97% specificity when compared to cultures. Considering this, as well as its speed and specificity for *M. tuberculosis*, this test could possibly show promise for use in countries with a high TB prevalence, especially when other diagnostic tests are not suitable (Albert *et al.*, 2002; Tortoli & Palomino, 2007).

Considering its mode of action, the phage-based assay could potentially be used as a fairly handy tool to follow treatment progression, especially in combination with other techniques. However, a major disadvantage to this is that this test requires approximately a five times larger volume of patient sample as compared to that required for direct microscopy and culture methods (Tortoli & Palomino, 2007).

2.3.3 Metabolomics biomarkers for tuberculosis diagnostics

Considering the above-mentioned advantages and disadvantages of current TB diagnostic techniques, no single test meets all the specifications of sensitivity, specificity, speed, safety, robustness, training simplicity, and cost. Thus, there is a need for new TB diagnostic techniques, which will overcome these limitations and can easily be implemented in low-income, high-burden counties. One of the first steps towards achieving this goal is to address this problem from a different perspective and to identify new TB biomarkers, which will lead to the development of improved diagnostic strategies.

Although numerous single biomarkers have been implemented successfully for use in the diagnosis and prognosis of various disease states, the emergence of the more recent “omics” technologies has unleashed the possibility of using biosignatures, which are profiles of combined biomarkers, towards this end. These biosignatures are especially useful when individual biomarkers do not exist or cannot be identified. The term “omics” is used to describe research models aimed at acquiring large-scale data from each sample in a sample group, for the purpose of identifying disease biomarkers and/or elucidating novel functional or pathological mechanisms (Wheelock *et al.*, 2013). These datasets can comprise of a

collection of genes (genomics), products of gene expression (transcriptomics), proteins (proteomics), or metabolites (metabolomics). Due to the largely untargeted means by which these datasets are usually generated, they may contain hundreds or even thousands of variables, and therefore, biomarkers and/or biosignatures are mined from the data using advanced biostatistical approaches. The validation and clinical performance of these biosignatures are evaluated by reporting the sensitivity, specificity and receiver operator characteristic (ROC) curves, amongst others, according to the Standards for the Reporting of Diagnostic Accuracy Studies (STARD) guidelines (Bossuyt *et al.*, 2003; Xia *et al.*, 2013).

To date, several metabolomics studies have been done with the aim of identifying specific single metabolite biomarkers and/or biosignatures which can be used in either the initial screening or subsequent speciation phase of TB diagnostics. As indicated in Tables 2-2 to 2-6, these approaches include the use of various sample matrices which were analysed using a variety of analytical approaches. The identified biomarkers not only include the well-known TB biomarkers; tuberculostearic acid (TBSA), branched chain fatty acids, and other cell wall components, but also novel compounds, never before associated with TB infection. Some of the compounds identified as potential biomarkers were detected in more than one study and in different sample matrices (indicated in bold text in Tables 2-2 to 2-6), the most significant of which will be discussed below.

2.3.3.1 Metabolomics biomarkers detected using *Mycobacterium* cultures

Although the metabolite profiles of the various TB-causing mycobacteria grown *in vitro* differs to some extent from that grown *in vivo*, some researchers still prefer to use bacteriological cultures in the initial stages of diagnostic biomarker identification. Biomarkers identified in this manner are indicative of those organism-specific compounds which may be present in more complex patient-collected diagnostic samples, such as sputum or blood, and this pre-identification may assist in their detection in these patient samples where they may occur in diluted quantities and/or be masked by the matrix background.

Pavlou *et al.* (2004) used a metabolomics approach to distinguish between *M. tuberculosis*, *M. avium*, *P. aeruginosa* and a mixed infection (*M. tuberculosis* and *M. scrofulaceum*). They achieved a 96% prediction value when analysing the headspace above the cultured samples of the aforementioned infectious *Mycobacterium* species, using an electronic nose (EN). The underlying principle of the method is as follows: the EN contains 14 conducting-polymer sensor arrays which interact with the volatile compounds liberated in the headspace of the cultured *Mycobacterium* samples. This interaction results in a change in electrical resistance, producing characteristic signals of multiple sensor parameters. These variables collected for each culture were then used to construct a data matrix which was statistically

analysed via discriminant function analysis (DFA) to differentiate the sample groups. When this methodology was applied to 46 patient sputum samples, it successfully differentiated sputum containing *M. tuberculosis*, *M. avium*, *P. aeruginosa*, a mixed infection and eight control samples (TB-negative patients admitted to the clinic) (Pavlou *et al.*, 2004). The method was then further validated using 330 TB-positive and TB-negative patient collected sputum samples, and correctly identified TB with a sensitivity of 89% and specificity of 91%, at a detection limit of 1×10^4 bacteria/mL sputum (Fend *et al.*, 2006). When investigating the potential clinical application of this EN methodology, two different off-the-shelf EN devices were tested in a real-life clinical setting. The first device could diagnose TB with a sensitivity of 68%, specificity of 69% and accuracy of 69%, while the second showed a sensitivity of 75%, specificity of 67% and accuracy of 69%. Although these results show promise, they do suggest that the EN technology is not yet sensitive, specific or accurate enough for clinical application towards TB diagnostics (Kolk *et al.*, 2010). Furthermore, the chemical characteristics and identities of the signals detected cannot be characterised or quantified as yet and, therefore, the characteristic compounds causing these signals were not identified.

Mycobacterial cell wall components, specifically the characteristic mycolic acids, potentially serve as good indicators for the presence of this pathogen in a biological sample. Numerous studies have focused on the presence or absence, as well as the ratios, of these compounds, as a means to characterise, identify or differentiate the various infectious *Mycobacterium* species from one another, for application to TB diagnostics (Butler *et al.*, 1991; Freeman *et al.*, 1994; Glickman *et al.*, 1994). To date, the Sherlock™ Mycobacteria ID system (MYCO-LCS) is the only commercially available high-performance liquid chromatography (HPLC) based, pattern recognition method for mycobacterial speciation, and functions by identifying variations in the infecting mycobacteria's mycolic acid content. This method claims to have the ability to identify and differentiate 25 mycobacteria species and 35 non-mycobacteria species from pure culture, in less than 30 minutes, at a cost of approximately $6 USD per analysis (Midi Inc, 2015). In 2012, Olivier and Loots applied a similar approach, identifying the characteristic lipid profiles of *M. tuberculosis*, *M. bovis*, *M. kansasii* and *M. avium,* using a GC-MS metabolomics research methodology, to fully automate this classification model. Twelve lipid biomarkers were identified (Table 2-2) and used to build a multivariate discriminant model, which could correctly assign unknown samples to their respective species groups with probabilities ranging from 72 to 100%. Although this method requires 16 hours to achieve a diagnostic result, only 1×10^3 cultured cells were needed for correct classification, making it more sensitive than the previously described HPLC approach (Olivier & Loots, 2012b). The research team however, later improved on this method by implementing the use of a mixer mill to assist in the extraction of these biomarkers, resulting in a five hour turnaround time and the need for smaller patient sample and solvent volumes. A further development was the application of an alternative

extraction approach, allowing for the detection of metabolites belonging to not only the lipid markers described, but also other compounds including amino acids, alcohols, organic acids, monosaccharides, alkenes, alkanes, purines, pyrimidines, etc. Nineteen metabolites were identified as biomarkers for differentiating *M. tuberculosis*, *M. avium*, *M. bovis*, *M. kansasii* and *P. aeruginosa* (Table 2-2). Indole-acetic acid, cadaverine, purine, putrescine, and two unknown compounds with molar masses of 343 and 373 were detected exclusively in *P. aeruginosa*, whereas inositol and myo-inositol were characteristic of the *Mycobacterium* species investigated. Succinic acid and an unknown compound with a mass of 268 were uniquely identified in *M. kansasii*, and another unknown compound with a mass of 541 was detected exclusively in *M. tuberculosis* (Olivier & Loots, 2012a). In addition to the capacity of the extraction method to identify a greater variety of metabolites, it further reduced the solvent volumes required, and improved the detection limit to only 250 bacterial cells per analysis (Du Preez & Loots, 2013a).

More recently, Lau *et al.* (2015) described an optimised ultra-high performance liquid chromatography–electrospray ionisation–quadruple time-of-flight mass spectrometry (UHPLC-ESI-Q-TOFMS) metabolomics approach to differentiate various *Mycobacterium* species from one another, by analysing their excreted metabolite profiles. Although the main aim of this investigation was to differentiate *M. tuberculosis* and non-tuberculous mycobacteria (NTM), clear separation between all species was achieved when the collected data were analysed using PCA and PLS-DA. Of the 24 biomarkers identified (metabolites detected in significantly higher concentrations in *M. tuberculosis* as compared to the other NTM strains), only seven could be annotated via MS/MS database matching (Table 2-2) (Lau *et al.*, 2015).

Furthermore, in the context of the rising incidence of drug-resistance, a successful TB diagnostic method should also be able to identify whether the infected *M. tuberculosis* strain is fully drug-susceptible or drug-resistant to either first- or second-line drugs. Although no metabolomics study to date has identified biomarkers for TB drug-susceptibility testing per se, differentiation of various drug-resistant strains from drug-susceptible parent strains has been demonstrated using the differences in the metabolite profiles of these organisms. In two related studies, GC-MS metabolomics were used to compare the fatty acid metabolome (Du Preez & Loots, 2012) and total metabolome (Loots, 2015) profiles of two *rpoB* mutant *M. tuberculosis* strains (S522L and S531L) to that of a fully drug-susceptible wild-type parent strain, to better characterise rifampicin-resistance. In each study, biomarkers for rifampicin-resistance were identified and used to elucidate previously unknown biological mechanisms related to the survival and adaptation of these mutant strains. Contributing to this, Lahiri *et al.* (2016) used a LC-MS based organism-wide screen to compare approximately 10 000 cell wall lipids occurring in rifampicin-resistant mutants, in order to identify the lipid alterations

induced by various *rpoB* mutations. Their results indicated altered mycobactin siderophores and acylated sulfoglycolipid concentrations, both of which are essential for bacterial growth. This study provided further evidence of characteristic cell wall lipid remodelling in rifampicin-resistant *M. tuberculosis* strains, and additionally suggests new diagnostic and therapeutic targets against drug-resistant strains of *M. tuberculosis* (Lahiri *et al.*, 2016). In a similar fashion, Loots (2014) investigated mono-resistance to isoniazid (resulting from mutations in the *katG* gene) in *M. tuberculosis*. From the differentiating metabolites identified, it was shown that the isoniazid-resistant strains are more susceptible to oxidative stress, and a survival adaptation to this is to increase the uptake and utilisation of alkanes and fatty acids as a carbon/energy source, and to synthesise antioxidant compounds, i.e. ascorbic acid, and its oxidation via an ascorbate degradation pathway (Loots, 2014). Although the diagnostic potential of the identified biomarkers were not investigated, these studies show that metabolomics can be implemented towards the identification of characteristic drug-resistant biomarkers.

Although the methods described above all required a time-consuming culturing step, these attempts at improved TB diagnostics using bacterial cultures prove the capacity of metabolomics for detecting new diagnostic biomarkers. Ideally, a TB diagnostic method, including the initial screening and speciation steps, should be done directly from the collected clinical samples (preferably collected in a non-invasive manner) to shorten the delay between disease presentation and treatment onset. The true clinical application of these biomarkers can therefore only be determined when investigating their diagnostic value in patient-collected samples, as was done by Syhre *et al.* (2009) and Phillips *et al.* (2010) (see section 2.3.3.5).

Table 2-2: Tuberculosis biomarkers identified in culture using a metabolomics research approach.

Sample cohort	Analytical apparatus	Biomarkers identified	Reference
M. tuberculosis	ATD-GC-MS	**1-methyl-naphthalene**; 3-heptanone; methylcyclododecane; 2,2,4,6,6-pentamethyl-heptane; trans-anti-1-methyldecahydronaphthalene; 2-hexene; 1-methyl-4-(1-ethylethyl)-benzene; **1,4-dimethyl-cyclohexane**; 3,5-dimethylamphetamine; 3-methyl-butanal	Phillips *et al.* (2007)
M. tuberculosis *M. bovis* *M. bovis BCG* *M. avium complex* *M. fortuitum* *M. chelonae* *M. abscessus*	SPME GC-MS	methyl phenylacetate; methyl p-anisate; **methyl nicotinate**; o-phenylanisole	Syhre and Chambers (2008)

(Continues on next page)

(Continues from previous page)

Sample cohort	Analytical apparatus	Biomarkers identified	Reference
M. tuberculosis *M. avium* *M. bovis* *M. kansasii*	GC-MS	isopropyl-tetradecanoic acid; **9-hexadecenoic acid**; **TBSA**; **10-heptadecenoic acid**; **heptadecanoic** acid; **11-eicosanoic acid; 13-docosenoic acid**; tricosane; C32 mycocerosic acid; **15-tetracosenoic acid**	Olivier and Loots (2012a)
M. tuberculosis *M. avium* *M. bovis* *M. kansasii* *P. aeruginosa*	GC-MS	**TBSA**; **docosanoic acid**; 2-octyl-cyclopropaneoctanoic acid; **heptadecanoic** acid; **13-docosenoic acid**; **15-tetracosenoic acid**; 2,4-dimethyltetradecanoic acid; **9-hexadecenoic acid**; **eicosanoic acid**; tetradecanoic acid; octacosanoic acid; pentacosanoic acid; 2-hexyl-cyclopropaneoctanoic acid; purine; **9-octadecenoic acid**; indole-acetic acid; **11-eicosanoic acid**; citric acid; **cadaverine**; inositol; 5'-adenylic acid; decanoic acid; 9-hexadecenoic acid (trans); **myo-inositol**; succinic acid; putrescine; erythritol; valeric acid	Olivier and Loots (2012b)
M. tuberculosis *M. avium complex* *M. bovis (BCG)* *M. chelonae* *M. fortuitum* *M. kansasii*	UHPLC-ESI-QTOFMS	dexpanthenol; Val-His-Glu-His; 1-tuberculosinyladenosine; 1-tuberculosinyl2'-deoxyadenosine; 1-tuberculosinyl derivative; 1-tuberculosinyl-O-acetyladenosine; phosphatidylglycerol	Lau *et al.* (2015)

Note: Compounds in bold have been identified as potential TB biomarkers in more than one metabolomics study.
Abbreviations: GC-MS, gas chromatography mass spectrometry; ATD, automated thermal desorption; SPME, Solid phase micro-extraction; ESI, electrospray ionisation; TOFMS, time-of-flight mass spectrometry; UHPLC-MS/MS, ultra-high performance liquid chromatography-tandem mass spectrometry.

2.3.3.2 Metabolomics biomarkers detected using sputum

The most commonly used sample matrix for diagnosing pulmonary TB to date is sputum. This mucus-like biofluid originates directly from the area of infection, i.e. the airways of the lungs, and is highly populated with *M. tuberculosis* in TB-positive patients. A metabolic profile of a TB patient's sputum would contain *Mycobacterium*-specific metabolites, due to the physical presence of the organism in this sample, and also various altered, disease-induced host metabolites, thereby expanding the possibility of identifying diagnostic biomarkers. In addition to diagnostics, these biomarkers could also be implemented towards the elucidation of disease mechanisms, seeing that they are a true reflection of adaptations to the metabolome of *M. tuberculosis* due to *in vivo* growth and the host response to infection/disease. Despite this, however, very few metabolomics studies have been focused on biomarker identification for improved TB diagnostics using patient-collected sputum. The reason for this is most likely linked to the complexities in using this sample matrix, such as its viscosity and uneven consistency, and also the likelihood of possible TB infection when handling these samples in a standard analytical laboratory not equipped for these purposes. The aforementioned viscosity and uneven consistency of sputum bring about the need for additional, time-consuming sample pre-processing steps before metabolomics analyses can

commence. These methods were not always standardised, and even after successfully applying the necessary extraction procedures and analyses of the samples, the metabolic profiles obtained are complex and contain many compounds from a variety of compound classes. Hence, comprehensive statistical analyses to sift through the masses of data are required to identify potential biomarkers (Du Preez & Loots, 2013b).

In 2012, specifically for metabolomics TB biomarker identification, Schoeman *et al.* developed a new sputum pre-processing method, in conjunction with a global metabolite extraction approach and GCxGC-TOFMS analysis, to differentiate sputum spiked with *M. tuberculosis* bacilli and control sputum (not spiked). The compounds identified as markers mainly represented those associated with the cell wall of *M. tuberculosis* (Table 2-3) (Schoeman *et al.*, 2012). This method was applied to differentiate culture-confirmed TB-positive (n=34) and TB-negative (n=61) patient sputum, and the 22 compounds (Table 2-3) best explaining the variation between the groups, were once again identified. Various new disease mechanisms were described, including the presence of a citramalate cycle in *M. tuberculosis*, and the interaction of this cycle with an upregulated glyoxylate cycle and increased fatty acid oxidation during *in vivo* growth in the human host. Furthermore, these biomarkers also shed light on an additional process by which the host produces hydrogen peroxide (H_2O_2) via glucose oxidation, as a means to eliminate the infected bacteria. The elevated concentrations of various neurotransmitters associated with the TB-infection provided added information explaining many of the clinical symptoms that TB patients experience (Du Preez & Loots, 2013b). Although the objective of this study was to better characterise TB, the differentiation of the groups based on underlying metabolite differences suggest the potential for this type of metabolomics approach to be used for diagnostic purposes.

Although not using a traditional metabolomics approach, research groups have also investigated the possible use of individual sputum metabolites, such as the aforementioned TBSA, for diagnosing TB. In 2009, Cha *et al.* indicated that TBSA can be detected directly from patient sputum samples using solid-phase micro-extraction (SPME) and post-derivatisation coupled to GC-MS. The group reported that this procedure is more sensitive than smear microscopy and requires only five hours per analysis, thereby proving that the use of single metabolite biomarkers can be useful in the initial screening phase of TB diagnostics. However, similar to a positive smear microscopy result, the detection of TBSA in these samples is due to the presence of the mycolic acids, a common component in the cell walls of all mycobacteria, and can therefore not be used for speciation or the detection of drug-resistance (Cha *et al.*, 2009).

Table 2-3: Tuberculosis biomarkers identified in human sputum using a metabolomics research approach.

Sample cohort	Analytical apparatus	Biomarkers identified	Reference
6 x TB- spiked with *M. tuberculosis* 6 x TB- (Not spiked)	GCxGC-TOFMS	**nonadecanoic acid**; **TBSA**; hexacosanoic acid; acetohydroxamic acid; **myo-inositol**; **eicosanoic acid á-D-glucopyranoside**; D-glycero-L-manno-heptonic acid; palmitoleic acid; furan; tetracosanoic acid; arachidonic acid; octadecanoic acid; **D-glucose**; propane; **heptadecanoic acid**; **docosanoic acid**; **9-octadecenoic acid**; **D-mannose**; D-galactose; **D-glucosamine**; glycerol; **uridine**; D-fructose; 2-O-glycerol-à-D-galactopyranoside; **cadaverine**; á-D-galactopyranoside; á-D-xylopyranose; arabinofuranose; D-erythro-pentitol; **L-threonine**; á-D-galactofuranose; phenylethanolamine	Schoeman *et al.* (2012)
61 x TB+ 34 x TB-	GCxGC-TOFMS	à-D-glucopyranose-2-acetylamino-2-deoxy; **à-D-glucopyranoside**; á-L-mannopyranose; N-acetyl-glucosamine; à-D-mannopyranoside; D-galactose-6-deoxy; à-D-galactopyranose; **D-glucosamine**; D-citramalic acid; methyl-17-methyl-octadecanoic acid; D-gluconic acid; d-lactone; **10-heptadecenoic acid**; **9-octadecenoic acid**; **nonadecanoic acid**; 2-deoxy-D-erythro-pentitol; glutaric acid; sebacic acid; ethane; butanal; g-aminobutyric acid; 3,4-dihydroxybutanoic acid; normetanephrine	Du Preez and Loots (2013b)

Note: Compounds in bold have been identified as potential TB biomarkers in more than one metabolomics study.
Abbreviations: GC-MS, gas chromatography mass spectrometry; TOFMS, time-of-flight mass spectrometry; TB+, positively diagnosed for active TB via culture; TB-, patients portraying TB-related symptoms, but tested negative for TB via culture.

2.3.3.3 Metabolomics biomarkers detected using blood and tissue

Blood samples (plasma and serum) are regarded as being comparatively more homogeneous in composition, less viscose and easier to process than sputum. Furthermore, although very low concentrations of potential metabolites originating directly from the infectious bacteria may be detected in the blood, it reflects those metabolites altered in the host due to the infection, and is therefore better suited to identify host pathological and defence mechanisms (Weiner *et al.*, 2012), to examine disease progression (Somashekar *et al.*, 2012), and/or for diagnostic and therapeutic monitoring purposes (Zhou *et al.*, 2013).

When analysing serum samples collected from healthy controls, individuals with latent *M. tuberculosis* infection, and patients with active TB, using GC-MS and UHPLC-MS/MS, Weiner *et al.* (2012) were able to detect alterations in amino acids, lipids and nucleotides, due to the anti-inflammatory changes that typically occur during TB disease progression. The altered metabolite profiles also revealed reduced phospholipase activity, increased indoleamine 2,3-dioxygenase 1 (IDO1) activity, an increased abundance of adenosine metabolic products, and indicators of fibrotic lesions in those patients with active TB, comparatively. Additionally, 20 metabolites with potential diagnostic value were identified

(Table 2-4) (Weiner *et al.*, 2012). Similarly, using serum-based NMR metabolomics, Zhou *et al.* (2013) identified 17 metabolites with possible diagnostic value (Table 2-4), in addition to better describing the TB disease state in terms of an observed increased glycolysis, lipid degradation, nucleotide biosynthesis, energy consumption, and a modified protein metabolism in TB patients. Expanding on this, they also investigated the metabolite changes induced by other disease states, including diabetes, malignancy and community-acquired pneumonia, and subsequently indicated that each of the changes to the metabolome induced by these perturbations is, in fact, disease specific (Zhou *et al.*, 2013). With the main aim of identifying TB diagnostic biomarkers, Feng *et al.* (2015) also included patients with diseases other than TB in their patient cohort. The serum metabolome profiles of patients with active TB, chronic obstructive pulmonary disease (COPD), pulmonitis, bronchiectasis, lung cancer and healthy controls, were compared using UHPLC-MS analyses. By employing OPLS-DA, they were able to differentiate TB patients from healthy controls and patients with other lung conditions, and identified a TB-specific serum biosignature. A set of 12 metabolites, mostly fatty acids, amino acids and lipids (Table 2-4), were identified as biomarkers for active TB, and a combination of four of these compounds (i.e. lysophosphatidylcholine (18:0), behenic acid, threoninyl-γ-glutamate and presqualene diphosphate) allowed for discrimination between active TB and control samples with an area under the curve (AUC) value of 0.99 (Feng *et al.*, 2015).

With the aim of explaining pathogen-induced changes in the host, in addition to using serum, Shin *et al.* (2011) compared the spleen, lung, and liver tissue metabolome data from *M. tuberculosis*-infected and healthy control mice using ^{1}H NMR. Based on the metabolite profiles, clear differentiation between TB-positive and control mice were evident for all tissue and serum samples collected, and the most characteristic biomarkers were identified (Table 2-4). These compounds indicated that precursors of membrane phospholipids, i.e. phosphocholine and phosphoethanolamine, as well as the oxidative stress response, glycolysis, amino acid metabolism and nucleotide metabolism, are altered due to the TB disease state (Shin *et al.*, 2011). Similarly, Somashekar *et al.* (2012) applied a ^{1}H NMR metabolomics approach using lung and serum samples collected from guinea pigs, 30 and 60 days after infection, with the aim of studying the effects of disease progression on the host metabolome. Sixteen lung metabolites were identified as TB biomarkers (Table 2-4), four of which (aspartate, glutathione, betaine and trimethylamine N-oxide) were uniquely detected in the active TB group as a result of increased oxidative stress and glutaminolysis in the lung lesions. Furthermore, several other metabolites, including lactate, choline-containing compounds, ethanolamine, phosphocreatine, nicotinamide and glutamate, were detected in reduced amounts in the serum of the infected animals comparatively, whereas formate and acetate were significantly increased after infection. The authors concluded that the metabolite variation detected in both lung tissue and serum indicate metabolic changes

associated with necrotic disease processes, including anaerobic glycolysis, glutaminolysis, and gluconeogenesis (Somashekar *et al.*, 2012).

In an attempt to find possible diagnostic markers for active TB, not influenced by anti-TB drug administration, Che *et al.* (2013) used a GC-TOFMS metabolomics approach to compare serum samples collected from healthy controls and TB patients, before and after anti-TB treatment. They indicated that cholesterol, galactose and malonic acid were significantly elevated in the TB patient samples, whereas phenylalanine, glycerol 3-phospate, ornithine, inositol, lactic acid and 5-oxoproline were comparatively downregulated. Of these nine potential TB biomarkers, 5-oxoproline was determined to remain unaffected in the TB-positive patient sample group despite commencing treatment using first-line anti-TB drugs. Hence, this biomarker was validated using a bigger sample cohort. The importance of this study is that it indicates the capacity for metabolomics to identify useful biomarkers for early prediction of treatment outcome, and for possible use in new anti-TB drug development (Che *et al.*, 2013).

Table 2-4: Tuberculosis biomarkers identified in blood and tissue using a metabolomics research approach.

Sample matrix	Sample cohort	Analytical apparatus	Biomarkers identified	Reference
Lung Tissue (Guinea pigs)	40 x TB+ 10 x HC	HRMAS NMR	**aspartate**; **glutathione**; **betaine**; trimethylamine N-oxide; **lactate**; **alanine**; **acetate**; **creatine**; **phosphocholine**; **myo-inositol**; glycerophosphocholine	Somashekar *et al.* (2012)
Lung Tissue (Mice)	10 x TB+ 10 x HC	^{1}H NMR	**AMP**; **creatine**; **lactate**; **glucose**; NAD+; **NADP+**; **oxaloacetate**; **fumarate**; **succinate**; **alanine**; **aspartate**; **glutamate**; **leucine**; **lysine**; **isoleucine**; **phenylalanine**; **tyrosine**; **glutamine**; **betaine**; **phosphocholine**; **niacinamide**; phosphoethanolamine; acetaldehyde; **taurine**; **UDP-glucose**; **uracil**; **uridine**; xanthine; **glutathione**; itaconic acid	Shin *et al.* (2011)
Spleen tissue (Mice)	10 x TB+ 10 x HC	^{1}H NMR	**AMP**; **lactate**; **NADP+**; **oxaloacetate**; **fumarate**; **succinate**; **alanine**; **leucine**; **isoleucine**; **tyrosine**; **glutamine**; **phosphocholine**; **niacinamide**; **uridine**; **glutathione**	Shin *et al.* (2011)
Liver tissue (Mice)	10 x TB+ 10 x HC	^{1}H NMR	**AMP**; **creatine**; **lactate**; glycogen; **glucose**; **NADP+**; **oxaloacetate**; **succinate**; **uracil**; **aspartate**; **glutamate**; **leucine**; **lysine**; **isoleucine**; **uridine**; **phenylalanine**; **tyrosine**; **glutamine**; **phosphocholine**; **niacinamide**; **taurine**; **UDP-glucose**	Shin *et al.* (2011)
Serum (Mice)	10 x TB+ 10 x HC	^{1}H NMR	**glucose**; **leucine**; **isoleucine**; **phenylalanine**; **formate**	Shin *et al.* (2011)

(Continues on next page)

(Continues from previous page)

Sample matrix	Sample cohort	Analytical apparatus	Biomarkers identified	Reference
Serum (Guinea pigs)	40 x TB+ 10 x HC	HRMAS NMR	**lactate**; **choline-containing compounds**; ethanolamine; phosphocreatine; nicotinamide; **glutamate**; **formate**; **acetate**	Somashekar *et al.* (2012)
Serum (Human)	10 x TB+ 6 x TB+ (before and after treatment) 10 x TB-	GC-TOFMS	5-oxoproline	Che *et al.* (2013)
Serum (Human)	44 x TB+ 46 x latent TB+ (asymptomatic) 46 x healthy controls	UHPLC-MS/MS (positive mode) UHPLC-MS/MS (negative mode) GC-MS	histidine; cysteine; **threonine**; citrulline; cysteine-gluathione disulphide; urea; N-acetylneuraminate; glycocholenate sulphate; inosine; **tryptophan**; **mannose**; 3-carboxy-4-methyl-5-propyl-2-furanpropanoate; **phenylalanine**; pyroglutamine; taurocholenate sulphate; **glutamine**; octadecanedioate; gamma-glutamylglutamine; glycylvaline; asparate	Weiner *et al.* (2012)
Serum (Human)	38 x TB+ 39 x healthy controls	^{1}H NMR	1-methylhistidine; acetoacetate; acetone; **glutamate**; **glutamine**; **isoleucine**; **lactate**; **lysine**; glycine; nicotinate; **phenylalanine**; pyruvate; **tyrosine**; **alanine**; **formate**; glycerolphosphocholine; low-density lipoproteins	Zhou *et al.* (2013)
Serum (Human)	120 x TB+ 105 x healthy controls 51 x lung cancer 45 x pneumonia 28 x COPD 22 x bronchiectasis	UHPLC-MS	hexadecanoic acid; lysoPC (C16:0); lysoPC (C18:0); 3D, 7D, 11D-phytanic acid; **docosanoic acid**; phytal; threoninyl-γ-glutamate; kynurenine; **quinolinic acid**; presqualene diphosphate; lysoPC (P-18:1 (9Z)); lysoPC (P-16:0)	Feng *et al.* (2015)
Serum (Human)	20 x TB+ 18 x healthy controls	LC-MS	99 biomarkers identified - related to glycerophospholipid and arachidonic acid metabolism	Zhong *et al.* (2016)
Plasma (Human)	17 x TB+ 17 x healthy controls	LC-MS	61 biomarkers identified, including anti-TB drugs, glutamate, choline derivatives, *M. tuberculosis*-derived cell wall glycolipids (trehalose-6-mycolate and phosphatidylinositol), lipid mediators of inflammation	Frediani *et al.* (2014)

Note: Compounds in bold have been identified as potential TB biomarkers in more than one metabolomics study.
Abbreviations: GC-MS, gas chromatography mass spectrometry; LC-MS, liquid chromatography mass spectrometry; TOFMS, time-of-flight mass spectrometry; HRMAS, high resolution magic angle spinning; ^{1}H NMR, nuclear magnetic resonance; UHPLC-MS/MS, ultra-high performance liquid chromatography-tandem mass spectrometry; TB+, positively diagnosed for active TB via culture; TB-, patients portraying TB-related symptoms, but tested negative for TB via culture; Healthy control, healthy with no TB-related symptoms and tested negative for TB via culture.

In some instances, due to the complexity of the blood metabolome and the increasing sensitivity of some of the analytical equipment used for the analysis thereof, an extremely large number of biomarkers can potentially be identified. In two such studies using LC-MS, Frediani *et al.* (2014) and Zhong *et al.* (2016) identified 99 and 61 biomarkers respectively when comparing the blood metabolomes of TB-positive patients and healthy controls

(Frediani *et al.*, 2014; Zhong *et al.*, 2016). Although the disease-induced alterations to the host's metabolic pathways were discussed in these studies (Table 2-4), neither investigated the potential diagnostic capacity of the identified biomarkers.

From a diagnostic perspective, although standardised protocols are available for the collection of blood samples, these techniques are considered invasive, and in many patients the fear of needles can induce a stress response, altering the metabolic profile and inducing many of the markers previously described for the TB disease state (Dunn *et al.*, 2005).

2.3.3.4 Metabolomics biomarkers detected using urine

Similar to blood, little or no mycobacteria are expected to be present in the urine of TB patients, and therefore, very low concentrations of those metabolites associated with or unique to *M. tuberculosis* would likely be detected. However, urine collected from TB patients would contain higher concentrations of the altered host-derived metabolites induced by the TB disease state, when compared to that of other patient sample matrixes. This may be considered an additional advantage for biomarker identification for better disease characterisation and diagnostic applications, and therefore many metabolomics research investigations have been done using urine as the preferred sample matrix.

Banday *et al.* (2011) analysed the volatile organic compounds (VOCs) in patient-collected urine for the purpose of discriminating TB patients from healthy controls, and from patients with other diseases such as COPD and lung cancer, using headspace GC-MS. They found significantly elevated levels of o-xylene and isopropyl acetate, accompanied by reduced levels of 3-pentanol, dimethylstyrene and cymol in the TB-positive patient group comparatively (Table 2-5) (Banday *et al.*, 2011). Similarly, Luies and Loots (2016) used a GCxGC-TOFMS approach to investigate the urinary TB metabolite markers induced by infection-related adaptations of the host metabolome and/or host–pathogen interactions. Clear PCA differentiation was achieved between the TB-positive and healthy control group based on their respective metabolite profiles. The identified metabolite biomarkers (Table 2-5) indicated TB-induced changes resulting in abnormal host fatty acid and amino acid metabolism (phenylalanine in particular), mediated through changes to IFN-γ and possibly insulin (Luies & Loots, 2016). These studies demonstrate the potential of metabolomics to identify TB biomarkers using patient-collected urine.

In addition to disease-induced metabolic changes, urine also directly reflects metabolites excreted by the liver and kidneys, and therefore provides additional information regarding changes to an individual's metabolome as a result of a toxic insult. To this end, Das *et al.* (2015) identified unique urinary TB biomarkers and determined the effects of TB treatment on the levels of these compounds. By comparing the GC-MS analysed urine metabolic

profiles of TB-positive patients to that of TB-negative individuals (with TB-like symptoms), they detected 42 features with potential diagnostic value, of which six could accurately be annotated compound names (Table 2-5). Most of the metabolites were related to the phenylalanine metabolic pathway (similar to what was observed by Luies and Loots (2016)), and when using them to plot a ROC curve, the AUC was calculated as 0.85. The influence of first-line TB treatment on these metabolic profiles was investigated by analysing samples obtained from TB patients at various time points during the treatment regimen. A treatment dependent trend was observed, and those patients considered successfully cured after treatment completion, showed a metabolic profile similar to that of healthy controls (Das *et al.*, 2015). This study indicates that in addition to pre-treatment diagnostics, urinary biomarkers also have the capacity to monitor disease progression and to potentially monitor treatment response and predict treatment outcome. This information would be useful in new TB drug trials, since these markers could serve as surrogate endpoints for treatment outcome, long before the six month treatment regimen is completed.

Table 2-5: Tuberculosis biomarkers identified in human urine using a metabolomics research approach.

Sample cohort	Analytical apparatus	Biomarkers identified	Reference
117 x TB+ 37 x healthy controls 7 x COPD 5 x other pulmonary diseases	GC-MS	o-xylene; isopropyl acetate; 3-pentanol; dimethylstyrene; cymol	Banday *et al.* (2011)
21 x TB+ 21 x TB-	GC-MS	lactic acid; norepinephrine; hydroquinone; malic acid; 4-hydroxybenzoic acid; **glucose***	Das *et al.* (2015)
30 x healthy controls 46 x TB+	GCxGC-TOFMS	2-C-methylglycerol; 2-octenoic acid; 5-hydroxyhexanoic acid; 5-hydroxyhydantoin; glycerol monostearate; indole-3-carboxylic acid; kynurenic acid; L-rhamnulose; oxalic acid; phenylacetic acid; **quinolinic acid**; ribitol; homovanillic acid; 5-hydroxyindoleacetic acid; **tryptophan**; phenyllactic acid; **N-acetyltyrosine (indicating elevated tyrosine)**	Luies and Loots (2016)

**Of the 42 biomarkers, the identities of only these six compounds were confirmed.*

Note: Compounds in bold have been identified as potential TB biomarkers in more than one metabolomics study.

Abbreviations: COPD, chronic obstructive pulmonary disease; GC-MS, gas chromatography mass spectrometry; TOFMS, time-of-flight mass spectrometry; TB+, positively diagnosed for active TB via culture; TB-, patients portraying TB-related symptoms, but tested negative for TB via culture; Healthy control, healthy with no TB-related symptoms and tested negative for TB via culture.

2.3.3.5 Metabolomics biomarkers detected using breath

Due to the recent advancements in breath collection techniques, this sample matrix has become an attractive option for diagnostic biomarker identification for TB. Breath originates

directly from the infected lung, and would be expected to contain the characteristic host and bacterial VOCs. Additionally, it can be collected in a non-invasive manner (Scott-Thomas *et al.*, 2013), even from children and HIV-positive individuals, where sputum collection is problematic.

Using an automated thermal desorption (ATD) GC-MS, a set of 130 compounds (mostly alkanes, naphthalene and benzene derivatives) were uniquely detected in the headspace of *M. tuberculosis* cultures, compared to those produced in the headspace of sterile culture media. In a pilot study, the group applied this knowledge to patient-collected breath samples, and successfully differentiated healthy controls (n=59) from hospitalised patients with a suspicion of pulmonary TB, with a 100% sensitivity and specificity. Using only the most abundant compounds detected in the headspace of the *M. tuberculosis* cultures (Table 2-6) as a diagnostic biosignature, they were also able to classify patients with negative (n=19) and positive (n=23) sputum culture results, with a specificity of 100% and sensitivity of 82.6% (Phillips *et al.*, 2007). This approach was then applied to a larger sample cohort of 226 patients and successfully identified the TB-positive patients (those for whom the sputum culture, smear, and chest radiograph were all positive) with a sensitivity of 84% and specificity of 64.7% (Phillips *et al.*, 2010).

Syhre and Chambers (2008) analysed the compounds in the headspace of cultures of various *Mycobacterium* species (*M. tuberculosis*, *M. fortuitum*, *M. bovis*, *M. bovis* BCG, *M. abscessus*, and *M. avium* complex) using SPME and GC-MS. The aim of their investigation was not to identify biomarkers for speciating infection, but rather to identify specific markers which can differentiate *M. tuberculosis* and *M. bovis* from all other mycobacteria, for later application to patient breath sample diagnostics. Two compounds, methyl p-anisate and methyl nicotinate, were identified exclusively in *M. tuberculosis* and *M. bovis*, whereas methyl phenylacetate was identified as a common biomarker in *M. tuberculosis*, *M. bovis* and *M. avium* complex (Table 2-6) (Syhre & Chambers, 2008). An *in vivo* validation for methyl nicotinate as a TB-specific biomarker in patient-collected breath showed significant differences in the concentration when comparing TB-positive patients to healthy controls, proving that it may potentially be valuable in a clinical environment (Syhre *et al.*, 2009). However, a true validation of this would need to be done in breath collected from TB-positive patients and clinically sick controls (those with TB-like symptoms, but testing negative for TB via culture).

In a related study, Kolk *et al.* (2012) later also compared the GC-MS profiles of breath samples collected from 50 TB-positive patients and 50 hospitalised TB-negative patients. Seven VOCs (Table 2-6), including one unknown compound, were identified as potential biomarkers and were used to build a classification model which could predict TB with a sensitivity of 72%, specificity of 86% and accuracy of 79%. The model was further validated

using a different set of samples (21 TB-positive and 50 hospitalised TB-negative), and correctly classified TB with a sensitivity of 62%, specificity of 84% and accuracy of 77% (Kolk *et al.*, 2012).

Considering the above mentioned evidence, breath biomarkers show huge potential for their application in a non-invasive TB diagnostic method, however, further validation is necessary in order to determine which of these compounds are present due to a general disease state, oxidative stress or other external confounders such as smoking or diet, and which of these are, in fact, truly TB-specific.

Table 2-6: Tuberculosis biomarkers identified in human breath using a metabolomics research approach.

Sample cohort	Analytical apparatus	Biomarkers identified	Reference
10 x TB+ 10 x healthy controls	SPME GC-MS	**methyl nicotinate**	Syhre *et al.* (2009)
23 x TB+ 19 x TB- 59 x healthy controls	ATD-GC-MS	1,3-dimethyl-trans-cyclohexane; 1,4-dichloro-benzene; **1,4-dimethyl-cyclohexane**; 2-butyl-1-octanol; **1-methyl-naphthalene**; camphene; 4-methyl-decane; 3-ethyl-2-methyl-heptane; 2,6-dimethyl-octane; 1,2,3,4-tetramethyl-benzene; 3,6,6-trimethyl-bicyclo-3,1,1-hept-2-ene; 1-ethyl-4-methyl-trans-cyclohexane; l-beta-pinene; ethyl-benzene; methyl-benzene; propyl-benzene; 3-methyl-heptane; 2-methoxy-2-me-propane; 1-octene; cyclohexane; heptanal; 2,4-dimethyl-heptane; 4-methyl-heptane; nonanal; 2-methyl-pentane; styrene; **tridecane**	Phillips *et al.* (2007)
30 x TB+ 196 x TB-	ATD-GC-MS	3-(1-methylethyl)-oxetane; 4-methyl-dodecane; hexyl-cyclohexane; Bis-(3,5,5-trimethylhexyl) phthalate; 1,3,5-trimethyl-benzene; 3,7-dimethyl-decane; **tridecane**; 4,6,8-trimethyl-1-nonene; 5-ethyl-2-methyl-heptane; 4-methyl-1-hexene	Phillips *et al.* (2010)
50 x TB+ 50 x TB-	GC-MS	dodecane; 3-heptafluorobutyroxypentadecane; octanol; 5-hexenoic acid; 2-ethyl-1-hexanol	Kolk *et al.* (2012)

Note: Compounds in bold have been identified as potential TB biomarkers in more than one metabolomics study.
Abbreviations: GC-MS, gas chromatography mass spectrometry; ATD, automated thermal desorption; SPME, Solid phase micro-extraction; TB-, patients portraying TB-related symptoms, but tested negative for TB via culture; Healthy control, healthy with no TB-related symptoms and tested negative for TB via culture.

2.3.3.6 Diagnostic validity of the metabolomics tuberculosis biomarkers identified to date

A more comprehensive summary of the 49 metabolite biomarkers detected in more than one study (those markers indicated in bold text in Tables 2-2 to 2-6) is given in Appendix A of this thesis (see Table A-1). The biomarkers identified from cultured *Mycobacterium* (9-hexadecenoic acid, TBSA, heptadecanoic acid, 11-eicosenoic acid, 13-docosenoic acid,

15-tetracosenoic acid, eicosanoic acid and cadaverine) were all detected more than once in different studies conducted by the same research group (Olivier & Loots, 2012a; Olivier & Loots, 2012b; Schoeman *et al.*, 2012), aimed at differentiating various infectious *Mycobacterium* species from one another. These eight compounds are associated with the unique cell wall structure of these organisms, which is already well-known to vary between *Mycobacterium* species, therefore these markers would be expected to have diagnostic value, especially for speciation (Midi Inc, 2015). Additionally, 10-heptadecenoic acid, 9-octadecenoic acid, nonadecanoic acid, á-D-glucopyranoside and D-glucosamine were detected more than once as biomarkers for cultured *M. tuberculosis* (Olivier & Loots, 2012b; Schoeman *et al.*, 2012), and were confirmed when comparing TB-positive and TB-negative patient sputum (Du Preez & Loots, 2013b). These biomarkers are also associated with the unique cell wall of *M. tuberculosis* and therefore show promise for possible use in new TB diagnostic methods using patient-collected sputum. An additional 14 compounds characterising TB were repeatedly detected in different tissue samples (liver, spleen and lung tissue) (Table 2-4). Interestingly, both myo-inositol and uridine were confirmed as characteristic compounds in TB-positive tissue from mice and guinea pigs (Shin *et al.*, 2011; Somashekar *et al.*, 2012), and were also identified as biomarkers in experiments using cultured *M. tuberculosis* (Olivier & Loots, 2012b; Schoeman *et al.*, 2012). The use of tissue samples in a clinical environment is, however, impractical, and the presence of these biomarkers in human blood or urine, for instance, needs to be validated before evaluating their diagnostic capacity. Considering this, eight compounds identified in TB-positive host tissue were in fact also identified as TB biomarkers in serum and urine samples collected from TB-positive patients (Luies & Loots, 2016; Weiner *et al.*, 2012; Zhou *et al.*, 2013). Of these, lysine, isoleucine, phenylalanine and tyrosine were detected in elevated concentrations in all of the TB-positive sample groups analysed, irrespective of whether tissue, serum or urine was used, which serves as confirmation for their potential use as TB biomarkers. However, discrepancies in the concentrations of many other metabolite markers do leave room for debate as to their potential application for these purposes. Lactate and glutamate, for example, were seen to accumulate in various TB-positive host tissue (Shin *et al.*, 2011) and TB-positive patient serum (Zhou *et al.*, 2013) samples, but were found to be reduced in serum collected from TB-infected guinea pigs (Somashekar *et al.*, 2012). Additionally, glutamine was detected in reduced amounts in TB-positive patient serum by one study group (Weiner *et al.*, 2012), and then later in an elevated abundance by a different group, both using patient-collected serum (Zhou *et al.*, 2013). Similarly, elevated levels of formate were identified in the serum collected from *M. tuberculosis*-infected mice (Shin *et al.*, 2011) and guinea pigs (Somashekar *et al.*, 2012), however, this metabolite was indicated to be reduced in serum collected from TB-positive patients when compared to healthy controls (Zhou *et al.*, 2013). However, this may be due to the challenge posed by the fact that experimental models allow for very tight control of time (post infection), whereas clinical TB

manifests at an indeterminate stage post infection. Thus, this could simply represent different disease stages, which are themselves dynamic.

1-Methyl-naphthalene, 1,4-dimethyl-cyclohexane (Phillips *et al.*, 2007) and methyl nicotinate (Syhre & Chambers, 2008; Syhre *et al.*, 2009) were identified as characteristic markers when using headspace analyses of *M. tuberculosis* cultures, and were also confirmed as biomarkers from TB-positive patient-collected breath samples. The use of methyl nicotinate as a possible TB diagnostic marker was further validated in a follow-up study using a larger sample cohort of collected breath samples from TB-positive and TB-negative patients (Syhre *et al.*, 2009). Although this compound is used in a niacin test to distinguish cultured *M. tuberculosis* from other related species, its quantities in patient-collected breath may be influenced by other confounding factors such as smoking, chewing betel nuts and a suppressed immune system (for example in HIV-positive patients) (Syhre *et al.*, 2009). Another TB breath biomarker, tridecane, was verified using two independent sample cohorts (Phillips *et al.*, 2010). However, since this marker is thought to originate from increased oxidative stress in TB patients (Phillips *et al.*, 2010), it would most probably not be specific to a TB disease state, since elevated oxidative stress is a common occurrence in most diseases. Similarly, the biomarker quinolinic acid, identified in TB patient-collected serum (Feng *et al.*, 2015) and urine (Luies & Loots, 2016), originates from the kynurenine pathway, induced by several pro-inflammatory cytokines and chemokines involved in immune reactions related to a number of disease states, and is thus probably also not specific enough for TB.

The challenges experienced to date for identifying a single, disease-specific metabolite for diagnostic application, could potentially be solved by using a combination of metabolites/ biomarkers simultaneously. Docosanoic acid (behenic acid), for example, was detected as one of the most abundant fatty acids in *M. tuberculosis* (Olivier & Loots, 2012b), but its diagnostic capacity using TB-positive patient samples was rather poor. However, when used in combination with three other compounds (lysophosphatidylcholine (18:0) threoninyl-γ-glutamate, and presqualene diphosphate), it could differentiate TB patients from healthy controls as well as patients with other lung diseases with an AUC of 0.99 (Feng *et al.*, 2015).

Regardless of the aim of the study, all biomarkers or biosignatures identified should be validated both statistically and biologically (Broadhurst & Kell, 2006). Although all of the above mentioned metabolomics investigations identified potentially valuable biomarkers, the majority of these were used to better characterise the disease and the pathogen-induced metabolic changes in the host. For this reason, only four of the 21 studies described in this section, and summarised in Tables 2-2 to 2-6, actually validated the identified biomarkers for their potential diagnostic applications, using either an independent test or validation sample set. Another point to note is that many of the TB characterisation and diagnostic studies

described above used rather strict inclusion and exclusion criteria for sample selection. Although this serves well to eliminate variation in the sample cohort to better characterise the disease mechanisms, it could potentially bias the outcome towards better diagnostic outcomes, which would not necessarily be attained in a true clinical setting. Additionally, only two of the diagnostic studies described (Banday *et al.*, 2011; Feng *et al.*, 2015), compared the identified TB biomarkers to those identified in patients with related diseases, such as COPD and lung cancer, while only one other research group (Weiner *et al.*, 2012) included asymptomatic individuals with latent TB infection. In the immediate future, it would be beneficial to see increased efforts towards validating these newly identified biomarkers for use in a clinical context, especially towards improved TB diagnostics.

2.3.3.7 From benchtop to clinical application: Future prospects and challenges of metabolomics biomarkers

Considering the above, over the last ten years, metabolomics has contributed exponentially to the number of new TB biomarkers identified, which have not only shed light on new disease mechanisms but also improved TB diagnostics. Ideally, a TB diagnostic method should be sensitive, specific, fast, inexpensive, and easy to use without extensive training or clinical facilities. Such a method should be able to accurately diagnose TB, and be capable of identifying the species of *Mycobacterium* causing the disease, while also indicating whether or not it is drug-resistant. The method should additionally be able to predict and/or monitor treatment outcome since this would determine the most appropriate treatment approach. Although these ideal diagnostic requirements seem impossible considering the limitations of current TB diagnostic approaches, with the recent advancements in metabolomics and the new TB biomarkers identified to date, it may be achievable. Improvements to the sensitivities and specificities of these approaches can easily be attained using other detection technologies or targeted analytical approaches on the same metabolomics MS-based equipment. The analyses times can also be significantly reduced to only a few minutes, or even seconds, when searching for the identified biomarkers in a targeted manner. Furthermore, metabolomics has identified characteristic biomarkers of active TB using patient-collected sputum, blood, urine and breath (hence the application of these to less-invasive diagnostics), and also for speciation, characterising drug-resistance, and predicting/monitoring treatment outcome. The latter is not only important in the sense that one could potentially predict a patient's response to first-line treatment at time of diagnosis (i.e. before the medication is administrated) or well before the six month treatment regimen is completed, but these markers could accelerate the time-consuming process of new TB medication efficacy testing, and elucidate their mechanisms of action.

The major limitations of the ideal diagnostic requirements we propose above regarding the current metabolomics strategies, would be that a metabolomics approach typically uses expensive analytical equipment, operated by highly trained staff, in a laboratory environment. One should, however, bear in mind that the general idea behind metabolomics is to use these high-end analytical apparatus only for the initial biomarker(s) identification and validation phase. Once these markers have been identified, other technologies (like nanotechnology, or other detection technologies) can be used to develop strategies to increase the sensitivity and specificity of detection and to develop an inexpensive kit for rapid point-of-care diagnostics, with minimal training to the user. Alternatively, kits can also be developed for high-throughput laboratory based TB diagnostics on 96 well plates, using inexpensive laboratory equipment, such as a basic spectrophotometer. Another limitation is that no research group to date has reported a single, unique biomarker present in only TB-positive patients (with the exception of the well-known TBSA). That is not to say that such biomarkers have not in fact been identified, and/or are being developed for diagnostic applications. Using a biosignature rather than a single characteristic marker to diagnose a disease does complicate the matter to a certain extent, however, with the aid of statistical or mathematical prediction models, or pattern recognition software, this can easily be overcome and the diagnosis automated.

2.4 TUBERCULOSIS TREATMENT

The Bacille Calmette-Guerin (BCG) strain, an avirulent *M. bovis* strain, is currently used to manufacture the only TB vaccine available for TB prevention. Administration of this vaccine shortly after birth is mandatory in high-incidence TB countries, and is largely considered effective (Delogu & Fadda, 2009), although this is a huge point of debate, since it is reportedly only effective in a mere 50% of all adult cases (González-Martín *et al.*, 2010). Currently, sixteen new vaccines for TB prevention are in clinical trials, with a few approaching or currently in proof-of-concept field studies, with numerous others in various stages of preclinical development. However, the results of the phase III trials will only be known within a few years (World Health Organization, 2015b).

Anti-TB drugs are categorised as either first-line or second-line drugs. First-line drugs are typically used to treat drug-susceptible TB, and include isoniazid (INH), rifampicin (RIF), pyrazinamide (PZA) and ethambutol (EMB), used in accordance with the Directly Observed Therapy Short-course (DOTS) program, as recommended by the WHO. The DOTS treatment program is a recommended six month therapy regimen, consisting of an initial two month (intensive) treatment phase using a combination of INH, RIF, PZA and EMB, followed by a four month maintenance phase, consisting of only INH and RIF (Olivier & Loots, 2011;

Raviglione *et al.*, 1992). Second-line drugs, on the other hand, are typically only used to treat drug-resistant *M. tuberculosis* strains, and include the fluoroquinolones, thioamides, aminoglycosides, polypeptides, D-cycloserine, terizidone and aminosalicylic acid (Caminero *et al.*, 2010; Parida & Kaufmann, 2010).

When the DOTS program is fully adhered to, it has a high successful treatment rate in patients with drug-susceptible TB (World Health Organization, 2010). Despite this, however, almost 15% of the 10.4 million new TB cases reported annually, are treated unsuccessfully (World Health Organization, 2015a). Several factors, including varying individual metabolism and HIV co-infection, have been associated with unsuccessful treatment outcomes. Unsuccessful treatment, and the consequent development of drug-resistant TB, can also be ascribed to disruptions and/or early termination of this first-line treatment by the patient due to the severe side-effects associated with these drugs, or perhaps as a result of the limited access to the medical facilities where the drugs are administered in many third world countries (De Villiers & Loots, 2013). For MDR-TB cases, defined as resistance to both RIF and INH, the WHO recommends the addition of the more expensive and more toxic second-line drugs to this DOTS program, thereby extending the total therapy time to 24 months (known as DOTS plus) (World Health Organization, 2010). However, despite these treatment recommendations, globally only 125 000 of the 580 000 (20%) patients with confirmed MDR-TB actually received treatment in 2015, and of these, only 52% were reportedly treated successfully (World Health Organization, 2016). For the sake of brevity, only the first-line drugs will be discussed in this thesis, considering its relevance to the current investigation searching for metabolomics biomarkers related to explaining and predicting first-line treatment failure mechanisms in Chapters 5 and 6, respectively.

2.4.1 Isoniazid

Isoniazid (INH), also known as isonicotinylhydrazide, is a nicotinamide analogue with a structure comprising a pyridine ring and hydrazine group. This drug has been used since 1952 and has a minimum inhibitory concentration (MIC) of 0.02–0.20 µg/mL. Its mode of action is bactericidal on rapidly growing bacilli, with a diminished effect on slow- and intermittently growing bacilli (Arbex *et al.*, 2010). The mechanism of action of INH action is based on the production of various highly reactive compounds, including reactive oxygen species, reactive organic species, and certain electrophilic species, which act on numerous targets in *M. tuberculosis* (Jena *et al.*, 2014). INH penetrates host cells and passively diffuses across the *M. tuberculosis* membrane. Since it is a prodrug, it requires catalytic oxidative activation by the *M. tuberculosis* catalase-peroxidase enzyme, encoded by the *M. tuberculosis KatG* gene, using nicotinamide adenine dinucleotide (NAD), manganese ions and oxygen as cofactors in this reaction. INH activation produces organic free radicals and

oxygen-derived free radicals, namely superoxide, hydrogen peroxide, peroxynitrite, and an isonicotinoyl radical or anion (Fernandes *et al.*, 2017), the latter of which inhibits the FASII NADH-dependent 2-transenoyl acyl carrier protein (ACP) reductase (InhA). Inhibition of this enzymatic complex results in long-chain fatty acid accumulation, thereby inhibiting mycolic acid biosynthesis, which leads to DNA damage and *M. tuberculosis* death (Jena *et al.*, 2014; Timmins & Deretic, 2006; Vilcheze & Jacobs, 2007). Additionally, the various highly reactive compounds and active metabolites of INH may inhibit other vital *M. tuberculosis* cellular pathways, including NAD metabolism, as well as nucleic acid and phospholipid biosynthesis (Arbex *et al.*, 2010; Caminero *et al.*, 2010; Chetty *et al.*, 2016; Kolyva & Karakousis, 2012). A variety of mutations (insertions, deletions and point mutations) in the *KatG* gene and the intergenic region *mabA-inhA* of *M. tuberculosis* results in INH-resistance (50–80%), decreasing INH activity and preventing conversion into its active metabolite. Although many known mutations may result in catalase-peroxidase inactivation, Ser315Thr is the most commonly reported to date (Arbex *et al.*, 2010; Caminero *et al.*, 2010; Zhang & Yew, 2015).

In the patient, INH is metabolised in the liver by N-acetyltransferase 2 (NAT 2) acetylation, producing acetylisoniazid. This metabolite may be hydrolysed into two products namely; isonicotinic acid and acetylhydrazine, the latter of which is further hydrolysed into hydrazine, or acetylated into diacetylhydrazine (Tostmann *et al.*, 2008a). The half-life of INH varies between 1–5 hours (depending on the genetic characteristic of each patient), after which most of the drug (70–96%) is excreted by the kidneys, predominantly producing inactive metabolites, and a small portion excreted in faecal matter (Arbex *et al.*, 2010). Various adverse side-effects are associated with INH ingestion, including anti-TB drug-induced hepatotoxicity (ATDH)/hepatitis, rash, increased urination, severe dizziness/nausea, vomiting, fever, headache, drug-induced lupus erythematous, high anion gap metabolic acidosis, seizures (status epilepticus), mood changes (confusion, psychosis, depression), and pyridoxine (vitamin B_6) depletion (resulting in CNS, peripheral neuropathy, effects and sideroblastic anaemia) (De Villiers & Loots, 2013).

2.4.2 Rifampicin

Rifampicin (RIF), a lipophilic ansamycin with a MIC for *M. tuberculosis* of 0.05–0.50 µg/mL, has been used for treating TB since 1966. RIF, like INH, also has bactericidal activity, and is considered the most important of all anti-TB drugs due to its ability to kill metabolically active *M. tuberculosis*, as well as those in the stationary phase (Arbex *et al.*, 2010). Due to its lipophilic profile, RIF can easily diffuse across the *M. tuberculosis* cell membranes (Kolyva & Karakousis, 2012), after which it binds with high affinity to the β-subunit of DNA-dependent ribonucleic acid (RNA) polymerase, encoded by *rpoB*, inhibiting gene transcription of the infecting mycobacteria, preventing messenger RNA transcription and protein synthesis,

resulting in cell death (Ahmad & Mokaddas, 2009; Arbex *et al.*, 2010; Chetty *et al.*, 2016). Resistance to RIF occurs mainly due to single point mutations in the 81-bp "hotspot" region of the *rpoB* gene (predominantly in codons 531, 526 and 516), which results in the ineffective binding of the drug to this region. Since mono-resistance to RIF is rare (90% of all RIF isolates are also resistant to INH), RIF-resistance has been used as a surrogate marker for detecting MDR-TB (Ahmad & Mokaddas, 2009; Arbex *et al.*, 2010; Chetty *et al.*, 2016; Kolyva & Karakousis, 2012; Zhang & Yew, 2015).

In the patient, approximately 85% of RIF is metabolised by the microsomal enzymes of the liver cytochrome P450 (CYP450) system, producing desacetylrifampicin, which in turn undergoes hydrolysis to produce 3-formyl rifampicin. RIF has a half-life of 2–4 hours, after which it is excreted via the biliary tract (60–65%), with about 6–15% excreted in its unmetabolised form, which may be reabsorbed in the intestine, and an even smaller portion of the unmetabolised RIF excreted in the urine (Arbex *et al.*, 2010; Tostmann *et al.*, 2008a). Various adverse side-effects are associated with RIF, including ATDH/hepatitis, upregulation of hormonal hepatic metabolism (decreasing hormone levels), flushing, pruritus, rash, breathlessness, cramps, diarrhoea, nausea, vomiting, fever, headache, arthralgia, malaise, dysphoria, immunological reactions, and colour change in bodily fluid (urine, tears, sweat) (De Villiers & Loots, 2013).

2.4.3 Pyrazinamide

Pyrazinamide (PZA), or pyrazoic acid amide, is an amide derivative of pyrazine-2-carboxylic acid and a nicotinamide analogue with a similar molecular structure to that of INH. This drug has been used to treat TB since 1952 and has a MIC for *M. tuberculosis* of 6.25–50.0 µg/mL and functions at a pH of 5.5. PZA has a potent sterilising effect with a unique ability to target and eliminate semi-dormant bacilli populations that reside in acidic environments, i.e. within macrophages and other sites of acute inflammation (Arbex *et al.*, 2010; Kolyva & Karakousis, 2012). Although the mechanism of action of PZA is not yet fully understood, it is proposed to enter *M. tuberculosis* via an adenosine triphosphate (ATP)-dependent transport system, and is then converted to its active form, pyrazinoic acid, via the bacterial enzyme pyrazinamidase, encoded for by *pncA*. This metabolite accumulates in the bacterial cytoplasm due to an inefficient efflux system unique to *M. tuberculosis*, decreasing the intracellular pH of the infecting bacteria, thereby inactivating fatty acid synthase I, inhibiting mycolic acid and cell wall structure biosynthesis, reducing the pathogen's membrane potential, inhibiting its ATP synthesis and ultimately resulting in bacterial death (Ahmad & Mokaddas, 2009; Arbex *et al.*, 2010; Kolyva & Karakousis, 2012; Zhang & Yew, 2015). In 2011, Shi *et al.* identified *M. tuberculosis* RpsA (ribosomal protein S1) as the target of pyrazinoic acid, which inhibits trans-translation, a crucial process for freeing rare ribosomes

in non-replicating bacilli (Shi *et al.*, 2011). *M. tuberculosis* PZA-resistance is due to mutations in its *pncA* gene, encoding for the enzyme pyrazinamidase, which is responsible for converting PZA into its active form. Most of these mutations have been reported to occur in the 561-bp and 82-bp genome regions (Arbex *et al.*, 2010; Kolyva & Karakousis, 2012). Additionally, Gopal *et al.* (2016) suggested that PZA-resistance may be caused by two distinct mechanisms, namely (a) mutations in aspartate decarboxylate *panD* that prevent coenzyme A depletion, and (b) the loss of virulence factor synthesis due to mutations in the phthiocerol dimycocerosate (PDIM) genes; *mas* (mycocerosic acid synthase) and *ppsA-E* (phenolthiocerol synthesis type I polyketide synthases) (Gopal *et al.*, 2016).

In the patient, PZA is easily absorbed and distributed throughout the body, following oral administration. PZA is metabolised in the liver, and has a half-life of 9–10 hours, after which it is excreted in the urine through the glomerulus (70%, of which 3% is in its unmetabolised form) (Arbex *et al.*, 2010). Various adverse side-effects are associated with PZA, including ATDH/hepatitis, pruritus (skin irritation), rash and/or urticarial (hives), dysuria (painful urination), interstitial nephritis (kidney inflammation), nausea, vomiting, fever, arthralgia (due to decreased uric acid secretion), malaise, and sideroblastic anaemia (De Villiers & Loots, 2013).

2.4.4 Ethambutol

Ethambutol (EMB), also known as dextro-2,2-(ethylenediimino)-di-1-butanol, has been used since 1966 as an anti-TB drug due to its bacteriostatic effect mainly on rapidly growing intra- and extracellular bacilli, with an MIC for *M. tuberculosis* of 1–5 µg/mL (Arbex *et al.*, 2010; Kolyva & Karakousis, 2012). EMB interferes with the biosynthesis of arabinogalactan, the primary polysaccharide in the mycobacterial cell wall, via inhibition of three homologous, membrane-associated arabinosyl-transferase enzymes (encoded for by the *embCAB* genes), which mediate the polymerisation of arabinose into arabinogalactan (Ahmad & Mokaddas, 2009; Arbex *et al.*, 2010). Furthermore, EMB has also been reported to inhibit RNA metabolism, mycolic acid transfer into the cell wall, phospholipid synthesis, and spermidine biosynthesis (Kolyva & Karakousis, 2012). *In vitro* resistance to EMB develops slowly and is usually associated with point mutations in the *embCAB* operon, with mutations in *embB* (in codons 306, 406 and 497) being most prevalent. Previous studies have linked the *EmbA* and *EmbB* proteins with the formation of the proper terminal hexa-arabinofuranoside motif during arabinogalactan synthesis, while EmbC is associated with lipoarabinomannan synthesis (Ahmad & Mokaddas, 2009; Kolyva & Karakousis, 2012; Zhang & Yew, 2015).

Following oral administration, 75–80% of the EMB is absorbed in the gastrointestinal tract of the patient and distributed throughout the body (with the exception of the cerebrospinal fluid

(CSF) in patients without meningitis). Approximately 8–15% of the total dose of EMB administered is metabolised (oxidised) in the liver to form an aldehyde intermediate, which is then converted into a hydrophilic dicarboxylic acid; 2,2'-(ethylenediimino)-di-butyric acid. The half-life of EMB is 3–4 hours, after which the majority of the drug (50–80%) is excreted in the urine, and approximately 20% excreted in faecal matter (Arbex *et al.*, 2010). Various adverse side-effects are associated with EMB, including ATDH/hepatitis, milk skin reaction, retrobulbar/optic neuritis, red-green colour blindness, vertical nystagmus, hyperuricaemia, arthralgia, and peripheral neuropathy (De Villiers & Loots, 2013).

2.4.5 Using metabolomics to better explain anti-tuberculosis drug action and drug metabolism

Although the DOTS and DOTS plus TB treatment regimens are considered effective when fully adhered to by the patient, the exact mechanisms of action and xenobiotic metabolism of each of these drugs, in addition to the molecular mechanisms responsible for mycobacterial dormancy (suspending or temporarily stopping all physiological functions, i.e. latent infection), persistence (inability rid the host of metabolically active bacteria despite treatment) and drug-resistance, remain poorly understood (Lipworth *et al.*, 2016; Van den Boogaard *et al.*, 2009). By using a global metabolomics approach, hundreds of different compounds can be analyzed and identified simultaneously, allowing for the discovery of new enzyme activities and metabolic pathways (Prosser *et al.*, 2014). A number of research groups to date have applied this approach towards a better understanding of first/second-line TB drug action and metabolism, as well as to investigate the therapeutic characteristics of new TB drug candidates, with the hope that such insights may eventually lead to the development of more effective, less toxic, shortened treatment regimes.

Li *et al.* (2011) applied an ultra-performance liquid chromatography time-of-flight mass spectrometry (UPLC-TOFMS) metabolomics approach to characterize INH metabolism in humans. When comparing the urinary metabolite profiles obtained from healthy controls to those from INH-administered healthy volunteers, eight known and seven new INH-derived metabolites were identified, the latter of which comprised of two new hydroxylated INH metabolites (2-oxo-1,2-dihydro-pyridine-4-carbohydrazine and isoniazid N-oxide) and five new hydrazones (INH conjugated to 4-methyl-2-oxopentanoic acid and/or methyl-2-oxopentanoic acid; INH conjugated to 2-oxohexanedioic acid; INH conjugated to 3-(4-hydroxyphenyl)-2-oxopropanoic acid; INH conjugated to 2-oxo-3-phenylpropanoic acid; and INH conjugated to 3-(1H-indol-3-yl)-2-oxopropanoic acid). The hydrazones are intermediates of the leucine and/or isoleucine, lysine, tyrosine, tryptophan and phenylalanine metabolic pathways, suggesting a disruption in the essential amino acid metabolism as an additional mechanism by which INH inhibits *M. tuberculosis* growth (Li *et al.*, 2011). Similarly,

Nandakumar *et al.* (2014) used LC-TOFMS metabolic profiling to identify metabolic changes in replicating *M. tuberculosis* populations induced by a sub-lethal dose of INH, RIF and streptomycin, respectively. By applying unsupervised hierarchical clustering analysis, a subset of drug-induced metabolome changes common to all three antibiotics was identified. These changes represented the tricarboxylic acid (TCA) cycle, glyoxylate metabolism and biosynthetic pathways of various amino acids. With regards to the TCA intermediates, pyruvate, succinate, fumarate and malate levels were seen to be elevated subsequent to drug exposure, whereas a-ketoglutarate levels was reduced, and citrate/isocitrate levels remained unchanged. This result suggests that despite the previously known diverse bacterial targets of these anti-TB drugs, all three also result in the activation of isocitrate lyases (ICLs). These findings were further substantiated by fact that the ICL-deficient *M. tuberculosis* strains are significantly more sensitive to these anti-TB drugs. Additionally, this increased susceptibility could be chemically restored using antioxidants, confirming that the observed TCA changes are indicative of an adaptive respiratory slowdown caused by an antibiotic-triggered increase in the production of respiratory-derived reactive oxygen intermediates. This metabolomics approach therefore identified a new mechanism by which first-line TB drugs may act on *M. tuberculosis*, thereby suggesting an alternative strategy which may be exploited in the development of more effective anti-TB drugs (Nandakumar *et al.*, 2014).

With application of metabolomics to investigating second-line TB drug action, Halouska *et al.*, (2007) used a NMR metabolomics approach to identify biomarkers which may explain the mechanisms by which D-cycloserine (DCS), a MDR anti-TB drug, functions. Although DCS was previously reported to inhibit both D-alanine racemase (Alr) and D-alanine-D-alanine ligase (Ddl) activity *in vitro*, its *in vivo* activity against *M. tuberculosis* was not yet confirmed. Therefore, they compared the NMR metabolite profiles from a wild-type *M. smegmatis* strain to that of an *alr* null mutant, its complemented derivatives, and two DCS-resistant mutant strains. The metabolite changes identified when comparing the complemented derivatives, indicated that a second D-alanine pathway exists in the *alr* null mutants, which appears to persist after exposure to DCS. PCA differentiation of the *alr* null mutant and its complemented derivatives was, however, lost after DCS exposure, which suggests that the multiple DCS targets (Alr and Ddl) determined *in vitro,* are also present *in vivo*. Clustering of the DCS-resistant strains with the complemented derivative strains suggests that DCS treatment also results in inhibition of the cell wall enzymes, with minimal anti-microbial impact however, and that Ddl appears to be the primary *in vivo* lethal target of DCS (Halouska *et al.*, 2007). In a follow-up metabolomics study, the same group confirmed these *in vivo* drug targets by indicating that cell growth is inhibited when D-alanyl-D-alanine production is halted (Halouska *et al.*, 2013). Prosser and De Carvalho (2013) later confirmed these results using a more direct *ex vivo* stable isotope metabolomics approach, identifying Ddl as a possible

target for future drug development initiatives against *M. tuberculosis* (Prosser & de Carvalho, 2013).

Although there are currently a number of new potential anti-TB drugs undergoing phase II and III preclinical trials, delamanid and bedaquiline are the only newly approved anti-TB drugs in over 50 years. These drugs, however, are currently only used for treating adults with MDR-TB when no other alternatives show effect (Zumla *et al.*, 2013). Considering the urgent need for new anti-TB drugs and alternative TB treatment approaches, De Carvalho *et al.* (2011) used a targeted LC-MS metabolic profiling method to investigate the potential mechanisms of drug action of nitazoxanide (NTZ) and tizoxanide (TIZ) (two broad spectrum anti-parasitic drugs) against *M. tuberculosis*. The results of their metabolomics investigation proved that both compounds are capable of penetrating and accumulating in *M. tuberculosis* cells, and act against replicating and non-replicating populations by disrupting the membrane potential and pH homeostasis. This mechanism of action is identical to that of an anti-mycobacterial drug with a similar chemical structure, niclosamide (NCS), which is an oxidative phosphorylation uncoupler. This study shows how metabolomics can be used in the development of more potent and bioavailable anti-TB drug analogues of NTZ and/or TIZ (De Carvalho *et al.*, 2011), and similarly to both Halouska *et al.*, (2007) and Prosser and De Carvalho (2013), they proved the capability for using metabolomics to successfully elucidate previously unknown mechanisms of drug action.

A comprehensive review of this topic and how metabolomics has and can be used towards improved TB treatment strategies was written by the author, and attached in Appendix D. One of the objectives of this study, however, was to use metabolomics to identify biomarkers better describing and predicting treatment failure, hence a comprehensive background to this will be given in the next section of this literature chapter.

2.5 TUBERCULOSIS TREATMENT FAILURE

TB treatment failure can be defined as the occurrence of persistently positive sputum smears despite treatment (Mukherjee *et al.*, 2004). Treatment failure may be attributed to a long list of factors, including: (a) irregular or inadequate anti-TB drug supplies by health providers, (b) a lack of patient TB-education, (c) poverty and poor life quality (further resulting in transport problems to the sites of treatment administration), as well as (d) long treatment duration. Additionally, and more relevant to this investigation, are the various biochemically related factors contributing to a poor treatment outcome, including: (a) variable individual metabolism of the anti-TB drugs (including variation in xenobiotic metabolism, drug absorption and drug-drug interactions), (b) drug-resistance and pathogenicity by the infectious organism, and (c) non-adherence to the treatment program, especially due to the

associated side-effects (Adhvaryu & Vakharia, 2011; Du Preez & Loots, 2012; Mukonzo, 2011; Quy *et al.*, 2003; Wares *et al.*, 2003). These proposed mechanisms are, however, not yet fully understood. Considering this, new research approaches, such as metabolomics, are currently being used to increase our understanding of the above mentioned biochemical mechanisms related to treatment failure, and could contribute significantly towards improved TB control, and will be discussed in greater detail below.

2.5.1 Variable individual metabolism

2.5.1.1 Xenobiotic metabolism

Xenobiotic metabolism refers to a natural response by which the body modifies a foreign compound (such as drugs) by enzymatic/biochemical reactions (biotransformation), in order to detoxify the body, by eventually eliminating these exogenous substances. During this process, many drugs may, however, be converted to their active states. Furthermore, the duration and intensity of drug action may also be influenced by the rate of its metabolism/biotransformation, which may vary between individuals. These metabolised compounds also serve as "signals" for various pathophysiological outcomes (Chasseaud & Hawkins, 1990; Chen *et al.*, 2007; Liska, 1998), such as explaining or predicting a response to the medication and of treatment outcome. Consequently, identifying xenobiotic metabolites and their related metabolic pathways are considered essential to understanding drug action and their associated side-effects. Variability in individual xenobiotic metabolism has remained a major challenge for pharmacology research. It is important to understand why individual drug responses differ, and to what extent this variability is responsible for the differences we observe in therapeutic efficacy and tendencies towards adverse reactions (Lu, 1998). To date, research has shown that individual variability of xenobiotic metabolism and their associated side-effects are influenced by a number of factors, including genetics, race, ethnicity, gender and external stimuli such as diet, alcohol consumption, drug-drug interactions, etc. (Matthews, 1995).

The current knowledge pool pertaining to the functionality of anti-TB drugs is considered insufficient, as their exact mechanisms of action, especially in the human host, and on the bacteria, are still largely unclear (Böttger, 2011; Olivier & Loots, 2011). This is most probably a major contributor to the lack of development of new anti-TB drugs over the last five decades (only two new drugs have been approved by the Food and Drug Administration (FDA) since 1962), and the rising threat of drug-resistant strains. Thus, it is clear that new strategies are urgently required for better understanding the current *in vivo* host metabolism, anti-TB drug mechanisms and drug-drug interactions (Halouska *et al.*, 2012), which when elucidated, would undoubtedly lead to improved patient prognosis, as measures could then

be put into place, using the knowledge gained in order to prevent or lower the prevalence of treatment failure. To date, a number of examples of this already exists, where metabolomics is not only being used for the identification of new metabolite markers and altered patient metabolite profiles, shedding light on new mechanisms of TB disease and anti-TB drug action, but when combined with other "omics" approaches, allows for a more in-depth and holistic understanding of the research question (Olivier & Loots, 2011). This approach could potentially also be used for identifying markers which could predict individual drug response phenotypes (pharmacometabolomics), as well as in diagnostic applications for the early prediction of treatment failure.

Three approaches are suggested, focusing on individual xenobiotic metabolism research, aimed at resolving practical issues in drug metabolism and toxicology. Firstly, metabolomics could typically be used to better elucidate drug metabolism by identifying metabolites differentiating two groups, one receiving the xenobiotic and another not, after which the biological samples (urine, blood, tissue, saliva, faeces, etc.) are collected for comparison of the resulting difference in the metabolite profiles using a selected analytical method. When comparing these groups, in theory, all the additional compounds or alterations to the metabolite profile in the xenobiotic group in comparison with the control, would be present because of the exogenous compound ingested/injected (Chen *et al.*, 2007). These differences can broadly be visualised using a multivariate, unsupervised scores plot (such a PCA) and then, if a natural differentiation is achieved, those compounds/xenobiotic metabolites best explaining or contributing to the differences, identified using a variety of supervised multivariate (such as PLS-DA) in conjunction with other univariate methods (such as *t*-test and effect sizes) (Loots *et al.*, 2013). For example, Loots *et al.* (2005) conducted a rat-model metabolomics study, based on a similar study design, in order to evaluate the effect of combined anti-TB drug (Rifater) therapy, on the free radical and organic acid profiles of Sprague-Dawley rats, to get a better understanding of their impact on normal metabolism and potentially better explain their associated side-effects. These metabolic effects were also compared to an additional treatment group receiving a combined treatment of Rifater and melatonin. In their study, they determined that the combined anti-TB drug Rifater results in oxidative stress and a metabolite profile associated with individuals suffering from a multiple acyl-CoA dehydrogenase deficiency (MADD), and eventually linked this to the drug causing inhibition of the electron transport flavoproteins, in addition to an explanation of a number of side-effects, which could be corrected for by the co-administration of melatonin (Loots *et al.*, 2005). This approach may however be limiting if the differences detected are due to both exogenous and endogenous xenobiotic metabolites, as was the case in Loots *et al.*'s study. One is, however, able to sift out these differences using libraries specific towards identifying either the drug metabolites, or the altered natural metabolome of the organism. The advantage of this approach is that it may potentially describe alternative pathways

influenced by the drugs, better explaining the side-effects. Considering this, Chen *et al.* (2007) proposed a stable isotope-labelled xenobiotic treatment group comparative to an unlabelled group, as an alternative to this approach, and in so doing, one would be able to track the drug specific metabolites and their associated metabolism and also elucidate which metabolites are directly linked to the xenobiotic metabolism, and which are due to secondary toxicity (Chen *et al.*, 2007; VandenHeuvel, 1986). Lastly, Chen *et al.* (2007) propose a study design to identify human xenobiotic-metabolising enzyme (XME) polymorphisms responsible for adverse drug reactions (ADRs), which can be used for investigating anti-TB drug side-effects, or an individual's response to anti-TB drugs. ADR may also occur due to various factors relating to a patient's general health status, various environmental contaminants and as a result of drug-drug interactions with other co-administered drugs requiring the same enzyme systems for metabolic processing (Chen *et al.*, 2007). In these cases, ADR may largely be caused by genetic polymorphisms of XMEs, or other factors influencing enzymatic activities or their expression, such as an individual's general health, environmental contaminant exposure, cultural and/or genetic factors, and the previously mentioned drug-drug interactions. If ADRs are observed in a patient group, suspected to be due to XME polymorphisms, an investigation into the metabolite profiles and pharmacokinetics would be of great importance, and could potentially identify the metabolic pathways affected and pinpoint the genetic factors causing these ADRs. This in turn may lead to the necessary adaptations to the medication and ultimately individualised treatment procedures (Chen *et al.*, 2007; Evans & Relling, 1999; Matthews, 1995).

2.5.1.2 Drug malabsorption

Although rare, malabsorption of anti-TB drugs may also lead to a poor treatment outcome. Probable reasons for malabsorption, although poorly characterised, include comorbidities of TB with HIV/AIDS infection, hypoalbuminemia in the malnourished, infectious gastroenteritis, gastric achlorydria, gastrointestinal and/or malabsorption diseases, thyroid and end-stage liver or renal diseases. Although either a single or combination of these factors may contribute to drug malabsorption, HIV/AIDS infection is considered by far the most common risk factor associated with malabsorption of anti-TB drugs (Bento *et al.*, 2010; Gurumurthy *et al.*, 2004).

It has previously been reported in the literature that HIV/AIDS infection may lead to anti-TB drug malabsorption and acquired drug-resistance (Gurumurthy *et al.*, 2004; Wells *et al.*, 2007). This was based on the observation that 68% of all patients with reduced serum RIF and INH levels were co-infected with HIV (Bento *et al.*, 2010; Wells *et al.*, 2007). As both RIF and INH reach maximum serum concentrations 2 hours after oral administration, Bento *et al.* (2010) proposed taking two measurements; 2 hours and 6 hours after administration, in

order to determine if a patient has malabsorption (indicated by persistently low concentrations) or delayed absorption (slowly increase concentrations) (Bento *et al.*, 2010), for the purpose of adapting the treatment protocols. However, this approach may be costly and have significant health policy implications. Alternatively, it was proposed that urinary RIF and INH levels could be monitored, providing important information related to drug absorption. Considering this, Gurumurthy *et al.* (2004) investigated the absorption of RIF and INH in patients with HIV/AIDS and diarrhoea, TB/HIV co-infection, pulmonary TB alone, and healthy subjects, by analysing the urinary excretion of these drugs and the associated metabolites. Their investigation indicated that TB-positive patients, without an underlying disease or illnesses, absorb anti-TB drugs reliably, while those with gastrointestinal disorders or HIV/AIDS infection have less optimal absorption (Gurumurthy *et al.*, 2004).

Although no metabolomics investigations on this topic have been done to date, this research approach could be used/contribute to it in the following manner. A comparison of the metabolomes of those individuals grouped as good absorbers vs. poor absorbers, using blood/urine collected at time of diagnosis (prior to commencing treatment), may better explain the underlying mechanisms associated with treatment outcome, be it by underlying disease mechanisms or genetic factors. This information could be used to better explain this phenomenon, or alternatively be used diagnostically to identify malabsorbers prior to treatment, so that the necessary adaptations can be made, i.e. personalised treatment approaches. Similarly, these groups can be compared during treatment using a number of methodological approaches and possibly lead to a better understanding of potentially varying drug metabolism between these individuals.

2.5.1.3 Drug-drug interactions

Drug-drug interactions are especially prevalent in TB patients co-infected with HIV/AIDS, as INH and RIF in particular have inauspicious drug interactions with many of the antiretroviral drugs used to treat HIV/AIDS (Adhvaryu & Vakharia, 2011). These unwanted drug-drug interactions (inhibition or induction) are thought to result in (a) lowered efficacy of co-administered drugs, leading to a poor treatment outcome and drug-resistance by the organism, and also (b) increased risk of toxicity and hence, treatment interruptions (Dean *et al.*, 2002), also leading to drug-resistance.

These drug-drug interactions are thought to occur due to these different medications all being metabolised by the same CYP450 enzyme system (Wong *et al.*, 2013). The CYP450 enzymes refer to a group of important XMEs, and are especially important for the metabolic activation and metabolism of several drugs. INH is known to permanently inhibit several of the CYP450 enzymes, and this state persists until new CYP450 enzymes are synthesised.

Side-effects due to INH drug interactions may include an increased risk of hepatotoxicity, hypertension, flushing and tachycardia. Excess catecholamine (a dopamine precursor), and elevated levels of other medications, may also be present due to inefficient metabolism of these drugs when INH is present (Cheung & Gonzalez, 2008; King & Brucker, 2010).

On the other hand, an increased hepatic blood flow or CYP450 synthesis, leads to enzyme induction, as is the case with RIF administration (Michalets, 1998), where this reaction is mediated through the activation of the pregnane-X receptor by RIF (Ma *et al.*, 2010), reducing the therapeutic effects of numerous other co-ingested medications metabolised through this enzyme system. This may further lead to complicated health consequences when RIF is taken in combination with other common drugs, such as oral contraceptives, warfarin (Coumadin) or sulfonylureas. Furthermore, RIF is a known inducer of cytochrome P450 (CYP) 3A, reflected by independent regulation of CYP3A expression at various sites, with large inter-individual variation observed (Gorski *et al.*, 2003). Gorski *et al.* (2003) investigated CYP3A susceptibility to RIF induction, as a function of age and gender, with midazolam, a CYP3A substrate. RIF significantly increased both the systemic and oral clearance of midazolam and reduced the oral midazolam availability by 88%. Interestingly, gender also had a significant effect on induction of oral midazolam clearance, as this was found to be greater in males when compared to females, largely due to the fact that these gender-related differences in response to CYP3A inducers are substrate-dependent (Gorski *et al.*, 2003). RIF related drug-drug interactions are also reported to lead to an increased risk of aplastic anaemia and hypoglycaemia. Lowered levels of other medications in the blood are commonly reported, thus dosages of these may need to be increased to avoid lowered efficacy when given in combination to RIF (King & Brucker, 2010; Michalets, 1998). For example, verapamil is shown to be decreased 10–20 fold when co-administered with RIF. On the other hand, however, recent studies have shown the potential of verapamil to be used as an effective efflux inhibitor in many organisms, including *M. tuberculosis*, thereby reducing the duration of anti-TB treatment and the emergence of drug-resistance (Gupta *et al.*, 2013). These drug-drug interactions may persist, depending either on the half-life of the inducer or the rate of enzyme degradation and new enzyme synthesis. Considering the short half-life of RIF and the 1–6 day turnover time of CYP450 enzymes, RIF elimination is more rapid than that of excess CYP450, thus interaction duration would be dependent on enzyme turnover (Michalets, 1998).

Bedaquiline (TMC207), a diarylquinoline, was recently introduced as a new anti-TB drug for treating pulmonary MDR-TB in adults. Bedaquiline has specific ATP synthase inhibition activity against *M. tuberculosis*, affecting the energy production required for cell viability, acting as a bactericidal and sterilising agent against *M. tuberculosis* and various other mycobacterial species, including, amongst others, the *M. avium* complex, *M. chelonae* and

M. kansasii (Palomino & Martin, 2013). Based on clinical evaluations of its safety, tolerance and efficacy, the process for approving bedaquiline for treating pulmonary MDR-TB in adults has been accelerated (Matteelli *et al.*, 2010; Palomino & Martin, 2013). Bedaquiline's long half-life and unique mechanism of action potentially makes it suitable for intermittent administration, which is a major advantage compared with RIF, for instance. However, one of the major concerns of this drug is its drug-drug interactions with RIF, as co-administration with RIF has been shown to halve the maximum bedaquiline concentration. As bedaquiline is metabolised by CYP3A4, producing the less-active N-monodesmethyl bedaquiline, any CYP3A4 inducer, such as RIF, will affect the final concentration of bedaquiline (Palomino & Martin, 2013). Another concern is the possible drug-drug interactions between bedaquiline and antiretroviral agents (ARV's), when treating TB/HIV co-infected patients. As current studies devoted to the use of bedaquiline in these cases are still underway, no clinically significant findings have been suggested to date (Matteelli *et al.*, 2010; Palomino & Martin, 2013).

As anti-TB drug-induced toxicity may have a major influence on other medications used simultaneously, the need to predict ADRs as early as possible during the drug treatment regimen is crucial. Metabolomics can provide the means for a global overview of individual variation and the prediction of potential drug-drug interactions, via analyses of urinary metabolites after drug administration, which can serve as a signature for organ-specific drug toxicity, general activity or inactivation (Chen *et al.*, 2007; Clayton *et al.*, 2006; Idle & Gonzalez, 2007). Metabolomic profiling has already been shown to predict drug responses in animal models (Clayton *et al.*, 2006), however, as xenobiotic metabolism differs between different species, particularly due to molecular variation in receptors and XMEs, it presents a problem when it comes to extrapolating data obtained from rodent models (especially rats and mice) to humans (Gonzalez & Yu, 2006; Idle & Gonzalez, 2007). Thus, more reliable *in vivo* systems, such as humanised mouse models, need to be generated in order to study and potentially predict human xenobiotic responses (Gonzalez & Yu, 2006). Currently, the development of animal models as a tool to simulate human xenobiotic metabolism is expanding rapidly (Tateno *et al.*, 2004). Considering this, humanised transgenic mice can be generated by introducing a specific human gene into the rodent genome, of which many genetically-modified mouse model systems for several important nuclear receptors (including AhR, PXR, CAR) have been established to date, which not only offer a better animal system for studying the influence of drug-drug interactions on xenobiotic metabolism, and the consequent prediction of human xenobiotic responses, but also facilitates the elucidation of the related underlying drug mechanisms (Chen *et al.*, 2007; Gonzalez & Yu, 2006). An additional example of this is the Kramnik mouse model (mice harboring the C3HeB/FeJ genotype), in which necrotic lung lesions are generated. Following *M. tuberculosis* infection, animals may develop necrotic lung lesions, where drug-resistant bacilli are typically present,

persisting in an extracellular micro-environment within these lesions (Driver *et al.*, 2012; Kaushal, 2012). Using this model, Driver *et al.* (2012) examined the efficacy of drug treatment by comparing the formation of colliquative necrotic lesions in Kramnik mice to that of BALB/c mice, with non-necrotic lesions, following aerosol challenge. This was accomplished by infecting both the Kramnik and BALB/c mice with *M. tuberculosis* bacilli, followed by 7–8 weeks of monotherapy using drugs with varying modes of action. Drug efficacy was quantified by determining the bacterial load, and disease progression visualised via staining methods. In the Kramnik mice, results indicated the formation of fibrous encapsulated lung lesions with central colliquative necrosis having abundant extracellular bacilli, in the late stages of infection, while the BALB/c mice formed non-necrotic lesions, primarily infected with intracellular bacilli. Furthermore, the Kramnik mice had indications of hypoxia within the necrotic lesions, and significantly more anti-TB drug-resistant colonies, especially for PZA, compared to the BALB/c mice (Driver *et al.*, 2012). As the Kramnik mouse model is considered a reliable representation of human TB, these results not only indicate the suitability of the Kramnik mouse model for testing anti-TB drug efficacy, especially for drug-resistant *M. tuberculosis* strains, allowing for validation of possible candidate anti-TB drugs and vaccine selection, but also the capacity to develop a TB mouse model with a disease state more closely representing that of human TB. This model is advantageous when wanting to elucidate the mechanisms of *M. tuberculosis* pathogenesis (Driver *et al.*, 2012; Kaushal, 2012), and could also potentially also be used in various metabolomics investigations.

To date, xenobiotic receptor-mouse models have also been developed to closely resemble the human metabolome, and have been used to better elucidate the metabolic pathways leading to drug toxicity and drug-drug interactions, aiding in the identification of potential drug biomarkers in humans. In these instances, the unique combination of humanised mouse models subjected to metabolomic analysis, in conjunction with multivariate analysis, is providing answers to key questions related to human health and disease questions (Chen *et al.*, 2007; Cheung & Gonzalez, 2008; Idle & Gonzalez, 2007). Most ideal, however, is the use of metabolomics to investigate these potential drug-drug interactions, using clinical samples of patients already being treated for one or more co-infections. This will allow for the identification of drug-drug interaction markers relating to drug efficacy and/or side-effects in humans.

2.5.2 Drug-resistance and pathogenicity by the infectious organism

As previously mentioned, the DOTS program is based on administering four of the first-line anti-TB drugs, namely INH, RIF, PZA and EMB, simultaneously. To summarise briefly, INH has high bactericidal activity and is used to efficiently and specifically inhibit active growing

M. tuberculosis (Carlton & Kreutzberg, 1966). RIF is considered a powerful sterilising drug, responsible for eliminating active and semi-dormant *M. tuberculosis* from the host (Ahmad & Mokaddas, 2009), as its primary mechanism of action is transcription inhibition (Ahmad & Mokaddas, 2009; Gumbo *et al.*, 2009). PZA is also classified as a sterilising drug and functions by eliminating semi-dormant *M. tuberculosis* bacilli located within host macrophages (Ahmad & Mokaddas, 2009; Gumbo *et al.*, 2009). EMB is classified as a bacteriostatic agent and typically adds additional coverage to reduce the development of drug-resistance (Ahmad & Mokaddas, 2009; King & Brucker, 2010), and is considered effective against both intra- and extracellular *M. tuberculosis* bacteria (Olivier & Loots, 2011).

Considering the limited information surrounding mechanisms by which drug-resistance occur, one of the few metabolomics investigations published regarding drug-resistance to anti-TB drugs was done in 2012, by Du Preez and Loots. Using GC-MS, they investigated an altered fatty acid metabolome of two *rpoB* mutant *M. tuberculosis* strains by comparison to that of an isogenic drug-sensitive wild-type parent strain, in order to characterise RIF-resistance due to mutations in the *rpoB* gene. Metabolite markers best describing the variation were identified using multivariate statistical approaches, namely PCA and PLS-DA. This study was the first to identify a link between an altered fatty acid metabolism and RIF-resistance, *rpoB* mutations, and the β-subunit of RNA polymerase, present in *M. tuberculosis*. In short, they determined that the mutation in the *rpoB* gene of *M. tuberculosis*, resulting in RIF-resistance, influences mycolic acid synthesis by disruption of the mRNA:NTP (ribonucleoside 5'-triphosphotases) equilibrium. This was confirmed by the lack of detection of 10-methyl branched chain fatty acids in the RIF-resistant mutants, accompanied by the accumulation of their fatty acid substrates. Additionally, these mutants displayed greater dependence on fatty acids and acetate as altered energy sources, and the resulting upregulation of the glyoxylic acid cycle. This study not only indicated the potential of pharmacometabolomics for the identification of metabolites associated with drug-resistance, but also mapped its potential for global health and personalised medicine (Du Preez & Loots, 2012). In view of the above, the methodology described by Du Preez and Loots (2012), could furthermore be applied to investigating the mechanisms of drug-resistance of *M. tuberculosis* to the other first-line, or even second-line drugs. The elucidation of xenobiotic-related metabolites in these organisms, as opposed to those previously described for in the host, by comparison of drug-sensitive and drug-resistant strains, with low doses of these drugs in the growth media, could provide further clues pertaining to the biotransformation of xenobiotics in these organisms, and shed light on additional metabolic pathways (and ultimately genes) involved in drug-resistance.

In the following year, Chakraborty *et al.* (2013) conducted a metabolomics approach proving that para-aminosalicylic acid (PAS) does not inhibit *M. tuberculosis* growth by its inhibiting

action on dihydropteroate synthase (DHPS). PAS was previously thought to act as a competitive inhibitor of DHPS by acting as an alternative substrate, in competition with *p*-aminobenzoate (PABA), in folate metabolism. They additionally indicated that other sulfonamides are ineffective against *M. tuberculosis* due to inactivation by bacterial metabolism, rather than inadequate uptake. As new approaches are needed to improve our understanding of drug action and its efficacy for drug development, this discovery highlights that catalysis by an enzyme, rather than the inhibition thereof, could possibly be the key to new drug development (Chakraborty *et al.*, 2013).

Furthermore, from a growth and pathogenesis perspective, bacteria are known to utilise catabolite repression in order to maximise growth, via consumption of host carbon substrates, in a preferred sequence. A metabolomics study conducted by de Carvalho *et al.* (2010) revealed that *M. tuberculosis* is capable of compartmentalised/simultaneous co-catabolism of multiple carbon sources/substrates, enhancing its growth. These carbon sources can be catabolised through the glycolytic, tricarboxylic acid, and/or pentose phosphate pathways, a phenomenon related to a previously undescribed bacterial metabolic network, and possibly linked to the pathogenicity of *M. tuberculosis*. Thus, using metabolomics, new knowledge was obtained, reshaping our basic understanding of *M. tuberculosis* and its ability to adapt to the host for survival (De Carvalho *et al.*, 2010).

Additionally, a GCxGC-TOFMS metabolomics study conducted by Meissner-Roloff *et al.* (2012) compared the metabolomes of a hyper- and hypo-virulent Beijing *M. tuberculosis* strain, attempting to characterise the varying degrees of virulence within the *M. tuberculosis* species, and consequently identified metabolite markers best differentiating these two strains. This pilot study was the first of its kind to identify metabolites associated with an increased metabolic activity, cell wall synthesis, replication rates, and an altered anti-oxidant mechanism, which is thought to result in the increased pathogenicity and survival of this hyper-virulent strain of *M. tuberculosis* (Meissner-Roloff *et al.*, 2012).

In light of the above, a better understanding of the metabolome of the infectious organism, be it drug-sensitive or drug-resistant, in response to a drug, may potentially lead to improved treatment approaches, and to either preventing drug-resistance or to the development of alternative drugs for treating the disease (Chen *et al.*, 2007). With the emergence of MDR-, XDR-, and more recently even TDR-TB, the robustness of our approaches to discover and develop new, more effective anti-TB drugs is questionable. The discovery and development of new anti-TB drugs are largely dependent on our understanding of the complex host-pathogen interactions (Duncan, 2003), xenobiotic metabolism by both the host and bacteria, the specific drug action, their side-effects, etc.

Furthermore, another shortcoming is that of translating basic research into drug discovery programs. To date, various genetic tools have been developed for manipulating *M. tuberculosis* to identify specific drug targets, but few candidates have emerged. Once validated drug targets are identified, lead compounds acting against these targets should be established, in order for medicinal chemistry to utilise these leads in the development of new drugs. Additionally, mycobacterial persistence/resistance models need be characterised for rapid analyses of compounds with potent sterilising activity against *M. tuberculosis*. Finally, the identification of metabolite markers providing an early indication or prediction of treatment outcome could facilitate clinical trials tremendously (Duncan, 2003).

Unfortunately, drug discovery and development is a lengthy process, especially in the case of those developed against *M. tuberculosis*, and resources remain limiting. Although the estimated cost of drug development may vary, it is estimated that the full development of a new anti-TB drug, could cost up to approximately 100 million US dollars. In 2000, The Global Alliance for TB Drug Development (GATB) was established, in order to address these issues. The aims of this alliance is to overcome natural barriers in the field of TB drug development, by forming partnerships between various diverse organisations, such as academic institutions, the pharmaceutical industry, government research laboratories, and non-governmental organisations (Duncan, 2003). The objectives of the GATB include: (a) reducing the current duration and complexity of treatment and/or the doses needed to cure active, drug-sensitive TB, (b) improving the efficacy of MDR-TB treatment, (c) reducing the treatment period of latent TB infection, and (d) simplifying the treatment of TB/HIV co-infected patients, by lowering the incidence of drug-drug interactions, especially those associated with ARV's (Duncan, 2003; Ginsberg & Spigelman, 2007). These will not only result in reduced side-effects, better patient adherence and lower the costs of delivering DOTS, but will also ultimately have a major impact on treatment outcome. The challenges for reaching these goals include: (a) elucidating the biological mechanisms of mycobacterial resistance and latency, (b) discovery and development of new anti-TB drugs with novel mechanisms of action, (c) developing and validating accurate animal models that can reliably be translated to humans, (d) developing and validating metabolite markers for predicting treatment outcome, possibly reducing clinical trial duration, (e) identifying optimum drug combinations, and (f) the means to conduct clinical trials in high-burden TB countries (Ginsberg & Spigelman, 2007). Thus, in order to reach these GATB goals, a significantly increase in effort and resources is required at every stage of the drug discovery and development process. Furthermore, optimisation of these programs will ensure that valuable research is not wasted and potential new anti-TB drugs missed, due to attrition in the development process (Duncan, 2003).

2.5.3 Non-adherence due to the associated side-effects

According to numerous studies, reasons for non-adherence of patients to the anti-TB treatment program includes, amongst others, drug-related side-effects (especially nausea, vomiting, giddiness), initial relief from TB symptoms, alcohol consumption, work or domestic related problems, being too ill to get to treatment points, or the stigma associated with having TB (Jaggarajamma *et al.*, 2007). However, the anti-TB drug side-effects are most often associated with non-adherence, leading to treatment failure and drug-resistance (Duncan, 2003), and thus this aspect of non-adherence is of biological mechanistic relevance, topical for this thesis.

According to a 2008 WHO bulletin, the side-effects associated with anti-TB drug use include hepatotoxicity/hepatitis (liver inflammation), jaundice, exanthema (rash), dyspepsia (indigestion) and arthralgia (joint pain), and are considered the major factors leading to non-adherence. According to this bulletin, 23% of all TB patients terminated or defaulted treatment due to one or more of these side-effects. Furthermore, up to 86% of all patients may develop these side-effects, especially those co-infected with HIV or with a history of hepatitis (Awofeso, 2008). TB treatment regimens require a high (>90%) compliance rate in order to achieve a successful treatment outcome (Amuha *et al.*, 2009). Amuha *et al.* (2008) conducted a study aimed at determining the prevalence and factors associated with non-adherence to anti-TB drugs. An interviewer administered questionnaire was used for qualitative data collection, and revealed that most patients tend to terminate treatment when they start to feel better, or due to the severe side-effects and pill burden, as these drugs may be considered by the patient to be too many to take at once (Amuha *et al.*, 2009). Although the aim of TB control is that trained health staff is able to recognise and manage these side-effects, it would be best if these could be eliminated entirely. Of further concern, however, are the additional expenses which may be incurred by supplying complementary medication to alleviate these associated side-effects (Awofeso, 2008).

To the best of our knowledge, only one such metabolomics study has been done to date. As described earlier, in 2005, Loots *et al.* conducted a metabolomics investigation into the side-effects associated with the oral administration of the combined anti-TB drug Rifater, and alleviating them using combination therapy with melatonin, in a Sprague-Dawley rat model. Rifater was seen to not only result in increased free radicals, but also induce increased organic acid levels characteristic of a MADD metabolic profile, which includes elevated levels of isovalerylglycine, ethylmalonic acid, butyrylglycine, 2-methylbutyrylglycine and suberic acid. On the basis of this, the similarity in symptoms, and the known mechanism by which this abnormal metabolite profile exists in MADD, a better understanding of how these anti-TB drugs cause these associated side-effects was proposed: by inhibition of the electron

transport flavoproteins. Therefore, melatonin, a well-known antioxidant previously proposed to increase the efficacy of INH, was investigated as a means of alleviating these symptoms, when administered in combination to these drugs. The result was alleviation of both the oxidative stress and MADD metabolite profile (Loots *et al.*, 2005).

Despite the limited metabolomics data, a number of studies have been done investigating the general metabolites of first-line anti-TB drugs formed in the host, and their associated side-effects. These are summarised in Table 2-7.

Table 2-7: A summary of the general metabolites of first-line anti-TB drugs and their associated side-effects.

Anti-TB drug (Year discovered)	Associated metabolites	Associated side-effects
Isoniazid (INH) (1952): (Arbex *et al.*, 2010; Ohno *et al.*, 2000; Roy *et al.*, 2008; Tostmann *et al.*, 2008a; Tostmann *et al.*, 2008b)	Ammonia; Diacetylhydrazine (DiAcHZ), Hydrazine (HZ); Hydrazones with pyruvic (INH-PA) and ketoglutaric acids (INH-KA); Isonicotinic acid (INA); Monoacetylhydrazine (AcHZ); N1-Acetyl-N2-isonicotinylhydrazide (AcINH); Oxidising free radicals	• Hepatotoxic: ATDH/hepatitis • Cutaneous: rash • Nephrology: increased urination • Abdominal: severe dizziness/nausea, vomiting • Flu-like symptoms: fever, headache • Other: drug-induced lupus erythematous, high anion gap metabolic acidosis, seizures (status epilepticus), mood changes (confusion, psychosis, depression), pyridoxine (vitamin B_6) depletion (resulting in central nervous system (CNS), peripheral neuropathy, effects and sideroblastic anaemia)
Pyrazinamide (PZA) (1952): (Arbex *et al.*, 2010; Heifets *et al.*, 1989; Olivier & Loots, 2011; Tostmann *et al.*, 2008a)	Pyrazinoic acid (PA); 5-Hydroxypyrazinoic acid (5-OH-PA)	• Hepatotoxic: ATDH/hepatitis • Cutaneous: pruritus (skin irritation), rash and/or urticarial (hives) • Nephrology: dysuria (painful urination), interstitial nephritis (kidney inflammation) • Abdominal: nausea, vomiting • Flu-like symptoms: fever, arthralgia (due to decreased uric acid secretion), malaise • Other: sideroblastic anaemia
Ethambutol (EMB) (1966): (Arbex *et al.*, 2010; Becker *et al.*, 2008)	2,2'-(Ethylenediimino)-di-butyraldehyde; 2,2'-(Ethylenediimino)-di-butyric acid (EDA)	• Hepatotoxic: ATDH/hepatitis • Cutaneous: milk skin reaction • Optical: retrobulbar/optic neuritis, red-green colour blindness, vertical nystagmus • Nephrology: hyperuricaemia • Flu-like symptoms: arthralgia • Other: peripheral neuropathy
Rifampicin (RIF) (1966): (Arbex *et al.*, 2010; Ohno *et al.*, 2000; Tostmann *et al.*, 2008a)	Desacetylrifampicin; 3-Formyl rifampicin	• Hepatotoxic: ATDH/hepatitis, upregulation of hormonal hepatic metabolism (decreasing hormone levels) • Cutaneous: flushing, pruritus, rash • Respiratory: breathlessness • Abdominal: cramps, diarrhea, nausea, vomiting • Flu-like symptoms: fever, headache, arthralgia, malaise, dysphoria • Other: immunological reactions, colour change in bodily fluid (urine, tears, sweat)

From the literature, however, very little information exist linking metabolism of drugs and the resulting metabolites to the aforementioned side-effects, except for that of the hepatotoxicity which is relatively well described. All four first-line drugs used as part of the DOTS program (EMB to a lesser extent) are associated with anti-TB drug-induced hepatotoxicity (ATDH). ATDH generally occurs within the first two months of treatment, and is thought to be dose-dependent. Hence, decreasing the drug dosage may lighten the symptoms. Although the occurrence of ATDH varies in different populations, studies have shown the most common factors to determine a predisposition towards this include genetics, advanced age, gender, malnutrition, alcoholism, hepatitis B or C, and HIV/AIDS infection, which may in turn lead to increased morbidity, mortality, non-adherence, treatment failure/relapse and the consequential development of MDR-TB. Genetics is considered the most common cause of ATDH and polymorphisms in various drug metabolising loci, such as N-acetyl transferase 2 (NAT 2), cytochrome P450 oxidase (CYP2E1) and glutathione S-transferase Mu 1 (GSTM1), contributes to the inter-individual variations and varying pharmacological responses to these drugs (Arbex *et al.*, 2010; Roy *et al.*, 2008). Although the exact mechanism of ATDH remains unknown, the toxic metabolites associated with the first-line anti-TB drugs play a crucial role in ATDH development (Tostmann *et al.*, 2008a), as these drugs are metabolised by the liver. INH, RIF and PZA (especially when combined) are potentially hepatotoxic drugs and mostly associated with an increased risk of developing ATDH if the patient is unable to clear the toxic metabolites formed in liver due to any of the aforementioned predispositions. Ammonia, diacetylhydrazine (DiAcHZ), hydrazine (HZ), hydrazones with pyruvic (INH-PA) and ketoglutaric acids (INH-KA), isonicotinic acid (INA), monoacetylhydrazine (AcHZ), N1-acetyl-N2-isonicotinylhydrazide (AcINH) and oxidising free radicals, are the toxic metabolites produced when INH is metabolised, either via NAT 2 or hydrolysis (Ohno *et al.*, 2000; Roy *et al.*, 2008; Tostmann *et al.*, 2008b). These reactions may further be facilitated with the co-administration of RIF, via its induction of the hepatic microsomal enzymes of CYP450. Furthermore, as RIF competes with bilirubin absorption and conjugation for inhibition of glucuronosyltransferase, jaundice and increased serum bilirubin and/or hepatic enzyme levels may occur (Arbex *et al.*, 2010). Additionally, pyrazinoic acid (PA) and 5-hydroxypyrazinoic acid (5-OH-PA), toxic metabolites associated with PZA metabolism, have also recently been implicated to be responsible for PZA-induced hepatotoxicity (Shih *et al.*, 2013).

2.6 PREDICTING TUBERCULOSIS TREATMENT OUTCOME

The WHO has compiled a specific list defining possible treatment outcomes for TB patients using first-line medication, i.e. individuals with drug-susceptible TB (see Table 2-8) (World Health Organization, 2014).

Table 2-8: Treatment outcome for tuberculosis patients on first-line medication, as defined by the World Health Organization (World Health Organization, 2014).

Treatment Outcome	Definition
Cured	A pulmonary TB patient with bacteriologically confirmed TB at the beginning of treatment who was smear- or culture-negative in the last month of treatment and on at least one previous occasion.
Treatment completed	A TB patient who completed treatment without evidence of failure, but with no record to show that sputum smear or culture results in the last month of treatment and on at least one previous occasion were negative, either because tests were not done or because results are unavailable.
Treatment failed	A TB patient whose sputum smear or culture is positive at month 5 or later during treatment.
Died	A TB patient who dies for any reason before starting or during the course of treatment.
Lost to follow-up (defaulter)	A TB patient who did not start treatment or whose treatment was interrupted for 2 consecutive months or more.
Relapse	A TB patient who was previously treated and declared cured or treatment completed at the end of their most recent course of treatment, and is now diagnosed with a recurrent episode of TB (either a true relapse or a new episode of TB caused by reinfection).
Treatment success	The sum of cured and treatment completed.

Current TB control aims to prevent TB morbidity and mortality, as well as anti-TB drug-resistance via the early diagnosis of active TB and the effective treatment thereof (Antoine *et al.*, 2007). TB treatment is considered effective when the patient is completely cured from the active disease state (thus preventing transfer of the infectious organism to other healthy individuals) and also from latent *M. tuberculosis* infection (preventing relapse). Developing an effective method to monitor treatment progression and/or accurately predict treatment outcome or potential relapse, well before the end of the treatment regimen, would therefore be regarded as a major breakthrough since alternative treatment approaches can be considered earlier. Considering this, various metabolomics studies have been done to identify biomarkers for this purpose.

Mahapatra *et al.* (2014) used LC-MS metabolomics to identify metabolic biosignatures characterising treatment outcome using patient-collected urine. By comparing the metabolite profiles of urine collected at time of diagnosis (before treatment) to those collected one month after treatment has commenced, they were able to identify 45 metabolites as potential biomarkers differentiating these sample groups. After reducing these metabolites to a set of six statistically significant molecular features, they were able to classify the urine as either baseline or one month treatment samples, with an error rate of 11.8% (Mahapatra *et al.*, 2014). In future, these metabolites may serve as biomarkers for evaluating drug response, which could potentially accelerate anti-TB drug clinical trials. In the following year, Das *et al.* (2015) investigated the altered urinary metabolite profiles of active TB-positive patients on first-line anti-TB drugs by comparing samples collected at various time points during the treatment regimen, with those collected from healthy controls. A time-dependent treatment trend was observed, and the successfully cured patients were shown to have a metabolic profile similar to that of asymptomatic healthy controls. Although this study indicates the potential of metabolomics towards broadening our understanding of TB treatment progression, the main aim of the study was to identify diagnostic biomarkers and, therefore, no metabolite biomarkers explaining this trend were actually identified or discussed (Das *et al.*, 2015). The same group also investigated individual variation of anti-TB drug metabolism, which can be associated with treatment outcome, drug-related side-effects and the development of drug-resistance. They applied an untargeted GC-TOFMS metabolomics approach and analysed urine samples collected from newly diagnosed TB patients at various intervals (i.e. 2h, 6h, 12h, 24h, 36h and 48h) after first-line oral anti-TB drug administration. The majority of the first-line drugs (INH, RIF and PZA) administered was detected, as well their known metabolites. This metabolomics study not only allowed for a comprehensive description of the drug metabolism phenotype associated with first-line anti-TB drugs (by interpretation of the variation in drug abundance and their known metabolites), but also identified a previously unknown drug metabolite of EMB (Das *et al.*, 2016). In a similar study, Tientcheu *et al.* (2015) compared the transcriptomic and metabolic profiles of patients with either a *M. africanum* or *M. tuberculosis*-infection after receiving the same anti-TB treatment, in order to identify host biomarkers associated with lineages-specific pathogenesis and treatment response. The serum metabolic profiles showed a decline in pro-inflammatory metabolites after treatment, which was more pronounced in the *M. tuberculosis*-infected patients, comparatively. They concluded that differences in the transcriptomic and serum metabolic profiles associated with these two lineages may be an indication of intrinsic host factors associated with susceptibility to TB and/or efficacy of standard anti-TB treatment (Tientcheu *et al.*, 2015).

Over the past few years, progress has also been made in the development and application of electronic nose (sensor array) technologies toward metabolomics research. An electronic nose contains conducting-polymer sensor arrays that interact with volatile compounds, resulting in electrical resistance producing characteristic signals of multiple sensor parameters (Pavlou *et al.*, 2004; Wilson, 2016). In 2016, Zetola *et al.* applied this technology towards the development of improved TB diagnostics and monitoring TB treatment efficacy. Breath samples were collected from healthy, HIV-uninfected controls and pulmonary TB-positive patients, of which 61% were co-infected with HIV, at time of diagnosis and at various intervals during the treatment regimen. The electronic nose signals obtained from the VOCs in these breath samples indicated that the device had a sensitivity of 94.1% and a specificity of 90.0% for differentiating TB-positive patients (before treatment onset) and healthy controls. Changes in these breath signals were also observed longitudinally (at 0, 2, 7, 14, and 30 days following TB treatment onset), indicating its potential application for monitoring treatment progression (Zetola *et al.*, 2016).

Considering the above, metabolomics can be used as an extremely powerful approach to identify markers predicting or monitoring treatment outcome, which can also be used as surrogate endpoints for evaluating the efficacy of new anti-TB drugs during clinical trials.

To date, however, no metabolomics studies have been done using patient urine, collected over the entire six month DOTS treatment regimen, comparing those who were eventually cured to those whose treatment failed.

2.7 REFERENCES

Adhvaryu, M. & Vakharia, B. 2011. Drug-resistant tuberculosis: emerging treatment options. *Clinical Pharmacology: Advances and Applications*, 3:51-67.

Ahmad, S. & Mokaddas, E. 2009. Recent advances in the diagnosis and treatment of multidrug-resistant tuberculosis. *Respiratory Medicine*, 103(12):1777-1790.

Albert, H., Heydenrych, A., Brookes, R., Mole, R.J., Harley, B., Subotsky, E., Henry, R. & Azevedo, V. 2002. Performance of a rapid phage-based test, FASTPlaqueTB, to diagnose pulmonary tuberculosis from sputum specimens in South Africa. *The International Journal of Tuberculosis and Lung Disease*, 6(6):529-537.

Aligent Technologies. 2007. Considerations for selecting GC/MS or LC/MS for metabolomics.

Almstetter, M.F., Oefner, P.J. & Dettmer, K. 2012. Comprehensive two-dimensional gas chromatography in metabolomics. *Analytical and Bioanalytical Chemistry*, 402(6):1993-2013.

Amuha, M.G., Kutyabami, P., Kitutu, F.E., Odoi-Adome, R. & Kalyango, J.N. 2009. Non-adherence to anti-TB drugs among TB/HIV co-infected patients in Mbarara Hospital Uganda: Prevalence and associated factors. *African Health Sciences*, 9(2):8-15.

Antoine, D., French, C.E., Jones, J. & Watson, J.M. 2007. Tuberculosis treatment outcome monitoring in England, Wales and Northern Ireland for cases reported in 2001. *Journal of Epidemiology and Community Health*, 61(4):302-307.

Arbex, M.A., Varella, M.d.C.L., Siqueira, H.R.d. & Mello, F.A.F.d. 2010. Antituberculosis drugs: Drug interactions, adverse effects, and use in special situations-part 1: first-line drugs. *Jornal Brasileiro de Pneumologia*, 36(5):626-640.

Awofeso, N. 2008. Anti-tuberculosis medication side-effects constitute major factor for poor adherence to tuberculosis treatment. *Bulletin of the World Health Organization*, 86(3):B-D.

Bakhsi, S. 2006. Tuberculosis in the United Kingdom: A tale of two nations: Troubador Publishing Ltd.

Banday, K.M., Pasikanti, K.K., Chan, E.C.Y., Singla, R., Rao, K.V.S., Chauhan, V.S. & Nanda, R.K. 2011. Use of urine volatile organic compounds to discriminate tuberculosis patients from healthy subjects. *Analytical Chemistry*, 83(14):5526-5534.

Becker, C., Dressman, J.B., Amidon, G.L., Junginger, H.E., Kopp, S., Midha, K.K., Shah, V.P., Stavchansky, S. & Barends, D.M. 2008. Biowaiver monographs for immediate release solid oral dosage forms: Ethambutol dihydrochloride. *Journal of Pharmaceutical Sciences*, 97(4):1350-1360.

Behr, M.A. & Waters, W.R. 2014. Is tuberculosis a lymphatic disease with a pulmonary portal? *The Lancet Infectious Diseases*, 14(3):250-255.

Bento, J., Duarte, R., Brito, M.C., Leite, S., Lobato, M.R., do Carmo Caldeira, M. & Carvalho, A. 2010. Reminder of important clinical lesson: Malabsorption of antimycobacterial drugs as a cause of treatment failure in tuberculosis. *BMJ Case Reports*, 2010.

Bernabe-Ortiz, A., Carcamo, C.P., Sanchez, J.F. & Rios, J. 2011. Weight variation over time and its association with tuberculosis treatment outcome: a longitudinal analysis. *PloS one*, 6(4):e18474.

Blakemore, R., Nabeta, P., Davidow, A.L., Vadwai, V., Tahirli, R., Munsamy, V., Nicol, M., Jones, M., Persing, D.H. & Hillemann, D. 2011. A multisite assessment of the quantitative capabilities of the Xpert MTB/RIF assay. *American Journal of Respiratory and Critical Care Medicine*, 184(9):1076-1084.

Boehme, C.C., Nicol, M.P., Nabeta, P., Michael, J.S., Gotuzzo, E., Tahirli, R., Gler, M.T., Blakemore, R., Worodria, W. & Gray, C. 2011. Feasibility, diagnostic accuracy, and effectiveness of decentralised use of the Xpert MTB/RIF test for diagnosis of tuberculosis and multidrug resistance: a multicentre implementation study. *The Lancet*, 377(9776):1495-1505.

Bossuyt, P.M., Reitsma, J.B., Bruns, D.E., Gatsonis, C.A., Glasziou, P.P., Irwig, L.M., Moher, D., Rennie, D., De Vet, H.C.W. & Lijmer, J.G. 2003. The STARD statement for reporting studies of diagnostic accuracy: Explanation and elaboration. *Annals of Internal Medicine*, 138(1):W1-12.

Böttger, E.C. 2011. Drug resistance in *Mycobacterium tuberculosis*: Molecular mechanisms and laboratory susceptibility testing. *Antituberculosis chemotherapy. Basel: Karger Medical and Scientific*:128-144.

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

Broadhurst, D.I. & Kell, D.B. 2006. Statistical strategies for avoiding false discoveries in metabolomics and related experiments. *Metabolomics*, 2(4):171-196.

Butler, W.R., Jost, K.C., Jr. & Kilburn, J.O. 1991. Identification of mycobacteria by high-performance liquid chromatography. *Journal of Clinical Microbiology*, 29(11):2468-2472.

Caminero, J.A., Sotgiu, G., Zumla, A. & Migliori, G.B. 2010. Best drug treatment for multidrug-resistant and extensively drug-resistant tuberculosis. *The Lancet Infectious Diseases*, 10(9):621-629.

Carlton, W.W. & Kreutzberg, G. 1966. Isonicotinic acid hydrazide-induced spongy degeneration of the white matter in the brains of Pekin ducks. *The American Journal of Pathology*, 48(1):91.

Cha, D., Cheng, D.e., Liu, M., Zeng, Z., Hu, X. & Guan, W. 2009. Analysis of fatty acids in sputum from patients with pulmonary tuberculosis using gas chromatography–mass spectrometry preceded by solid-phase microextraction and post-derivatization on the fiber. *Journal of Chromatography A*, 1216(9):1450-1457.

Chakraborty, S., Gruber, T., Barry, C.E., Boshoff, H.I. & Rhee, K.Y. 2013. Para-Aminosalicylic Acid Acts as an Alternative Substrate of Folate Metabolism in *Mycobacterium tuberculosis*. *Science*, 339(6115):88-91.

Chasseaud, L.F. & Hawkins, D.R. 1990. Biotransformation of drugs. *Encyclopedia of pharmaceutical technology, edited by J. Swarbrick and J.C. Boylan*:129-157.

Che, N., Cheng, J., Li, H., Zhang, Z., Zhang, X., Ding, Z., Dong, F. & Li, C. 2013. Decreased serum 5-oxoproline in TB patients is associated with pathological damage of the lung. *Clinica Chimica Acta*, 423:5-9.

Chen, C., Gonzalez, F.J. & Idle, J.R. 2007. LC-MS-based metabolomics in drug metabolism. *Drug Metabolism Reviews*, 39(2-3):581-597.

Chetty, S., Ramesh, M., Singh-Pillay, A. & Soliman, M.E. 2016. Recent advancements in the development of anti-tuberculosis drugs. *Bioorganic & Medicinal Chemistry Letters*.

Cheung, C. & Gonzalez, F.J. 2008. Humanized mouse lines and their application for prediction of human drug metabolism and toxicological risk assessment. *Journal of Pharmacology and Experimental Therapeutics*, 327(2):288-299.

Clayton, T.A., Lindon, J.C., Cloarec, O., Antti, H., Charuel, C., Hanton, G., Provost, J.-P., Le Net, J.-L.c., Baker, D. & Walley, R.J. 2006. Pharmaco-metabonomic phenotyping and personalized drug treatment. *Nature*, 440(7087):1073-1077.

Cruciani, M., Scarparo, C., Malena, M., Bosco, O., Serpelloni, G. & Mengoli, C. 2004. Meta-analysis of BACTEC MGIT 960 and BACTEC 460 TB, with or without solid media, for detection of mycobacteria. *Journal of Clinical Microbiology*, 42(5):2321-2325.

Das, M.K., Arya, R., Debnath, S., Debnath, R., Lodh, A., Bishwal, S.C., Das, A. & Nanda, R.K. 2016. Global urine metabolomics in patients treated with first-line tuberculosis drugs and identification of a novel metabolite of ethambutol. *Antimicrobial Agents and Chemotherapy*, 60(4):2257-2264.

Das, M.K., Bishwal, S.C., Das, A., Dabral, D., Badireddy, V.K., Pandit, B., Varghese, G.M. & Nanda, R.K. 2015. Deregulated tyrosine–phenylalanine metabolism in pulmonary tuberculosis patients. *Journal of Proteome Research*, 14(4):1947-1956.

De Carvalho, L.P.S., Darby, C.M., Rhee, K.Y. & Nathan, C. 2011. Nitazoxanide disrupts membrane potential and intrabacterial pH homeostasis of *Mycobacterium tuberculosis*. *ACS Medicinal Chemistry Letters*, 2(11):849-854.

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

De Villiers, L. & Loots, D.T. 2013. Using metabolomics for elucidating the mechanisms related to tuberculosis treatment failure. *Current Metabolomics*, 1(4):306-317.

Dean, G.L., Edwards, S.G., Ives, N.J., Matthews, G., Fox, E.F., Navaratne, L., Fisher, M., Taylor, G.P., Miller, R. & Taylor, C.B. 2002. Treatment of tuberculosis in HIV-infected persons in the era of highly active antiretroviral therapy. *Aids*, 16(1):75-83.

Delogu, G. & Fadda, G. 2009. The quest for a new vaccine against tuberculosis. *The Journal of Infection in Developing Countries*, 3(01):005-015.

Dhingra, V.K., Aggarwal, N., Rajpal, S., Aggarwal, J.K. & Gaur, S.N. 2003. Validity and reliability of sputum smear examination as diagnostic and screening test for tuberculosis. *Indian Journal of Allergy, Asthma and Immunology*, 17(2):67-69.

Driver, E.R., Ryan, G.J., Hoff, D.R., Irwin, S.M., Basaraba, R.J., Kramnik, I. & Lenaerts, A.J. 2012. Evaluation of a mouse model of necrotic granuloma formation using C3HeB/FeJ mice for testing of drugs against *Mycobacterium tuberculosis*. *Antimicrobial Agents and Chemotherapy*, 56(6):3181-3195.

Du Preez, I. & Loots, D.T. 2012. Altered fatty acid metabolism due to rifampicin-resistance conferring mutations in the rpoB gene of *Mycobacterium tuberculosis*: Mapping the potential of pharmaco-metabolomics for global health and personalized medicine. *OMICS: A Journal of Integrative Biology*, 16(11):596-603.

Du Preez, I. & Loots, D.T. 2013a. Detection limit for differentiating between various *Mycobacterium* species and *Pseudomonas aeruginosa* using gas chromatography-mass spectrometry (GC-MS) metabolomics: A comparison of two extraction methods. *African Journal of Microbiology Research*, 7(9):797-801.

Du Preez, I. & Loots, D.T. 2013b. New sputum metabolite markers implicating adaptations of the host to *Mycobacterium tuberculosis*, and vice versa. *Tuberculosis*, 93(3):330-337.

Duncan, K. 2003. Progress in TB drug development and what is still needed. *Tuberculosis*, 83(1):201-207.

Dunn, W.B., Bailey, N.J.C. & Johnson, H.E. 2005. Measuring the metabolome: Current analytical technologies. *Analyst*, 130(5):606-625.

Evans, W.E. & Relling, M.V. 1999. Pharmacogenomics: Translating functional genomics into rational therapeutics. *Science*, 286(5439):487-491.

Fend, R., Kolk, A.H., Bessant, C., Buijtels, P., Klatser, P.R. & Woodman, A.C. 2006. Prospects for clinical application of electronic-nose technology to early detection of *Mycobacterium tuberculosis* in culture and sputum. *Journal of Clinical Microbiology*, 44(6):2039-2045.

Feng, S., Du, Y.Q., Zhang, L., Zhang, L., Feng, R.R. & Liu, S.Y. 2015. Analysis of serum metabolic profile by ultra-performance liquid chromatography-mass spectrometry for biomarkers discovery: Application in a pilot study to discriminate patients with tuberculosis. *Chinese Medical Journal*, 128(2):159-168.

Fennelly, K.P. & Jones-López, E.C. 2015. Quantity and quality of inhaled dose predicts immunopathology in tuberculosis. *Frontiers in Immunology*, 6:13.

Fernandes, G.F.d.S., Salgado, H.R.N. & Santos, J.L.d. 2017. Isoniazid: A Review of Characteristics, Properties and Analytical Methods. *Critical Reviews in Analytical Chemistry*:1-11.

Ferrara, G., Losi, M., D'Amico, R., Roversi, P., Piro, R., Meacci, M., Meccugni, B., Dori, I.M., Andreani, A. & Bergamini, B.M. 2006. Use in routine clinical practice of two commercial blood tests for diagnosis of infection with *Mycobacterium tuberculosis*: A prospective study. *The Lancet*, 367(9519):1328-1334.

Frediani, J.K., Jones, D.P., Tukvadze, N., Uppal, K., Sanikidze, E., Kipiani, M., Tran, V.T., Hebbar, G., Walker, D.I. & Kempker, R.R. 2014. Plasma metabolomics in human pulmonary tuberculosis disease: A pilot study. *PloS one*, 9(10):e108854.

Freeman, R., Goodacre, R., Sisson, P.R., Magee, J.G., Ward, A.C. & Lightfoot, N.F. 1994. Rapid identification of species within the *Mycobacterium tuberculosis* complex by artificial neural network analysis of pyrolysis mass spectra. *Journal of Medical Microbiology*, 40(3):170-173.

Frieden, T.R. & Driver, C.R. 2003. Tuberculosis control: Past 10 years and future progress. *Tuberculosis*, 83(1):82-85.

Garrigó, M., Aragón, L.M., Alcaide, F., Borrell, S., Cardenosa, E., Galán, J.J., Gonzalez-Martín, J., Martin-Casabona, N., Moreno, C. & Salvado, M. 2007. Multicenter laboratory evaluation of the MB/BacT Mycobacterium detection system and the BACTEC MGIT 960 system in comparison with the BACTEC 460TB system for susceptibility testing of *Mycobacterium tuberculosis*. *Journal of Clinical Microbiology*, 45(6):1766-1770.

Ginsberg, A.M. & Spigelman, M. 2007. Challenges in tuberculosis drug research and development. *Nature*, 13(3):290-294.

Glickman, S.E., Kilburn, J.O., Butler, W.R. & Ramos, L.S. 1994. Rapid identification of mycolic acid patterns of mycobacteria by high-performance liquid chromatography using pattern recognition software and a Mycobacterium library. *Journal of Clinical Microbiology*, 32(3):740-745.

Golden, M.P. & Vikram, H.R. 2005. Extrapulmonary tuberculosis: An overview. *American Family Physician*, 72(9):1761-1768.

González-Martín, J., García-García, J.M., Anibarro, L., Vidal, R., Esteban, J., Blanquer, R., Moreno, S. & Ruiz-Manzano, J. 2010. Consensus document on the diagnosis, treatment and prevention of tuberculosis. *Archivos de Bronconeumología (English Edition)*, 46(5):255-274.

Gonzalez, F.J. & Yu, A.-M. 2006. Cytochrome P450 and xenobiotic receptor humanized mice. *Annual Review of Pharmacology and Toxicology*, 46:41.

Gopal, P., Yee, M., Sarathy, J., Low, J.L., Sarathy, J.P., Kaya, F., Dartois, V., Gengenbacher, M. & Dick, T. 2016. Pyrazinamide resistance is caused by two distinct mechanisms: Prevention of coenzyme a depletion and loss of virulence factor synthesis. *ACS Infectious Diseases*, 2(9):616-626.

Gorski, J.C., Vannaprasaht, S., Hamman, M.A., Ambrosius, W.T., Bruce, M.A., Haehner-Daniels, B. & Hall, S.D. 2003. The effect of age, sex, and rifampin administration on intestinal and hepatic cytochrome P450 3A activity. *Clinical Pharmacology & Therapeutics*, 74(3):275-287.

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

Gumbo, T., Dona, C.S.W.S., Meek, C. & Leff, R. 2009. Pharmacokinetics-pharmacodynamics of pyrazinamide in a novel in vitro model of tuberculosis for sterilizing effect: A paradigm for faster assessment of new antituberculosis drugs. *Antimicrobial Agents and Chemotherapy*, 53(8):3197-3204.

Gupta, S., Tyagi, S., Almeida, D.V., Maiga, M.C., Ammerman, N.C. & Bishai, W.R. 2013. Acceleration of Tuberculosis Treatment by Adjunctive Therapy with Verapamil as an Efflux Inhibitor. *American Journal of Respiratory and Critical Care Medicine*(ja).

Gurumurthy, P., Ramachandran, G., Kumar, A.K.H., Rajasekaran, S., Padmapriyadarsini, C., Swaminathan, S., Venkatesan, P., Sekar, L., Kumar, S. & Krishnarajasekhar, O.R. 2004. Malabsorption of rifampin and isoniazid in HIV-infected patients with and without tuberculosis. *Clinical Infectious Diseases*, 38(2):280-283.

Halouska, S., Chacon, O., Fenton, R.J., Zinniel, D.K., Barletta, R.G. & Powers, R. 2007. Use of NMR metabolomics to analyze the targets of D-cycloserine in mycobacteria: Role of D-alanine racemase. *Journal of Proteome Research*, 6(12):4608-4614.

Halouska, S., Fenton, R.J., Barletta, R.l.G. & Powers, R. 2012. Predicting the in vivo mechanism of action for drug leads using NMR metabolomics. *Acs Chemical Biology*, 7(1):166-171.

Halouska, S., Fenton, R.J., Zinniel, D.K., Marshall, D.D., Barletta, R.l.G. & Powers, R. 2013. Metabolomics analysis identifies d-Alanine-d-Alanine ligase as the primary lethal target of d-Cycloserine in mycobacteria. *Journal of Proteome Research*, 13(2):1065-1076.

Heifets, L.B., Flory, M.A. & Lindholm-Levy, P.J. 1989. Does pyrazinoic acid as an active moiety of pyrazinamide have specific activity against *Mycobacterium tuberculosis*? *Antimicrobial Agents and Chemotherapy*, 33(8):1252-1254.

Hillemann, D., Rüsch-Gerdes, S., Boehme, C. & Richter, E. 2011. Rapid molecular detection of extrapulmonary tuberculosis by the automated GeneXpert MTB/RIF system. *Journal of Clinical Microbiology*, 49(4):1202-1205.

Idle, J.R. & Gonzalez, F.J. 2007. Metabolomics. *Cell Metabolism*, 6(5):348-351.

Jaggarajamma, K., Sudha, G., Chandrasekaran, V., Nirupa, C., Thomas, A., Santha, T., Muniyandi, M. & Narayanan, P.R. 2007. Reasons for non-compliance among patients treated under Revised National Tuberculosis Control Programme (RNTCP), Tiruvallur district, south India. *Indian Journal of Tuberculosis*, 54(3):130-135.

Jena, L., Waghmare, P., Kashikar, S., Kumar, S. & Harinath, B.C. 2014. Computational approach to understanding the mechanism of action of isoniazid, an anti-TB drug. *International Journal of Mycobacteriology*, 3(4):276-282.

Kaddurah-Daouk, R., Kristal, B.S. & Weinshilboum, R.M. 2008. Metabolomics: A global biochemical approach to drug response and disease. *Annual Review of Pharmacology and Toxicology*, 48:653-683.

Kaushal, D. 2012. Eicosanoids, prostaglandins, and the progression of tuberculosis. *Journal of Infectious Diseases*:jis611.

Kawabata, Y., Semba, I., Hirayama, Y., Koga, T., Nagao, S. & Takada, H. 1998. Wax D of *Mycobacterium tuberculosis* induced osteomyelitis accompanied by reactive bone formation in Buffalo rats. *FEMS Immunology & Medical Microbiology*, 22(4):293-302.

Kik, S.V., Franken, W.P.J., Mensen, M., Cobelens, F.G.J., Kamphorst, M., Arend, S.M., Erkens, C., Gebhard, A., Borgdorff, M.W. & Verver, S. 2010. Predictive value for progression to tuberculosis by IGRA and TST in immigrant contacts. *European Respiratory Journal*, 35(6):1346-1353.

King, T. & Brucker, M. 2010. Pharmacology for Women's Health. Ontario, Canada: Jones & Bartlett Publishers.

Knechel, N.A. 2009. Tuberculosis: pathophysiology, clinical features, and diagnosis. *Critical Care Nurse*, 29(2):34-43.

Kolk, A., Hoelscher, M., Maboko, L., Jung, J., Kuijper, S., Cauchi, M., Bessant, C., van Beers, S., Dutta, R. & Gibson, T. 2010. Electronic-nose technology using sputum samples in diagnosis of patients with tuberculosis. *Journal of Clinical Microbiology*, 48(11):4235-4238.

Kolk, A., Van Berkel, J., Claassens, M., Walters, E., Kuijper, S., Dallinga, J. & Van Schooten, F. 2012. Breath analysis as a potential diagnostic tool for tuberculosis. *The International Journal of Tuberculosis and Lung Disease*, 16(6):777-782.

Kolyva, A.S. & Karakousis, P.C. 2012. Old and new TB drugs: Mechanisms of action and resistance. Understanding tuberculosis - new approaches to fighting against drug resistance. *Intech*:978-953.

Konstantinos, A. 2010. Testing for tuberculosis. *Australian Prescriber*, 33(1):12-18.

Lahiri, N., Shah, R.R., Layre, E., Young, D., Ford, C., Murray, M.B., Fortune, S.M. & Moody, D.B. 2016. Rifampin Resistance Mutations Are Associated with Broad Chemical Remodeling of *Mycobacterium tuberculosis*. *Journal of Biological Chemistry*, 291(27):14248-14256.

Lau, S.K.P., Lam, C.-W., Curreem, S.O.T., Lee, K.-C., Lau, C.C.Y., Chow, W.-N., Ngan, A.H.Y., To, K.K.W., Chan, J.F.W. & Hung, I.F.N. 2015. Identification of specific metabolites in culture supernatant of *Mycobacterium tuberculosis* using metabolomics: Exploration of potential biomarkers. *Emerging Microbes & Infections*, 4(1):e6.

Lee, J.-C., Yu, F.-L., Lin, M.-H., Huang, G.-S., Chang, C.-Y., Cheng, C.-L. & Wang, G.-C. 2010. Utility of immunochromatographic assay for detecting *Mycobacterium tuberculosis* from positive BACTEC MGIT 960 cultures. *Journal of Biomedical and Laboratory Sciences*, 22(64):e9.

Lei, Z., Huhman, D.V. & Sumner, L.W. 2011. Mass spectrometry strategies in metabolomics. *Journal of Biological Chemistry*, 286(29):25435-25442.

Leung, A.N. 1999. Pulmonary tuberculosis: The essentials. *Radiology*, 210(2):307-322.

Li, F., Miao, Y., Zhang, L., Neuenswander, S.A. & Douglas, J.T. 2011. Metabolomic analysis reveals novel isoniazid metabolites and hydrazones in human urine. *Drug Metabolism and Pharmacokinetics*, 26(6):569-576.

Lipworth, S., Hammond, R., Baron, V., Hu, Y., Coates, A. & Gillespie, S. 2016. Defining dormancy in mycobacterial disease. *Tuberculosis*, 99:131-142.

Liska, D.J. 1998. The detoxification enzyme systems. *Alternative Medicine Review*, 3(3):187-198.

Loots, D.T. 2014. An altered *Mycobacterium tuberculosis* metabolome induced by katG mutations resulting in isoniazid resistance. *Antimicrobial Agents and Chemotherapy*, 58(4):2144-2149.

Loots, D.T. 2015. New insights into the survival mechanisms of rifampicin-resistant *Mycobacterium tuberculosis*. *The Journal of Antimicrobial Chemotherapy*, 71(3):655-660.

Loots, D.T., Meissner-Roloff, R.J., Newton-Foot, M. & Van Pittius, N.C.G. 2013. A metabolomics approach exploring the function of the ESX-3 type VII secretion system of *M. smegmatis*. *Metabolomics*, 9(3):631-641.

Loots, D.T., Wiid, I.J., Page, B.J., Mienie, L.J. & Helden, P.D. 2005. Melatonin prevents the free radical and MADD metabolic profiles induced by antituberculosis drugs in an animal model. *Journal of Pineal Research*, 38(2):100-106.

Lu, A.Y.H. 1998. Drug-metabolism research challenges in the new millennium: Individual variability in drug therapy and drug safety. *Drug Metabolism and Disposition*, 26(12):1217-1222.

Luies, L. & Loots, D. 2016. Tuberculosis metabolomics reveals adaptations of man and microbe in order to outcompete and survive. *Metabolomics*, 12(3):1-9.

Ma, Z., Lienhardt, C., McIlleron, H., Nunn, A.J. & Wang, X. 2010. Global tuberculosis drug development pipeline: The need and the reality. *The Lancet*, 375(9731):2100-2109.

Mahapatra, S., Hess, A.M., Johnson, J.L., Eisenach, K.D., DeGroote, M.A., Gitta, P., Joloba, M.L., Kaplan, G., Walzl, G. & Boom, W.H. 2014. A metabolic biosignature of early response to anti-tuberculosis treatment. *BMC Infectious Diseases*, 14(1):53.

Matteelli, A., Carvalho, A.C.C., Dooley, K.E. & Kritski, A. 2010. TMC207: The first compound of a new class of potent anti-tuberculosis drugs. *Future Microbiology*, 5(6):849-858.

Matthews, H.W. 1995. Racial, ethnic and gender differences in response to medicines. *Drug Metabolism and Drug Interactions*, 12(2):77-92.

Mazurek, G.H., LoBue, P.A., Daley, C.L., Bernardo, J., Lardizabal, A.A., Bishai, W.R., Iademarco, M.F. & Rothel, J.S. 2001. Comparison of a whole-blood interferon γ assay with tuberculin skin testing for detecting latent *Mycobacterium tuberculosis* infection. *JAMA: the journal of the American Medical Association*, 286(14):1740.

Meissner-Roloff, R.J., Koekemoer, G., Warren, R.M. & Loots, D.T. 2012. A metabolomics investigation of a hyper-and hypo-virulent phenotype of Beijing lineage *M. tuberculosis*. *Metabolomics*, 8(6):1194-1203.

Michalets, E.L. 1998. Update: Clinically significant cytochrome P-450 drug interactions. *Pharmacotherapy*, 18(1):84-112.

Midi Inc. 2015. Sherlock™ Mycobacteria Identification System. http://www.midi-inc.com/ Date of access: 2016.

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

Moore, D.F. & Curry, J.I. 1995. Detection and identification of *Mycobacterium tuberculosis* directly from sputum sediments by Amplicor PCR. *Journal of Clinical Microbiology*, 33(10):2686-2691.

Mukherjee, J.S., Rich, M.L., Socci, A.R., Joseph, J.K., Virú, F.A., Shin, S.S., Furin, J.J., Becerra, M.C., Barry, D.J. & Kim, J.Y. 2004. Programmes and principles in treatment of multidrug-resistant tuberculosis. *The Lancet*, 363(9407):474-481.

Mukonzo, J.K.B. 2011. Pharmacogenetic aspects of HIV/AIDS, tuberculosis and malaria: Emphasis on Ugandan population: Inst för laboratoriemedicin/Dept of Laboratory Medicine.

Nandakumar, M., Nathan, C. & Rhee, K.Y. 2014. Isocitrate lyase mediates broad antibiotic tolerance in *Mycobacterium tuberculosis*. *Nature Communications*, 5:4306.

Nunes-Alves, C., Booty, M.G., Carpenter, S.M., Jayaraman, P., Rothchild, A.C. & Behar, S.M. 2014. In search of a new paradigm for protective immunity to TB. *Nature Reviews Microbiology*, 12(4):289-299.

Ohno, M., Yamaguchi, I., Yamamoto, I., Fukuda, T., Yokota, S., Maekura, R., Ito, M., Yamamoto, Y., Ogura, T. & Maeda, K. 2000. Slow N-acetyltransferase 2 genotype affects the incidence of isoniazid and rifampicin-induced hepatotoxicity. *The International Journal of Tuberculosis and Lung Disease*, 4(3):256-261.

Olivier, I. & Loots, D.T. 2011. An overview of tuberculosis treatments and diagnostics. What role could metabolomics play. *Journal of Cell and Tissue Research*, 11(1):2655-2671.

Olivier, I. & Loots, D.T. 2012a. A comparison of two extraction methods for differentiating and characterising various *Mycobacterium* species and *Pseudomonas aeruginosa* using GC-MS metabolomics. *African Journal of Microbiology Research*, 6(13):3159-3172.

Olivier, I. & Loots, D.T. 2012b. A metabolomics approach to characterise and identify various *Mycobacterium* species. *Journal of Microbiological Methods*, 88(3):419-426.

Orešič, M. 2009. Metabolomics, a novel tool for studies of nutrition, metabolism and lipid dysfunction. *Nutrition, Metabolism and Cardiovascular Diseases*, 19(11):816-824.

Pai, M., Behr, M.A., Dowdy, D., Dheda, K., Divangahi, M., Boehme, C.C., Ginsberg, A., Swaminathan, S., Spigelman, M., Getahun, H., Menzies, D. & Raviglione, M. 2016. Tuberculosis. *Nature Reviews Disease Primers*, 2:23.

Pai, M., Flores, L.L., Pai, N., Hubbard, A., Riley, L.W. & Colford Jr, J.M. 2003. Diagnostic accuracy of nucleic acid amplification tests for tuberculous meningitis: A systematic review and meta-analysis. *The Lancet Infectious Diseases*, 3(10):633-643.

Pai, M., Kalantri, S. & Dheda, K. 2006a. New tools and emerging technologies for the diagnosis of tuberculosis: Part I. Latent tuberculosis. *Expert Review of Molecular Diagnostics*, 6(3):413-422.

Pai, M., Kalantri, S. & Dheda, K. 2006b. New tools and emerging technologies for the diagnosis of tuberculosis: Part II. Active tuberculosis and drug resistance. *Expert Review of Molecular Diagnostics*:423-432.

Palomino, J.C. & Martin, A. 2013. TMC207 becomes bedaquiline, a new anti-TB drug. *Future Microbiology*, 8(9):1071-1080.

Parida, S.K. & Kaufmann, S.H.E. 2010. The quest for biomarkers in tuberculosis. *Drug Discovery Today*, 15(3):148-157.

Pavlou, A.K., Magan, N., Jones, J.M., Brown, J., Klatser, P. & Turner, A.P.F. 2004. Detection of *Mycobacterium tuberculosis* (TB) in vitro and in situ using an electronic nose in combination with a neural network system. *Biosensors and Bioelectronics*, 20(3):538-544.

Philips, J.A. & Ernst, J.D. 2012. Tuberculosis pathogenesis and immunity. *Annual Review of Pathology: Mechanisms of Disease*, 7:353-384.

Phillips, M., Basa-Dalay, V., Bothamley, G., Cataneo, R.N., Lam, P.K., Natividad, M.P.R., Schmitt, P. & Wai, J. 2010. Breath biomarkers of active pulmonary tuberculosis. *Tuberculosis*, 90(2):145-151.

Phillips, M., Cataneo, R.N., Condos, R., Erickson, G.A.R., Greenberg, J., La Bombardi, V., Munawar, M.I. & Tietje, O. 2007. Volatile biomarkers of pulmonary tuberculosis in the breath. *Tuberculosis*, 87(1):44-52.

Piersimoni, C., Mustazzolu, A., Giannoni, F., Bornigia, S., Gherardi, G. & Fattorini, L. 2013. Prevention of false resistance results obtained in testing the susceptibility of *Mycobacterium*

tuberculosis to pyrazinamide with the Bactec MGIT 960 system using a reduced inoculum. *Journal of Clinical Microbiology*, 51(1):291-294.

Pineda, P.R., Leung, A., Muller, N.L., Allen, E.A., Black, W.A. & FitzGerald, J.M. 1993. Intrathoracic paediatric tuberculosis: A report of 202 cases. *Tubercle and Lung Disease*, 74(4):261-266.

Porcel, J.M. 2009. Tuberculous pleural effusion. *Lung*, 187(5):263.

Pouchot, J., Grasland, A., Collet, C., Coste, J., Esdaile, J.M. & Vinceneux, P. 1997. Reliability of tuberculin skin test measurement. *Annals of Internal Medicine*, 126(3):210-214.

Powers, R. 2009. NMR metabolomics and drug discovery. *Magnetic Resonance in Chemistry*, 47(S1):S2-S11.

Prosser, G.A. & de Carvalho, L.P. 2013. Metabolomics reveal d-alanine: d-alanine ligase as the target of d-cycloserine in *Mycobacterium tuberculosis*. *ACS Medicinal Chemistry Letters*, 4(12):1233-1237.

Prosser, G.A., Larrouy-Maumus, G. & de Carvalho, L.P.S. 2014. Metabolomic strategies for the identification of new enzyme functions and metabolic pathways. *EMBO Reports*, 15(6):657-669.

Quy, H.T.W., Lan, N.T.N., Borgdorff, M.W., Grosset, J., Linh, P.D., Tung, L.B., Van Soolingen, D., Raviglione, M., Co, N.V. & Broekmans, J. 2003. Drug resistance among failure and relapse cases of tuberculosis: Is the standard re-treatment regimen adequate? *The International Journal of Tuberculosis and Lung Disease*, 7(7):631-636.

Raviglione, M.C., Narain, J.P. & Kochi, A. 1992. HIV-associated tuberculosis in developing countries: Clinical features, diagnosis, and treatment. *Bulletin of the World Health Organization*, 70(4):515.

Rhee, K.Y., De Carvalho, L.P.S., Bryk, R., Ehrt, S., Marrero, J., Park, S.W., Schnappinger, D., Venugopal, A. & Nathan, C. 2011. Central carbon metabolism in *Mycobacterium tuberculosis*: An unexpected frontier. *Trends in Microbiology*, 19(7):307-314.

Roy, P.D., Majumder, M. & Roy, B. 2008. Pharmacogenomics of anti-TB drugs-related hepatotoxicity. *Pharmacogenomics*, 9(3):311-321.

Schoeman, J.C., Du Preez, I. & Loots, D.T. 2012. A comparison of four sputum pre-extraction preparation methods for identifying and characterising M. tuberculosis using GCxGC-TOFMS metabolomics. *Journal of Microbiological Methods*, 91:301-311.

Schoeman, J.C. & Loots, D.T. 2011. Improved disease characterisation and diagnostics using metabolomics: A review. *Journal of Cell and Tissue Research*, 11:2673-2683.

Scott-Thomas, A., Epton, M. & Chambers, S. 2013. Validating a breath collection and analysis system for the new tuberculosis breath test. *Journal of Breath Research*, 7(3):037108.

Sharma, M.P. & Bhatia, V. 2004. Abdominal tuberculosis. *Indian Journal of Medical Research*, 120:305-315.

Shenai, S., Ronacher, K., Malherbe, S., Stanley, K., Kriel, M., Winter, J., Peppard, T., Barry, C.E., Wang, J. & Dodd, L.E. 2016. Bacterial loads measured by the Xpert MTB/RIF assay as markers of culture conversion and bacteriological cure in pulmonary TB. *PloS one*, 11(8):e0160062.

Shi, W., Zhang, X., Jiang, X., Yuan, H., Lee, J.S., Barry, C.E., Wang, H., Zhang, W. & Zhang, Y. 2011. Pyrazinamide inhibits trans-translation in *Mycobacterium tuberculosis*. *Science*, 333(6049):1630-1632.

Shih, T.-Y., Pai, C.-Y., Yang, P., Chang, W.-L., Wang, N.-C. & Hu, O.Y.-P. 2013. A Novel Mechanism Underlies the Hepatotoxicity of Pyrazinamide. *Antimicrobial Agents and Chemotherapy*, 57(4):1685-1690.

Shin, J.-H., Yang, J.-Y., Jeon, B.-Y., Yoon, Y.J., Cho, S.-N., Kang, Y.-H., Ryu, D.H. & Hwang, G.-S. 2011. 1H NMR-based metabolomic profiling in mice infected with *Mycobacterium tuberculosis*. *Journal of Proteome Research*, 10(5):2238-2247.

Shulaev, V. 2006. Metabolomics technology and bioinformatics. *Briefings in Bioinformatics*, 7(2):128-139.

Small, P.M. & Pai, M. 2010. Tuberculosis diagnosis—time for a game change. *New England Journal of Medicine*, 363(11):1070-1071.

Somashekar, B.S., Amin, A.G., Tripathi, P., MacKinnon, N., Rithner, C.D., Shanley, C.A., Basaraba, R., Henao-Tamayo, M., Kato-Maeda, M. & Ramamoorthy, A. 2012. Metabolomic signatures in guinea pigs infected with epidemic-associated W-Beijing strains of *Mycobacterium tuberculosis*. *Journal of Proteome Research*, 11(10):4873-4884.

Steingart, K.R., Flores, L.L., Dendukuri, N., Schiller, I., Laal, S., Ramsay, A., Hopewell, P.C. & Pai, M. 2011. Commercial serological tests for the diagnosis of active pulmonary and extrapulmonary tuberculosis: An updated systematic review and meta-analysis. *Plos medicine*, 8(8):e1001062.

Syhre, M. & Chambers, S.T. 2008. The scent of *Mycobacterium tuberculosis*. *Tuberculosis*, 88(4):317-323.

Syhre, M., Manning, L., Phuanukoonnon, S., Harino, P. & Chambers, S.T. 2009. The scent of *Mycobacterium tuberculosis*–part II breath. *Tuberculosis*, 89(4):263-266.

Tateno, C., Yoshizane, Y., Saito, N., Kataoka, M., Utoh, R., Yamasaki, C., Tachibana, A., Soeno, Y., Asahina, K. & Hino, H. 2004. Near completely humanized liver in mice shows human-type metabolic responses to drugs. *The American Journal of Pathology*, 165(3):901-912.

Tientcheu, L.D., Maertzdorf, J., Weiner, J., Adetifa, I.M., Mollenkopf, H.-J., Sutherland, J.S., Donkor, S., Kampmann, B., Kaufmann, S.H. & Dockrell, H.M. 2015. Differential transcriptomic and metabolic profiles of *M. africanum*-and *M. tuberculosis*-infected patients after, but not before, drug treatment. *Genes and Immunity*, 16(5):347-355.

Timmins, G.S. & Deretic, V. 2006. Mechanisms of action of isoniazid. *Molecular Microbiology*, 62(5):1220-1227.

Todar, K. 2009. *Mycobacterium tuberculosis* and Tuberculosis. *Textbook of Bacteriology 2008. Available from: URL: http://www.textbookofbacteriology.net/tuberculosis.html*.

Tortoli, E. & Palomino, J.C. 2007. Tuberculosis 2007: From basic science to patient care (Chapter 14: New diagnostic methods). 1. Brazil.

Tostmann, A., Boeree, M.J., Aarnoutse, R.E., De Lange, W., Van Der Ven, A.J.A.M. & Dekhuijzen, R. 2008a. Antituberculosis drug-induced hepatotoxicity: Concise up-to-date review. *Journal of Gastroenterology and Hepatology*, 23(2):192-202.

Tostmann, A., Boeree, M.J., Peters, W.H.M., Roelofs, H.M.J., Aarnoutse, R.E., van der Ven, A.J.A.M. & Dekhuijzen, P.N. 2008b. Isoniazid and its toxic metabolite hydrazine induce in vitro pyrazinamide toxicity. *International Journal of Antimicrobial Agents*, 31(6):577-580.

Trivedi, M. 2015. An Impact of Treatment: Antimycobacterial Susceptibility Potential Using BACTEC 460/MGIT-TB System. *Mycobacterial Diseases*, 5(4).

Van den Boogaard, J., Kibiki, G.S., Kisanga, E.R., Boeree, M.J. & Aarnoutse, R.E. 2009. New drugs against tuberculosis: Problems, progress, and evaluation of agents in clinical development. *Antimicrobial Agents and Chemotherapy*, 53(3):849-862.

Van der Werf, M.J., Overkamp, K.M., Muilwijk, B., Coulier, L. & Hankemeier, T. 2007. Microbial metabolomics: Toward a platform with full metabolome coverage. *Analytical Biochemistry*, 370(1):17-25.

VandenHeuvel, W.J.A. 1986. Drug metabolite identification: Stable isotope methods. *The Journal of Clinical Pharmacology*, 26(6):427-434.

Vilcheze, C. & Jacobs, J., William R. 2007. The mechanism of isoniazid killing: clarity through the scope of genetics. *Annual Review of Microbiology*, 61:35-50.

Wallace, A.M., Gibson, S., De La Hunty, A., Lamberg-Allardt, C. & Ashwell, M. 2010. Measurement of 25-hydroxyvitamin D in the clinical laboratory: Current procedures, performance characteristics and limitations. *Steroids*, 75(7):477-488.

Want, E.J., Nordström, A., Morita, H. & Siuzdak, G. 2007. From exogenous to endogenous: The inevitable imprint of mass spectrometry in metabolomics. *Journal of Proteome Research*, 6(2):459-468.

Wares, D.F., Singh, S., Acharya, A.K. & Dangi, R. 2003. Non-adherence to tuberculosis treatment in the eastern Tarai of Nepal. *The International Journal of Tuberculosis and Lung Disease*, 7(4):327-335.

Warner, D.F. 2014. *Mycobacterium tuberculosis* metabolism. *Cold Spring Harbor perspectives in medicine*, 5(4):10.1101/cshperspect.a021121.

Weckwerth, W. & Morgenthal, K. 2005. Metabolomics: From pattern recognition to biological interpretation. *Drug Discovery Today*, 10(22):1551-1558.

Weiner, J., Parida, S.K., Maertzdorf, J., Black, G.F., Repsilber, D., Telaar, A., Mohney, R.P., Arndt-Sullivan, C., Ganoza, C.A. & Fae, K.C. 2012. Biomarkers of inflammation, immunosuppression and stress with active disease are revealed by metabolomic profiling of tuberculosis patients. *PloS one*, 7(7):e40221.

Wells, C.D., Cegielski, J.P., Nelson, L.J., Laserson, K.F., Holtz, T.H., Finlay, A., Castro, K.G. & Weyer, K. 2007. HIV infection and multidrug-resistant tuberculosis—the perfect storm. *Journal of Infectious Diseases*, 196(Supplement 1):S86-S107.

Wheelock, C.E., Goss, V.M., Balgoma, D., Nicholas, B., Brandsma, J., Skipp, P.J., Snowden, S., Burg, D., D'Amico, A., Horvath, I., Chaiboonchoe, A., Ahmed, H., Ballereau, S., Rossios, C., Chung, K.F., Montuschi, P., Fowler, S.J., Adcock, I.M., Postle, A.D., Dahlen, S.E., Rowe, A., Sterk, P.J., Auffray, C., Djukanovic, R. & Group, U.B.S. 2013. Application of 'omics technologies to biomarker discovery in inflammatory lung diseases. *The European Respiratory Journal*, 42(3):802-825.

Willey, J.M. 2008. Prescott, Harley, and Klein's Microbiology-7th international ed./Joanne M. Willey, Linda M. Sherwood, Christopher J. Woolverton: New York [etc.]: McGraw-Hill Higher Education.

Wilson, A.D. 2016. Biomarker metabolite signatures pave the way for electronic-nose applications in early clinical disease diagnoses. *Current Metabolomics*, 4(2):12.

Wong, E.B., Cohen, K.A. & Bishai, W.R. 2013. Rising to the challenge: New therapies for tuberculosis. *Trends in Microbiology*, 21(9):493-501.

World Health Organization. 2006. Diagnostics for tuberculosis: Global demand and market potential. Geneva, Switzerland (WHO Press).

World Health Organization. 2010. The global plan to stop TB 2011-2015: Transforming the fight towards elimination of tuberculosis. Geneva, Switzerland (WHO Press).

World Health Organization. 2011. Tuberculosis serodiagnostic tests policy statement 2011. Geneva, Switzerland (WHO Press).

World Health Organization. 2014. Definitions and reporting framework for tuberculosis – 2014 revision. Geneva, Switzerland (WHO Press).

World Health Organization. 2015a. Global tuberculosis report 2015. Geneva, Switzerland (WHO Press).

World Health Organization. 2015b. Immunization, Vaccines and Biologicals: Tuberculosis vaccine development. Geneva, Switzerland (WHO Press).

World Health Organization. 2016. Global tuberculosis report 2016. Geneva, Switzerland (WHO Press).

Xia, J., Broadhurst, D.I., Wilson, M. & Wishart, D.S. 2013. Translational biomarker discovery in clinical metabolomics: An introductory tutorial. *Metabolomics*, 9(2):280-299.

Zetola, N.M., Modongo, C., Matsiri, O., Tamuhla, T., Mbongwe, B., Matlhagela, K., Sepako, E., Catini, A., Sirugo, G. & Martinelli, E. 2016. Diagnosis of pulmonary tuberculosis and assessment of treatment response through analyses of volatile compound patterns in exhaled breath samples. *Journal of Infection*, 74(4):367-376.

Zhang, Y. & Yew, W. 2015. Mechanisms of drug resistance in *Mycobacterium tuberculosis*: Update 2015. *The International Journal of Tuberculosis and Lung Disease*, 19(11):1276-1289.

Zhao, P., Fang, F., Yu, Q., Guo, J., Zhang, J.-h., Qu, J. & Liu, Y. 2014. Evaluation of BACTEC MGIT 960 system for testing susceptibility of *Mycobacterium tuberculosis* to first-line drugs in China. *PloS one*, 9(9):e99659.

Zhong, L., Zhou, J., Chen, X. & Yin, Y. 2016. Serum metabolomic study for the detection of candidate biomarkers of tuberculosis. *International Journal of Clinical and Experimental Pathology*, 9(3):3256-3266.

Zhou, A., Ni, J., Xu, Z., Wang, Y., Lu, S., Sha, W., Karakousis, P.C. & Yao, Y.-F. 2013. Application of 1H NMR spectroscopy-based metabolomics to sera of tuberculosis patients. *Journal of Proteome Research*, 12(10):4642-4649.

Zumla, A., Nahid, P. & Cole, S.T. 2013. Advances in the development of new tuberculosis drugs and treatment regimens. *Nature Reviews Drug Discovery*, 12(5):388-404.

CHAPTER 3: METABOLOMICS METHODOLOGY AND REPEATABILITY

3.1 INTRODUCTION

To date, the majority of research explaining *M. tuberculosis* virulence and growth mechanisms was generated from experiments using *M. tuberculosis* cultures. It has, however, been well established that various growth and virulence mechanisms are only activated once in the host's growth milieu and/or challenged by the host's immune system. Considering this, host–microbe interactions and adaptations has become a new research interest over the last 5–10 years, and studying these organisms isolated from patient sputum, or studying the metabolome from patient-collected sputum, blood or urine, has become an increasingly important research approach towards better understanding these mechanisms. Similar to blood, little or no mycobacteria are expected to be present in the urine of TB patients and, therefore, very low concentrations of those metabolites associated with or unique to *M. tuberculosis* would likely be detected. However, urine collected from TB patients would contain higher concentrations of the altered host-derived metabolites induced by the TB disease state, when compared to that of other patient sample matrixes. In addition to disease-induced metabolic changes, urine also directly reflects metabolites excreted by the liver and kidneys, and therefore provides additional information regarding changes to an individual's metabolome as a result of a toxic insult by the infectious pathogen or treatment thereof. Urine sample collection is also considered non-invasive, with larger amounts of sample material readily available, which can be collected and analysed with far lower chance of transferring disease, and easily transported and stored.

Since TB metabolomics using patient-collected urine is a new approach, a basic method validation of the experimental procedures and the data generated for later metabolomics biomarker identification are required, and is described in this chapter.

3.2 EXPERIMENTAL DESIGN

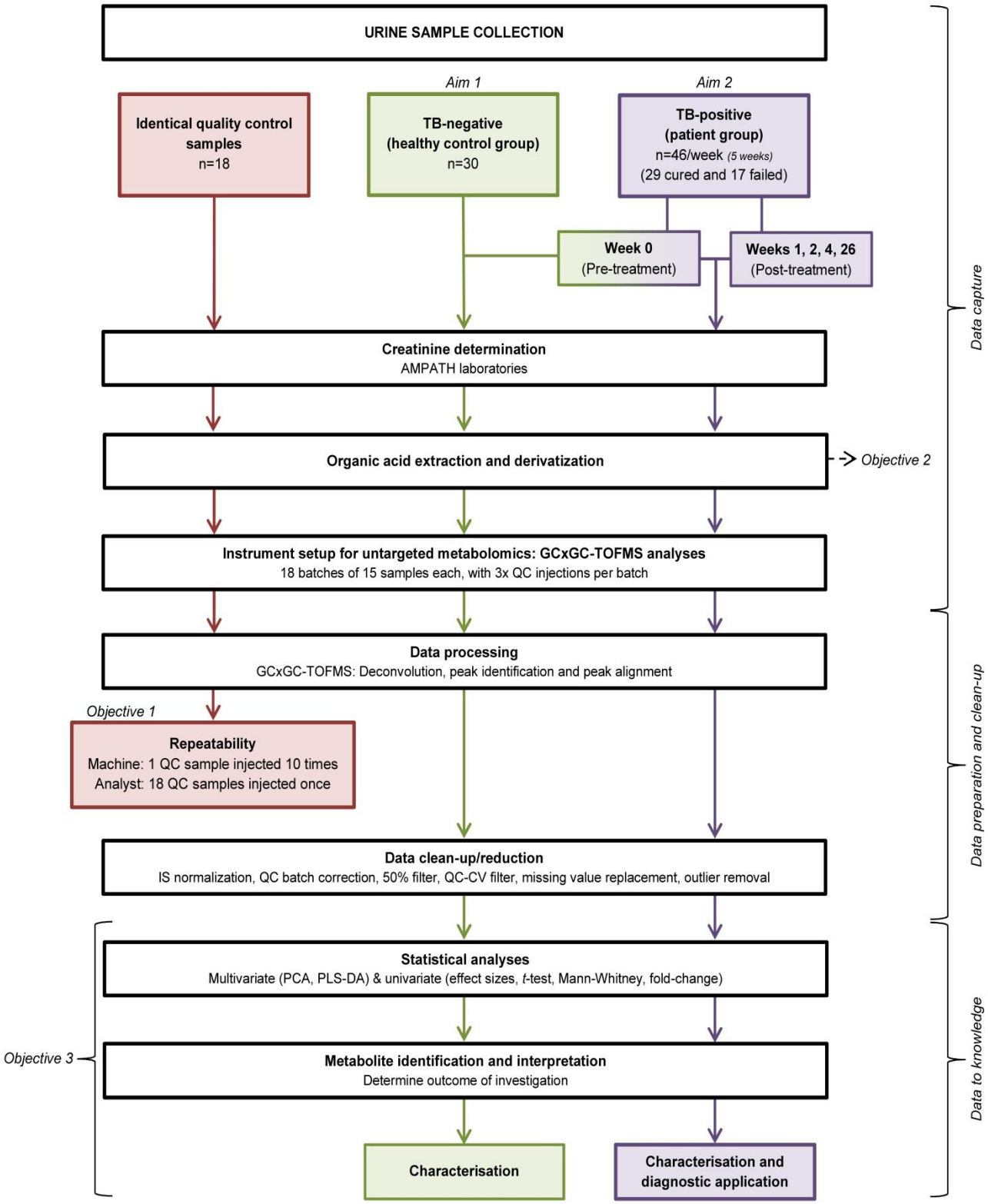


Figure 3-1: Experimental design. A schematic representation of the experimental design of this metabolomics investigation, aimed at determining repeatability and identifying urinary metabolite markers better characterising TB (Aim 1), treatment failure thereof (Aim 2a), and also for the early prediction of an unsuccessful treatment outcome (Aim 2b).

The broad outline of this experiment can be divided into three separate sections, as shown in Figure 3-1, indicating the experimental design proposed to answer the aims and objectives of this investigation.

For this study, urine samples from 38 culture-confirmed active pulmonary TB patients were collected at baseline (time of diagnosis), during the course of treatment (weeks 1, 2 and 4) and two weeks after treatment completion (week 26) (n=210), as well as from a healthy TB-negative group of individuals (n=30) from the same community (more detail on these samples will be given in the next section). A quality control (QC) sample stock solution was compiled by combining equal volumes of all collected urine samples, and aliquots of this were used as identical QC samples throughout the analysis. Since a total of 210 participant samples were collected, these were divided into smaller batches (n=18) for analyses. Samples were randomly assigned to these batches and included one QC aliquot per batch. Each batch of samples were analysed by performing a creatinine determination, organic acid extraction and derivatisation, followed by GCxGC-TOFMS analyses.

In order to achieve the first objective, the repeatability of the analyst/methodology and that of the GCxGC-TOFMS apparatus was evaluated using data collected from the QC samples, as shown in Figure 3-1 (indicated in red). For determining the analyst/methodology repeatability, the GCxGC-TOFMS data of all the QC sample aliquots (n=18) across all sample batches were statistically compared. The repeatability of the analytical apparatus was determined by comparing the GCxGC-TOFMS data collected from a randomly selected QC aliquot, injected and analysed repeatedly (n=10).

Once the GCxGC-TOFMS data was validated and its accuracy determined for metabolomics biomarker identification, the data was processed and cleaned prior to statistical analyses, and the metabolite markers were identified by comparing the appropriate groups to one another, as stipulated in the aims for this study (also indicated in green and purple in Figure 3-1).

3.3 MATERIALS AND METHODS

3.3.1 Urine sample collection and storage

This study was conducted according to the Declaration of Helsinki and International Conference on Harmonisation Guidelines. The necessary ethical approval for this investigation was obtained from the Ethics Committee of the North-West University, South Africa (reference number NWU-00127-11-A1), Stellenbosch University Health Research Ethics Committee (reference number 99/039) and Cape Town City Health. All participants

gave written informed consent for study participation and HIV testing. Furthermore, samples were anonymised. The clinical information received with regards to the urine samples included age, gender, body mass index (BMI), and treatment outcome. Additionally, information regarding the bacteriological cultures at the time of diagnosis and at the end of the treatment programme (week 26) was also obtained.

All urine samples were collected in standard urine collection vials from individuals in the same geographical area, after which they were immediately placed on ice, transported to the laboratory, and frozen at -20°C until analysis. These samples were collected from 30 TB-negative healthy controls and 38 culture-confirmed active TB-positive South African patients (19 males and 19 females, between the age of 17 and 58) at baseline (time of diagnosis, thus prior to initiation of treatment), during the course of treatment with the DOTS strategy (weeks 1, 2 and 4) and two weeks after treatment completion (week 26). All TB-positive patients included in this study had drug-susceptible TB, were HIV-seronegative, not pregnant, and with no other diseases (including diabetes, malignancy, lung cancer, chronic bronchitis and sarcoidosis). Furthermore, the TB-positive samples were divided into successful (n=27) and unsuccessful (n=11) treatment outcome groups, which was determined by bacteriological culture at 26 weeks (two weeks after treatment).

3.3.1.1 Quality control samples

The data generated from the quality control (QC) samples not only allow for the monitoring of the analytical procedure, but also allow for data validation and the correction of any potential batch effects which may have occurred during sample analysis. A single pooled QC sample was compiled by combining 0.2 mL from each of the collected urine samples in a new, clean vial and mixed. From this pooled QC sample stock solution, smaller QC aliquots of 1 mL were prepared and stored at -20°C until analysis. One QC aliquot was extracted, derivatised, and analysed with each batch and injected three times, i.e. in the beginning, middle and at the end of each batch. As it is necessary to equilibrate the machine, the first QC sample was injected five times prior to any analysis.

3.3.2 Reagents and chemicals

Ethyl acetate and diethyl ether were from Honeywell International Inc. (Muskegon, MI, USA). These organic solvents were ultra-pure Burdick and Jackson brands. Hydrochloric acid (HCl), pyridine, anhydrous sodium sulphate (Na_2SO_4) and trimethylchlorosilane (TMCS) were from Merck (Darmstadt, Germany). 3-Phenylbutyric acid, hexane and bis(trimethylsilyl)-trifluoroacetamide (BSTFA) were from Sigma-Aldrich (St. Louis, MO, USA).

3.3.3 Organic acid extraction procedure and derivatisation

Creatinine values for all urine samples, including QC samples, were determined prior to an organic acid extraction, using a creatinine enzyme kit (Thermo Scientific; reference number 981845) and analysed on an Indiko Clinical Analyser, Type 863 (Thermo Scientific). These creatinine values are used to normalise metabolite concentrations, as well as to determine the volume of urine, internal standard, BSTFA, TMCS and pyridine used for each extraction. For acidification, 5N HCl (six drops) was added to the determined amount of urine, followed by the addition of the internal standard, 3-phenylbutyric acid (25 µmol/mg creatinine). Thereafter, ethyl acetate (6 mL) was added, followed by mixing and centrifugation (3000 rpm for 3 minutes) to achieve solvent phase separation, after which the organic phase was aspirated into a clean tube. Diethyl ether (3 mL) was then added to the remaining aqueous phase, again followed by mixing, centrifugation and aspirating the organic solvent into the tube containing the previously collected phase. Hereafter, Na_2SO_4 (approximately 3 g) was added to remove any water still present in the sample, followed by a brief vortex and centrifugation. The organic phase was then decanted from the pellet and evaporated to complete dryness under a stream of nitrogen at 37°C. This was followed by a 60 minute derivatisation period at 60°C with BSTFA (22.6 µL/µmol creatinine), TMCS (4.5 µL/µmol creatinine) and pyridine (4.5 µL/µmol creatinine).

3.3.4 GCxGC-TOFMS analyses

Each derivatised sample was transferred to a 1.5 mL GC-MS sample vial, capped and placed into the auto-sampler tray of a Gerstel Multi-Purpose Sampler (Gerstel GmbH and co. KG, Mülheim an der Ruhr, Germany), coupled to a Pegasus 4D GCxGC-TOFMS (LECO Africa (Pty) Ltd, Johannesburg, South Africa), equipped with an Agilent 7890A gas chromatograph (Agilent, Atlanta, USA), and a TOFMS (LECO Africa). All samples (1 µL) were injected in random sequence, along with QC samples, at regular intervals throughout the analytical run, using a 1:12 split ratio. Compound separation was achieved using a Restek Rxi-5Sil MS primary capillary column (30 m, 0.25 µm film thickness and 250 µm internal diameter), and a Rxi-17 secondary capillary column (0.9 m, 0.1 µm film thickness and 100 µm internal diameter). Cryomodulation was achieved with a hot pulse of nitrogen gas for 0.7 s, every 3 s. The injector temperature was maintained at a constant 280°C for the entire run and ultra-high purity helium was used as the carrier gas at a constant flow of 1 mL/minute. The temperature program for the primary column was 55°C for 1 minute, after which it was increased at a rate of 7°C/minute to a temperature of 285°C, at which it was maintained for a further 4 minutes, followed by a temperature increase of 20°C/minute, to a final temperature of 305°C, where it was maintained for a minute. The secondary oven's

temperature ramp was programed identically to that of the primary column, with the exception of a +5°C deviation at all of the time points. The filament was switched off for the first 8 minutes of each sample run, as this period was considered to correspond to a solvent delay during which no mass spectra were recorded. However, this 8 minute interval was included on the first column's time axis to reflect accurate retention times. The transfer line temperature was held at a constant 270°C and the ion source temperature at a constant 200°C for the entire run. The detector voltage was 1600 V and the filament bias -70 eV. Mass spectra were collected over the range 50–600 m/z at an acquisition rate of 200 spectra per second.

3.3.5 Data processing

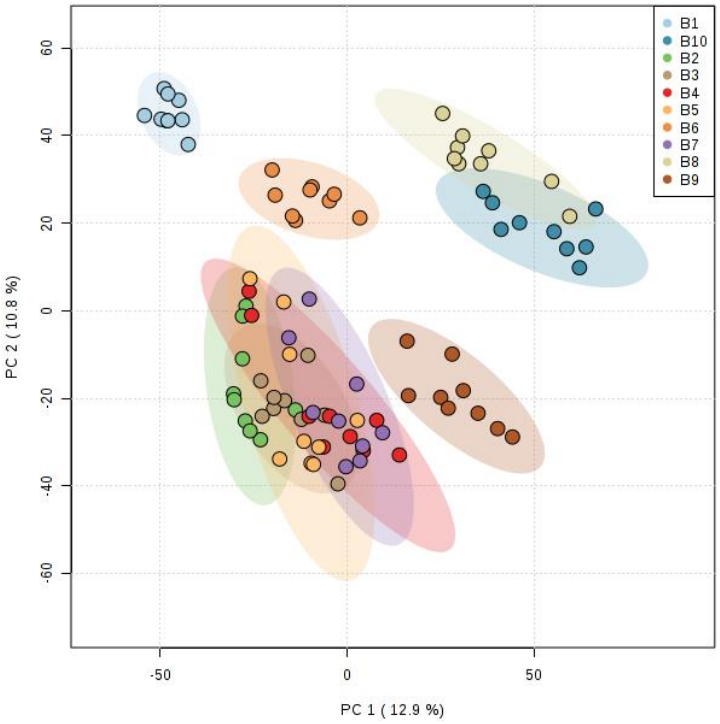
Mass spectral deconvolution, peak identification/annotation, and peak alignment was conducted using ChromaTOF software (version 4.32), from the LECO Corporation. Mass spectral deconvolution is used to reveal peaks with mixed mass spectra of co-eluting compounds and assigns an accurate mass spectrum to each individual analyte. Deconvolution was performed at a signal-to-noise (S/N) ratio of 300, with a minimum of three apexing peaks. Peak alignment functions by using the retention times and mass spectra to align peaks between a variety of samples, compiling a data matrix that contains the relative concentrations of all detected compounds, from all samples analysed, aligned with one another. Hence, possible retention time shifts were corrected across all samples by aligning identical mass spectra displaying similar retention times. Peaks were identified by comparison of their characteristic mass fragmentation patterns and retention times with various libraries compiled from previously injected standards.

3.3.6 Batch effect

A batch effect refers to a gradual change in the accuracy or precision of the results, which may exist due to non-biological systematic differences that occur due to small changes or variation during the extraction and analytical procedures. This effect has been well documented in large-scale metabolomic studies. Although the necessary precautions should always be taken to minimise possible batch effects, these may be inevitable, and thus need to be corrected for. Considering this, a drift correction step was applied, using quantile equating, to correct for any potential batch effects (Wang *et al.*, 2012). As shown in Figure 3-2, using the data collected from the QC samples analysed, a PCA (of which a brief description is given in section 4.1.3) was done both before and after batch correction, to visualise whether or not this was successful. According to this illustration, a batch effect is evident (Figure 3-2a), and after correction of the entire dataset (i.e. QC, patient and control

samples), better overlap between the batches occur (Figure 3-2b). Therefore, any differentiation of the patient/control groups investigated from this point onwards, and the metabolite makers identified, can be ascribed to variation due to a biological effect and not a batch effect.

(a)



(b)

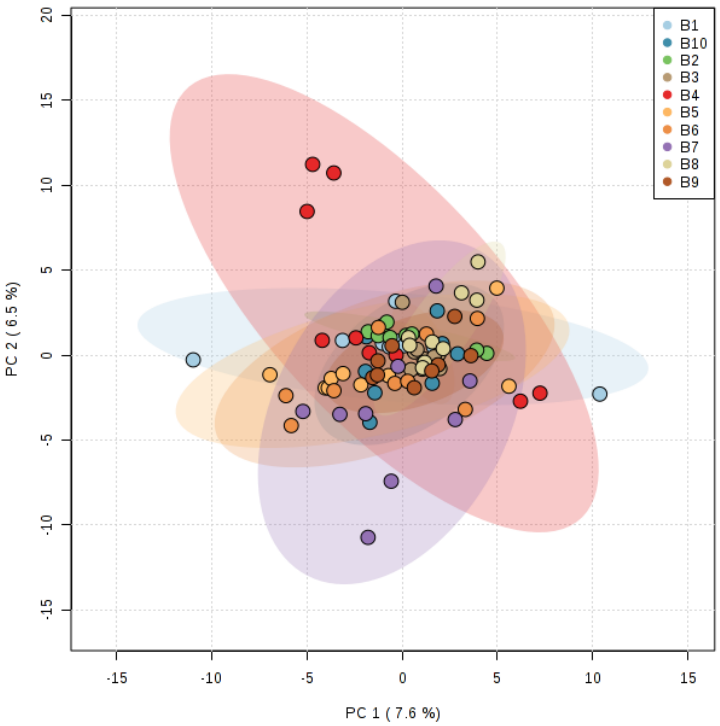


Figure 3-2: Batch correction. Two dimensional principle component analysis (PCA) scores plot of the GCxGC-TOFMS quality control data (a) before and (b) after a correcting the observed batch effect, using quantile equating.

3.4 RESULTS AND DISCUSSION

3.4.1 Repeatability

In order to achieve the first objective (section 1.2.2), the repeatability of the GCxGC-TOFMS apparatus used (analytical machine repeatability) and that of the methodology (analyst/extraction methodology repeatability), was determined.

The analytical machine repeatability was determined and validated by extracting and derivatising one randomly selected QC sample (as described in section 3.3.3), and injecting and analysing this sample repeatedly (n=10) using a GCxGC-TOFMS (see section 3.3.4). The data were processed (as described in section 3.3.5) and, using this dataset of ten injection repeats, a coefficient of variation (CV) value was determined for all detected compounds as well as for ten compounds occurring in the highest concentrations throughout the chromatographic run. Similarly, the repeatability of the analyst/extraction methodology was also determined. However, in this instance, all of the QC samples (n=18) were separately extracted, injected and analysed in triplicate, using the GCxGC-TOFMS, over 18 days (one sample per day) followed by data processing. Once again, the CV values for all the detected compounds, and for the same ten compounds occurring in the highest concentrations throughout the total chromatographic run, were determined.

The CV value is a parameter for measuring repeatability, and can be mathematically defined as $CV = (s \div \bar{x}) \times 100$, where "s" is the compound standard deviation and "$\bar{x}$" is the compound average. The CV value of a compound directly correlates with its detected concentration, thus a compound detected in high concentrations will inherently have a low CV value, and vice versa. The FDA recommends that the CV value be below 20% for GC-MS analyses (Food and Drug Administration, 2001), however, as per definition, metabolomics considers all compounds in an unbiased manner, and should therefore, in theory, also be unbiased towards compounds occurring in lower concentrations. For this reason, a hypothetical CV cut-off value of 50% was shown to be optimal for metabolomics investigations, in order to include compounds occurring in lower concentrations, in the final data matrix, for answering the biological questions (Schoeman *et al.*, 2012).

When considered from this perspective, 97.44% of all the detected compounds from the ten QC sample repeats fell below the hypothetical 50% CV value. Interestingly, 78.77% of all CV values were below 20%, further confirming the excellent GCxGC-TOFMS repeatability and its application for metabolomics investigations. Similarly, when evaluating the repeatability of the analyst/extraction methodology using the same approach, 91.61% of all detected compounds fell below the hypothetical CV cut-off of 50%, however far fewer of these had a

CV below 20%, which can be expected since each step of the extraction procedure could be a possible source for variation. Figure 3-3 graphically depicts the distribution of the CV values for all the detected compounds, following an organic acid extraction, derivatisation and GCxGC-TOFMS analyses, for both the analytical machine and analyst/extraction methodology repeatability.

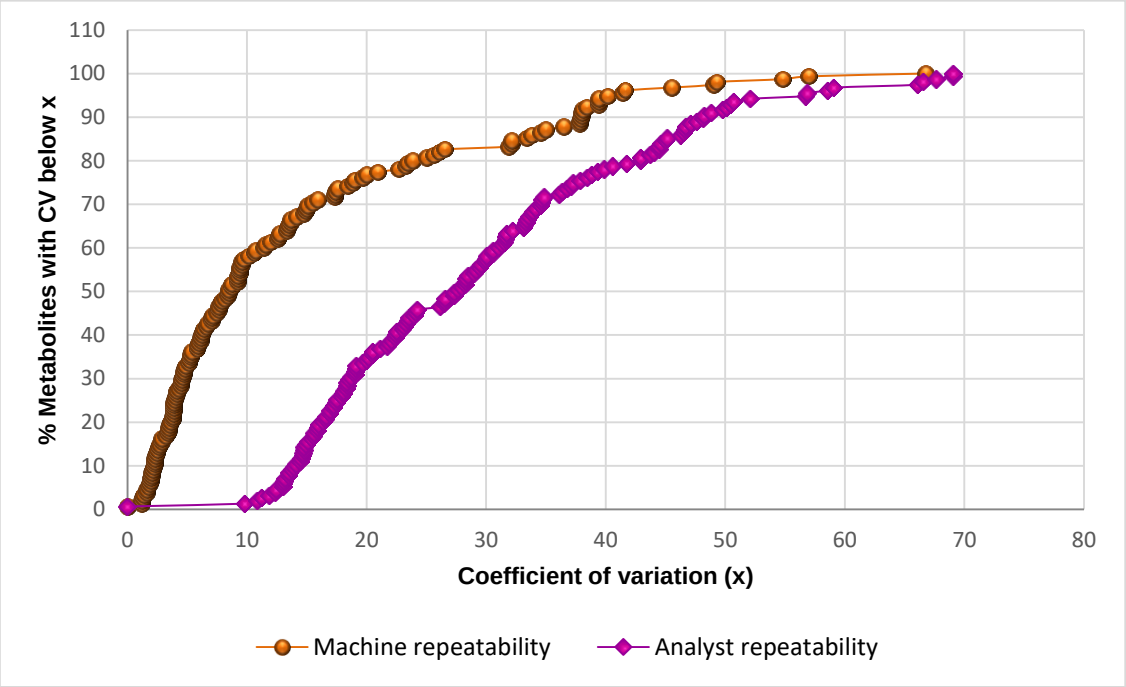


Figure 3-3: Distribution of the coefficient of variation (CV) values for the analytical machine (GCxGC-TOFMS) and analyst/extraction methodology repeatability.

As previously mentioned, since the concentration of a compound can influence its CV, for the purpose of showing the repeatability of this approach from a more traditional perspective, the ten compounds occurring in the highest concentrations throughout the chromatographic run with their respective relative concentrations and CV values are given in Table 3-1. From this it can be seen that all of the compounds fell below the FDA's recommended 20% cut-off value when considering the analytical machine and analyst/extraction methodology repeatability.

Table 3-1: Average concentration and coefficient of variation (CV) values, calculated for the ten most abundant compounds throughout the chromatographic run.

Metabolite name	Average concentration (mg/g creatinine) (standard error of the mean)	Machine repeatability (CV value)	Analyst/ extraction methodology repeatability (CV value)
(R*,R*)-2,3-Dihydroxybutanoic acid	59.978 (1.293)	1.62%	12.41%
1-Propene-1,2,3-tricarboxylic acid	203.971 (3.359)	4.89%	19.13%
2-Hydroxy-1,2,3-Propanetricarboxylic acid	238.054 (5.371)	2.67%	14.95%
2-Methyl-2-hydroxypropanoic acid	77.834 (1.946)	3.86%	17.17%
4-Hydroxybenzeneacetic acid	190.744 (2.708)	2.88%	15.41%
Butanedioic acid	146.171 (3.248)	1.22%	10.84%
Citramalic acid	4.283 (0.672)	1.17%	17.52%
Ethanedioic acid	92.393 (4.974)	5.15%	19.98%
Hydroxyacetic acid	94.050 (2.306)	4.66%	18.70%
N-Benzoylglycine	932.037 (13.99)	2.82%	15.11%

3.5 CONCLUSION

After a batch correction step using quantile equating, the generated GCxGC-TOFMS data showed excellent batch overlap of the data generated from the QC samples and, furthermore, acceptable repeatability, validating the methodology and data for use in further metabolomics investigations and addressing the proposed aims of this study (see section 1.2.1).

3.6 REFERENCES

Food and Drug Administration. Services, U.S.D.O.H.a.H. 2001. Guidance for industry: Bioanalytical method validation. Rockville: Center for Drug Evaluation and Research. 25 pages.

Schoeman, J.C., Du Preez, I. & Loots, D.T. 2012. A comparison of four sputum pre-extraction preparation methods for identifying and characterising M. tuberculosis using GCxGC-TOFMS metabolomics. *Journal of Microbiological Methods*, 91:301-311.

Wang, S.-Y., Kuo, C.-H. & Tseng, Y.J. 2012. Batch normalizer: A fast total abundance regression calibration method to simultaneously adjust batch and injection order effects in liquid chromatography/time-of-flight mass spectrometry-based metabolomics data and comparison with current calibration methods. *Analytical Chemistry*, 85(2):1037-1046.

CHAPTER 4: ADAPTATIONS OF MAN AND MICROBE IN ORDER TO OUTCOMPETE AND SURVIVE

This chapter has been published:

- Luies, L. & Loots, DT. (2016). Tuberculosis metabolomics reveals adaptations of man and microbe in order to outcompete and survive. *Metabolomics*, 12(3), 1-9.

4.1 ABSTRACT

Despite numerous research efforts to control TB, it is still regarded as a global pandemic. It is clear that the infectious agent responsible and its associated disease mechanisms remain poorly understood. Alternative research strategies are therefore urgently needed to better characterise active-TB, especially the adaptations of the host and microbe as they compete to survive. Using a GCxGC-TOFMS metabolomics approach, we identified new urinary TB metabolite markers induced by adaptations of the host metabolome and/or host–pathogen interactions. The most significant of these were the TB-induced changes resulting in abnormal host fatty acid and amino acid metabolism, in particular to tryptophan, phenylalanine and tyrosine, inducing a metabolite profile similar to that of patients suffering from phenylketonuria, mediated through changes to IFN-γ and possibly insulin. This also explains some of the symptoms associated with TB and provides clues to better treatment approaches.

4.2 INTRODUCTION

TB, a highly contagious bacterial disease caused by *M. tuberculosis*, is considered the leading cause of death globally from a single bacterial pathogen. This disease affects approximately one-third of the world's population, either in its active or latent form. The latest reports indicate 9 million new TB cases per annum, resulting in an estimated 1.8 million deaths (3 800 deaths a day) (World Health Organization, 2014).

TB transmission occurs through small infectious aerosol droplets that contain live *M. tuberculosis* bacilli. When inhaled into the lungs, they undergo rapid replication, before being engulfed by host macrophages. Following macrophage internalisation, *M. tuberculosis* avoids death by blocking phagolysosomal fusion, allowing the bacilli to persist in a non- or slowly-replicating state within the relatively hospitable niche of these phagosomes, where they can survive for decades (Philips & Ernst, 2012; Warner, 2014). Thus, although the host immune response is unable to eradicate *M. tuberculosis*, a fully immune-competent host can suppress the infection. This asymptomatic and non-infectious state is referred to as latent TB. However, when the immune system becomes compromised, latent TB develops into active TB — the symptomatic and highly infectious state of the disease. Common symptoms associated with active pulmonary TB include a bad cough with mucus, pleurisy, haemoptysis, dyspnoea, wheezing, weakness/fatigue, weight-loss, loss of appetite, chills/fever, and night sweats (Long *et al.*, 2002).

Despite numerous research efforts to date, since the discovery of *M. tuberculosis* in 1882, it is clear that this pathogen and its underlying disease mechanisms remain poorly understood. Thus, alternative research approaches, giving a different perspective of the disease, are urgently needed to better characterise TB, especially those investigating host–microbe interactions and each of their adaptations, as they compete with each other in order to survive. To this end, metabolomics has the potential to identify new disease markers, describing previously unidentified disease mechanisms, adaptations of the pathogen to environmental changes and how these react to antimicrobial treatments (Drapal *et al.*, 2014). "Metabolomics" refers to the unbiased identification and quantification of all metabolites present in a biological system, utilising highly sensitive and selective analytical methods, in conjunction with biostatistics for metabolite marker selection. This "omics" approach allows for a holistic view of a patient's metabolic state/metabolome, detecting alterations due to a specific perturbation, such as TB (De Villiers & Loots, 2013). Using this approach on sputum collected from TB patients, Du Preez and Loots (2013) were able to identify new metabolites from both *M. tuberculosis* and the infected host, thereby revealing previously unknown metabolic pathways and adaptations of each. These included: (a) the presence of a

citramalate cycle in *M. tuberculosis*; (b) interaction of this cycle with an upregulated glyoxylate cycle; and (c) increased utilisation of fatty acids and glutamate as additional carbon sources by *M. tuberculosis* during pulmonary infection. Adaptations of the human host included: (a) an alternative mechanism for H_2O_2 synthesis via glucose oxidation, in order to counteract the bacterial infection more efficiently; (b) inhibition of oxidative phosphorylation due to pronounced oxidative stress; and (c) elevated concentrations of various neurotransmitters and other metabolites related to some of the symptoms associated with TB (Du Preez & Loots, 2013).

In the current investigation, we applied a GCxGC-TOFMS metabolomics approach, to compare and differentiate culture-confirmed active TB-positive (n=38) and TB-negative healthy control (n=30) groups, based on the comparative detected changes to their urinary metabolomes. Apart from the fact that urine from TB patients is readily available, easily stored and can be collected by non-invasive techniques, it serves as an ideal matrix for identifying metabolites related to an altered TB host metabolome, since it contains metabolic signatures of many pathways that may change as a result of infection or disease.

4.3 MATERIALS AND METHODS

A summary of the experimental design is given in Figure 4-1, and described in detail below.

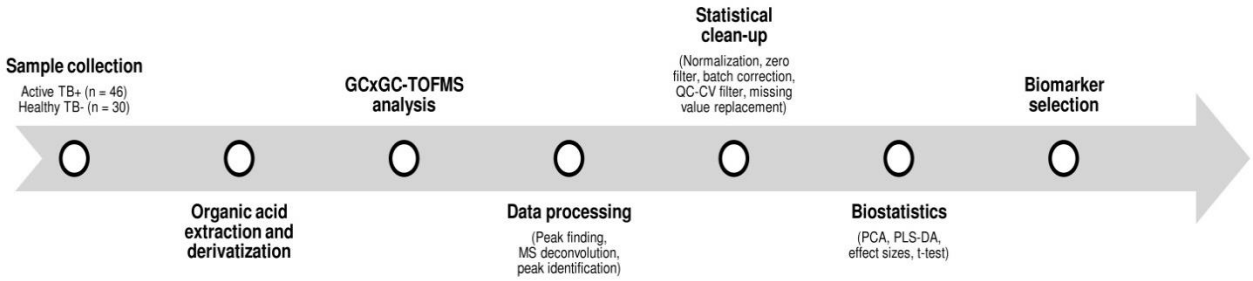


Figure 4-1: Summary of the experimental design. Subsequent to sample collection, an organic acid extraction and derivatisation of each sample extract was performed, followed by GCxGC-TOFMS analysis, data processing, statistical data clean-up and biostatistics. Hereafter metabolite markers were selected on the basis of various predefined statistical "cut offs".

4.3.1 Urine sample collection and storage

Urine samples were collected in standard vials, from 30 TB-negative healthy controls and 38 culture-confirmed active TB-positive patients from the same geographical area, prior to any treatment, after which they were placed on ice, immediately transported to the laboratory, and frozen at -20°C until analysis.

4.3.2 Sample analysis

Creatinine values for all urine samples were determined using a creatinine enzyme kit (Thermo Scientific; reference number 981845) and analysed on an Indiko Clinical Analyser, Type 863 (Thermo Scientific). The creatinine values are used to normalise metabolite concentrations, and to determine the volume of urine, internal standard, BSTFA, TMCS and pyridine needed for each extraction. An organic acid extraction of the collected patient urine samples were performed, analysed on a GCxGC-TOFMS and processed following the methods described in sections 3.3.3, 3.3.4 and 3.3.5, respectively.

4.3.3 Statistical data analyses

Prior to statistical data analyses, a standard metabolomics data clean-up procedure was applied. All compounds were normalised relative to the internal standard by calculating the relative concentration of each. Following this, variables showing no variation between the groups were removed, and a data filter was applied to each variable to eliminate those with more than 50% zero values in each group (Smuts *et al.*, 2013). Using the QC samples, quantile equating was applied to correct for any batch effects (Wang *et al.*, 2012). Hereafter, a 50% QC–CV filter was applied. Lastly, as most zero values present in the dataset were due to low abundance rather than being absent, they were replaced with a value calculated as half of the smallest value (detection limit) present in the dataset (Schoeman *et al.*, 2012).

Various multivariate (PCA; PLS-DA) and univariate (effect sizes; unpaired *t*-test) statistical analyses were applied using MetaboAnalyst, a metabolomics web-based server based on the statistical package “R” (version 2.10.0). PCA is an unsupervised method used to determine whether or not a natural grouping or differentiation exists between sample groups. PCA is based on a mathematical procedure that transforms possibly related variables (metabolites) into a smaller number of unrelated variables, known as the principal components (PCs). PC 1 accounts for the most variance in the data and each following PC (PC 2, PC 3, etc.) accounts for the next highest variance of the remaining data. Using one PC per axis, a PCA is visualised as a scores plot, and the variables/metabolites best describing this differentiation are ranked according to their respective modelling powers (Olivier & Loots, 2012). PLS-DA, on the other hand, is a supervised method used to determine group membership of an individual sample, based on its metabolic profile. It utilises multivariate regression techniques to extract, via linear combination of original variables (X), the information that can predict the class or group membership (Y). Thus, it identifies those variables/metabolites that best characterise the differentiated sample groups. Each metabolite is ranked according to the variables’ influence on the projection (VIP)

parameter, which directly corresponds to its importance (Du Preez & Loots, 2013). Effect sizes are used to compare each variable individually between the groups, without taking sample size into account, and indicate practical/clinical significance. For this parametric dataset, an effect size >0.5 indicates a moderate effect. The unpaired *t*-test establishes statistical significance (*P*-value) by determining whether or not the averages of two groups differ. Conventionally, a *P*-value <0.05 is considered significant.

4.4 RESULTS AND DISCUSSION

Figure 4-2 shows clear PCA differentiation of the TB-positive and TB-negative healthy control sample groups on the basis of the GCxGC-TOFMS metabolomics data collected. The total amount of variance explained by the first two principal components (PCs) (R_2X cum) was 71.0%, of which PC1 accounted for 62.5%, and PC2 accounted for 8.5%. Hereafter, a PLS-DA model was built and showed a modelling parameter R^2Y (cum) of 93.6%, indicative of the total explained variation of the response Y, and a Q^2 (cum), corresponding to the cross-validated variation accounted for by the response Y, of 81.0%.

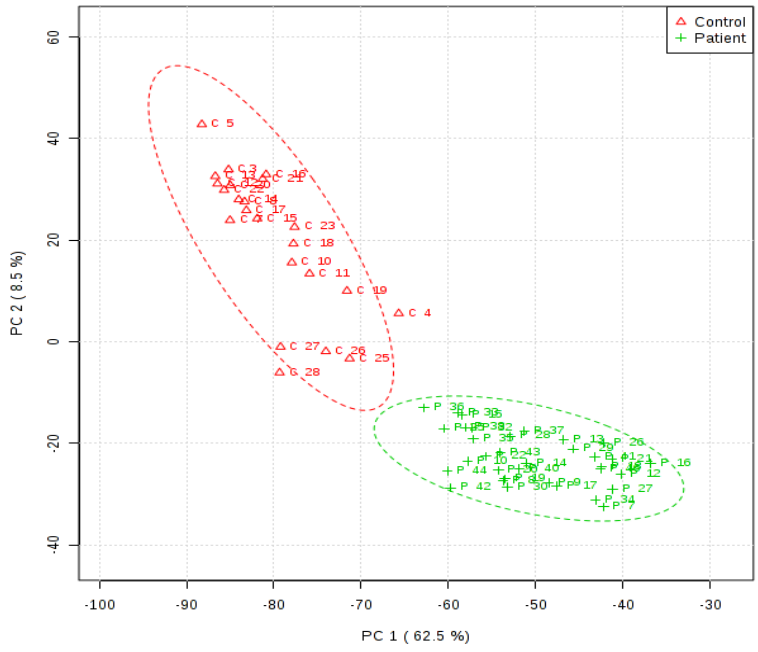


Figure 4-2: Principle component analysis (PCA) differentiation of GCxGC-TOFMS organic acid analysis data, of urine samples collected from active TB-positive and TB-negative patients. A plot of principle component 1 versus principle component 2 of the TB-negative healthy control (denoted by C) and active TB-positive patient (denoted by P) groups, showing clear differentiation of these two groups as a result of variation in their metabolite profiles. The variances accounted for are indicated in parenthesis.

As indicated in Figure 4-3, 12 compounds with a PCA modelling power >0.5 (Brereton, 2003), a PLS-DA VIP >1.0 (Smuts *et al.*, 2013), an effect size >0.5, and a *t*-test with a *P*-value <0.05 (Ellis & Steyn, 2003) were selected as the metabolite markers that best explained the variation between the analysed groups (listed in Table 4-1). Box-and-whisker plots of these 12 compounds are provided in Appendix B. This multi-statistical approach is based on the assumption that the different statistical tests correct/compensate for each other's flaws with the elimination of false-positive compounds (Venter *et al.*, 2015).

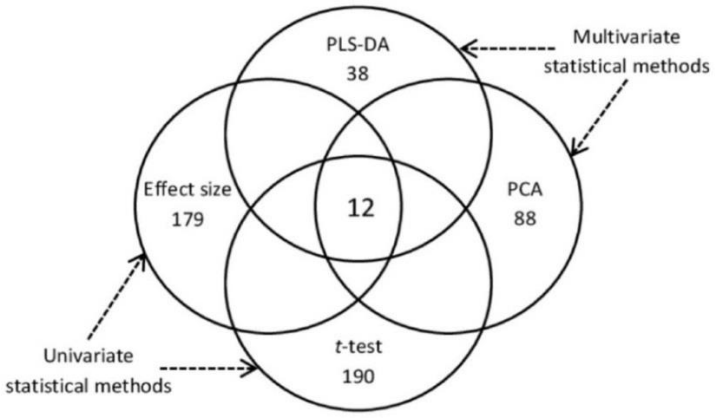


Figure 4-3: Venn-diagram indicating compound selection using a multi-statistical approach. A total of 12 compounds were identified as most significant for describing the variation between the TB-negative healthy control and TB-positive patient groups.

Table 4-1: The 12 urinary metabolite markers that best explain the variation detected in the urine samples of the TB-negative healthy control and TB-positive patients, ranked alphabetically.

Metabolite name	TB-negative healthy controls: Average concentration (mg/g creatinine) (standard error of the mean)	Active TB-positive patients: Average concentration (mg/g creatinine) (standard error of the mean)	PCA (Power)	PLS-DA (VIP)	Effect sizes (*d*-value)	*t*-test (*P*-value)
2-C-Methylglycerol	0.071 (0.020)	0.727 (0.084)	0.666	2.432	1.319	<0.001
2-Octenoic acid	0.011 (0.007)	0.114 (0.015)	0.568	2.366	1.150	<0.001
5-Hydroxyhexanoic acid	0.045 (0.021)	0.594 (0.127)	0.534	2.312	0.730	<0.001
5-Hydroxyhydantoin	0.049 (0.019)	0.287 (0.045)	0.522	1.961	0.902	<0.001
Glycerol monostearate	0 (0)	1.079 (0.247)	0.688	3.804	0.737	<0.001
Indole-3-carboxylic acid	0.137 (0.026)	0.929 (0.130)	0.642	1.048	1.032	<0.001
Kynurenic acid	46.055 (0.812)	23.312 (3.872)	0.701	1.017	0.993	<0.001
L-Rhamnulose	0.028 (0.021)	0.302 (0.041)	0.718	2.838	1.146	<0.001
Oxalic acid	0.030 (0.009)	0.143 (0.024)	0.821	2.147	0.813	<0.001
Phenylacetic acid	0.214 (0.034)	1.137 (0.114)	0.719	1.028	1.372	<0.001
Quinolinic acid	0.615 (0.046)	6.422 (1.042)	0.743	1.151	0.942	<0.001
Ribitol	1.089 (0.276)	4.768 (0.745)	0.663	2.694	0.838	<0.001

Most of the identified metabolites can be explained by changes in the host metabolome in response to TB, supporting previously speculated mechanisms and shedding light on hitherto unknown mechanisms related to *M. tuberculosis* pathogenesis, and host–microbe interactions and adaptations, as will be described below.

M. tuberculosis infection results in a host immune response, signalling macrophages to engulf the invading bacteria. Accompanying this is an inflammatory response resulting in the release of tumour necrosis factor alpha (TNF-α) and IFN-γ (Philips & Ernst, 2012), the latter of which has been shown to elevate indoleamine 2,3-dioxygenase 1 (IDO1), which in turn is also associated with the prognosis of various pathological conditions, including TB (Suzuki *et al.*, 2012). This IDO1-upregulating host response is thought to deplete the bacterial environment of tryptophan, an amino acid essential to *M. tuberculosis* growth (Warner, 2014), and also expected to result in alterations in the downstream metabolites in this pathway (see Figure 4-4). Quinolinic acid, for instance, a neurotoxic metabolite formed via tryptophan catabolism, has been linked to inflammatory diseases, and is known to be elevated through IDO1 activation (Heyes *et al.*, 1992). A study by Asp *et al.* (2011) indicated that TNF-α and IFN-γ reduced the levels of kynurenic acid (a neuroprotective agent), by inhibiting kynurenine aminotransferases (KATs) synthesis (Asp *et al.*, 2011). For these reasons, one would expect elevations in quinolinic acid and reductions in kynurenic acid in the TB-infected host, which is what we see from our metabolomics data. Additionally, indole-3-carboxylic acid, the dehydrogenation product of tryptophan (Lübbe *et al.*, 1983), was elevated approximately nine times in the TB-positive group, which may serve as an additional means by which the host attempts to reduce the much-needed tryptophan for *M. tuberculosis* survival. Interestingly, although not detected as one of the 12 metabolite markers using the multi-statistical method described, tryptophan was found to be elevated in the TB-positive group (0.667 vs. 0.096 mmol/L; P=0.041), either because *M. tuberculosis* can synthesise tryptophan *de novo* (Warner, 2014), or from another phenomenon explained below.

Another important observation is the presence of more than five times the normal levels of phenylacetic acid detected in the urine of TB-positive patients. Phenyllactic acid, although not identified as a metabolite marker, was also significantly elevated in the TB patients (4.107 vs. 0.009 mmol/L; P<0.0001). Elevations in these two metabolites are indicative of phenylalanine accumulation in the TB-positive patients, similarly to what is seen in phenylketonuria (PKU) patients. Normally, phenylalanine is metabolised to tyrosine via phenylalanine hydroxylase (PAH), using tetrahydrobiopterin (BH_4) as a cofactor. If PAH or BH_4 are deficient for some reason, phenylalanine will accumulate and undergo transamination to produce phenylpyruvic acid, which is then either oxidised to phenylacetic acid or reduced to phenyllactic acid (Puri, 2006). Since the usual tryptophan

(5-hydroxyindole-acetic acid (0.155 vs. 0.917 mmol/L; P>0.05)), and tyrosine (N-acetyltyrosine (0.014 vs. 0.117 mmol/L; P<0.0001); homovanillic acid (7.356 vs. 19.642 mmol/L; P<0.0001)) catabolism metabolites were elevated in the TB-positive group — and considering that both of these amino acid catabolic pathways require BH_4 — a respective catabolic enzyme defect or BH_4 deficiency can be ruled out. Elevated concentrations of these amino acids are more likely due to compromised insulin production, which further supports earlier evidence for the association between TB and diabetes (Dooley & Chaisson, 2009). This is further substantiated by the recently described detection of D-gluconic acid δ-lactone in the sputum of TB-positive patients, thought to be induced by elevated glucose levels, in addition to elevated levels of normetanephine, a metabolite of norepinephrine, associated with this response and also formed from tyrosine (Du Preez & Loots, 2013). Insulin not only influences carbohydrate but also amino acid and protein metabolism, as a means to preserve amino acids for other essential anabolic processes. Since tryptophan and phenylalanine, as well as tyrosine to a certain extent, are essential amino acids, their concentrations in the human systemic circulation is dependent on protein consumption and uptake from the portal circulation system. These amino acids can also be released from tissue reservoirs or circulating proteins, which is typically what happens during wasting, a well-known occurrence in patients suffering from active TB, and also linked to reduced insulin secretion. The quantities of dietary phenylalanine, tyrosine and tryptophan entering systemic circulation are regulated by their catabolic hepatic enzymes, namely PAH, tyrosine aminotransferase (TAT), and tryptophan dioxygenase (TDO), respectively, which are induced by various factors, including reduced insulin. Although recent developments have shed light on the association between compromised insulin secretion and altered amino acid metabolism, the exact mechanism by which this occurs remains unknown (Adams, 2011; Cansev & Wurtman, 2007). In a previous metabolomics investigation, in which insulin-sensitive and insulin-resistant subjects were compared, both elevated phenylalanine and tyrosine concentrations were factors differentiating these groups (Tai *et al.*, 2010) and serves as further evidence for the role of reduced insulin in the elevation of these amino acids. Moreover, reduced insulin secretion also results in free fatty acid oxidation *in vivo* (Adams, 2011). When the body uses fat as the primary energy source, it is broken down and released into the bloodstream, resulting in increased urinary fatty acids and their breakdown products, as was detected in our study. These fatty acids and breakdown products included glycerol monostearate, 2-octenoic acid and 5-hydroxyhexanoic acid. This could potentially explain the weight-loss associated with TB if these fatty acids are sourced from the body's fat stores (Kairamkonda *et al.*, 2003; Niu *et al.*, 2012). This host response may additionally be driven by various bacterial mechanisms. Upon infection, *M. tuberculosis* initially utilises host glucose and triacylglycerides as primary carbon sources, under aerobic conditions. However, with increased IFN-γ production, the glucose-deficient macrophages become

hypoxic, and access to iron and its usual carbon sources become limited, forcing *M. tuberculosis* to induce an iron-scavenging reaction and utilise host cholesterol and fatty acids exclusively as the primary carbon sources. Hence, *M. tuberculosis* is able to adapt its metabolic activity in order to survive and persist within its human host, as well as to maximise energy production from alternative nutrient sources (Philips & Ernst, 2012; Rhee *et al.*, 2011; Singh *et al.*, 2012). *M. tuberculosis* can also co-catabolise many carbon sources simultaneously via compartmentalisation using multiple pathways (De Carvalho *et al.*, 2010).

The elevated 5-hydroxyhydantoin detected in the TB-positive group is associated with DNA damage induced by oxidative stress, a common occurrence in TB patients, primarily due to H_2O_2 production by the macrophages in an attempt to overcome the bacterial invasion (Du Preez & Loots, 2013). The ascorbic acid (vitamin C) oxidation product, oxalic acid (Robitaille *et al.*, 2009), was observed in elevated concentrations in the TB-positive group. This is most likely due to ascorbic acid intake by the TB patients as this vitamin is not naturally synthesised by humans. Apart from this, foodborne moulds of *Aspergillus* species proliferate in the gastrointestinal tract of the host, as a secondary condition to the primary TB disease state (Singh & Toskes, 2004) and are also known to produce oxalic acid (Alam *et al.*, 2002).

The ribitol, L-rhamnulose and 2-C-methylglycerol (2MG) detected in elevated concentrations in the TB patients, are most likely derived directly from *M. tuberculosis,* as these are components of its cell walls (Mikusova *et al.*, 1996; Silhavy *et al.*, 2010; Velagapudi *et al.*, 2010). Furthermore, 2MG is also an intermediate of the 2-C-methyl-D-erythritol-4-phosphate (MEP) pathway, unique to *M. tuberculosis*, enabling this pathogen to synthesise essential metabolites derived from isoprenoid compounds, and is considered a target for the development of new anti-TB drugs, since homologous enzymes are absent in humans (Eoh *et al.*, 2007; Kholodar & Murkin, 2013).

Finally, previous studies have proposed that the nutritional imbalance (resulting in wasting) caused by *M. tuberculosis* trigger autophagy, a homeostatic intracellular degradation process targeting intracellular *M. tuberculosis*. During this process, macromolecules and organelles are fused with lysosomes and degraded, in order to sustain cellular anabolic processes (Choi *et al.*, 2013; Deretic & Levine, 2009; Goletti *et al.*, 2013; Gutierrez *et al.*, 2004). Once these macromolecules have been degraded, monomeric units, such as amino acids, fatty acids and DNA components, are either exported, excreted or reused (Goletti *et al.*, 2013; Mizushima, 2007). Various studies have also shown that reduced insulin secretion (Mao *et al.*, 2011) or increased glucagon (Deter & De Duve, 1967) may induce autophagy as an additional host response to TB, contributing to the elevated amino acids, fatty acids, DNA breakdown products, and bacterial components detected in the TB patients.

Apart from the novel contribution of this discovery towards better understanding host–microbe interactions/adaptations, these markers may help explain some of the symptoms associated with TB, including neurological abnormalities (since quinolinic acid is a N-methyl-D-aspartate receptor agonist) (Heyes *et al.*, 1992), vomiting, nausea, diarrhoea, drowsiness, fatigue, loss of appetite and weight-loss (associated with elevated levels of phenylacetic acid) (Sherwin & Kennard, 1919). Co-administration of anti-TB drugs with melatonin may assist in reducing elevated quinolinic acid (Cabrera *et al.*, 2000), since melatonin has also been shown to lower a number of toxic metabolites associated with the side-effects of anti-TB drugs and also increase its efficacy for eliminating *M. tuberculosis* (Loots *et al.*, 2005). Furthermore, BH_4 is given as a treatment for PKU patients in an attempt to reduce phenylacetic acid (McInnes *et al.*, 1984), hence co-administration of anti-TB drugs with this merits consideration. Lastly, considering the association between TB and diabetes, co-administration of anti-TB drugs with antidiabetic drugs, such as metformin, could also be investigated. Singhal *et al.* (2014) showed how treatment of TB patients with metformin reduced intracellular *M. tuberculosis* growth by inducing reactive oxygen species production, facilitated phagolysosomal fusion by enhancing the host immune response, reduced the TB-induced tissue pathology and inflammation, decreased TB severity, increased conventional anti-TB drug efficacy, improved clinical outcome, and reduced the incidence of latent TB (Singhal *et al.*, 2014).

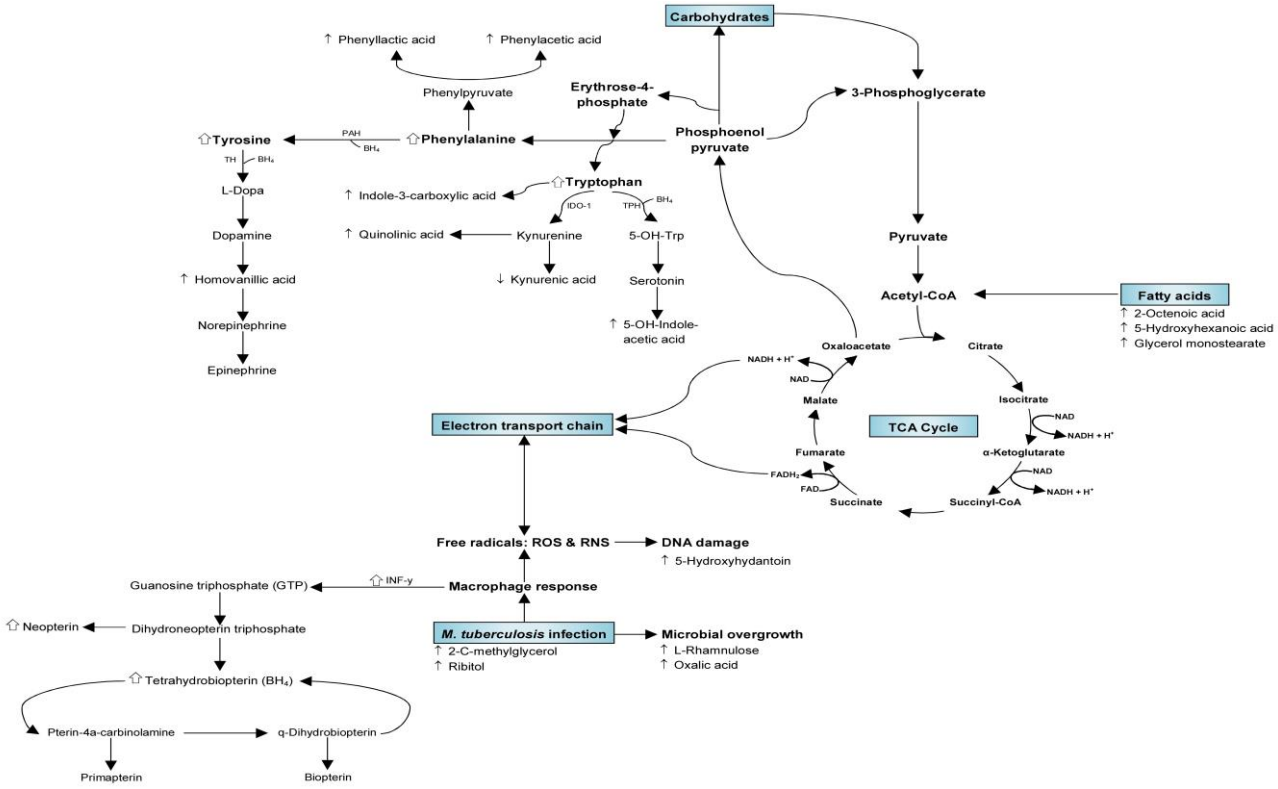


Figure 4-4: Altered host metabolome induced by M. tuberculosis. A schematic representation of the metabolite markers (and their associated metabolism) identified in the urine of TB-positive patients, which are either increased (↑) or decreased (↓) relative to the healthy TB-negative group. Metabolites hypothesised/previously indicated in the literature to be elevated, are indicated with open arrows.

4.5 CONCLUSION

Metabolomics is considered a multistage process, often aimed at collecting data for the purpose of formulating new hypotheses and elucidating previously unknown disease mechanisms and/or host–microbe interactions and adaptations. This metabolomics study compared a healthy TB-negative group to that of an active TB-positive patient group, to identify the host and microbe interactions and adaptations, in response to TB infection. The 12 identified urinary metabolites, best explaining the differences occurring between these groups, could directly be linked to an abnormal host fatty acid and amino acid metabolism, especially those affecting tryptophan, phenylalanine and tyrosine, and inducing a metabolite profile similar to that of a PKU patient. This is due to changes to IFN-γ and possibly insulin, and may also explain some of the TB-associated symptoms, providing clues to better treatment approaches.

4.6 REFERENCES

Adams, S.H. 2011. Emerging perspectives on essential amino acid metabolism in obesity and the insulin-resistant state. *Advances in Nutrition (Bethesda, Md.)*, 2(6):445-456.

Alam, S., Khalil, S., Ayub, N. & Rashid, M. 2002. In vitro solubilization of inorganic phosphate by phosphate solubilizing microorganisms (PSM) from maize rhizosphere. *International Journal of Agriculture and Biology*, 4(4):454-458.

Asp, L., Johansson, A.S., Mann, A., Owe-Larsson, B., Urbanska, E.M., Kocki, T., Kegel, M., Engberg, G., Lundkvist, G.B. & Karlsson, H. 2011. Effects of pro-inflammatory cytokines on expression of kynurenine pathway enzymes in human dermal fibroblasts. *Journal of Inflammation (London, England)*, 8:25-9255-9258-9225.

Brereton, R.G. 2003. Chemometrics: Data analysis for the laboratory and chemical plant. West Sussex, England: John Wiley & Sons.

Cabrera, J., Reiter, R.J., Tan, D.-X., Qi, W., Sainz, R.M., Mayo, J.C., Garcia, J.J., Kim, S.J. & El-Sokkary, G. 2000. Melatonin reduces oxidative neurotoxicity due to quinolinic acid: In vitro and in vivo findings. *Neuropharmacology*, 39(3):507-514.

Cansev, M. & Wurtman, R.J. 2007. Aromatic amino acids in the brain. Handbook of neurochemistry and molecular neurobiology. Springer. p. 59-97).

Choi, A.M.K., Ryter, S.W. & Levine, B. 2013. Autophagy in human health and disease. *New England Journal of Medicine*, 368(7):651-662.

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

De Villiers, L. & Loots, D.T. 2013. Using metabolomics for elucidating the mechanisms related to tuberculosis treatment failure. *Current Metabolomics*, 1(4):306-317.

Deretic, V. & Levine, B. 2009. Autophagy, immunity, and microbial adaptations. *Cell Host & Microbe*, 5(6):527-549.

Deter, R.L. & De Duve, C. 1967. Influence of glucagon, an inducer of cellular autophagy, on some physical properties of rat liver lysosomes. *The Journal of Cell Biology*, 33(2):437-449.

Dooley, K.E. & Chaisson, R.E. 2009. Tuberculosis and diabetes mellitus: Convergence of two epidemics. *The Lancet Infectious Diseases*, 9(12):737-746.

Drapal, M., Perez-Fons, L., Wheeler, P.R. & Fraser, P.D. 2014. The application of metabolite profiling to *Mycobacterium spp*.: Determination of metabolite changes associated with growth. *Journal of Microbiological Methods*, 106:23-32.

Du Preez, I. & Loots, D.T. 2013. New sputum metabolite markers implicating adaptations of the host to *Mycobacterium tuberculosis*, and vice versa. *Tuberculosis*, 93(3):330-337.

Ellis, S.M. & Steyn, H.S. 2003. Practical significance (effect sizes) versus or in combination with statistical significance (p-values): Research note. *Management Dynamics: Journal of the Southern African Institute for Management Scientists*, 12(4):p. 51-53.

Eoh, H., Brown, A.C., Buetow, L., Hunter, W.N., Parish, T., Kaur, D., Brennan, P.J. & Crick, D.C. 2007. Characterization of the *Mycobacterium tuberculosis* 4-diphosphocytidyl-2-C-methyl-D-erythritol synthase: Potential for drug development. *Journal of Bacteriology*, 189(24):8922-8927.

Goletti, D., Petruccioli, E., Romagnoli, A., Piacentini, M. & Fimia, G.M. 2013. Autophagy in *Mycobacterium tuberculosis* infection: A passepartout to flush the intruder out? *Cytokine & Growth Factor Reviews*, 24(4):335-343.

Gutierrez, M.G., Master, S.S., Singh, S.B., Taylor, G.A., Colombo, M.I. & Deretic, V. 2004. Autophagy is a defense mechanism inhibiting BCG and *Mycobacterium tuberculosis* survival in infected macrophages. *Cell*, 119(6):753-766.

Heyes, M.P., Saito, K. & Markey, S.P. 1992. Human macrophages convert L-tryptophan into the neurotoxin quinolinic acid. *Biochemical Journal*, 283:633-635.

Kairamkonda, V., Dalzell, M., Losty, P.D. & Davidson, C. 2003. Perforated duodenal ulcer disclosing medium chain acyl-CoA dehydrogenase deficiency. *Archives of Disease in Childhood*, 88(1):88-89.

Kholodar, S.A. & Murkin, A.S. 2013. DXP reductoisomerase: Reaction of the substrate in pieces reveals a catalytic role for the nonreacting phosphodianion group. *Biochemistry*, 52(13):2302-2308.

Long, N.H., Diwan, V.K. & Winkvist, A. 2002. Difference in symptoms suggesting pulmonary tuberculosis among men and women. *Journal of Clinical Epidemiology*, 55(2):115-120.

Loots, D.T., Wiid, I.J., Page, B.J., Mienie, L.J. & Helden, P.D. 2005. Melatonin prevents the free radical and MADD metabolic profiles induced by antituberculosis drugs in an animal model. *Journal of Pineal Research*, 38(2):100-106.

Lübbe, C., Van Pée, K.-H., Salcher, O. & Lingens, F. 1983. The metabolism of tryptophan and 7-chlorotryptophan in Pseudomonas pyrrocinia and Pseudomonas aureofaciens. *Hoppe-Seyler´ s Zeitschrift für physiologische Chemie*, 364(1):447-454.

Mao, F., Chen, T., Zhao, Y., Zhang, C., Bai, B., Zhao, S., Xu, Z. & Shi, C. 2011. Insulin resistance: A potential marker and risk factor for active tuberculosis? *Medical Hypotheses*, 77(1):66-68.

McInnes, R.R., Kaufman, S., Warsh, J.J., Van Loon, G.R., Milstien, S., Kapatos, G., Soldin, S., Walsh, P., MacGregor, D. & Hanley, W.B. 1984. Biopterin synthesis defect. Treatment with L-dopa and 5-hydroxytryptophan compared with therapy with a tetrahydropterin. *The Journal of Clinical Investigation*, 73(2):458-469.

Mikusova, K., Mikus, M., Besra, G.S., Hancock, I. & Brennan, P.J. 1996. Biosynthesis of the linkage region of the mycobacterial cell wall. *The Journal of Biological Chemistry*, 271(13):7820-7828.

Mizushima, N. 2007. Autophagy: Process and function. *Genes & Development*, 21(22):2861-2873.

Niu, J., Pi, Z., Yue, H., Wang, Y., Yu, Q. & Liu, S. 2012. Effect of ginseng polysaccharide on the urinary excretion of type 2 diabetic rats studied by liquid chromatography–mass spectrometry. *Journal of Chromatography B*, 907:7-12.

Olivier, I. & Loots, D.T. 2012. A metabolomics approach to characterise and identify various *Mycobacterium* species. *Journal of Microbiological Methods*, 88(3):419-426.

Philips, J.A. & Ernst, J.D. 2012. Tuberculosis pathogenesis and immunity. *Annual Review of Pathology: Mechanisms of Disease*, 7:353-384.

Puri, D. 2006. Textbook of Medical Biochemistry: Clinical cases, Chapter 13.2. Textbook of medical biochemistry. New-Delhi, India: Elsevier. p. 376).

Rhee, K.Y., De Carvalho, L.P.S., Bryk, R., Ehrt, S., Marrero, J., Park, S.W., Schnappinger, D., Venugopal, A. & Nathan, C. 2011. Central carbon metabolism in *Mycobacterium tuberculosis*: An unexpected frontier. *Trends in Microbiology*, 19(7):307-314.

Robitaille, L., Mamer, O.A., Miller Jr, W.H., Levine, M., Assouline, S., Melnychuk, D., Rousseau, C. & Hoffer, L.J. 2009. Oxalic acid excretion after intravenous ascorbic acid administration. *Metabolism*, 58(2):263-269.

Schoeman, J.C., Du Preez, I. & Loots, D.T. 2012. A comparison of four sputum pre-extraction preparation methods for identifying and characterising M. tuberculosis using GCxGC-TOFMS metabolomics. *Journal of Microbiological Methods*, 91:301-311.

Sherwin, C.P. & Kennard, K.S. 1919. Toxicity of phenylacetic acid. *Journal of Biological Chemistry*, 40(2):259-264.

Silhavy, T.J., Kahne, D. & Walker, S. 2010. The bacterial cell envelope. *Cold spring harbor Perspectives in Biology*, 2(5):a000414.

Singh, V., Jamwal, S., Jain, R., Verma, P., Gokhale, R. & Rao, K.V.S. 2012. *Mycobacterium tuberculosis*-driven targeted recalibration of macrophage lipid homeostasis promotes the foamy phenotype. *Cell Host & Microbe*, 12(5):669-681.

Singh, V.V. & Toskes, P.P. 2004. Small bowel bacterial overgrowth: Presentation, diagnosis, and treatment. *Current Treatment Options in Gastroenterology*, 7(1):19-28.

Singhal, A., Jie, L., Kumar, P., Hong, G.S., Leow, M.K., Paleja, B., Tsenova, L., Kurepina, N., Chen, J., Zolezzi, F., Kreiswirth, B., Poidinger, M., Chee, C., Kaplan, G., Wang, Y.T. & De Libero, G. 2014. Metformin as adjunct antituberculosis therapy. *Science Translational Medicine*, 6(263):263ra159.

Smuts, I., Van Der Westhuizen, F.H., Louw, R., Mienie, L.J., Engelke, U.F.H., Wevers, R.A., Mason, S., Koekemoer, G. & Reinecke, C.J. 2013. Disclosure of a putative biosignature for respiratory chain disorders through a metabolomics approach. *Metabolomics*, 9(2):379-391.

Suzuki, Y., Suda, T., Asada, K., Miwa, S., Suzuki, M., Fujie, M., Furuhashi, K., Nakamura, Y., Inui, N., Shirai, T., Hayakawa, H., Nakamura, H. & Chida, K. 2012. Serum indoleamine 2,3-dioxygenase activity predicts prognosis of pulmonary tuberculosis. *Clinical and Vaccine Immunology: Cvi*, 19(3):436-442.

Tai, E.S., Tan, M.L.S., Stevens, R.D., Low, Y.L., Muehlbauer, M.J., Goh, D.L.M., Ilkayeva, O.R., Wenner, B.R., Bain, J.R. & Lee, J.J.M. 2010. Insulin resistance is associated with a metabolic profile of altered protein metabolism in Chinese and Asian-Indian men. *Diabetologia*, 53(4):757-767.

Velagapudi, V.R., Hezaveh, R., Reigstad, C.S., Gopalacharyulu, P., Yetukuri, L., Islam, S., Felin, J., Perkins, R., Boren, J., Oresic, M. & Backhed, F. 2010. The gut microbiota modulates host energy and lipid metabolism in mice. *Journal of Lipid Research*, 51(5):1101-1112.

Venter, L., Lindeque, Z., Van Rensburg, P.J., Van Der Westhuizen, F., Smuts, I. & Louw, R. 2015. Untargeted urine metabolomics reveals a biosignature for muscle respiratory chain deficiencies. *Metabolomics*, 11(1):111-121.

Wang, S.-Y., Kuo, C.-H. & Tseng, Y.J. 2012. Batch normalizer: A fast total abundance regression calibration method to simultaneously adjust batch and injection order effects in liquid chromatography/time-of-flight mass spectrometry-based metabolomics data and comparison with current calibration methods. *Analytical Chemistry*, 85(2):1037-1046.

Warner, D.F. 2014. *Mycobacterium tuberculosis* metabolism. *Cold Spring Harbor perspectives in medicine*, 5(4):10.1101/cshperspect.a021121.

World Health Organization. 2014. Tuberculosis: WHO global tuberculosis report 2014. Geneva, Switzerland (WHO Press).

CHAPTER 5: URINARY METABOLITE MARKERS CHARACTERISING TUBERCULOSIS TREATMENT FAILURE

This chapter has been submitted for publication:

- Luies, L., Mienie, J., Motshwane, C., Ronacher, K., Walzl, G. & Loots, DT. (2017). Urinary metabolite markers characterising tuberculosis treatment failure. Submitted for publication to *Metabolomics* (Manuscript number: MEBO-D-17-00069).

5.1 ABSTRACT

Considering that approximately 15% of the nine million new TB cases reported per annum are not treated successfully, new, distinctive and specific biomarkers are needed to better characterise the biological basis of a poor treatment outcome. Urine samples from 36 active pulmonary TB patients were collected at baseline (time of diagnosis), during treatment (weeks 1, 2 and 4) and two weeks after treatment completion (week 26). These samples were divided into successful (cured) and unsuccessful (failed) treatment outcome groups and analysed using a GCxGC-TOFMS metabolomics research approach. The metabolite data collected showed clear differentiation of the cured and failed treatment outcome groups using the samples collected at the time of diagnosis, i.e. before any treatment was administered. The treatment failure group was characterised by an imbalanced gut microbiome, in addition to elevated levels of metabolites associated with abnormalities in the long-chain fatty acid β-oxidation pathway, accompanied by reduced L-carnitine and short-chain fatty acids, indicative of a mitochondrial trifunctional protein defect in particular. Furthermore, an altered amino acid metabolism was also observed in these patients, which confirms previous findings and associations to increased IFN-γ due to the host's immune response to *M. tuberculosis* and a compromised insulin secretion.

5.2 INTRODUCTION

TB remains the world's foremost cause of death from a single bacterial agent, which is alarming as it is considered curable. Nearly 10.4 million new TB cases are reported per annum, of which almost 15% are not treated successfully, resulting in approximately 1.8 million deaths globally (3 800 deaths a day) (World Health Organization, 2016). TB treatment failure may be attributed to, amongst others: (a) irregular or inadequate anti-TB drug supplies to rural areas and third-world countries, (b) poor patient TB-education, (c) poor socio-economic circumstances, (d) the prolonged treatment duration, (e) treatment non-adherence, (f) anti-TB drug-resistance, as well as (g) various biological/biochemical factors (De Villiers & Loots, 2013).

The WHO recommends a six-month multi-drug treatment regimen, in which a combination of four drugs (INH, RIF, PZA and EMB) is used for treating active TB. In patients with drug-susceptible TB, this regimen reportedly has a 1–4% failure rate, and only 7% of those patients with a successful treatment outcome relapse within 24 months (Dye *et al.*, 2005). When considering that every active TB patient can potentially infect an additional 10–15 individuals per annum (World Health Organization, 2016), it becomes evident that TB treatment failure and relapse are important considerations in achieving the millennium goals pertaining to the eradication of TB. Despite the many TB research efforts to date, the biological mechanisms associated with anti-TB drug response remains poorly understood. It is also unclear if certain patients have more efficient mechanisms for eliminating this disease and responding to the treatment thereof, and to what extent lifestyle and environmental factors may contribute to this (De Villiers & Loots, 2013).

Early disease diagnosis and effective treatment protocols are the two primary objectives of TB control, aimed at reducing mortality and morbidity while also preventing the development of drug-resistance (Antoine *et al.*, 2007). The means to accurately predict prognosis from monitoring disease progression early in the treatment regimen would be regarded as a major breakthrough. This would instruct alternative treatment approaches in a timely fashion in order to reduce treatment failure and development of drug-resistance (Horne *et al.*, 2010; Walzl *et al.*, 2008). Considering this, there is a need for new, sensitive and specific biomarkers, not only for use in the early prediction of treatment failure, but also for better characterising and explaining the underlying mechanisms related to this occurrence, so that alternative treatment approaches can be developed. Once identified, these biomarkers need to be sufficiently validated for use as surrogate endpoints of treatment failure, early in the treatment regimen, ideally even before treatment begins. To date, no such urinary biomarkers have been identified with absolute certainty, hence the aim of this study was to

use a GCxGC-TOFMS metabolomics approach to identify biomarkers differentiating individuals with a successful (n=27) and unsuccessful (n=11) treatment outcome, as early as possible during the treatment regimen, which would also better characterise and explain the biological mechanisms related to TB treatment failure. Urine was selected as the preferred sample for addressing the above mentioned aim, since large quantities can easily be obtained and less complex sample preparation is required for analyses, as compared to sputum for instance. Furthermore, in recent years, the need for a holistic approach to metabolism has led to the development of urinary metabolomics for biomarker discovery in various diseases (Mahapatra *et al.*, 2014; Ryan *et al.*, 2011).

5.3 METHODS

5.3.1 Clinical samples

Anonymised archived urine samples were procured from the Faculty of Medicine and Health Sciences, NRF/DST Centre of Excellence for Biomedical Research, at the Stellenbosch University/MRC Centre for TB Research, from where they were transported to the North-West University (NWU), Human Metabolomics: Infectious Disease Laboratory, for metabolomic analysis. These samples were selected from a prospective observational cohort study of individuals with active pulmonary TB, diagnosed using smear microscopy and bacteriological cultures (Hesseling *et al.*, 2010). From these original samples, all treatment failure patients were included and matched by age, gender and extent of disease on chest X-rays to cured patients. Hence, the samples included were from 38 culture-confirmed active TB-positive South African patients (19 males and 19 females, between the age of 17 and 58) at baseline (time of diagnosis, thus prior to initiation of treatment), during the course of treatment with the DOTS strategy (weeks 1, 2 and 4) and two weeks after treatment completion (week 26). These patients had drug-susceptible TB, were HIV-seronegative, not pregnant, and with no other diseases (including diabetes, malignancy, lung cancer, chronic bronchitis and sarcoidosis). The samples were divided into successful (n=27) and unsuccessful (n=11) treatment outcome groups. According to the WHO, treatment failure can be defined as a patient whose sputum smear or culture is positive at month five or later during treatment (World Health Organization, 2014).

5.3.2 Sample analysis

Creatinine values for all urine samples, including QC samples, were determined using a creatinine enzyme kit (Thermo Scientific; reference number 981845) and analysed on an Indiko Clinical Analyser, Type 863 (Thermo Scientific). These creatinine values are used to normalise metabolite concentrations, and to determine the volume of urine, internal standard, BSTFA, TMCS and pyridine needed for each extraction and derivatisation. Organic acid extractions of the patient-collected urine samples were performed, analysed on a Pegasus 4D GCxGC-TOFMS (LECO Africa (Pty) Ltd, Johannesburg, South Africa) along with the necessary QC samples, and processed (as described in sections 3.3.3, 3.3.4 and 3.3.5, respectively) to identify all compounds.

5.3.3 Statistical data analyses

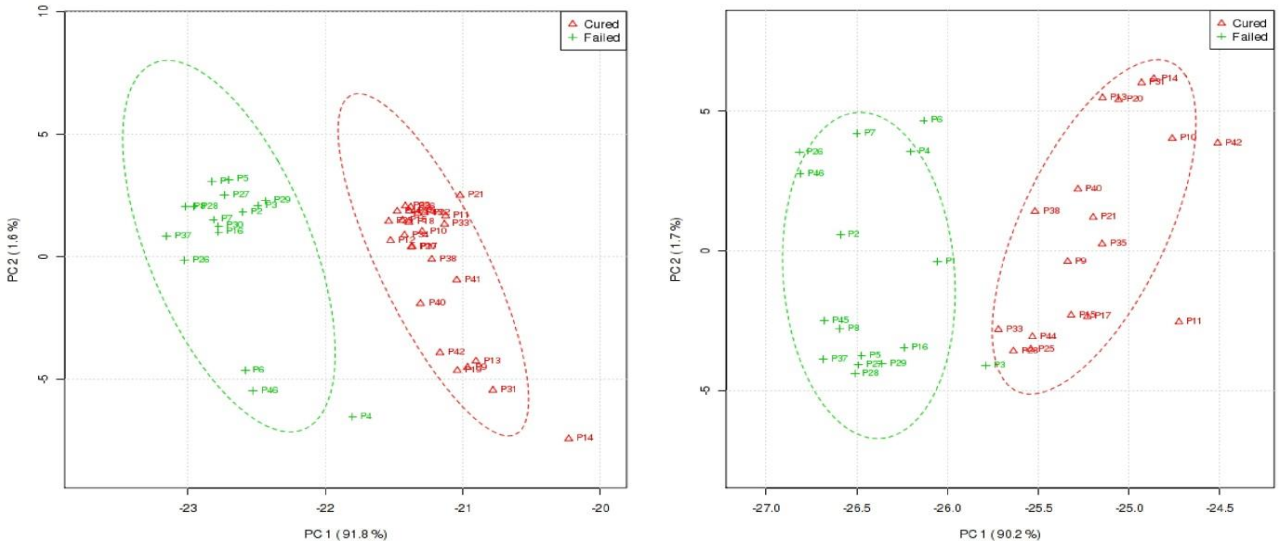
A standard metabolomics data clean-up procedure was applied before statistical data analyses. All metabolites were normalised relative to the internal standard by calculating the relative concentration of each, and a 50% zero filter was applied to each variable (Smuts *et al.*, 2013) to eliminate those compounds with more than 50% zero values within both groups. Quantile equating was applied to correct for any batch effects (Wang *et al.*, 2012), followed by a 50% QC–CV filter. Lastly, all zero values were replaced with a value calculated as half of the lowest detected value present in the entire dataset, as these may be due to low abundance rather than being absent (Schoeman *et al.*, 2012). GCxGC-TOFMS analysis yielded 782 compounds of which only 241 remained for further statistical analyses after the above mentioned data clean-up were completed.

The data was analysed via a variety of multi- and univariate statistical methods, using MetaboAnalyst (based on the statistical package “R”; version 2.10.0), which included PCA (Wold *et al.*, 1987), PLS-DA (Westerhuis *et al.*, 2008), fold change (Du Preez & Loots, 2013) and Mann-Whitney test (Pallant, 2001).

5.4 RESULTS AND DISCUSSION

Using the GCxGC-TOFMS data generated, a PCA was done to determine at which time point individuals who responded to the TB treatment can be differentiated from those who were not cured from TB. The PCA scores plots show clear differentiation of the successful and unsuccessful treatment outcome groups, using the urine collected at time of diagnosis and again at week 26 (Figure 5-1, a and b respectively), the latter of which can be expected

since the cured individuals are TB culture negative, and those with a poor treatment outcome are still TB culture positive at this point in time (Luies & Loots, 2016). No PCA differentiation was achieved for the other time intervals investigated (see Appendix C), most likely due to the effects of the anti-TB medication on the human metabolome, masking any underlying differences which occur in the metabolism initially differentiating the successful and unsuccessful treatment outcome groups at time of diagnosis. For the differentiation achieved at time of diagnosis (Figure 5-1a), the total amount of variance explained by the first two PCs (R^2X cum) was 93.4%, of which PC1 accounted for 91.8%, and PC2 accounted for 1.6%. For the differentiation achieved at week 26 (Figure 5-1b), the total amount of variance explained by the first two PCs (R^2X cum) was 91.9%, of which PC1 accounted for 90.2% and PC2 accounted for 1.7%.



(a) (b)

Figure 5-1: Principal components analysis (PCA) scores plots of principal component 1 versus principal component 2 of the successful and unsuccessful treatment outcome groups. Clear differentiation of these two groups was achieved at (a) time of diagnosis and (b) week 26, due to variation in their underlying metabolite profiles. The variances accounted for are indicated in parenthesis.

Since the aim of this study was to identify biomarkers for early prediction of treatment response and explain the mechanisms associated with treatment failure, only the PCA differentiation achieved at time of diagnosis was of further interest in this investigation. The separation of the groups at week 26 are mostly due to metabolic changes since one group is TB-positive and the other TB-negative, which was already investigated on a previous occasion (Luies & Loots, 2016). Considering the above, a PLS-DA model was built for the time of diagnosis data and showed a modelling parameter R^2Y (cum) of 93.48%, indicative of the total explained variation of the response Y.

Clear separation of the successful and unsuccessful treatment outcome groups is remarkable considering the possible predictive value that these may have for identifying individuals who will or will not respond to conventional treatment, even before treatment commences. However, before this biosignature or the individual metabolite markers can be utilised for this purpose, further validation is needed using new subjects to ensure these findings can be generalised. In the context of this study, the biosignature/metabolite markers were used to better describe the biological mechanisms as to why certain individuals with drug-sensitive TB are not successfully treated. Those compounds with a PLS-DA VIP >1.0 (Smuts *et al.*, 2013), or a fold change >|2|, or a Mann-Whitney with a *P*-value <0.05 (Du Preez & Loots, 2013), were selected as the metabolite markers that best explained the variation between the analysed groups. Of the 72 characteristic metabolite markers selected, 50 could be annotated using libraries compiled from previously injected standards, and are listed in Table 5-1.

Table 5-1: The 50 urinary metabolite markers identified at time of diagnosis that best explain the variation detected between the successful and unsuccessful treatment outcome groups.

Metabolite name	Successful treatment outcome: Average concentration (mg/g creatinine) (standard error of the mean)	Unsuccessful treatment outcome: Average concentration (mg/g creatinine) (standard error of the mean)	PLS-DA (VIP)	Fold change (absolute value)	Mann-Whitney test (*P*-value)
Amino acid metabolites:					
2-Ketovaleric acid	2.093 (0.428)	3.799 (1.302)	1.30	1.82	0.1318
2-Methyl-3-hydroxybutyric acid	3.777 (0.489)	5.487 (0.708)	2.30	1.45	0.0273
2-Piperidinecarboxylic acid (Pipecolic acid)	0.893 (0.230)	0.428 (0.104)	1.24	2.09	0.0197
3-Hydroxy-3-methylglutaric acid	9.437 (0.456)	12.179 (0.945)	1.08	1.29	0.0077
3-Hydroxyvaleric acid	7.691 (1.341)	8.896 (1.861)	1.05	1.16	0.3301
4-Methylcatechol	6.916 (1.150)	10.299 (1.854)	1.36	1.49	0.1022
Glyceric acid	0.268 (0.098)	0.356 (0.091)	1.14	1.33	0.0062
Methylsuccinic acid	3.912 (0.448)	6.581 (1.020)	2.60	1.68	0.0030
N-Tiglylglycine	0.834 (0.247)	1.567 (0.371)	1.91	1.88	0.0094
o-Hydroxyphenylacetic acid	3.093 (0.384)	5.293 (0.602)	2.50	1.71	0.0030
Oxalic acid	14.250 (1.308)	19.933 (2.116)	2.29	1.40	0.0191
Phenylacetic acid	0.305 (0.157)	0.903 (0.273)	2.14	2.97	0.0306
Quinolinic acid	5.311 (0.854)	9.448 (1.403)	5.50	1.78	0.0040
Vanillylmandelic acid	17.088 (0.878)	22.269 (1.830)	1.27	1.30	0.0040
***M. tuberculosis* cell wall components:**					
Ribitol	1.910 (0.627)	2.455 (0.523)	1.84	1.29	0.0708

(Continues on next page)

(Continues from previous page)

Gut microbiota imbalance metabolites:					
2,3-Butanediol	17.831 (4.351)	16.301 (5.137)	3.65	1.09	0.4450
3-(3-Hydroxyphenyl)-3-hydroxypropionic acid	20.030 (3.029)	28.087 (4.093)	5.40	1.40	0.0553
3-Hydroxy-3-(4-hydroxy-3-methoxyphenyl) propionic acid	2.912 (0.575)	8.366 (2.431)	6.45	2.87	0.0474
3-Hydroxyhippuric acid	17.580 (1.716)	23.662 (2.525)	2.37	1.35	0.0289
4-Hydroxybenzoic acid	4.642 (0.663)	9.966 (1.754)	4.92	2.15	0.0027
4-Hydroxyphenyllactic acid	5.695 (1.526)	6.724 (1.720)	1.21	1.18	0.1485
5-Hydroxymethyl-2-furoic acid	17.563 (4.279)	21.994 (6.053)	4.46	1.25	0.1795
Benzoic acid	0.936 (0.275)	1.367 (0.251)	1.56	1.46	0.0229
cis-4-Hydroxycyclohexanecarboxylic acid	1.428 (0.610)	0.666 (0.190)	0.87	2.15	0.0306
Citramalic acid	8.700 (1.608)	10.742 (3.124)	1.66	1.23	0.4550
Furan-2,5-dicarboxylic acid	8.755 (2.540)	10.156 (3.446)	1.14	1.16	0.3032
Furoylglycine	7.466 (1.443)	10.421 (2.241)	3.39	1.40	0.1115
Vanillic acid	7.731 (0.876)	11.249 (1.230)	2.23	1.46	0.0082
DNA damage and oxidative stress markers:					
2-Deoxyribonic acid	0.477 (0.113)	1.080 (0.245)	1.85	2.27	0.0169
Parabanic acid	2.984 (0.799)	3.584 (0.865)	2.68	1.20	0.0474
TCA and GABA metabolites:					
3,4-Dihydroxybutyric acid	9.160 (0.825)	11.730 (0.89)	1.03	1.28	0.0108
4-Hydroxybutyric acid	0.305 (0.097)	0.544 (0.158)	1.68	1.79	0.0088
Citric acid	2.730 (0.838)	3.526 (1.619)	1.34	1.29	0.2691
Malic acid	4.579 (2.712)	5.085 (2.668)	3.13	1.11	0.4351
Dicarboxylic acids and 3-hydroxy fatty acids:					
2,3-Dihydroxypentonic acid	4.632 (1.255)	6.118 (2.265)	2.56	1.32	0.3485
3,5-Dihydroxypentonic acid	3.956 (0.469)	5.698 (0.790)	1.29	1.44	0.0743
3-Hydroxydodecanedioic acid	4.946 (1.659)	2.313 (1.075)	2.27	2.14	0.2000
3-Hydroxysebacic acid	1.499 (0.386)	3.716 (1.140)	2.36	2.48	0.0582
3-Ketosebacic acid	2.419 (0.535)	6.451 (1.560)	6.96	2.67	0.0324
3-Methylhexanoic acid (3-Methyladipic acid)	7.723 (0.800)	12.631 (1.637)	3.30	1.64	0.0058
cis,*cis*-4,7-Decadiene-1,10-dioic acid	9.523 (1.259)	12.113 (1.359)	1.19	1.27	0.0743
cis-4-Decene-1,10-dioic acid	1.403 (0.316)	2.561 (0.649)	1.74	1.83	0.0499
Heptanedioic acid (Pimelic acid)	2.895 (0.434)	4.967 (0.684)	2.32	1.72	0.0024
Hexanedioic acid (Adipic acid)	0.713 (0.192)	1.369 (0.443)	1.38	1.92	0.0816
Nonanedioic acid (Azelaic acid)	4.596 (0.918)	6.465 (1.271)	2.01	1.41	0.0894
Octanedioic acid (Suberic acid)	3.452 (0.595)	4.577 (0.394)	1.68	1.33	0.0169
Other:					
1,2,3-Trihydroxybutane	1.962 (0.540)	2.842 (1.415)	1.22	1.45	0.4950
2-Methyl-1,2-dihydroxypropane	5.360 (1.768)	6.456 (2.985)	2.78	1.20	0.3672
4-Pentenoic acid	5.860 (0.898)	8.201 (1.289)	1.86	1.40	0.0611
Glucuronic acid	5.923 (0.442)	8.307 (0.757)	1.22	1.40	0.0062

When considering the metabolite markers best differentiating the successful and unsuccessful treatment outcome patient groups, the first important observation is the elevated levels of the metabolites associated with tryptophan (quinolinic acid (Heyes *et al.*, 1992)), phenylalanine (phenylacetic acid (Puri, 2006), o-hydroxyphenylacetic acid (Taniguchi & Armstrong, 1963)), and tyrosine (4-methylcatechol (Li *et al.*, 2007), vanillylmandelic acid (Eisenhofer *et al.*, 2004)) metabolism (see Figure 5-2). The same metabolic flux and the resulting accumulation of these amino acids, were also previously reported by Luies and Loots (2016) in TB-positive patient urine, and explained by elevations in IFN-γ due to the host's immune response to *M. tuberculosis* and a compromised insulin secretion by the host (Luies & Loots, 2016). Furthermore, markers directly associated with the presence of *M. tuberculosis* (ribitol (Silhavy *et al.*, 2010)), a gut microbiota imbalance (2,3-butanediol (Hong, 2011), 3-(3-hydroxyphenyl)-3-hydroxypropionic acid (Shaw, 2013), 3-hydroxy-3-(4-hydroxy-3-methoxyphenyl)propionic acid, 3-hydroxyhippuric acid (Gonthier *et al.*, 2003), 4-hydroxybenzoic acid (Russell *et al.*, 2013; Tomás-Barberán & Clifford, 2000), 4-hydroxyphenyllactic acid (Russell *et al.*, 2013), 5-hydroxymethyl-2-furoic acid (Jellum *et al.*, 1973), benzoic acid (Russell *et al.*, 2013), *cis*-4-hydroxycyclohexanecarboxylic acid (Kronick *et al.*, 1983), citramalic acid (Du Preez & Loots, 2013), furan-2,5-dicarboxylic acid (Jellum *et al.*, 1973; Pettersen & Jellum, 1972), furoylglycine (Pettersen & Jellum, 1972), vanillic acid (Tomás-Barberán & Clifford, 2000)), and DNA damage (2-deoxyribonic acid (Zhou & Greenberg, 2012)) due to oxidative stress (parabanic acid (Marklund *et al.*, 2000)), were also detected and correlate with those markers previously characterising TB, as described by Luies and Loots (2016). Considering the fact that these markers are not only indicative of a general TB disease state (Luies & Loots, 2016), but also more pronounced in the treatment failure group in this study, may indicate that these individuals who are unsuccessfully treated have an increased disease severity due to a microbiota imbalance and/or underlying host abnormality, which will be discussed in greater detail below.

The highly significant evidence from the almost three times elevated 3-hydroxy-3-(4-hydroxy-3methoxyphenyl)propionic acid and more than two times lower *cis*-4-hydroxycyclohexane-carboxylic acid in the treatment failure group, strongly testifies towards an imbalance in gut microbiota in the unsuccessfully treated patients, which is associated with a weaker immune system and inability to fight disease. Both the innate and adaptive immune response systems have evolved to rely on microbiota interactions, which not only promote immune cell maturation but also influence the normal development of immune functions (Clemente *et al.*, 2012). Additionally, gut microbiota is also well-known to interact with various drugs, influencing many factors relating to their absorption and plasma concentrations, and hence is considered an important contributor to xenobiotic/drug bioavailability and toxicity (Gonzalez *et al.*, 2011).

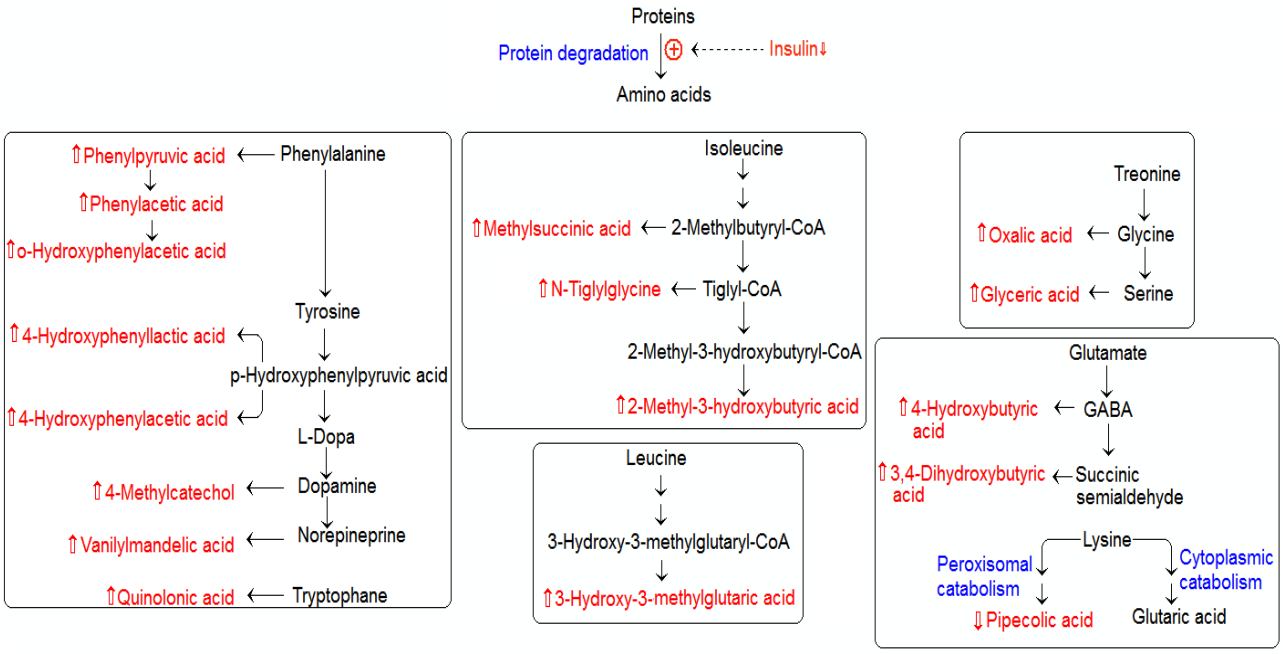


Figure 5-2: Amino acid metabolism of the identified urinary metabolite markers. These markers are indicated as either increased (↑) or decreased (↓) in the treatment failure group comparatively.

Furthermore, altered levels of various dicarboxylic acids, acylcarnitines and 3-hydroxy fatty acids were detected in the urine of the treatment failure group. Lipid molecules are increasingly recognised as having the potential to shape the immune response to infectious pathogens (Kaushal, 2012) since these are important bioactive mediators of cellular activity during pathophysiological processes, and affect various activities including cell apoptosis, monocyte adhesion, platelet aggregation, and regulation of immune responses (Hasanally *et al.*, 2014). Interestingly, however, the presence of these metabolites in urine is also used to distinguish between individuals with inherited defects of the following long-chain fatty acid (LCFA) transport-associated enzymes: (a) carnitine palmitoyltransferase 1 (CPT-1), (b) carnitine palmitoyltransferase 2 (CPT-2), (c) translocase (TL), (d) very-long-chain acyl-coenzyme A (CoA) dehydrogenase (VLCAD) or (e) any of the three enzymes making up the mitochondrial trifunctional protein (MTP); long chain 2-enoyl-Co A hydratase, long-chain 3-hydroxyacyl-CoA dehydrogenase and 3-ketoacyl-CoA thiolase (Garg & Dasouki, 2006; Sim *et al.*, 2002). A CPT-1 deficiency results in reduced C_{16}–C_{18} acylcarnitines and increased free carnitine, a CPT-2, TL or VLCAD deficiency results in both elevated long chain acylcarnitines and dicarboxylic acids with reduced free carnitine, and a MTP deficiency (on any of its three associated enzymes) results in elevated long chain acylcarnitines, dicarboxylic acids and 3-hydroxy fatty acids with reduced free carnitine (Garg & Dasouki, 2006). Since the organic acid extraction method used in this study was not suited for the extraction or detection of acylcarnitines, hence their absence in Table 5-1, we did a further ultra-performance liquid chromatography (UPLC) (Van Aardt *et al.*, 2016) urinary

acylcarnitine analysis (Venter *et al.*, 2015), and found significantly elevated palmitoylcarnitine (C_{16}) in the urine of the treatment failure group comparatively (136.19 vs. 58.22 mmol/L; P=0.0095). Additionally, although not statistically significant, the total short chain acylcarnitines were comparatively decreased in the treatment failure group (8.77 vs. 2.75 mmol/L; P=0.2298). Free carnitine (C_0) was also decreased (0.64 vs. 2.58 mmol/L; P=0.4329), as expected in such deficiencies, because these remain bound to the LCFAs and are unable to cross the mitochondrial membrane. Considering this diagnostic metabolite profile, a MTP deficiency or abnormality is suggested to occur in the treatment failure group (see Figure 5-3). MTP mutations are estimated at a prevalence of approximately 1:75 000 (Garg & Dasouki, 2006), and according to unpublished results of the Potchefstroom Laboratory for Inborn Errors of Metabolism (PLIEM), may be as high as 1 in 100 for a mutation in any of the seven enzyme/protein systems mentioned above. Additional evidence for this, is the serum carnitine deficiencies previously reported in 47.7% of all TB-positive patients (Hatamkhani *et al.*, 2014), which are also associated with these carnitine transporters (Flanagan *et al.*, 2010). Furthermore, L-carnitine has been shown to play a significant role in T-cell-dependent antibacterial activity in the host, and enhance the immune response via reduced macrophage and lymphocyte malfunction in TB patients (Jirillo *et al.*, 1993; Jirillo *et al.*, 1991), hence a deficiency in L-carnitine, due to an underlying enzyme/protein system deficiency, may explain why these individuals did not respond to treatment. Confirmation of a reduced capacity for mitochondrial β-oxidation due to such a deficiency, is the elevated concentrations of fatty acids with odd numbered carbons, such as azelaic acid and heptanedioic (pimelic) acid in the treatment failure group, due to the above mentioned accumulating fatty acids undergoing peroxisomal α-oxidation prior to peroxisomal β-oxidation (Figure 5-3) (Van Veldhoven, 2010).

Accumulation of fatty acyl-CoA derivatives, due to the defective LCFA transport, can also contribute to a compromised insulin secretion (Luies & Loots, 2016). Long-chain fatty acyl-CoA's are strong inhibitors of glutamate dehydrogenase (GDH) (Lai *et al.*, 1994), and since this enzyme is responsible for the conversion of glutamate to α-ketoglutarate, and an elevated flux of the TCA cycle, eventually resulting in insulin secretion (see Figure 5-3) (Newsholme *et al.*, 2006), suggests an additional means by which insulin secretion may be compromised in the treatment failure group. Further confirmation of this is the elevated levels of 3,4-dihydroxybutyric acid and 4-hydroxybutyric acid, indicating an increased metabolic flux in the gamma-aminobutyric acid (GABA) shunt pathway (Shinka *et al.*, 2002), in the opposite direction of the glutamate to α-ketoglutarate reaction previously mentioned (see Figure 5-2). Compromised insulin secretion has also been shown to result in an accumulation of various branched chain amino acids (BCAAs) (Lu *et al.*, 2013), which explains the elevated leucine (3-hydroxyvaleric acid, 3-hydroxy-3-methylglutaric acid) and

isoleucine (2-methyl-3-hydroxybutyric acid, methylsuccinic acid, N-tiglylglycine) metabolite intermediates observed in the treatment failure group comparatively.

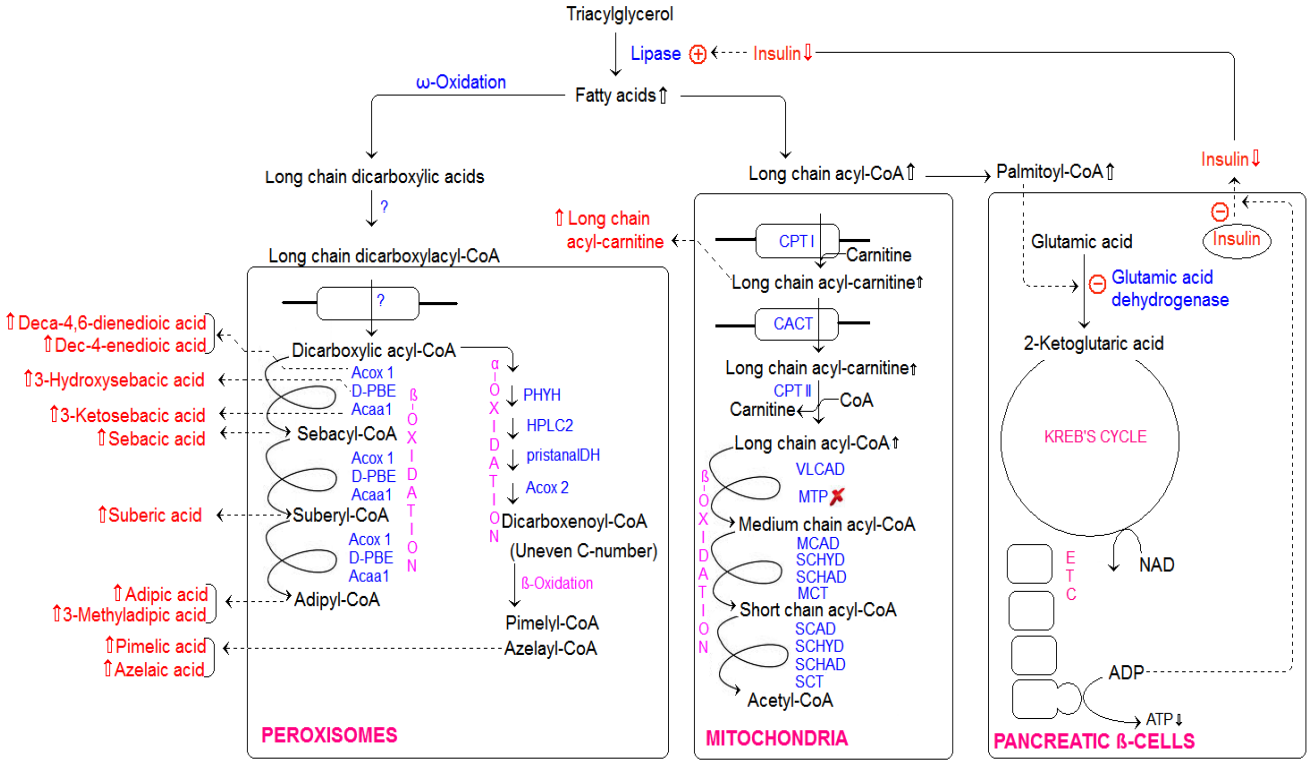


Figure 5-3: Fatty acid oxidation of the identified urinary metabolite markers, which are either increased (↑) or decreased (↓) in the treatment failure group comparatively. Abbreviations: CPT I, carnitine palmitoyltranferase 1; CPT II, carnitine palmitoyltransferase 2; CACT, carnitine-acylcarnitine translocase; VLCAD, very long chain acyl-CoA dehydrogenase; MCAD, medium chain acyl-CoA dehydrogenase; SCAD, short chain acyl-CoA dehydrogenase; M/SCHAD, medium/short chain hydroxyacyl-CoA dehydrogenase; MTP, mitochondrial trifunctional protein; SCYD, short chain enoyl-CoA hydratase; MCT, medium chain 3-ketoacyl-CoA thiolase; SCT, short chain 3-ketoacyl-CoA thiolase; ACOX 1, acyl-CoA oxidase 1; ACOX 2, acyl-CoA oxidase 2; D-PBE, peroxisomal bifunctional enzyme; ACAA1, 3-ketoacyl-CoA thiolase; PHYH, Phytanoyl-CoA hydroxylase; HPLC2, Unknown; pistanalDH, pristanaldehyde dehydrogenase.

Whatever the proposed mechanism, the patients in the treatment failure group show a reduced capacity to transport LCFA into the mitochondria for metabolism into short-chain fatty acids (SCFAs), which are well-known for their antimicrobial effects (i.e. increase the fluidity of the bacterial cell wall which negatively influence cell wall integrity (Royce *et al.*, 2013), and induce host defence peptide LL-37 (Mily *et al.*, 2015)). Considering this, apart from the reduced L-carnitine detected in the treatment failure group, the reduced SCFAs may be an additional explanation as to why these individuals may not be responding to the treatment and why the metabolite markers suggest elevated disease severity.

5.5 CONCLUSION

This study highlights the capacity of metabolomics to identify markers which predict a poor response to treatment, and also better characterise or propose previously unknown mechanisms resulting in TB treatment failure. The most significant observations in this metabolomics study were the elevated levels of those metabolites associated with an imbalance in the gut microbiome. Since this influence xenobiotic uptake and toxicity, the synchronous use of probiotics for optimising the microbiome during first-line anti-TB treatment, may improve treatment outcome, and could be a topic of further investigation. Furthermore, it may be interesting to compare these patients to a cohort of healthy humans with a normal intestinal biosis in order to firmly allocate the microbiota dysbiosis to the treatment failure cohort. Additionally, similar experimental undertakings with a larger sample cohort could certainly support this hypothesis. Also considering the fact that no PCA differentiation was achieved for the other time intervals investigated (see Appendix C), it would be interesting to learn whether the microbiota balance in the two comparative groups become equally disturbed or equally restored during treatment. The issue of whether the metabolites analysed in the urine of these patients are due to human host metabolism or whether they originate from the commensal intestinal bacteria can be addressed in future by experiments with germ-free and conventionalised *M. tuberculosis*-infected laboratory animals. Finally, another interesting observation was those metabolites associated with abnormalities in any of the three enzymes of the MTP complex in the treatment failure group. Since L-carnitine and SCFAs are also reduced in these individuals, and well-known for their anti-mycobacterial properties, this metabolic pathway may explain why these individuals have an increased disease severity and/or a poor response to TB treatment.

5.6 REFERENCES

Antoine, D., French, C.E., Jones, J. & Watson, J.M. 2007. Tuberculosis treatment outcome monitoring in England, Wales and Northern Ireland for cases reported in 2001. *Journal of Epidemiology and Community Health*, 61(4):302-307.

Clemente, J.C., Ursell, L.K., Parfrey, L.W. & Knight, R. 2012. The impact of the gut microbiota on human health: an integrative view. *Cell*, 148(6):1258-1270.

De Villiers, L. & Loots, D.T. 2013. Using metabolomics for elucidating the mechanisms related to tuberculosis treatment failure. *Current Metabolomics*, 1(4):306-317.

Du Preez, I. & Loots, D.T. 2013. New sputum metabolite markers implicating adaptations of the host to *Mycobacterium tuberculosis*, and vice versa. *Tuberculosis*, 93(3):330-337.

Dye, C., Watt, C.J., Bleed, D.M., Hosseini, S.M. & Raviglione, M.C. 2005. Evolution of tuberculosis control and prospects for reducing tuberculosis incidence, prevalence, and deaths globally. *Jama*, 293(22):2767-2775.

Eisenhofer, G., Kopin, I.J. & Goldstein, D.S. 2004. Catecholamine metabolism: A contemporary view with implications for physiology and medicine. *Pharmacological Reviews*, 56(3):331-349.

Flanagan, J.L., Simmons, P.A., Vehige, J., Willcox, M.D. & Garrett, Q. 2010. Role of carnitine in disease. *Nutrition & Metabolism*, 7(1):1.

Garg, U. & Dasouki, M. 2006. Expanded newborn screening of inherited metabolic disorders by tandem mass spectrometry: Clinical and laboratory aspects. *Clinical Biochemistry*, 39(4):315-332.

Gonthier, M.-P., Cheynier, V., Donovan, J.L., Manach, C., Morand, C., Mila, I., Lapierre, C., Rémésy, C. & Scalbert, A. 2003. Microbial aromatic acid metabolites formed in the gut account for a major fraction of the polyphenols excreted in urine of rats fed red wine polyphenols. *The Journal of Nutrition*, 133(2):461-467.

Gonzalez, A., Stombaugh, J., Lozupone, C., Turnbaugh, P.J., Gordon, J.I. & Knight, R. 2011. The mind-body-microbial continuum. *Dialogues in Clinical Neuroscience*, 13(1):55.

Hasanally, D., Chaudhary, R. & Ravandi, A. 2014. Role of phospholipases and oxidized phospholipids in inflammation. Phospholipases in health and disease. Springer. p. 55-72).

Hatamkhani, S., Khalili, H., Karimzadeh, I., Dashti-Khavidaki, S., Abdollahi, A. & Jafari, S. 2014. Carnitine for prevention of antituberculosis drug-induced hepatotoxicity: A randomized, clinical trial. *Journal of Gastroenterology and Hepatology*, 29(5):997-1004.

Hesseling, A., Walzl, G., Enarson, D., Carroll, N., Duncan, K., Lukey, P., Lombard, C., Donald, P., Lawrence, K. & Gie, R. 2010. Baseline sputum time to detection predicts month two culture conversion and relapse in non-HIV-infected patients. *The International Journal of Tuberculosis and Lung Disease*, 14(5):560-570.

Heyes, M.P., Saito, K. & Markey, S.P. 1992. Human macrophages convert L-tryptophan into the neurotoxin quinolinic acid. *Biochemical Journal*, 283:633-635.

Hong, Y.S. 2011. NMR-based metabolomics in wine science. *Magnetic Resonance in Chemistry*, 49(S1):S13-S21.

Horne, D.J., Royce, S.E., Gooze, L., Narita, M., Hopewell, P.C., Nahid, P. & Steingart, K.R. 2010. Sputum monitoring during tuberculosis treatment for predicting outcome: systematic review and meta-analysis. *The Lancet Infectious Diseases*, 10(6):387-394.

Jellum, E., Børresen, H.C. & Eldjarn, L. 1973. The presence of furan derivatives in patients receiving fructose-containing solutions intravenously. *Clinica Chimica Acta*, 47(2):191-201.

Jirillo, E., Altamura, M., Marcuccio, C., Tortorella, C., De Simone, C. & Antonaci, S. 1993. Immunological responses in patients with tuberculosis and in vivo effects of acetyl-L-carnitine oral administration. *Mediators of Inflammation*, 2(7):S17-S20.

Jirillo, E., Altamura, M., Munno, I., Pellegrino, N., Sabato, R., Fabio, S.D. & Simone, C.D. 1991. Effects of acetyl-L-carnitine oral administration on lymphocyte antibacterial activity and TNF-α levels in patients with active pulmonary tuberculosis. A randomized double blind versus placebo study. *Immunopharmacology and Immunotoxicology*, 13(1-2):135-146.

Kaushal, D. 2012. Eicosanoids, prostaglandins, and the progression of tuberculosis. *Journal of Infectious Diseases*:jis611.

Kronick, J., Mamer, O., Montgomery, J. & Scriver, C. 1983. Apparent gastrointestinal origin of cis-4-hydroxycyclohexanecarboxylic acid. *Clinica Chimica Acta*, 132(2):205-208.

Lai, J.C., Liang, B.B., Zhai, S., Jarvi, E.J. & Lu, D.R. 1994. Brain mitochondrial citrate synthase and glutamate dehydrogenase: differential inhibition by fatty acyl coenzyme A derivatives. *Metabolic Brain Disease*, 9(2):143-152.

Li, G., Zhang, H., Sader, F., Vadhavkar, N. & Njus, D. 2007. Oxidation of 4-methylcatechol: implications for the oxidation of catecholamines. *Biochemistry*, 46(23):6978-6983.

Lu, J., Xie, G., Jia, W. & Jia, W. 2013. Insulin resistance and the metabolism of branched-chain amino acids. *Frontiers of Medicine*, 7(1):53-59.

Luies, L. & Loots, D. 2016. Tuberculosis metabolomics reveals adaptations of man and microbe in order to outcompete and survive. *Metabolomics*, 12(3):1-9.

Mahapatra, S., Hess, A.M., Johnson, J.L., Eisenach, K.D., DeGroote, M.A., Gitta, P., Joloba, M.L., Kaplan, G., Walzl, G. & Boom, W.H. 2014. A metabolic biosignature of early response to anti-tuberculosis treatment. *BMC Infectious Diseases*, 14(1):53.

Marklund, N., Östman, B., Nalmo, L., Persson, L. & Hillered, L. 2000. Hypoxanthine, uric acid and allantoin as indicators of in vivo free radical reactions. Description of a HPLC method and human brain microdialysis data. *Acta Neurochirurgica*, 142(10):1135-1142.

Mily, A., Rekha, R.S., Kamal, S.M., Arifuzzaman, A.S.M., Rahim, Z., Khan, L., Haq, M.A., Zaman, K., Bergman, P. & Brighenti, S. 2015. Significant effects of oral phenylbutyrate and vitamin D3 adjunctive therapy in pulmonary tuberculosis: A randomized controlled trial. *PloS one*, 10(9):e0138340.

Newsholme, P., Brennan, L. & Bender, K. 2006. Amino acid metabolism, β-cell function, and diabetes. *Diabetes*, 55(Supplement 2):S39-S47.

Pallant, J. 2001. Manual SPSS Survival: A step by step guide to data analysis using SPSS (pp. 302): Open University Press, McGraw-Hill Education, Philadelphia, USA.

Pettersen, J.E. & Jellum, E. 1972. The identification and metabolic origin of 2-furoylglycine and 2, 5-furandicarboxylic acid in human urine. *Clinica Chimica Acta*, 41:199-207.

Puri, D. 2006. Textbook of Medical Biochemistry: Clinical cases, Chapter 13.2. Textbook of medical biochemistry. New-Delhi, India: Elsevier. p. 376).

Royce, L.A., Liu, P., Stebbins, M.J., Hanson, B.C. & Jarboe, L.R. 2013. The damaging effects of short chain fatty acids on Escherichia coli membranes. *Applied Microbiology and Biotechnology*, 97(18):8317-8327.

Russell, W.R., Hoyles, L., Flint, H.J. & Dumas, M.-E. 2013. Colonic bacterial metabolites and human health. *Current Opinion in Microbiology*, 16(3):246-254.

Ryan, D., Robards, K., Prenzler, P.D. & Kendall, M. 2011. Recent and potential developments in the analysis of urine: A review. *Analytica Chimica Acta*, 684(1):17-29.

Schoeman, J.C., Du Preez, I. & Loots, D.T. 2012. A comparison of four sputum pre-extraction preparation methods for identifying and characterising *M. tuberculosis* using GCxGC-TOFMS metabolomics. *Journal of Microbiological Methods*, 91:301-311.

Shaw, W. 2013. Increased urinary excretion of a 3-(3-hydroxyphenyl)-3-hydroxypropionic acid (HPHPA), an abnormal phenylalanine metabolite of Clostridia spp. in the gastrointestinal tract, in urine samples from patients with autism and schizophrenia. *Nutritional Neuroscience*.

Shinka, T., Inoue, Y., Ohse, M., Ito, A., Ohfu, M., Hirose, S. & Kuhara, T. 2002. Rapid and sensitive detection of urinary 4-hydroxybutyric acid and its related compounds by gas chromatography–mass spectrometry in a patient with succinic semialdehyde dehydrogenase deficiency. *Journal of Chromatography B*, 776(1):57-63.

Silhavy, T.J., Kahne, D. & Walker, S. 2010. The bacterial cell envelope. *Cold spring harbor perspectives in biology*, 2(5):a000414.

Sim, K.G., Hammond, J. & Wilcken, B. 2002. Strategies for the diagnosis of mitochondrial fatty acid β-oxidation disorders. *Clinica Chimica Acta*, 323(1):37-58.

Smuts, I., Van Der Westhuizen, F.H., Louw, R., Mienie, L.J., Engelke, U.F.H., Wevers, R.A., Mason, S., Koekemoer, G. & Reinecke, C.J. 2013. Disclosure of a putative biosignature for respiratory chain disorders through a metabolomics approach. *Metabolomics*, 9(2):379-391.

Taniguchi, K. & Armstrong, M.D. 1963. The enzymatic formation of o-hydroxyphenylacetic acid. *The Journal of Biological Chemistry*, 238:4091-4097.

Tomás-Barberán, F.A. & Clifford, M.N. 2000. Dietary hydroxybenzoic acid derivatives–nature, occurrence and dietary burden. *Journal of the Science of Food and Agriculture*, 80(7):1024-1032.

Van Aardt, W.J., le Roux, J.M., Lindeque, J.Z., Mason, S. & Louw, R. 2016. The effect of temperature on the respiration and metabolism of the African burrowing scorpion (Opistophthalmus latimanus). *Comparative Biochemistry and Physiology Part D: Genomics and Proteomics*, 20:50-56.

Van Veldhoven, P.P. 2010. Biochemistry and genetics of inherited disorders of peroxisomal fatty acid metabolism. *Journal of Lipid Research*, 51(10):2863-2895.

Venter, L., Lindeque, Z., Van Rensburg, P.J., Van Der Westhuizen, F., Smuts, I. & Louw, R. 2015. Untargeted urine metabolomics reveals a biosignature for muscle respiratory chain deficiencies. *Metabolomics*, 11(1):111-121.

Walzl, G., Ronacher, K., Siawaya, J.F.D. & Dockrell, H.M. 2008. Biomarkers for TB treatment response: Challenges and future strategies. *Journal of Infection*, 57(2):103-109.

Wang, S.-Y., Kuo, C.-H. & Tseng, Y.J. 2012. Batch normalizer: A fast total abundance regression calibration method to simultaneously adjust batch and injection order effects in liquid chromatography/time-of-flight mass spectrometry-based metabolomics data and comparison with current calibration methods. *Analytical Chemistry*, 85(2):1037-1046.

Westerhuis, J.A., Hoefsloot, H.C.J., Smit, S., Vis, D.J., Smilde, A.K., van Velzen, E.J.J., van Duijnhoven, J.P.M. & van Dorsten, F.A. 2008. Assessment of PLSDA cross validation. *Metabolomics*, 4(1):81-89.

Wold, S., Esbensen, K. & Geladi, P. 1987. Principal component analysis. *Chemometrics and Intelligent Laboratory Systems*, 2(1):37-52.

World Health Organization. 2014. Definitions and reporting framework for tuberculosis – 2014 revision. Geneva, Switzerland (WHO Press).

World Health Organization. 2016. Global tuberculosis report 2016. Geneva, Switzerland (WHO Press).

Zhou, C. & Greenberg, M.M. 2012. Histone-catalyzed cleavage of nucleosomal DNA containing 2-deoxyribonolactone. *Journal of the American Chemical Society*, 134(19):8090-8093.

CHAPTER 6: PREDICTING TUBERCULOSIS TREATMENT OUTCOME USING METABOLOMICS

This chapter has been submitted for publication:

- Luies, L., Van Reenen, M., Ronacher, K., Walzl, G. & Loots, DT. (2017). Predicting tuberculosis treatment outcome using metabolomics. Submitted for publication to *Biomarkers in Medicine* (Manuscript number: BMM-2017-0133).

6.1 ABSTRACT

Predicting a poor treatment outcome would offer significant benefits for patient care and for new drug development. Urine samples from TB-positive patients with a successful and unsuccessful treatment outcome were collected at baseline and analyzed. The identified metabolites were used in a forward logistic regression model, which achieved a ROC AUC of 0.94 (95% CI 0.84–1) and cross-validated well in a leave-one-out context, with an AUC of 0.89 (95% CI 0.7–1). Two possible predictors were identified, which are associated with a gut microbiota imbalance. Our findings show the capacity of metabolomics to predict treatment failure at time of diagnosis, which potentially offers significant benefits for use in new drug development clinical trials and individualized patient care.

6.2 INTRODUCTION

The latest reports indicate 10.4 million individuals are infected with *M. tuberculosis* globally, resulting in an estimated 1.8 million deaths from TB, annually. Clearly, current strategies to control this pandemic are inadequate, especially in high-burden countries with limited resources, where the treatment target goals, as recommended by the WHO, have not yet been met (World Health Organization, 2015).

Currently, the most effective anti-TB treatment strategy, DOTS, consists of an initial two month (intensive) phase using a combination of INH, RIF, PZA and EMB, followed by a four month (maintenance) phase, consisting of only INH and RIF (Olivier & Loots, 2011). The DOTS strategy, when fully adhered to, has a relatively high success rate in patients with drug-susceptible TB (World Health Organization, 2015). TB treatment failure using this approach, however, still occurs with drug-sensitive TB, and is defined as the occurrence of persistently positive sputum smears or cultures at the end of treatment (Mukherjee *et al.*, 2004). An additional concern is that these individuals, who are not treated successfully, remain infectious for prolonged periods of time, and continue to transmit the disease to others. Improved treatment outcomes and timely intervention is a priority for TB control programs and requires the identification of those factors associated with treatment failure (Alobu *et al.*, 2014). Previous studies suggest the following risk factors for drug-susceptible TB treatment failure and relapse: (a) a less than 5% gain in bodyweight (Krapp *et al.*, 2008); (b) age; (c) HIV co-infection; (d) diabetes type II, (e) illiteracy; (f) alcoholism and (g) prior TB treatment (Albuquerque *et al.*, 2007). The most prominent contributor to this occurrence, however, is a slow patient response to treatment. It is also important to note that a number of novel immunological biomarkers (i.e cytokines and chemokines) have been proposed for monitoring TB treatment response and predicting treatment outcome (Walzl *et al.*, 2008), including those directly associated with *M. tuberculosis*, such as lipoarabinomannan (Lawn *et al.*, 2017; Lawn *et al.*, 2012) or those associated with the overall inflammatory response, such as induced protein (IP)-10 (Cannas *et al.*, 2010; Petrone *et al.*, 2015; Petrone *et al.*, 2016).

Two diagnostic approaches are frequently used to detect active TB and monitor treatment outcome, namely sputum smear microscopy and bacteriological cultures, the latter of which is considered the gold standard. Response to treatment is usually determined based on sputum culture conversion after two months of treatment (Horne *et al.*, 2010). However, a period of two months or more before the first indications of treatment efficacy can be determined, is considered too long, since the bacteria have had the opportunity to adapt and create drug-resistant strains (Baumann *et al.*, 2012; Siawaya *et al.*, 2008), and also

considering that these patients may still transfer the disease to others during this time. Additionally, these treatment failure predictive methods require frequent laboratory monitoring, which is not only time consuming, but also costly. Thus, there is an urgent need to search for alternative approaches or biomarkers which can be used for the early detection of treatment failure (Namukwaya *et al.*, 2011).

Over the last 10 years, metabolomics has proved to be a very effective tool to search for and identify new diagnostic and prognostic biomarkers and/or disease risk factors (Dunn *et al.*, 2012). Metabolomics is defined as the unbiased identification and quantification of all intermediates of metabolism in a biological system, using specialised analytical techniques, in conjunction with biostatistical and mathematical analysis (De Villiers & Loots, 2013). In 2012, Olivier and Loots applied a GC-MS metabolomics approach to characterise and differentiate various infectious *Mycobacterium* species (i.e. *M. tuberculosis, M. bovis, M. kansasii and M. avium*) and *Pseudomonas aeruginosa* based on their characteristic lipid profiles. Twelve lipid biomarkers were identified and used to build a multivariate discriminant model, which could correctly assign unknown samples to their respective species groups with probabilities ranging from 72 to 100% (Olivier & Loots, 2012). Che *et al.* (2013) used a similar approach to compare serum samples collected from healthy controls and TB patients, before and after TB therapy, to find diagnostic markers for active TB, not influenced by anti-TB therapy. They indicated nine potential diagnostic TB biomarkers, one of which (i.e. 5-oxoproline) remained unaffected by first-line TB therapy (Che *et al.*, 2013). These studies indicate the capacity of metabolomics to not only differentiate *M. tuberculosis* from other *Mycobacterium* species, but also to identify useful biomarkers that can be used for the early prediction of treatment outcome, and for possible use in new anti-TB drug development. Considering this, we used an untargeted GCxGC-TOFMS urinary metabolomics approach to differentiate TB-positive patients with a successful (n=27) and unsuccessful (n=11) treatment outcome at time of diagnosis, and identified those metabolite markers which best differentiate the groups, in order to build a logistic regression model to possibly predict treatment failure before first-line anti-TB drug administration commences

6.3 MATERIALS AND METHODS

6.3.1 Clinical samples

Anonymised archived urine samples were procured from the Faculty of Medicine and Health Sciences, NRF/DST Centre of Excellence for Biomedical Research, at the Stellenbosch University/MRC Centre for TB Research, from where they were transported (frozen at -20ºC) to the North-West University (NWU), Human Metabolomics: Infectious Disease Laboratory,

for metabolomics analysis. These samples were part of a prospective observational cohort study of individuals with active pulmonary TB, diagnosed with *M. tuberculosis* infection using smear microscopy and bacteriological cultures and strain typing (Hesseling *et al.*, 2010). Representation of different *M. tuberculosis* strains was not significantly different between cured and failed participants. The samples were not collected at a specific time of day in order to ensure robustness of any identified metabolites. The samples were not collected at a specific time of day in order to ensure robustness of any identified metabolites. From the original samples, all treatment failure cases were included and matched by age, gender and extent of disease on chest X-rays to cured cases. Thus, the samples included were from 38 culture-confirmed active TB-positive South African patients (19 males and 19 females, aged between 17 and 58) at baseline (time of diagnosis, thus prior to initiation of treatment). These patients had drug-susceptible TB, were HIV-seronegative, not pregnant, and had no other diseases (i.e. diabetes, malignancy, lung cancer, chronic bronchitis and sarcoidosis), nor were they receiving any other medication, including antibiotics. These samples were divided into successful (n=27) and unsuccessful (n=11) treatment outcome groups.

6.3.2 Sample analysis

A creatinine value was determined for each urine sample collected, followed by an organic acid extraction and derivatisation (as described in section 3.3.3). Hereafter, the samples were analysed by randomly injecting each sample, along with the necessary QC samples, into a Pegasus 4D GCxGC-TOFMS and processed in order to identify compounds by comparison of their mass spectra to libraries generated from previously injected standards, as described in sections 3.3.4 and 3.3.5, respectively.

6.3.3 Statistical data analyses

It is standard practise in metabolomics investigations to perform various data clean-up procedures, prior to statistical data analyses (Van den Berg *et al.*, 2006). The concentrations of all the identified metabolites were calculated relative to the internal standard and expressed in relation to the creatinine value. A 50% zero filter was applied to each variable (Smuts *et al.*, 2013) to eliminate compounds with more than 50% zero values in both groups. Any batch effects were corrected for using quantile equating (Wang *et al.*, 2012), followed by a 50% quality control–coefficient of variation (QC–CV) filter. Finally, in order to select the most significant metabolite markers, various univariate statistical analyses (i.e. fold change, Mann-Whitney test and effect sizes) were applied to the untransformed data using MATLAB with Statistics and PLS Toolbox Release (2012; The MathWorks Inc., Natick, MA, USA) and SAS (2015; The SAS System for Windows Release 9.3 TS Level 1M0 Copyright© by SAS

Institute Inc., Cary, NC, USA). Fold change compares the absolute value changes of specific variables across two groups, and fold changes greater than two are considered significant (Du Preez & Loots, 2013). The Mann-Whitney test is considered the non-parametric alternative to the *t*-test, when assessing independent samples. This test determines significance by comparing the medians of variables across two groups, where a P-value below 0.05 is considered significant (Pallant, 2001). Effect sizes indicate practical significance by comparing variables individually between groups. For this non-parametric dataset, effect sizes were based on the associated Mann-Whitney z-values and an effect size above 0.3 (which indicates a moderate effect) were deemed practically relevant (Ellis & Steyn, 2003).

6.4 RESULTS AND DISCUSSION

As indicated in Figure 6-1, compounds with an absolute fold change exceeding two, a Mann-Whitney P-value of 0.05 or lower and an effect size larger than 0.3 were selected as the metabolite markers most capable of predicting treatment failure, and are listed in Table 6-1 (n=18). This multi-selection approach is based on the assumption that these different statistical methods compensate for each other's limitations and should reduce the rate of false discovery of unimportant compounds (Luies & Loots, 2016). A volcano plot was used to compare the size of the fold change to the statistical significance level by plotting the $\log_{10}$ scaled Mann-Whitney P-values against the $\log_2$ scaled fold change values (see Figure 6-2).

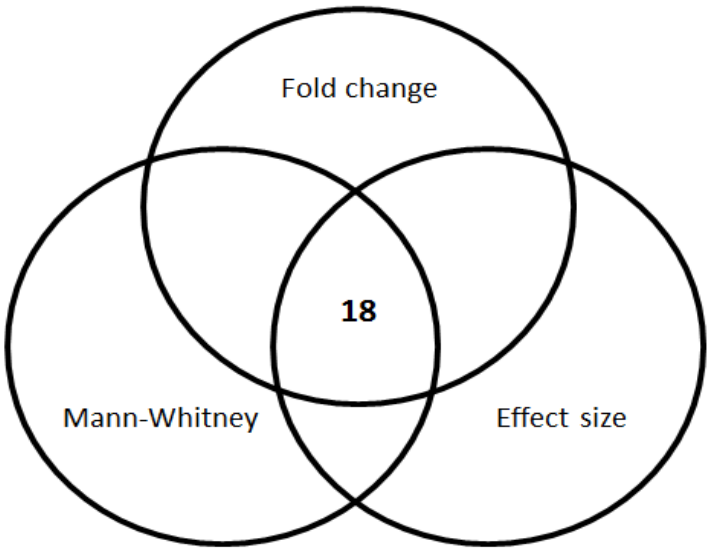


Figure 6-1: Venn-diagram indicating the multi-selection approach used to select the most significant metabolite markers for predicting treatment failure.

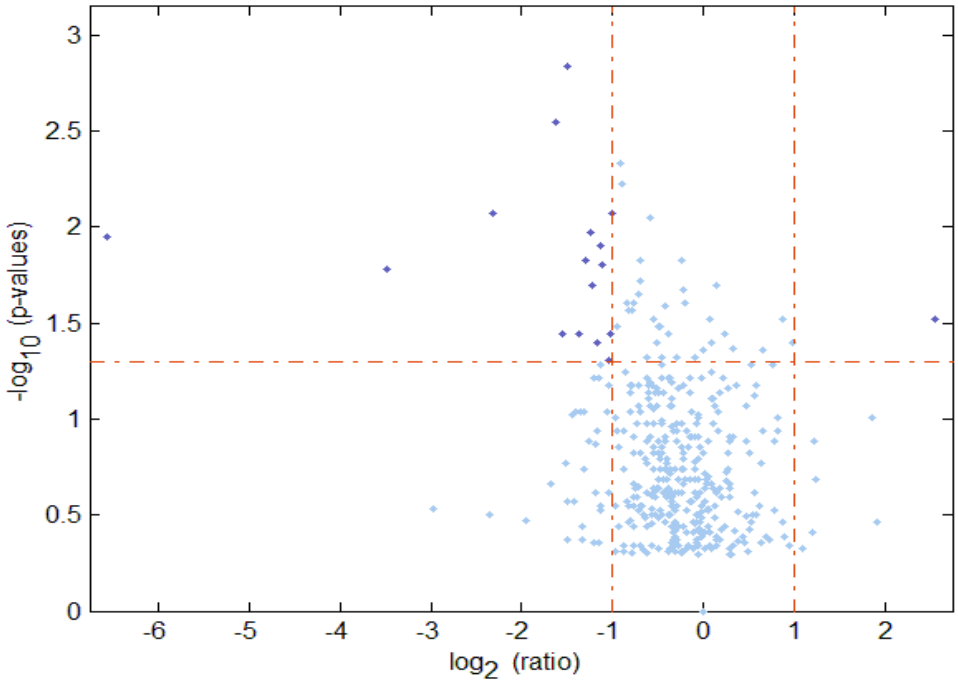


Figure 6-2: Volcano plot of the log_{10} scaled Mann-Whitney P-values against the log_2 scaled fold change values.

Table 6-1: The 18 univariate urine metabolite markers considered for potentially predicting treatment outcome at time of diagnosis.

Metabolite name	Retention time (RT) (s)	Successful treatment outcome: Average concentration (mg/g reatinine) (Standard deviation)	Samples detected (n=27)	Unsuccessful treatment outcome: Average concentration (mg/g creatinine) (Standard deviation)	Samples detected (n=11)	Fold change (absolute value)	Mann-Whitney test (P-value)	*Mann-Whitney Effect size (ω value)*
2,3,4,5-Tetrahydroxypentanoic acid-1,4-lactone	1396,1.59	17.509 (12.067)	95.20%	35.154 (35.4)	100%	2.008	0.008	0.429
2-Hydroxy-1,2,3-propanetricarboxylic acid	1573,1.52	13.388 (6.72)	100%	31.249 (36.752)	100%	2.334	0.02	0.368
2-Methoxy-5-methylphenol	1018,2.12	0.489 (0.35)	95.20%	5.458 (13.715)	100%	11.171	0.016	0.383
2-Octenoic acid	1021,1.69	0.576 (0.6)	95.20%	1.364 (1.018)	100%	2.367	0.011	0.414
3,5-Dihydroxybenzoic acid	1573,1.64	1.054 (0.49)	100%	2.374 (1.655)	100%	2.251	0.001	0.573
3-(4-Hydroxy-3-methoxyphenyl)propionic acid	1720,1.78	7.285 (4.019)	100%	22.472 (21.333)	100%	3.085	0.003	0.497
3-Hydroxy-hexanedioic acid	1444,1.48	1.737 (0.779)	100%	3.554 (3.813)	100%	2.046	0.05	0.296
3-Methoxy-4-hydroxyoxy-benzenepropanoic acid	1642,1.97	1.077 (1.799)	100%	2.649 (2.383)	100%	2.46	0.015	0.391
cis-4-Decene-1,10-dioic acid	1612,1.71	0.679 (0.706)	100%	1.759 (1.714)	100%	2.591	0.036	0.323
Phthalic acid mono-2-ethylhexyl ester	2128,2.53	11.231 (39.786)	100%	1.926 (1.077)	100%	5.832	0.03	0.338
tert-Butylhydroquinone	1363,1.58	0.45 (0.725)	90.50%	2.234 (2.357)	100%	4.969	0.008	0.429
Unknown 1	1489,1.7	0.43 (0.364)	100%	0.967 (0.858)	100%	2.251	0.04	0.315
Unknown 2	1648,1.77	0.158 (0.179)	95.20%	0.347 (0.232)	100%	2.201	0.013	0.402
Unknown 3	1753,1.59	0.087 (0.096)	100%	0.248 (0.21)	100%	2.838	0.001	0.535
Unknown 4	2050,1.84	98.212 (168.565)	100%	289.043 (416.387)	100%	2.943	0.036	0.323
Unknown 5	2101,1.85	8.294 (7.618)	100%	16.9 (13.627)	100%	2.038	0.036	0.323
Unknown 6	2116,2.05	0.129 (0.112)	95.20%	12.211 (36.959)	100%	94.334	0.011	0.41
Unknown 7	520,1.65	1.324 (0.511)	95.20%	2.874 (2.568)	100%	2.172	0.016	0.387

The predictive ability of these compounds (Table 6-1) were compared and combined by performing forward, backward and stepwise selection logistic regression. Therefore, all metabolites in Table 6-1 were given an equal opportunity to form part of the final model since all were provided to the logistic regression procedure. The procedure then considers various models, but penalizes models with higher numbers of predictors. Metabolites are added (forward), removed (backward) or both (stepwise) until a model is found which performs the best in terms of prediction and number of predictors. The final model was based on forward selection and was chosen as it showed the most predictive potential (i.e. had the highest area under the receiver operating characteristic) — it should be noted however that not all compounds with a high fold change or low Mann-Whitney p-value are necessarily good predictors (see Figure 6-3). The model contained two metabolite markers as potential predictors of treatment outcome, namely 3,5-dihydroxybenzoic acid and 3-(4-hydroxy-3-methoxyphenyl)propionic acid. The model showed acceptable fit (goodness-of-fit) with a non-significant Hosmer-Lemeshow statistic (P=0.444), as well as a reasonable relationship between treatment outcome and the above predictors with a maximum rescaled R-squared statistics of 0.75. Both predictors made significant contributions to the prediction, as indicated by the Wald criterion reported in Table 6-2.

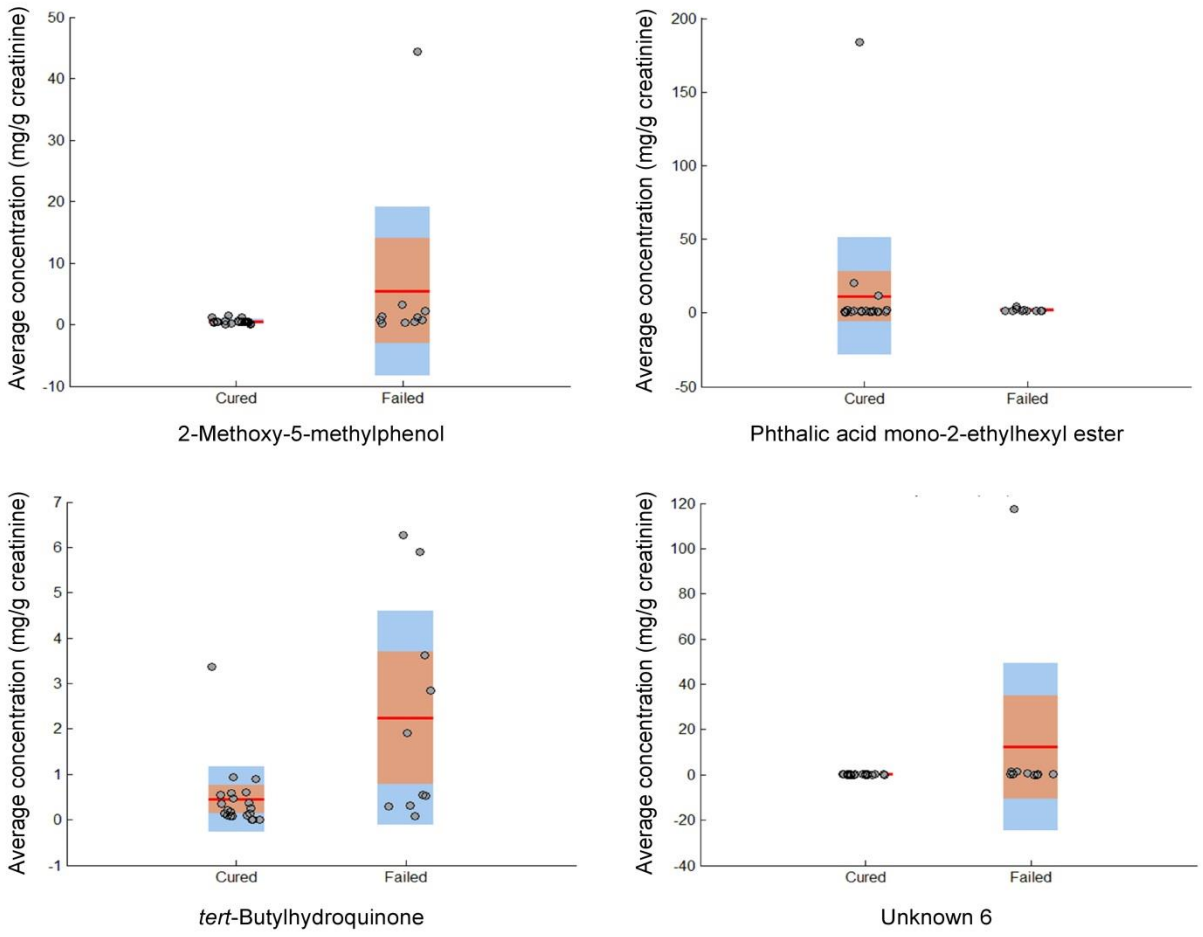


Figure 6-3: Box-like plots of the top four fold change metabolites, demonstrating that not all compounds are necessarily good predictors.

Table 6-2: The maximum likelihood parameter estimates and odds ratios.

Parameter	Estimate	Wald Chi-square	P-value	95% CI for Estimates	Odds Ratio (unit=1)	95% CI for Odds Ratio
3,5-Dihydroxybenzoic acid	3.24	5.71	0.017	1.2 – 6.8	25.6	3.2 – 886.1
3-(4-Hydroxy-3-methoxyphenyl) propionic acid	0.29	3.53	0.06	0.1 – 0.7	1.3	1.1 – 2.0

The odds ratio, reported for each of the possible predictors in Table 6-2, is a measure of the association between a specific feature (i.e. metabolite) and a treatment outcome (i.e. failure to respond to treatment) (Field, 2013). Thus, the odds ratio represents the increased or decreased probability that a patient will not be cured successfully (treatment failure) after anti-TB drug treatment, for every unit increase or decrease in the metabolite measured. An odds ratio of greater than 20 is considered significant for a diagnostic biomarker (Dunn *et al.*, 2012). An increase of one unit in the levels of 3,5-dihydroxybenzoic acid and 3-(4-hydroxy-3-methoxyphenyl)propionic acid are associated with an increased risk of treatment failure, i.e. the odds of treatment failure increase by 25.6 and 1.3, respectively, for each one unit increase. Although the odds ratio associated with 3-(4-hydroxy-3-methoxyphenyl)propionic acid are not indicative of a biomarker, the combination of multiple "weak" metabolite markers into a single model may also provide a high level of discrimination (Dunn *et al.*, 2012). The large confidence intervals (CIs) are due, as least in part, to the small groups used here and the relatively large variation within these groups. As with all newly developed models, further validation of these possible predictors is an absolute requirement prior to application.

The predictive ability of the model was assessed based on the receiver operating characteristic (ROC) curve, as shown in Figure 6-4. The model achieved an area under the ROC (AUC) value of 0.94 (95% CI 0.84 – 1) and cross-validated well in a leave-one-out context, with an AUC equal to 0.89 (95% CI 0.7 – 1). The predictive ability of a model is established if it achieves an AUC of 0.9 or higher (Dunn *et al.*, 2012), however, to truly validate the model will require a larger, independent sample cohort. Therefore, we focused our attention on interpreting the two possible predictors in the model.

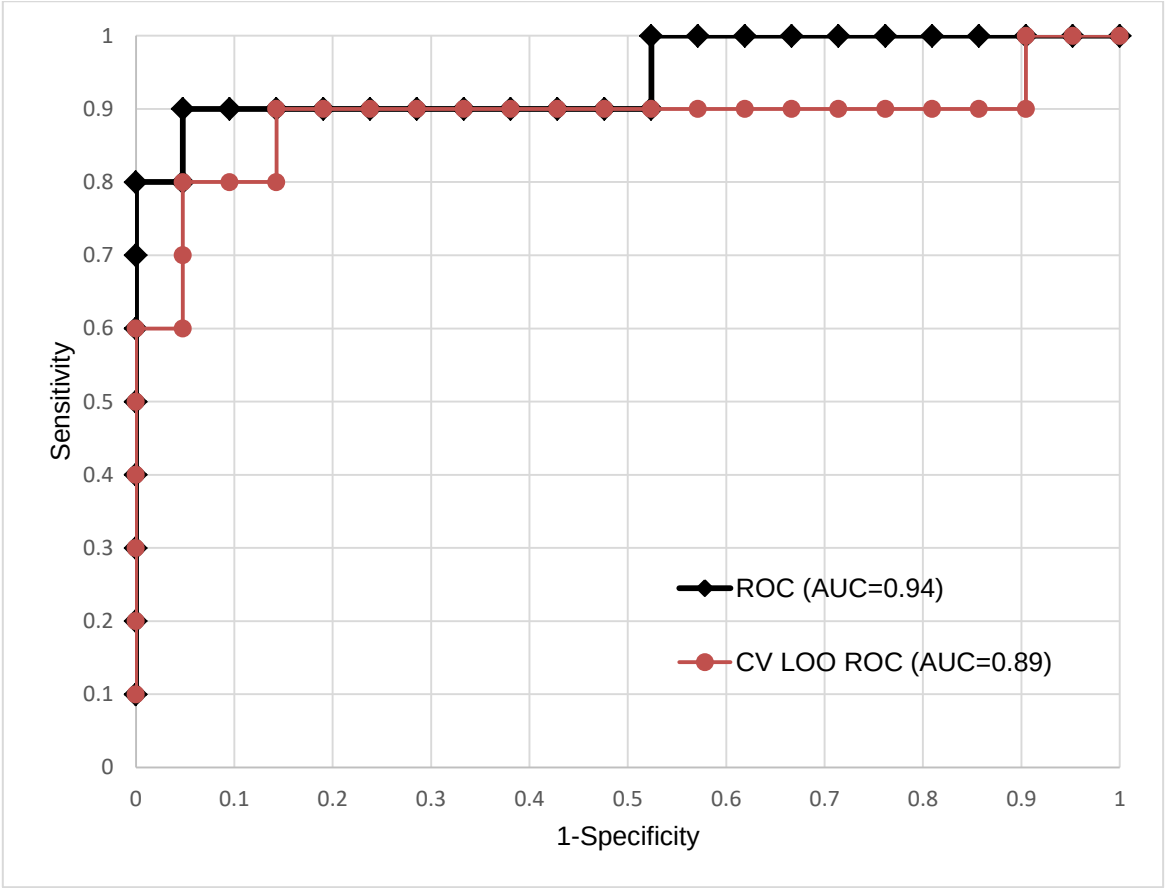


Figure 6-4: Receiver operating characteristic (ROC) curve of the final logistic regression model. The black line represents the ROC when all samples are used to access sensitivity and specificity of the model, while the red line represents the ROC when cross-validated using a leave-one-out approach.

The box-like plots of these two possible predictors (and their chemical structures), as illustrated in Figure 6-5, indicate the individual variation in the data. These plots were generated using notBoxPlot.m, a MATLAB function developed by Rob Campbell (http://www.mathworks.com/ matlabcentral/fileexchange/26508-raacampbell13-notboxplot). The raw data is scattered over a 95% confidence interval (red box) and one standard deviation (blue box) of the mean (red line). The individual observations (indicated by the grey dots) show a much larger variation in the unsuccessful treatment outcome group comparatively, as well as some overlap. It is therefore evident from these graphs that neither predictor will perform well in isolation. We also include a bivariate scatter plot (see Figure 6-6) to graphically illustrate the combined predictive ability of 3,5-dihydroxybenzoic acid and 3-(4-hydroxy-3-methoxyphenyl) propionic acid to discriminate between individuals with a successful and unsuccessful treatment outcome.

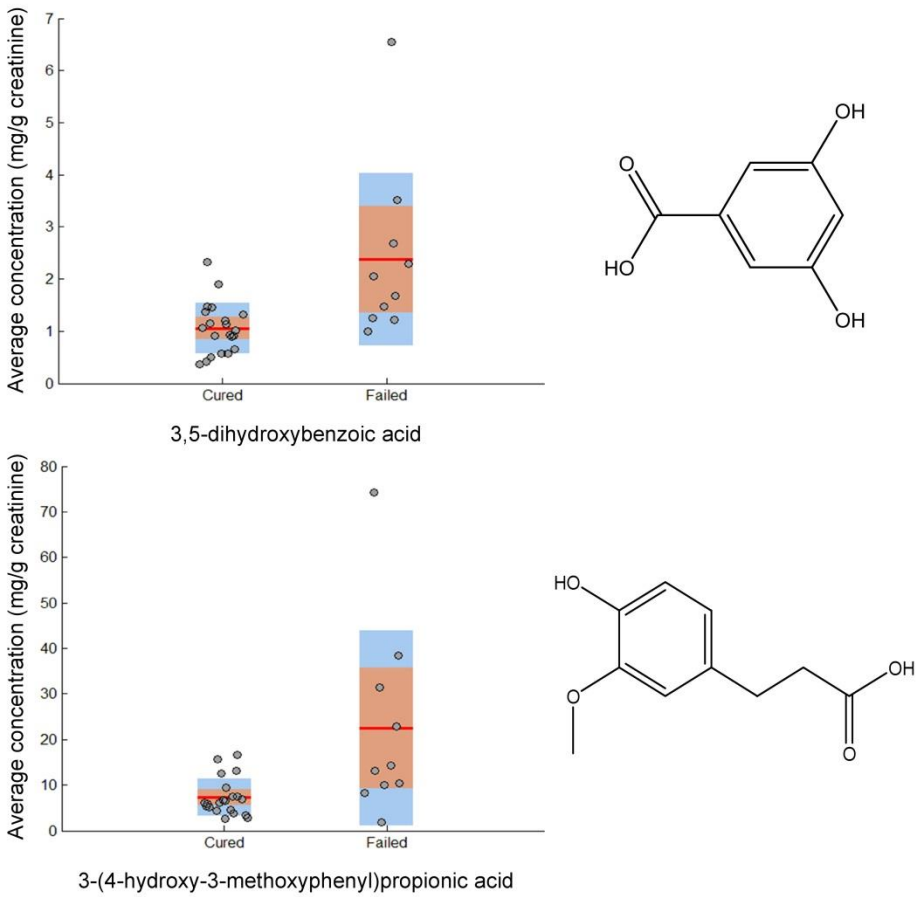


Figure 6-5: Box-like plots of the two identified possible predictors, with the accompanying chemical structures.

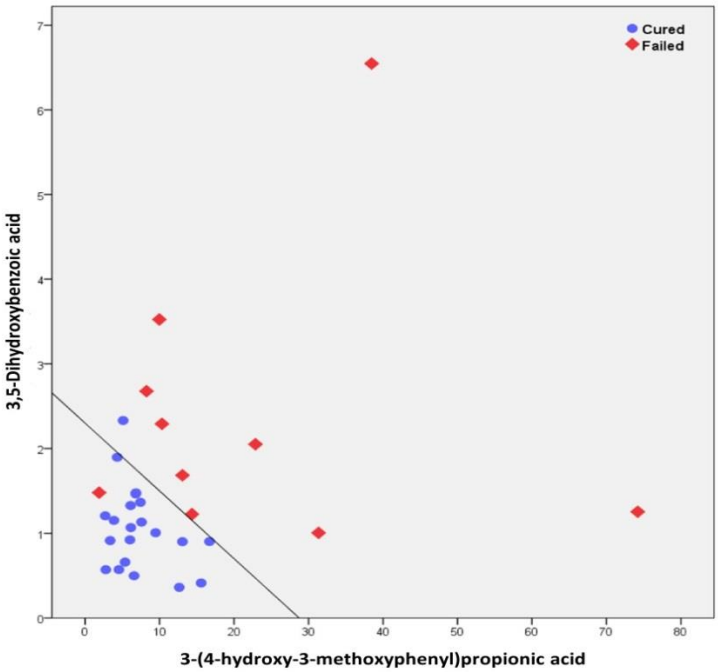


Figure 6-6: Bivariate scatter plot of the raw data generated for 3,5-dihydroxybenzoic acid versus 3-(4-hydroxy-3-methoxyphenyl)propionic acid. The line indicates clear separation between the two comparative groups based on the two metabolites selected as predictors.

When considering these two possible predictors from a biological mechanistic perspective, both 3,5-dihydroxybenzoic acid (α-resorcylic acid) (Liu *et al.*, 2012) and 3-(4-hydroxy-3-methoxyphenyl)propionic acid (dihydroferulic acid) (Rechner *et al.*, 2001) are compounds with aromatic rings associated with a microbial imbalance (Liu *et al.*, 2012; Rechner *et al.*, 2001), and also with a TB disease state (Luies & Loots, 2016; Singh & Toskes, 2004). Hence, an altered microbiome in the treatment failure group may be contributing to the lacking efficacy of the anti-TB drugs (Estudante *et al.*, 2013). These results also confirm previous findings by Luies *et al.*, (2017), who identified a number of metabolites related to a microbial imbalance, using a combination of multivariate and univariate statistical approaches, which characterised treatment failure to first-line drugs (Luies *et al.*, 2017). Since the innate and adaptive immune response systems rely on various microbiota interactions to promote immune cell maturation and function, a microbial imbalance results in a weakened immune system and an inability to fight disease (Clemente *et al.*, 2012). Additionally, the gut microbiota is considered an important contributor to xenobiotic/drug bioavailability and toxicity since it interacts with various drugs, influencing factors relating to their absorption, distribution and secretion, and hence plasma concentrations (Gonzalez *et al.*, 2011).

Compounds with aromatic rings (Estudante *et al.*, 2013; Vasiliou *et al.*, 2009; Warren *et al.*, 1975), including the majority of the anti-TB drugs used in this investigation (INH, RIF and PZA), and their phase II metabolites, are transported across cell membranes and eliminated via the ATP-binding cassette (ABC)-type multidrug transporters, which are highly expressed in human enterocytes of the liver, intestine, blood-brain barrier, blood-testis barrier, placenta and kidneys (Louw *et al.*, 2009; Vasiliou *et al.*, 2009) as well as on the cell surface of various intersinal microbiota (Mercado-Lubo, 2010). If for whatever reason (such as a microbial imbalance for instance), there is an overexpression/activity of these proteins on the host's cell surfaces, it would lead to an elevated excretion of these aromatics, explaining the elevated concentrations of the 3,5-dihydroxybenzoic acid and 3-(4-hydroxy-3-methoxyphenyl)propionic acid in the urine of the treatment failure patient group, suggesting that these aromatic anti-TB drugs would have suffered the same fate, possibly explaining treatment failure.

The genome of *M. tuberculosis* also contains 2.5% ABC transporters, which may be classified as either importers or exporters, determined by the direction of translocation of substrates (Louw *et al.*, 2009). The known mechanisms of drug-resistance to first-line anti-TB drugs, can be summarized as follows: INH-resistance is due to mutations in the *katG* and *inhA* genes; RIF-resistance results from alterations or point mutations in the *rpoB* gene of RNA polymerase; PZA-resistance may result from various mutations in the *pncA*, *rpsA* and

panD genes; and EMB-resistance is mainly due to mutations on the *embB* gene (Zhang & Yew, 2015). However, these mutations do not explain the mechanism of drug-resistance in all clinical isolates, suggesting that additional, unknown mechanisms may also be at play, such as the production of enzymes/proteins which modify/degrade drugs resulting in inactivation, cell wall impermeability and efflux mechanisms. Although such individual mechanisms may not result in clinical resistance, the combinations of these may confer high-level resistance, which may help explain the variation in resistance seen in *M. tuberculosis* isolates (Louw *et al.*, 2009).

Although, the scope of this investigation was only to identify metabolite markers and test their capacity for predicting treatment failure, using patient-collected urine samples, and not to elucidate the possible mechanisms for treatment failure, the above mentioned hypotheses are interesting topics for future investigation, and also serve as further validation for their possible use for predicting a failed treatment outcome to first-line anti-TB drugs. Furthermore, it may be interesting to compare these patients to healthy humans with a normal intestinal biosis in order to firmly allocate the microbiota dysbiosis to the treatment failure cohort. Additionally, the issue of whether the metabolites analysed in the urine of these patients are due to human host metabolism or whether they originate from the commensal intestinal bacteria can be addressed in future by experiments with germ-free and conventionalised *M. tuberculosis*-infected laboratory animals.

6.5 CONCLUSION

We were able to identify two urinary metabolite markers which possibly predict treatment failure of TB patients to first-line anti-TB drugs, before treatment onset. When these possible predictors were combined into a logistic regression model, they performed well with regards to sensitivity and specificity using leave-one-out cross-validation. It is important to note that the purpose of this model was to identify a combination of variables with high discriminatory ability containing as few variables as needed to achieve this, for potential diagnostic purposes. We propose that these two possible predictors warrant further investigation and validation using a bigger sample cohort. Once validated, the development of a prediction rule or predictive model can commence. Nonetheless, our findings show the capacity of metabolomics to identify patients at greater risk of a poor treatment outcome, at time of diagnosis, which potentially offers significant benefits for use in new drug development clinical trials and an individualized approach to patient care.

6.6 REFERENCES

Albuquerque, M.d.F.P.M., Ximenes, R.A.d.A., Lucena-Silva, N., Souza, W.V.d., Dantas, A.T., Dantas, O.M.S. & Rodrigues, L.C. 2007. Factors associated with treatment failure, dropout, and death in a cohort of tuberculosis patients in Recife, Pernambuco State, Brazil. *Cadernos de Saúde Pública*, 23(7):1573-1582.

Alobu, I., Oshi, D.C., Oshi, S.N. & Ukwaja, K.N. 2014. Profile and determinants of treatment failure among smear-positive pulmonary tuberculosis patients in Ebonyi, Southeastern Nigeria. *International Journal of Mycobacteriology*, 3(2):127-131.

Baumann, R., Kaempfer, S., Chegou, N.N., Nene, N.F., Veenstra, H., Spallek, R., Bolliger, C.T., Lukey, P.T., Van Helden, P.D. & Singh, M. 2012. Serodiagnostic markers for the prediction of the outcome of intensive phase tuberculosis therapy. *Tuberculosis*, 93:239-245.

Cannas, A., Calvo, L., Chiacchio, T., Cuzzi, G., Vanini, V., Lauria, F.N., Pucci, L., Girardi, E. & Goletti, D. 2010. IP-10 detection in urine is associated with lung diseases. *Bmc Infectious Diseases*, 10(1):333.

Che, N., Cheng, J., Li, H., Zhang, Z., Zhang, X., Ding, Z., Dong, F. & Li, C. 2013. Decreased serum 5-oxoproline in TB patients is associated with pathological damage of the lung. *Clinica Chimica Acta*, 423:5-9.

Clemente, J.C., Ursell, L.K., Parfrey, L.W. & Knight, R. 2012. The impact of the gut microbiota on human health: an integrative view. *Cell*, 148(6):1258-1270.

De Villiers, L. & Loots, D.T. 2013. Using metabolomics for elucidating the mechanisms related to tuberculosis treatment failure. *Current Metabolomics*, 1(4):306-317.

Du Preez, I. & Loots, D.T. 2013. New sputum metabolite markers implicating adaptations of the host to *Mycobacterium tuberculosis*, and vice versa. *Tuberculosis*, 93(3):330-337.

Dunn, W.B., Wilson, I.D., Nicholls, A.W. & Broadhurst, D. 2012. The importance of experimental design and QC samples in large-scale and MS-driven untargeted metabolomic studies of humans. *Bioanalysis*, 4(18):2249-2264.

Ellis, S.M. & Steyn, H.S. 2003. Practical significance (effect sizes) versus or in combination with statistical significance (p-values): Research note. *Management Dynamics: Journal of the Southern African Institute for Management Scientists*, 12(4):p. 51-53.

Estudante, M., Morais, J.G., Soveral, G. & Benet, L.Z. 2013. Intestinal drug transporters: an overview. *Advanced Drug Delivery Reviews*, 65(10):1340-1356.

Field, A. 2013. Discovering statistics using SPSS statistics. 3. London: Sage Publications Ltd.

Gonzalez, A., Stombaugh, J., Lozupone, C., Turnbaugh, P.J., Gordon, J.I. & Knight, R. 2011. The mind-body-microbial continuum. *Dialogues in Clinical Neuroscience*, 13(1):55.

Hesseling, A., Walzl, G., Enarson, D., Carroll, N., Duncan, K., Lukey, P., Lombard, C., Donald, P., Lawrence, K. & Gie, R. 2010. Baseline sputum time to detection predicts month two culture conversion and relapse in non-HIV-infected patients. *The International Journal of Tuberculosis and Lung Disease*, 14(5):560-570.

Horne, D.J., Royce, S.E., Gooze, L., Narita, M., Hopewell, P.C., Nahid, P. & Steingart, K.R. 2010. Sputum monitoring during tuberculosis treatment for predicting outcome: systematic review and meta-analysis. *The Lancet Infectious Diseases*, 10(6):387-394.

Krapp, F., Veliz, J., Cornejo, E., Gotuzzo, E. & Seas, C. 2008. Bodyweight gain to predict treatment outcome in patients with pulmonary tuberculosis in Peru. *The International Journal of Tuberculosis and Lung Disease*, 12(10):1153-1159.

Lawn, S.D., Kerkhoff, A.D., Burton, R., Schutz, C., Boulle, A., Vogt, M., Gupta-Wright, A., Nicol, M.P. & Meintjes, G. 2017. Diagnostic accuracy, incremental yield and prognostic value of Determine TB-LAM for routine diagnostic testing for tuberculosis in HIV-infected patients requiring acute hospital admission in South Africa: a prospective cohort. *BMC Medicine*, 15(1):67.

Lawn, S.D., Kerkhoff, A.D., Vogt, M. & Wood, R. 2012. Diagnostic accuracy of a low-cost, urine antigen, point-of-care screening assay for HIV-associated pulmonary tuberculosis before antiretroviral therapy: a descriptive study. *Lancet Infectious Diseases*, 12(3):201-209.

Liu, C., Kuei, C., Zhu, J., Yu, J., Zhang, L., Shih, A., Mirzadegan, T., Shelton, J., Sutton, S. & Connelly, M.A. 2012. 3,5-Dihydroxybenzoic acid, a specific agonist for hydroxycarboxylic acid 1, inhibits lipolysis in adipocytes. *Journal of Pharmacology and Experimental Therapeutics*, 341(3):794-801.

Louw, G., Warren, R., Van Pittius, N.G., McEvoy, C., Van Helden, P. & Victor, T. 2009. A balancing act: efflux/influx in mycobacterial drug resistance. *Antimicrobial Agents and Chemotherapy*, 53(8):3181-3189.

Luies, L. & Loots, D. 2016. Tuberculosis metabolomics reveals adaptations of man and microbe in order to outcompete and survive. *Metabolomics*, 12(3):1-9.

Luies, L., Mienie, J., Motshwane, C., Ronacher, K., Walzl, G. & Loots, D. Biochemistry. 2017. Urinary metabolite markers characterizing tuberculosis treatment failure (Unpublished manuscript).

Mercado-Lubo, R. 2010. The interaction of gut microbes with host ABC transporters. *Gut Microbes*, 1(5):301-306.

Mukherjee, J.S., Rich, M.L., Socci, A.R., Joseph, J.K., Virú, F.A., Shin, S.S., Furin, J.J., Becerra, M.C., Barry, D.J. & Kim, J.Y. 2004. Programmes and principles in treatment of multidrug-resistant tuberculosis. *The Lancet*, 363(9407):474-481.

Namukwaya, E., Nakwagala, F., Mulekya, F., Mayanja-Kizza, H. & Mugerwa, R. 2011. Predictors of treatment failure among pulmonary tuberculosis patients in Mulago hospital, Uganda. *African Health Sciences*, 11(3):105-111.

Olivier, I. & Loots, D.T. 2011. An overview of tuberculosis treatments and diagnostics. What role could metabolomics play. *Journal of Cell and Tissue Research*, 11(1):2655-2671.

Olivier, I. & Loots, D.T. 2012. A comparison of two extraction methods for differentiating and characterising various *Mycobacterium* species and *Pseudomonas aeruginosa* using GC-MS metabolomics. *African Journal of Microbiology Research*, 6(13):3159-3172.

Pallant, J. 2001. Manual SPSS Survival: A step by step guide to data analysis using SPSS (pp. 302): Open University Press, McGraw-Hill Education, Philadelphia, USA.

Petrone, L., Cannas, A., Aloi, F., Nsubuga, M., Sserumkuma, J., Nazziwa, R.A., Jugheli, L., Lukindo, T., Girardi, E. & Reither, K. 2015. Blood or urine IP-10 cannot discriminate between active

tuberculosis and respiratory diseases different from tuberculosis in children. *BioMedical Research International*, 2015.

Petrone, L., Cannas, A., Vanini, V., Cuzzi, G., Aloi, F., Nsubuga, M., Sserunkuma, J., Nazziwa, R.A., Jugheli, L., Lukindo, T., Girardi, E., Antinori, A., Pucci, L., Reither, K. & Goletti, D. 2016. Blood and urine inducible protein 10 as potential markers of disease activity. *International Journal of Tuberculosis and Lung Disease*, 20(11):1554-1561.

Rechner, A.R., Spencer, J.P.E., Kuhnle, G., Hahn, U. & Rice-Evans, C.A. 2001. Novel biomarkers of the metabolism of caffeic acid derivatives in vivo. *Free Radical Biology and Medicine*, 30(11):1213-1222.

Siawaya, J.F.D., Bapela, N.B., Ronacher, K., Veenstra, H., Kidd, M., Gie, R., Beyers, N., van Helden, P. & Walzl, G. 2008. Immune parameters as markers of tuberculosis extent of disease and early prediction of anti-tuberculosis chemotherapy response. *Journal of Infection*, 56(5):340-347.

Singh, V.V. & Toskes, P.P. 2004. Small bowel bacterial overgrowth: Presentation, diagnosis, and treatment. *Current Treatment Options in Gastroenterology*, 7(1):19-28.

Smuts, I., Van Der Westhuizen, F.H., Louw, R., Mienie, L.J., Engelke, U.F.H., Wevers, R.A., Mason, S., Koekemoer, G. & Reinecke, C.J. 2013. Disclosure of a putative biosignature for respiratory chain disorders through a metabolomics approach. *Metabolomics*, 9(2):379-391.

Van den Berg, R.A., Hoefsloot, H.C., Westerhuis, J.A., Smilde, A.K. & Van der Werf, M.J. 2006. Centering, scaling, and transformations: Improving the biological information content of metabolomics data. *BMC Genomics*, 7(1):1.

Vasiliou, V., Vasiliou, K. & Nebert, D.W. 2009. Human ATP-binding cassette (ABC) transporter family. *Human Genomics*, 3(3):281.

Walzl, G., Ronacher, K., Djoba Siawaya, J.F. & Dockrell, H.M. 2008. Biomarkers for TB treatment response: challenges and future strategies. *Journal of Infection*, 57(2):103-109.

Wang, X., Zhang, A. & Sun, H. 2012. Urine metabolomics. *Clinica Chimica Acta*, 414:65-69.

Warren, G., Houslay, M. & Metcalfe, J. 1975. Cholesterol is excluded from the phospholipid annulus surrounding an active calcium transport protein. *Nature*, 255:684-687.

World Health Organization. 2015. Global tuberculosis report 2015. Geneva, Switzerland (WHO Press).

Zhang, Y. & Yew, W. 2015. Mechanisms of drug resistance in *Mycobacterium tuberculosis*: Update 2015. *The International Journal of Tuberculosis and Lung Disease*, 19(11):1276-1289.

CHAPTER 7: FINAL CONCLUSIONS

Since the results were already discussed extensively in the previous chapters, this chapter will only summarise the main findings, in the context of the global investigation. Additional future recommendations/research prospects potentially emanating from this research investigation will also be discussed.

7.1 CONCLUDING SUMMARY

Despite the fervent research efforts to date focussing on the eradication of TB since the discovery of the causal agent in 1882 by Robert Koch, this disease still remains a global pandemic. One of the latest additions to the "omics" revolution, metabolomics, has allowed for the identification of new TB biomarkers, which has led to an exponential growth in the knowledge of TB disease mechanisms, as well as possible diagnostics. This research approach was also applied here to identify new TB biomarkers using patient-collected urine, and served well in explaining new host response mechanisms to active TB, as well as characterising and predicting treatment failure to first line anti-TB drugs.

Although only a few mycobacterial metabolites are expected to be present in the urine of TB-positive patients, it is considered a good sample matrix for identifying unique metabolites characterising the effects of TB on the host, i.e. host reaction and adaptations to the microbe, or altered host metabolome due to the disease. The fact that urine from TB patients can be easily collected, transported, stored, and carries low risk for infection, are additional advantages to using this sample matrix for research and possible diagnostic applications.

Before the data could be used, however, we needed to evaluate the repeatability of the analytical apparatus, in addition to that of the analyst/extraction methodology, to validate the generated data for use in the metabolomics applications pertaining to this study. After the necessary batch correction step was done on the entire GCxGC-TOFMS dataset, using the analysed QC samples, it was considered to be repeatable and reliable for metabolomics biomarker identification and the biological interpretation thereof, in the context of the aims of the study.

Metabolomics comparisons of a healthy TB-negative control group with an active TB-positive patient group led to the identification of new TB-associated urinary metabolite markers. These markers indicated that *M. tuberculosis*/TB induces an abnormal host fatty acid and amino acid profile (affecting tryptophan, phenylalanine and tyrosine in particular), most likely

due to the host IFN-γ response and possibly a compromised insulin production in these patients, substantiated by prior findings of our research team and others. These results also gave further clues explaining some of the TB-associated symptoms, providing suggestions regarding improved treatment approaches.

Next, we investigated the underlying mechanisms of TB treatment failure by comparing the urinary metabolome of TB-positive patients who received first-line anti-TB drugs and were successfully cured, to those with an unsuccessful treatment outcome, at baseline (time of diagnosis, thus prior to initiation of treatment), during the course of DOTS (i.e. weeks 1, 2 and 4) and two weeks post-treatment completion (week 26). The metabolite profiles of these TB-positive individuals showed clear differentiation of the groups at time of diagnosis. The treatment failure group was characterised by metabolite markers indicative of an imbalance in their gut microbiome, in addition to elevated levels of those metabolites associated with abnormalities in the long-chain fatty acid β-oxidation pathway, accompanied by reduced L-carnitine and SCFAs, indicative of a possible MTP protein defect in particular. The altered amino acid metabolites also observed in the treatment failure patients, confirms our previous findings and associations of treatment failure with an increased IFN-γ host immune response to *M. tuberculosis*, and also possibly compromised insulin secretion.

Since these results showed that a metabolomics approach can differentiate active TB-positive individuals based on treatment outcome, at baseline, this data was reanalysed using a univariate statistical approach. Two metabolites were identified as potential predictors of an unsuccessful treatment outcome, and when combined in a forward selection logistic regression model, performed well with regards to both sensitivity and specificity when using a leave-one-out cross-validation. Interestingly, these two predictors are also well-known to be associated with a gut microbiome imbalance, further substantiating this mechanism as the major contributor to an unsuccessful treatment outcome.

7.2 FUTURE RESEARCH PROSPECTS

As suggested in Chapter 4, co-administration of anti-TB drugs with melatonin may assist in reducing the elevated quinolinic acid, reduce a number of toxic metabolites associated with the side-effects of anti-TB drugs, and also increase its efficacy in eliminating *M. tuberculosis*. Furthermore, BH_4 is given as a treatment for PKU patients in an attempt to reduce phenylacetic acid; hence co-administration of anti-TB drugs with this also merits consideration and further research. Lastly, considering the association between TB and diabetes, co-administration of anti-TB drugs with antidiabetic drugs, such as metformin, could also be investigated.

Additionally, since the results obtained in Chapters 5 and 6 suggest an imbalance in the gut microbiome, the synchronous use of probiotics for optimising the microbiome during first-line anti-TB treatment may improve treatment outcome, and could be a topic of further investigation. It may also be interesting to compare these patients to a cohort of healthy humans with a normal intestinal biosis in order to firmly allocate the microbiota dysbiosis to the treatment failure cohort. The issue of whether the metabolites analysed in the urine of these patients are due to human host metabolism or whether they originate from the commensal intestinal bacteria can be addressed in future by experiments with germ-free and conventionalised *M. tuberculosis*-infected laboratory animals.

Considering the data already made available in this study, the same metabolomics biomarker identification statistics methodology could also be used to compare the data generated pre- and post-treatment administration in order to identify markers (a) characterising the host response to treatment, (b) elucidating the mechanism(s) of drug action/side-effects, and (c) identifying inter-individual variations in the host response to treatment and mechanisms of drug action from a metabolomics perspective, and therefore generate data better describing the underlying cause of successful or unsuccessful treatment outcomes.

Additionally, given the longitudinal nature of this investigation when comparing the urinary metabolomes of the successful and unsuccessful treatment outcomes patient groups at various time intervals, it may be interesting to explore the metabolic changes in these patients over time, to determine the capacity of these markers for monitoring treatment response.

Table A-1: A summary of those compounds detected as biomarkers for TB, by more than one research group, in more than one analytical sample media.

Biomarker	Sample media	Sample cohort	Corresponding reference
Biomarkers detected by the same research group in pilot and follow-up validation study (same sample media) (n=9)			
11-Eicosenoic acid	Culture	*Mycobacterium spp.*	Olivier and Loots (2012a) Olivier and Loots (2012b)
13-Docosenoic acid	Culture	*Mycobacterium spp.*	Olivier and Loots (2012a) Olivier and Loots (2012b)
15-Tetracosenoic acid	Culture	*Mycobacterium spp.*	Olivier and Loots (2012a) Olivier and Loots (2012b)
9-Hexadecenoic acid	Culture	*Mycobacterium spp.*	Olivier and Loots (2012a) Olivier and Loots (2012b)
Cadaverine	Culture **Sputum spiked with culture**	*Mycobacterium spp.* *M. tuberculosis*	Olivier and Loots (2012a) Schoeman *et al.* (2012)
Eicosanoic acid	Culture **Sputum spiked with culture**	*Mycobacterium spp.* *M. tuberculosis*	Olivier and Loots (2012a) Schoeman *et al.* (2012)
Heptadecanoic acid	Culture **Sputum spiked with culture**	*Mycobacterium spp.* *M. tuberculosis*	Olivier and Loots (2012a) Olivier and Loots (2012b) Schoeman *et al.* (2012)
Tridecane	**Breath**	TB+, TB-, HC TB+, TB-	Phillips *et al.* (2007) Phillips *et al.* (2010)
Tuberculostearic acid	Culture **Sputum spiked with culture**	*Mycobacterium spp.* *M. tuberculosis*	Olivier and Loots (2012a) Olivier and Loots (2012b) Schoeman *et al.* (2012)
Biomarkers detected by the same research group in pilot and follow-up validation study (different sample media) (n=6)			
10-Heptadecenoic acid	Culture **Sputum**	*Mycobacterium spp.* TB+, TB-	Olivier and Loots (2012b) Du Preez and Loots (2013)
9-Octadecenoic acid	Culture **Sputum spiked with culture** **Sputum**	*Mycobacterium spp.* *M. tuberculosis* TB+, TB-	Olivier and Loots (2012a) Schoeman *et al.* (2012) Du Preez and Loots (2013)
alpha-D-glucopyranoside	*Sputum spiked with culture* **Sputum**	*M. tuberculosis* TB+, TB-	Schoeman *et al.* (2012) Du Preez and Loots (2013)
D-glucosamine	**Sputum spiked with culture** **Sputum**	*M. tuberculosis* TB+, TB-	Schoeman *et al.* (2012) Du Preez and Loots (2013)
Methyl nicotinate	Culture **Breath**	*Mycobacterium spp.* TB+, HC	Syhre and Chambers (2008) Syhre *et al.* (2009)
Nonadecanoic acid	**Sputum spiked with culture** **Sputum**	*M. tuberculosis* TB+, TB-	Schoeman *et al.* (2012) Du Preez and Loots (2013)

Biomarkers detected in the same study, but in different sample media (n=13)			
1,4-Dimethyl-cyclohexane	**Culture** **Breath**	*M. tuberculosis* TB+, TB-, HC	Phillips *et al.* (2007)
1-Methyl-naphthalene	**Culture** **Breath**	*M. tuberculosis* TB+, TB-, HC	Phillips *et al.* (2007)
Acetate	**Lung tissue (Guinea pigs), Serum (Guinea pigs)**	TB+, HC	Somashekar *et al.* (2012)
AMP	**Lung tissue (Mice), Spleen tissue (Mice), Liver tissue (Mice)**	TB+, HC	Shin *et al.* (2011)
Fumarate	*Lung tissue (Mice), Spleen tissue (Mice)*	TB+, HC	Shin *et al.* (2011)
Leucine	**Lung tissue (Mice), Spleen tissue (Mice),Liver tissue (Mice), Serum (Mice)**	TB+, HC	Shin *et al.* (2011)
Nadp+	*Lung tissue (Mice), Spleen tissue (Mice)*	TB+, HC	Shin *et al.* (2011)
Niacinamide	**Lung tissue (Mice), Spleen tissue (Mice),** *Liver tissue (Mice)*	TB+, HC	Shin *et al.* (2011)
Oxaloacetate	*Lung tissue (Mice), Spleen tissue (Mice), Liver tissue (Mice)*	TB+, HC	Shin *et al.* (2011)
Succinate	**Lung tissue (Mice), Spleen tissue (Mice), Liver tissue (Mice)**	TB+, HC	Shin *et al.* (2011)
Taurine	**Lung tissue (Mice), Liver tissue (Mice)**	TB+, HC	Shin *et al.* (2011)
UDP-glucose	**Lung tissue (Mice), Liver tissue (Mice)**	TB+, HC	Shin *et al.* (2011)
Uracil	**Lung tissue (Mice),Liver tissue (Mice)**	TB+, HC	Shin *et al.* (2011)
Biomarkers detected by different research groups using the same sample media (n=2)			
Betaine	**Lung tissue (Mice)** **Lung tissue (Guinea pigs)**	TB+, HC TB+, HC	Shin *et al.* (2011) Somashekar *et al.* (2012)
Formate	**Serum (Mice)** **Serum (Guinea pigs)** *Serum (Human)*	TB+, HC TB+, HC TB+, HC	Shin *et al.* (2011) Somashekar *et al.* (2012) Zhou *et al.* (2013)
Biomarkers detected by different study groups using different sample media (n=19)			
Alanine	**Lung tissue (Mice), Spleen tissue (Mice)** **Lung tissue (Guinea pigs)** *Serum (Human)*	TB+, HC TB+, HC TB+, HC	Shin *et al.* (2011) Somashekar *et al.* (2012) Zhou *et al.* (2013)
Aspartate	**Lung tissue (Mice), Liver tissue (Mice)** **Lung tissue (Guinea pigs)**	TB+, HC TB+, HC	Shin *et al.* (2011) Somashekar *et al.* (2012)
Creatine	**Lung tissue (Mice), Liver tissue (Mice)** **Lung tissue (Guinea pigs)**	TB+, HC TB+, HC	Shin *et al.* (2011) Somashekar *et al.* (2012)
D-mannose	**Sputum spiked with culture** **Serum (Human)**	TB+, TB- TB+, TB-, HC	Schoeman *et al.* (2012) Weiner *et al.* (2012)
Docosanoic acid	Culture *Serum (Human)*	*Mycobacterium spp.* TB+, HC, OD	Olivier and Loots (2012a) Feng *et al.* (2015)
Glucose	*Lung tissue (Mice), Liver tissue (Mice*), **Serum (Mice)** **Urine**	TB+, HC TB+, TB-	Shin *et al.* (2011) Das *et al.* (2015)

Glutamate	**Lung tissue (Mice), Liver tissue (Mice)** *Serum (Guinea pigs)* **Serum (Human)**	TB+, HC TB+, HC TB+, HC	Shin *et al.* (2011) Somashekar *et al.* (2012) Zhou *et al.* (2013)
Glutamine	**Lung tissue (Mice), Spleen tissue (Mice), Liver tissue (Mice)** **Serum (Human)** *Serum (Human)*	TB+, HC TB+, HC TB+, TB-, HC	Shin *et al.* (2011) Zhou *et al.* (2013) Weiner *et al.* (2012)
Glutathione	**Lung tissue (Mice), Spleen tissue (Mice)** **Lung tissue (Guinea pigs)**	TB+, HC TB+, HC	Shin *et al.* (2011) Somashekar *et al.* (2012)
Isoleucine	**Lung tissue (Mice), Spleen tissue (Mice), Liver tissue (Mice), Serum (Mice)** **Serum (Human)**	TB+, HC TB+, HC	Shin *et al.* (2011) Zhou *et al.* (2013)
Lactate	**Lung tissue (Mice), Spleen tissue (Mice), Liver tissue (Mice)** **Lung tissue (Guinea pigs**), *Serum (Guinea pigs)* **Serum (Human)**	TB+, HC TB+, HC TB+ HC	Shin *et al.* (2011) Somashekar *et al.* (2012) Zhou *et al.* (2013)
L-threonine	**Sputum spiked with culture** *Serum (Human)*	TB+, TB- TB+, TB-, HC	Schoeman *et al.* (2012) Weiner *et al.* (2012)
Lysine	**Lung tissue (Mice), Liver tissue (Mice)** **Serum (Human)**	TB+, HC TB+, HC	Shin *et al.* (2011) Zhou *et al.* (2013)
Myo-inositol	Culture **Lung tissue (Guinea pigs)**	*Mycobacterium spp.* TB+, TB-	Olivier and Loots (2012a) Somashekar *et al.* (2012)
Phenylalanine	**Lung tissue (Mice), Liver tissue (Mice), Serum (Mice)** **Serum (Human)** **Serum (Human)** **Urine**	TB+, HC TB+, HC TB+, TB-, HC TB+, HC	Shin *et al.* (2011) Zhou *et al.* (2013) Weiner *et al.* (2012) Luies and Loots (2016)
Phosphocholine	**Lung tissue (Mice), Spleen tissue (Mice), Liver tissue (Mice)** **Lung tissue (Guinea pigs)**	TB+, HC TB+, HC	Shin *et al.* (2011) Somashekar *et al.* (2012)
Quinolinic acid	**Serum (Human)** **Urine**	TB+, HC, OD TB+, TB-	Feng *et al.* (2015) Luies and Loots (2016)
Tyrosine	**Lung tissue (Mice), Spleen tissue (Mice), Liver tissue (Mice)** **Serum (Human)** **Urine**	TB+, HC TB+, HC TB+, HC	Shin *et al.* (2011) Zhou *et al.* (2013) Luies and Loots (2016)
Uridine	**Sputum spiked with culture** **Lung tissue (Mice),** *Spleen tissue (Mice), Liver tissue (Mice)*	TB+, TB- TB+, HC	Schoeman *et al.* (2012) Shin *et al.* (2011)

Abbreviations: TB+: positively diagnosed for active TB via culture; TB-: patients portraying TB-related symptoms, but tested negative for TB via culture; Healthy control: HC, healthy control with no TB-related symptoms and tested negative for TB via culture, OD, other, related diseases

Bold = increased abundance in TB+ group as compared to TB- and/or HC

Italics = decreased abundance in TB+ group as compared to TB- and/or HC

Normal text = concentration variations between different Mycobacterium species

References:

Das, M.K., Bishwal, S.C., Das, A., Dabral, D., Badireddy, V.K., Pandit, B., Varghese, G.M. & Nanda, R.K. 2015. Deregulated tyrosine–phenylalanine metabolism in pulmonary tuberculosis patients. *Journal of Proteome Research*, 14(4):1947-1956.

Du Preez, I. & Loots, D.T. 2013b. New sputum metabolite markers implicating adaptations of the host to *Mycobacterium tuberculosis*, and vice versa. *Tuberculosis*, 93(3):330-337.

Feng, S., Du, Y.Q., Zhang, L., Zhang, L., Feng, R.R. & Liu, S.Y. 2015. Analysis of serum metabolic profile by ultra-performance liquid chromatography-mass spectrometry for biomarkers discovery: Application in a pilot study to discriminate patients with tuberculosis. *Chinese Medical Journal*, 128(2):159-168.

Luies, L. & Loots, D. 2016. Tuberculosis metabolomics reveals adaptations of man and microbe in order to outcompete and survive. *Metabolomics*, 12(3):1-9.

Olivier, I. & Loots, D.T. 2012a. A comparison of two extraction methods for differentiating and characterising various *Mycobacterium* species and *Pseudomonas aeruginosa* using GC-MS metabolomics. *African Journal of Microbiology Research*, 6(13):3159-3172.

Olivier, I. & Loots, D.T. 2012b. A metabolomics approach to characterise and identify various *Mycobacterium* species. *Journal of Microbiological Methods*, 88(3):419-426.

Phillips, M., Basa-Dalay, V., Bothamley, G., Cataneo, R.N., Lam, P.K., Natividad, M.P.R., Schmitt, P. & Wai, J. 2010. Breath biomarkers of active pulmonary tuberculosis. *Tuberculosis*, 90(2):145-151.

Phillips, M., Cataneo, R.N., Condos, R., Erickson, G.A.R., Greenberg, J., La Bombardi, V., Munawar, M.I. & Tietje, O. 2007. Volatile biomarkers of pulmonary tuberculosis in the breath. *Tuberculosis*, 87(1):44-52.

Schoeman, J.C., Du Preez, I. & Loots, D.T. 2012. A comparison of four sputum pre-extraction preparation methods for identifying and characterising *M. tuberculosis* using GCxGC-TOFMS metabolomics. *Journal of Microbiological Methods*, 91:301-311.

Shin, J.-H., Yang, J.-Y., Jeon, B.-Y., Yoon, Y.J., Cho, S.-N., Kang, Y.-H., Ryu, D.H. & Hwang, G.-S. 2011. 1H NMR-based metabolomic profiling in mice infected with *Mycobacterium tuberculosis*. *Journal of Proteome Research*, 10(5):2238-2247.

Somashekar, B.S., Amin, A.G., Tripathi, P., MacKinnon, N., Rithner, C.D., Shanley, C.A., Basaraba, R., Henao-Tamayo, M., Kato-Maeda, M. & Ramamoorthy, A. 2012. Metabolomic signatures in guinea pigs infected with epidemic-associated W-Beijing strains of *Mycobacterium tuberculosis*. *Journal of Proteome Research*, 11(10):4873-4884.

Syhre, M. & Chambers, S.T. 2008. The scent of *Mycobacterium tuberculosis*. *Tuberculosis*, 88(4):317-323.

Syhre, M., Manning, L., Phuanukoonnon, S., Harino, P. & Chambers, S.T. 2009. The scent of *Mycobacterium tuberculosis*–part II breath. *Tuberculosis*, 89(4):263-266.

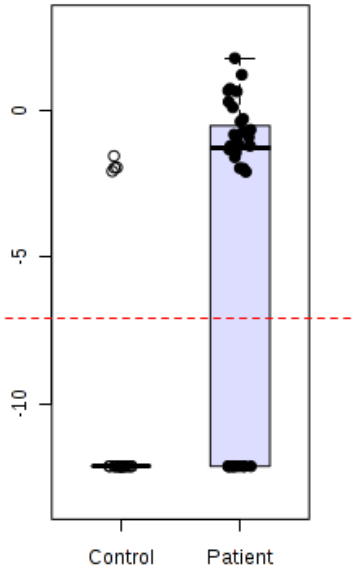
Weiner, J., Parida, S.K., Maertzdorf, J., Black, G.F., Repsilber, D., Telaar, A., Mohney, R.P., Arndt-Sullivan, C., Ganoza, C.A. & Fae, K.C. 2012. Biomarkers of inflammation, immunosuppression and stress with active disease are revealed by metabolomic profiling of tuberculosis patients. *PloS one*, 7(7):e40221.

Zhou, A., Ni, J., Xu, Z., Wang, Y., Lu, S., Sha, W., Karakousis, P.C. & Yao, Y.-F. 2013. Application of 1H NMR spectroscopy-based metabolomics to sera of tuberculosis patients. *Journal of Proteome Research*, 12(10):4642-4649.

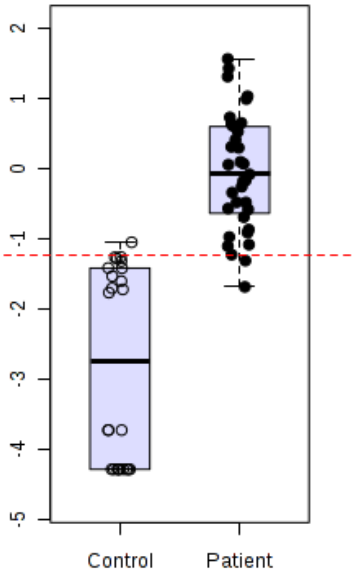
APPENDIX B

Box-and-whisker plots of the 12 compounds that best explained the variation between the TB-negative healthy control and TB-positive patients, selected on the basis of a PCA modelling power >0.5, a PLS-DA VIP >1.0, an effect size >0.5, and a P-value <0.05. These plots illustrate the data distribution around the median (thick horizontal line in centre of the box), the 25th and 75th percentiles/quartiles (bottom and top of the box, respectively), as well as the 5th and 95th percentiles (ends of the whiskers).

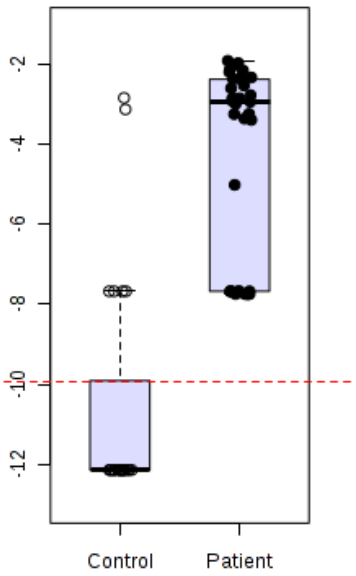
5-Hydroxyhexanoic acid



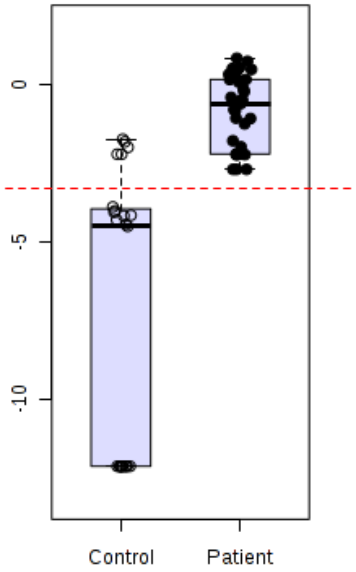
Phenylacetic acid



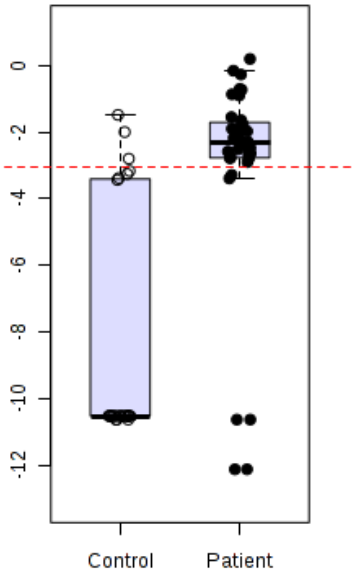
2-Octenoic acid



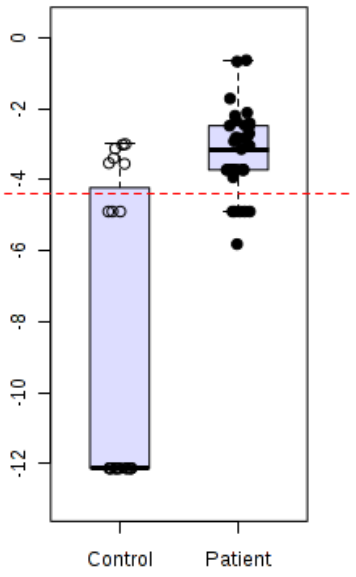
2-C-Methylglycerol



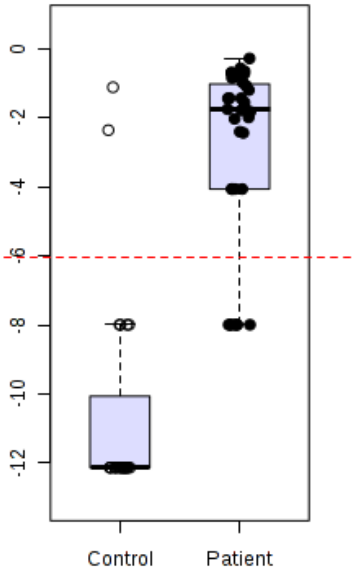
5-Hydroxyhydantoin



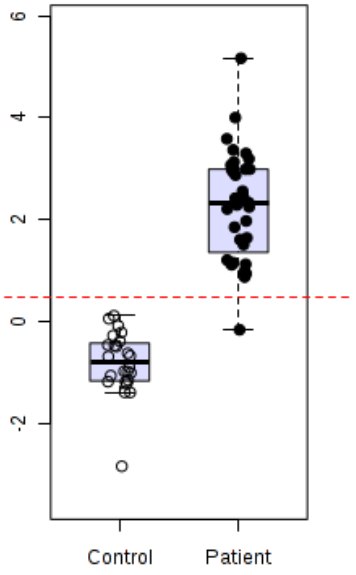
Oxalic acid



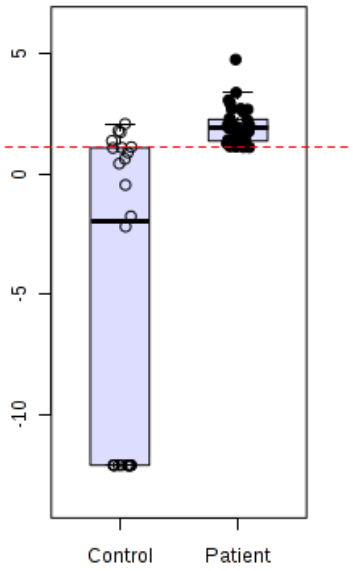
L-Rhamnulose



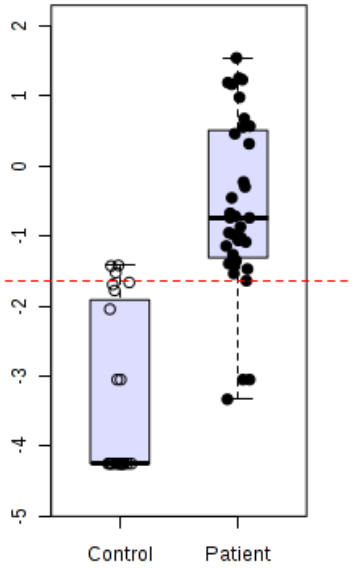
Quinolinic acid



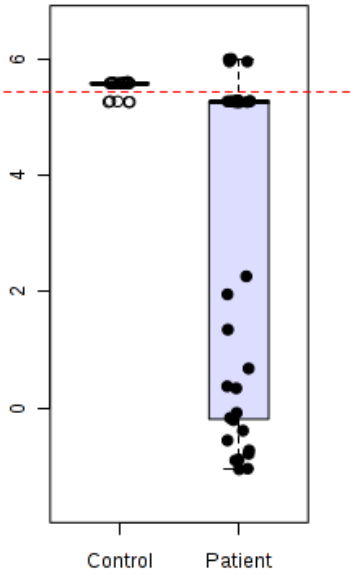
Ribitol



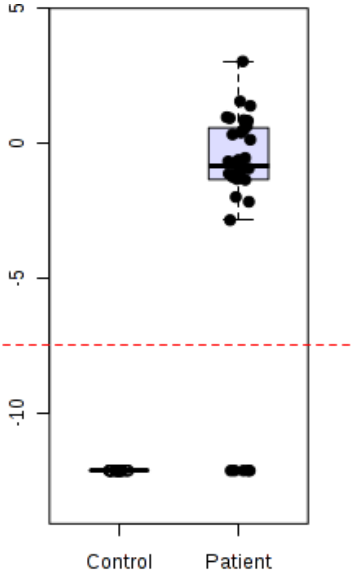
Indole-3-carboxylic acid



Kynurenic acid



Glycerol monostearate



APPENDIX C

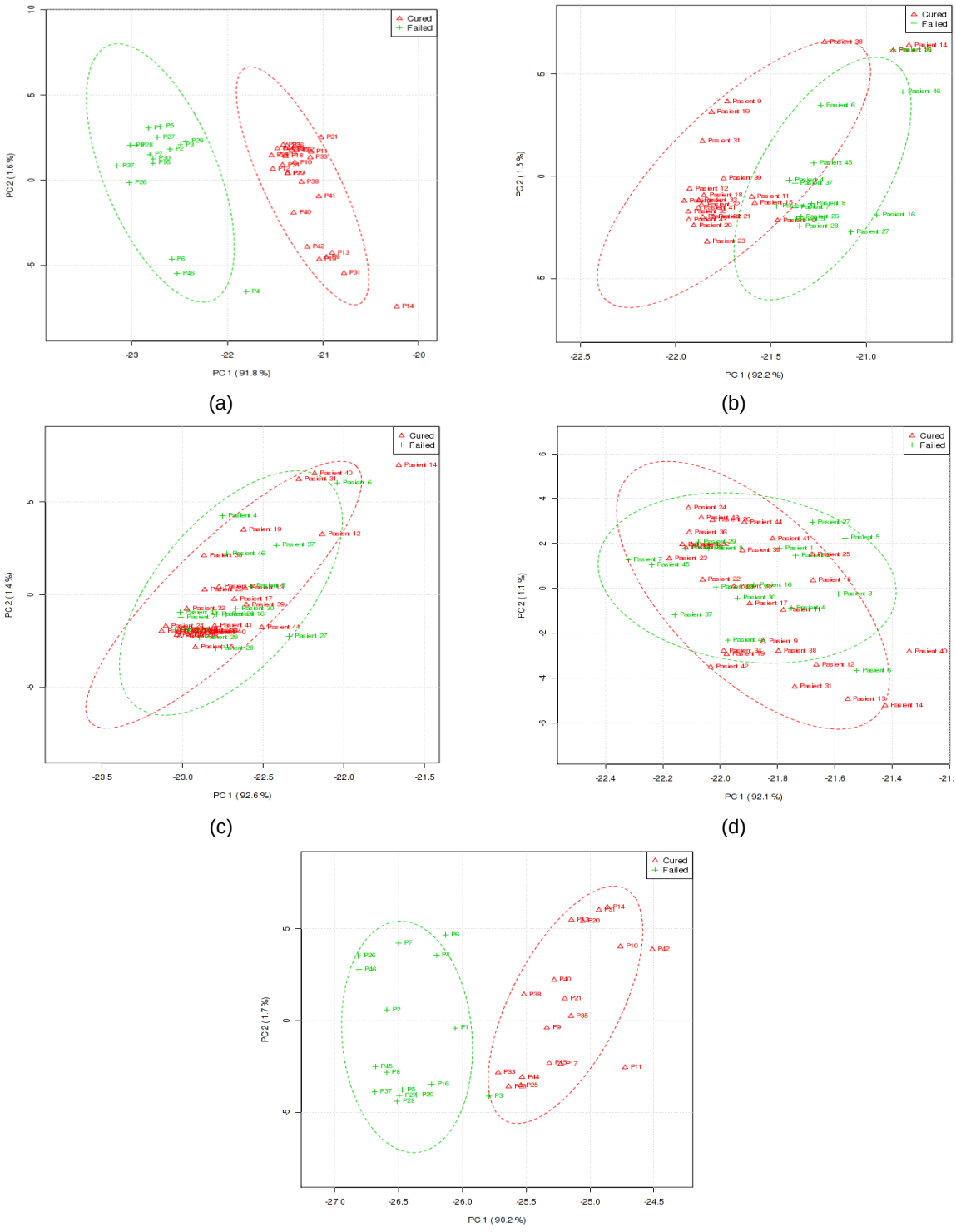


Figure A-1: Principal components analysis (PCA) scores plots of principal component 1 versus principal component 2 of the successful and unsuccessful treatment outcome groups, at (a) time of diagnosis, (b) week 1, (c) week 2, (d) week 4 of treatment and (e) two weeks after treatment completion (week 26), subsequent to an organic acid extraction and GCxGC-TOFMS analysis. The explained variances are indicated in parenthesis.

APPENDIX D

The following is a list of all publications originating from this metabolomics investigation:

- De Villiers, L. & Loots, DT. (2013). Using metabolomics for elucidating the mechanisms related to tuberculosis treatment failure. *Current Metabolomics*, 1(4): 306-317.

- Du Preez, I., Luies, L. & Loots, DT. (2017). Metabolomics biomarkers for tuberculosis diagnostics: Current status and future objectives. *Biomarkers in Medicine,* 11 (2): 179-194.

- Luies, L., Du Preez, I. & Loots, DT. (2017). The role of metabolomics in tuberculosis treatment research. Submitted for publication to *Biomarkers in Medicine* (Manuscript number: BMM-2017-0141).

- Luies, L. & Loots, DT. (2016). Tuberculosis metabolomics reveals adaptations of man and microbe in order to outcompete and survive. *Metabolomics*, 12(3): 1-9.

- Luies, L., Mienie, J., Motshwane, C., Ronacher, K., Walzl, G. & Loots, DT. (2017). Urinary metabolite markers characterising tuberculosis treatment failure. Submitted for publication in *Metabolomics* (Manuscript number: MEBO-D-17-00069).

- Luies, L., Van Reenen, M., Ronacher, K., Walzl, G. & Loots, DT. (2017). Predicting tuberculosis treatment outcome using metabolomics. Submitted for publication to *Biomarkers in Medicine* (Manuscript number: BMM-2017-0133).

Send Orders for Reprints to reprints@benthamscience.net

306 *Current Metabolomics,* **2013**, *1*, 306-317

Using Metabolomics for Elucidating the Mechanisms Related to Tuberculosis Treatment Failure

Laneke de Villiers and Du Toit Loots*

School for Physical and Chemical Sciences, Centre for Human Metabonomics, North-West University (Potchefstroom Campus), Private Bag x6001, Box 269, Potchefstroom, 2531, South Africa

Abstract: Tuberculosis (TB), caused by *Mycobacterium tuberculosis* is a highly infectious disease, responsible for 1.4 million deaths annually. Unfortunately, the latest reports of treatment outcomes for patients using the current anti-TB drugs are disappointing, considering the high treatment failure rates. This has been ascribed to a number of factors, including, amongst others: (1) variable individual metabolism (including xenobiotic metabolism, drug malabsorption and drug-drug interactions), (2) drug resistance by the infectious organism and (3) non-adherence to the treatment program, especially due to the associated drug side-effects. Despite this however, only two new drugs have been approved by the FDA since 1962. Thus, it is clear that new strategies are needed for the better elucidation of *in vivo* anti-TB drug mechanisms and their associated side-effects, and the mechanisms by which the infectious organism develops drug resistance. To this end, the relatively new research approach termed 'metabolomics', shows promising results through its capacity for identifying new drug markers, or metabolic pathways related to these contributing factors. Metabolomics refers to the unbiased identification and quantification of all metabolites (products of bio-molecular processes) present in a biological system, using highly selective and sensitive analytical methods. The application of metabolomics, for biomarker discovery, is based on the principle that an external stimulus, such as TB disease or infection, an anti-TB drug, or a mutation resulting in drug resistance, may disrupt normal metabolism, altering the overall physiological status of an organism or host, and these metabolic changes are specific to the perturbation investigated and not due to overall inflammation or disease process. Analyses of these altered metabolic pathways, may subsequently shed new light on the mechanisms associated with the causes of treatment failure, and ultimately lead to new treatment strategies which may most likely be aimed at targeting specific metabolic pathways in *M. tuberculosis*, and/or the genes/proteins associated with these.

Keywords: Drug-drug interactions, drug mechanisms, metabolomics, treatment failure, tuberculosis, xenobiotic metabolism.

INTRODUCTION

Tuberculosis (TB), caused by *Mycobacterium tuberculosis* (*M. tuberculosis*), is considered the world's foremost cause of death from a single infectious agent, which is alarming as it is widely believed to be curable [1]. TB infects approximately one-third of the world's population, either in its active (symptomatic) or the latent (asymptomatic) state [2, 3], and is responsible for as many as 1.4 million deaths annually [4]. Furthermore, the ever increasing TB prevalence reported, especially for African countries, have been attributed to, amongst others, the increased incidence of HIV, poor treatment compliance resulting in anti-TB drug resistance and poor socio-economic conditions (poverty, malnutrition and overcrowding) [5]. Subsequently, the DOTS (Directly Observed Therapy, Short course) treatment program was introduced in an attempt to reduce the global TB epidemic. DOTS is the recommended 6 month therapy regimen, consisting of an initial 2 month (intensive) treatment phase using a combination of isoniazid (INH), rifampicin (RIF), pyrazinamide (PZA) and ethambutol (EMB), followed by a four month maintenance phase, consisting of only INH and RIF [1, 6]. The aim of TB treatment is to cure the patient, prevent death, cease transmission and prevent drug resistance [7].

TB treatment failure can be defined as the occurrence of persistently positive sputum smears despite treatment [8]. Treatment failure may be attributed to a long list of factors, including: (1) irregular or inadequate anti-TB drug supplies by health providers, (2) a lack of patient TB-education, (3) poverty and poor life quality (further resulting in transport problems to the sites of treatment administration), as well as (4) long treatment duration. Additionally, and more relevant to this review, are the various biochemically related factors contributing to poor treatment outcome, including: (1) variable individual metabolism of the anti-TB drugs (including variation in xenobiotic metabolism, drug absorption and drug-drug interactions), (2) drug resistance and pathogenicity by the infectious organism, and (3) non-adherence to the treatment program, especially due to the associated side-effects [9-13]. These proposed mechanisms, are however not fully understood. Considering this, new research approaches, such as metabolomics, are currently being used to increase our understanding of the above mentioned

*Address correspondence to this author at the School for Physical and Chemical Sciences, Centre for Human Metabonomics, North-West University (Potchefstroom Campus), Private Bag x6001, Box 269, Potchefstroom, South Africa, 2531; Tel: +27 (0) 18 299 1818; Fax: +27 (0) 18 299 1823; E-mail: dutoit.loots@nwu.ac.za

2213-2368/13 $58.00+.00

biochemical mechanisms related to treatment failure, and could contribute significantly towards improved TB control.

Metabolomics can be defined as the unbiased identification and quantification of all metabolites present in a biological system, using highly selective and sensitive analytical methods [14-16], in conjunction with bioinformatics, for identifying new metabolite markers [1]. Metabolomics is the holistic study of an organisms' metabolism, and gives an overview of the complete metabolome, resulting from a specific perturbation, including a patho-physiological stimulus or genetic alteration [17], and is considered an important addition to systems biology, when the data generated is interpreted together with that generated using genomics, transcriptomics and proteomics [18]. Metabolomics is based on the concept that an individual's metabolic state is a representation of the overall physiological status. Therefore, it is possible to better understand the mechanisms of disease, identify diagnostic biomarkers, and predict individual/organism drug responses, using this research approach [17], on the basis of the new metabolite markers identified due to the perturbation. When seen from this perspective, TB disrupts the host metabolism, initiating metabolite changes which can be visualized as metabolic biosignatures or metabolite patterns [19, 20]. Metabolomics subsequently analyzes the end-products of these perturbations present in blood, sputum and/or urine (or any other fluid or tissue for that matter), and identifies metabolites that are both endogenous and exogenous to the perturbation, capturing information with regards to the mechanisms of disease or drug action [21].

Various chromatographic instruments are used as tools in metabolomics, e.g. gas chromatography (GC), liquid chromatography (LC), capillary electrophoresis (CE) and nuclear magnetic resonance (NMR) [17] for detecting these alterations. Usually, these instruments are combined with mass spectrometry (MS), of which a variety are available, including the quadrupole MS, ion trap MS and the time of flight MS (TOFMS), each of which having their own advantages and disadvantages pertaining to sensitivity, specificity and molecular preferences [19]. By combining these instruments, the best possible results and reliable metabolite profiles are acquired [16, 17]. The choice of the analytical instrumentation used would be dictated by the goal(s) of the study and the type of molecule or metabolite pathways one would expect to show the most variation [19], also considering the advantages and disadvantages of these different metabolomic approaches for these applications (see Table **1**). However, as per definition, metabolomics is the "study of all metabolites in a biological system", many approaches have been developed in order to analyze the entire metabolome, or as much of this as possible, using one or more analytical methods [13, 22-24]. In both of the aforementioned instances, and part and parcel of true metabolomics research, multivariate statistical analysis, such as principle component analysis (PCA), partial least squares discriminant analysis (PLS-DA) or orthogonal partial least-squares discriminant analysis (OPLSDA), are employed to extract information from the large metabolite datasets generated, in order to identify possible biomarkers [18, 25, 26].

Considering this, metabolomics could serve as a valuable tool for better understanding the proposed causes of treatment failure, from a biochemical perspective, considering the contributory causes, and will subsequently be discussed under the following headings in the remainder of this review.

1. VARIABLE INDIVIDUAL METABOLISM

1.1. Xenobiotic Metabolism

Xenobiotic metabolism refers to a natural response by which the body modifies a foreign compound (such as drugs) by enzymatic/biochemical reactions (biotransformation), in order to detoxify the body, by eventually eliminating these exogenous substances. During this process, many drugs may however be converted to their active states. Furthermore, the duration and intensity of drug action may also be influenced by the rate of its metabolism/biotransformation, which may vary between individuals. These metabolized compounds may also serve as "signals" for various patho-physiological outcomes [18, 32, 33], such as explaining or predicting a response to the medication and of treatment outcome. Consequently, identifying xenobiotic metabolites and their related metabolic pathways are considered essential to understanding drug action and their associated side-effects. Variability in individual xenobiotic metabolism has remained a major challenge for pharmacology research. It is important to understand why individual drug responses differ, and to what extent this variability is responsible for the differences we observe in therapeutic efficacy and tendencies towards adverse reactions [34]. To date, research has shown that individual variability of xenobiotic metabolism and their associated side-effects are influenced by a number of factors, including genetics, race, ethnicity, gender and external stimuli such as diet, alcohol consumption, drug-drug interactions, etc. [35].

Considering this, the current knowledge pool pertaining to the functionality of anti-TB drugs is considered insufficient, as their exact mechanisms of action, especially in the human host, and on the bacteria, are still largely unclear [1, 36]. This is most probably a major contributor to the lack of development of new anti-TB drugs over the last five decades (only two new drugs have been approved by the Food and Drug Administration (FDA) since 1962), and the rising threat of drug-resistant strains. Considering this, it is clear that new strategies are urgently required for better understanding the current *in vivo* host metabolism, anti-TB drug mechanisms and drug-drug interactions [25], which when elucidated, would undoubtedly lead to improved patient prognosis, as measures could then be put into place, using the knowledge gained to preventing or lowering the prevalence of treatment failure. To date, a number of examples of this already exists, where metabolomics is not only being used for the identification of new metabolite markers and altered patient metabolite profiles, shedding light on new mechanisms of TB disease and anti-TB drug action, but when combined with other 'omics' approaches, allows for a more in-depth and holistic understanding of the research question [1]. This approach, could potentially also be used for identifying markers which could be used to

Table 1. A summary of the advantages and disadvantages of the common techniques and tools used in metabolomics [27-31].

	Advantages	Disadvantages/Limitations
Direct MS	• Good combination of sensitivity and selectivity • Highly specific chemical information (accurate mass, isotope distribution patterns, characteristic fragment ions)	• Susceptible to ion suppression or enhancement • Data interpretation can be challenging • Inability to differentiate isomers
GC-MS	• Suited for volatile and non-volatile (after derivatization) compound analyses • Affordable, with relatively low running costs • High sensitivity, mass resolution and accuracy • Good dynamic range • Compound identification using mass spectral library matching • Provides additional and orthogonal data (i.e. retention time/factor/index) • Reproducible chromatographic separations • Identification of stereoisomers possible • Shorter run times • Lower bleed (thinner films)	• Limited to volatile, thermally stable, and energetically stable compounds • Requires additional sample preparation, such as derivatization, as this approach depends on the analytes being volatile and thermally stable, and few metabolites meet this requirement in their natural state • Less amenable to large, highly polar metabolites (poor volatility) • Co-eluting analytes in single dimension GC (corrected for using GCxGC) • Careful attention required for splitless injections • Slower scan rates/speed, unless coupled with TOFMS • Lower mass accuracy, unless coupled with TOFMS detectors
LC-MS	• Applicable for targeted and non-targeted metabolomics • Relatively low reagent cost • High analytical sensitivity, specificity and coverage depth • Able to analyze a wide range of compound classes • Ideal for highly polar and ionic compounds/metabolites • Requires minimum sample preparation • Low matrix effects and interferences • The application of both positive and negative ionization, hence more comprehensive metabolome coverage	• Lower chromatographic resolution compared to GC-MS • Comparatively higher running costs • Electrospray ionization (ESI) can suffer from ionization suppression • Retention time shifts are known to occur • Higher signal to noise (S/N) ratios compared to GC-MS • Low sample throughput
CE-MS	• Applicable for targeted and non-targeted metabolomics • Ideal for highly polar and ionic compounds/metabolites • Fast, relatively affordable, and highly efficient separation technique • Typically coupled with TOFMS for fast acquisition rates	• Lower sensitivity compared to other techniques • Poor reproducibility • Electrochemical reactions of metabolites • May lack the necessary robustness • Least suitable for analyzing complex biological samples
NMR-MS	• Highly selective • Non-destructive to sample material • Metabolite structural elucidation	• Lower sensitivity compared to other techniques

predict individual drug response phenotypes (pharmacometabolomics), and subsequently, in diagnostic applications for the early prediction of treatment failure.

Considering this, three approaches are suggested, focusing on individual xenobiotic metabolism research, aimed at resolving practical issues in drug metabolism and toxicology. Firstly, metabolomics could typically be used to better elucidate drug metabolism, by identifying metabolites differentiating between two groups, one receiving the xenobiotic and another not, after which the biological samples (urine, blood, tissue, saliva, feces, etc.) are collected for comparison of the resulting difference in the metabolite profiles using a selected analytical method. When comparing these groups, in theory, all the additional compounds or alterations to the metabolite profile in the xenobiotic group in comparison to the control, would be because of the exogenous compound ingested/injected [18]. Subsequently, these differences can broadly be visualized using a multivariate, unsupervised scores plot (such a PCA) and then, if a natural differentiation if achieved, those compounds/xenobiotic metabolites best explaining or contributing to the differences, identified using a variety of supervised multivariate (such as PLS-DA) in conjunction with other univariate methods (t-test and effect sizes) [37]. For example, Loots *et al.* (2005) conducted a rat-model metabolomics study, based on a similar study design, in order to evaluate the effect of combined anti-TB drug (Rifater) therapy, on the free radical and organic acid profiles of

Sprague-Dawley rats, in order to get a better understanding of their impact on normal metabolism and potentially better explain their associated side-effects. Additionally, these metabolic effects where compared to an additional treatment group receiving a combined treatment of Rifater and melatonin. In their study, they determined that the combined anti-TB drug Rifater results in oxidative stress and a metabolite profile associated with individuals suffering from a multiple acyl-CoA dehydrogenase deficiency (MADD), and eventually linked this to the drug causing inhibition of the electron transport flavoproteins, in addition to an explanation of a number of side-effects, which could be corrected for by the co-administration of melatonin [22]. This approach may however be limiting if the differences detected are due to both exogenous and endogenous xenobiotic metabolites, as was the case in Loots *et al.*'s study. One is, however, able to sift out these differences using libraries specific towards identifying either the drug metabolites, or the altered natural metabolome of the organism. The advantages of this approach is that it may potentially describe alternative pathways influenced by the drugs, better explaining the side-effects, which will be discussed in section 3 of this review. Considering this, Chen *et al.* (2007) proposed a stable isotope-labeled xenobiotic treatment group comparative to an unlabeled group, as an alternative to this approach, and in so doing, one would be able to track the drug specific metabolites and their associated metabolism and also elucidate which metabolites are directly linked to the xenobiotic metabolism, and which are due to secondary toxicity [18, 38]. Lastly, Chen *et al.* (2007) propose a study design to identify human xenobiotic-metabolizing enzyme (XME) polymorphisms responsible for adverse drug reactions (ADRs), which can be used for investigating anti-TB drug side-effects, or an individual's response to anti-TB drugs. ADR may also occur due to various factors relating to a patient's general health status, various environmental contaminants and as a result of drug-drug interactions with other co-administered drugs requiring the same enzyme systems for metabolic processing [18]. In these cases, ADR may largely be caused by genetic polymorphisms of XMEs, or other factors influencing enzymatic activities or their expression, such as an individual's general health, environmental contaminant exposure, cultural and/or genetic factors, and the previously mentioned drug-drug interactions. If ADRs are observed in a patient group, suspected to be due to XME polymorphisms, an investigation into the metabolite profiles and pharmacokinetics would be of great importance, and could potentially identify the metabolic pathways affected and pinpoint the genetic factors causing these ADRs. This in turn may lead to the necessary adaptions to the medication and ultimately individualized treatment procedures [18, 35, 39].

1.2. Drug Malabsorption

Although rare, malabsorption of anti-TB drugs may also lead to a poor treatment outcome. Probable reasons for malabsorption, although poorly characterized, include co-morbidities of TB with HIV/AIDS infection, hypoalbuminemia in the malnourished, infectious gastroenteritis, gastric achlorydria, gastrointestinal and/or malabsorption diseases, thyroid and end-stage liver or renal diseases. Although either a single or combination of these factors may contribute to drug malabsorption, HIV/AIDS infection is considered by far the most common risk factor associated with malabsorption of anti-TB drugs [40, 41].

It has previously been reported in the literature, that HIV/AIDS infection may lead to anti-TB drug malabsorption and acquired drug resistance [40, 42]. This was based on the observation that 68% of all patients with reduced serum RIF and INH levels, were co-infected with HIV [41, 42]. As both RIF and INH reach maximum serum concentrations 2h after oral administration, Bento *et al.* (2010) proposed taking two measurements; 2h and 6h after administration, in order to determine if a patient has malabsorption (indicated by persistently low concentrations) or delayed absorption (slowly increase concentrations) [41], for the purpose of adapting the treatment protocols. However, this approach may be costly and have significant health policy implications. Alternatively, it was proposed that urinary RIF and INH levels could be monitored, providing important information related to drug absorption [40]. Considering this, Gurumurthy *et al.* (2004) investigated the absorption of RIF and INH in patients with HIV/AIDS and diarrhea, HIV/AIDS and TB co-infection, pulmonary TB alone, and healthy subjects, by analyzing the urinary excretion of these drugs and the associated metabolites. Their investigation indicated that TB positive patients, without an underlying disease or illnesses, absorb anti-TB drugs reliably, while those with gastrointestinal disorders or HIV/AIDS infection have less optimal absorption [40].

Although no metabolomics investigations on this topic have been done to date, this research approach could be used/contribute to this in the following manner. A comparison of the metabolomes of those individuals grouped as good absorbers vs. poor absorbers, using blood/urine collected at time of diagnosis (prior to commencing treatment), may better explain the underlying mechanisms associated with treatment outcome, be it by underlying disease mechanisms or genetic factors. This information could be used to better explain this phenomenon, or alternatively be used diagnostically to identify malabsorpers prior to treatment, so that the necessary adaptions can be made i.e. personalized treatment approaches. Similarly, these groups can be compared during treatment, by one or more of the methodological approaches discussed in the previous section, and potentially lead to a better understanding of potentially varying drug metabolism between these individuals.

1.3. Drug-drug Interactions

Drug-drug interactions are especially prevalent in TB patients co-infected with HIV/AIDS, as INH and RIF in particular, have inauspicious drug interactions with many of the antiretroviral drugs used to treat HIV/AIDS [11]. These unwanted drug-drug interactions (inhibition or induction), are thought to result in (1) lowered efficacy of co-administered drugs, leading to a poor treatment outcome and drug resistance by the organism, and also (2) increased risk of toxicity and hence, treatment interruptions [43], also leading to drug resistance.

These drug-drug interactions are thought to occur due to these different medications, all being metabolized by the same cytochrome P450 (CYP450) enzyme system [44]. The CYP450 enzymes, refer to a group of important XMEs, and are especially important for the metabolic activation and metabolism of several drugs. INH is known to permanently inhibit several of the CYP450 enzymes, and this state persists until new CYP450 enzymes are synthesized. Side-effects due to INH drug interactions may include an increased risk of hepatotoxicity, hypertension, flushing and tachycardia. Excess catecholamine (a dopamine precursor) and elevated levels of other medications, may also be present due to inefficient metabolism of these drugs when INH is present [45, 46].

On the other hand, an increased hepatic blood flow or CYP450 synthesis, leads to enzyme induction, as is the case with RIF administration [47], where this reaction is mediated through the activation of the pregnane-X receptor by RIF [48], reducing the therapeutic effects of numerous other co-ingested medications metabolized through this enzyme system. This may further lead to complicated health consequences when RIF is taken in combination with other common drugs, such as oral contraceptives, warfarin (Coumadin) or sulfonylureas. Furthermore, RIF is a known inducer of cytochrome P450 (CYP) 3A, reflected by independent regulation of CYP3A expression at various sites, with large inter-individual variation observed [49]. Gorski *et al.* (2003) investigated CYP3A susceptibility to RIF induction, as a function of age and gender, with midazolam, a CYP3A substrate. RIF significantly increased both the systemic and oral clearance of midazolam and reduced the oral midazolam availability by 88%. Interestingly, gender also had a significant effect on induction of oral midazolam clearance, as this was found to be greater in males when compared to females, largely due to the fact that these gender-related differences in response to CYP3A inducers are substrate-dependent [49]. RIF related drug-drug interactions are also reported to lead to an increased risk of aplastic anaemia and hypoglycaemia. Lowered levels of other medications in the blood are commonly reported, thus dosages of these may need to be increased to avoid lowered efficacy when given in combination to RIF [46, 47]. For example, verapamil is shown to be decreased 10-20 fold when co-administered with RIF. On the other hand however, recent studies have shown the potential of verapamil to be used as an effective efflux inhibitor in many organisms, including *M. tuberculosis*, subsequently reducing the duration of anti-TB treatment and the emergence of drug resistance [50]. These drug-drug interactions may persist, depending either on the half-life of the inducer or the rate of enzyme degradation and new enzyme synthesis. Considering the short half-life of RIF and the 1-6 day turnover time of CYP450 enzymes, RIF elimination is more rapid than that of excess CYP450, thus interaction duration would be dependent on enzyme turnover [47].

Bedaquiline (TMC207), a diarylquinoline, was recently introduced as a new anti-TB drug for treating pulmonary MDR-TB in adults. Bedaquiline has specific ATP synthase inhibition activity against *M. tuberculosis*, affecting the energy production required for cell viability, acting as a bactericidal and sterilizing agent against *M. tuberculosis* and various other mycobacterial species, including, amongst others, the *M. avium* complex, *M. chelonae* and *M. kansasii* [51]. Based on clinical evaluations of its safety, tolerance and efficacy, the process for approving bedaquiline for treating pulmonary MDR-TB in adults, has been accelerated [51, 52]. Bedaquiline's long half-life and unique mechanism of action, potentially makes it suitable for intermittent administration, which is a major advantage compared with RIF for instance. However, one of the major concerns of this drug is its drug-drug interactions with RIF, as co-administration with RIF has been shown to half the maximum bedaquiline concentration. As bedaquiline is metabolized by CYP3A4, producing the less-active N-monodesmethyl bedaquiline, any CYP3A4 inducer, such as RIF, will affect the final concentration of bedaquiline [51]. Another concern is the possible drug-drug interactions between bedaquiline and antiretroviral agents (ARV's), when treating TB-HIV/AIDS co-infected patients. As current studies devoted to the use of bedaquiline in these cases are still underway, no clinical significant findings have been suggested to date [51, 52].

As anti-TB drug induced toxicity may have a major influence on other medications used simultaneously, the need to predict ADRs as early as possible during the drug treatment regimen is crucial. Metabolomics can provide the means for a global overview of individual variation and the subsequent prediction of potential drug-drug interactions, via analysis of urinary metabolites after drug administration, which can serve as a signature for organ-specific drug toxicity, general activity or inactivation [18, 53, 54]. Metabolomic profiling has already been shown to predict drug responses in animal models [53], however, as xenobiotic metabolism differs between different species, particularly due to molecular variation in receptors and XMEs, it presents a problem when it comes to extrapolating data obtained from rodent models (especially rats and mice) to humans [54, 55]. Thus, more reliable *in vivo* systems, such as humanized mouse models, need to be generated in order to study and potentially predict human xenobiotic responses [55]. Currently, the development of animal models as a tool to simulate human xenobiotic metabolism is expanding rapidly [56, 57]. Considering this, humanized transgenic mice can be generated by introducing a specific human gene into the rodent genome, of which many genetically-modified mouse model systems for several important nuclear receptors (including AhR, PXR, CAR) have been established to date, which not only offer a better animal system for studying the influence of drug-drug interactions on xenobiotic metabolism, and the consequent prediction of human xenobiotic responses, but also facilitates the elucidation of the related underlying drug mechanisms [18, 55]. An additional example of this is the Kramnik mouse model (mice harboring the C3HeB/FeJ genotype), in which necrotic lung lesions are generated. Following *M. tuberculosis* infection, animals may develop necrotic lung lesions, where drug-resistant bacilli are typically present, persisting in an extracellular micro-environment within these lesions [58, 59]. Using this model, Driver *et al.* (2012) examined the efficacy of drug treatment, by comparing the formation of colliquative necrotic lesions in Kramnik mice, to that of BALB/c mice, with non-necrotic lesions, following aerosol challenge. This was accomplished by infecting both the Kramnik and BALB/c mice with *M. tuberculosis* bacilli,

followed by 7 to 8 weeks of monotherapy, using drugs with varying modes of action. Drug efficacy was quantified by determining the bacterial load, and disease progression visualized via staining methods. In the Kramnik mice, results indicated the formation of fibrous encapsulated lung lesions with central colliquative necrosis having abundant extracellular bacilli, in the late stages of infection, while the BALB/c mice formed non-necrotic lesions, primarily infected with intracellular bacilli. Furthermore, the Kramnik mice had indications of hypoxia within the necrotic lesions, and significantly more anti-TB drug-resistant colonies, especially for PZA, compared to the BALB/c mice [58]. As the Kramnik mouse model is considered a reliable representation of human TB, these results not only indicate the suitability of the Kramnik mouse model for testing anti-TB drug efficacy, especially for drug-resistant *M. tuberculosis* strains, allowing for validation of possible candidate anti-TB drugs and vaccine selection, but also the capacity to develop a TB mouse model with a disease state more closely representing that of human TB. This model is not only advantageous when wanting to elucidate the mechanisms of *M. tuberculosis* pathogenesis [58, 59], but could potentially also be used in various metabolomics investigations.

To date, xenobiotic receptor-mouse models have also been developed to closely resemble the human metabolome, and have subsequently been used to better elucidate the metabolic pathways leading to drug toxicity and drug-drug interactions, aiding in the identification of potential drug biomarkers in humans. In these instances, the unique combination of humanized mouse models subjected to metabolomic analysis, in conjunction with multivariate analysis, is providing answers to key questions related to human health and disease questions [18, 45, 54]. Most ideally however, is the use of metabolomics to investigate these potential drug-drug interactions, using clinical samples of patients already being treated for one or more co-infections. This will allow for the identification of drug-drug interaction markers relating to drug efficacy and/or side-effects in humans.

2. DRUG RESISTANCE AND PATHOGENICITY of THE INFECTIOUS ORGANISM

As previously mentioned, the DOTS program is based on administering four of the first-line anti-TB drugs, namely INH, RIF, PZA and EMB, simultaneously. As an introduction to this section, we thought to give a brief overview on drug action first. As several reviews have been written on the topic, we will focus on the metabolomics investigations done to date. However, to summarize briefly, INH has high bactericidal activity and is used to efficiently and specifically inhibit active growing *M. tuberculosis* [60]. RIF is considered a powerful sterilizing drug, responsible for eliminating active and semi-dormant *M. tuberculosis* from the host [61], as its primary mechanism of action is transcription inhibition [61, 62]. PZA, is also classified as a sterilizing drug and functions by eliminating semi-dormant *M. tuberculosis* bacilli located within hosts macrophages [61, 62], and EMB, is classified as a bacteriostatic agent and typically adds additional coverage to reduce the chance of developing drug resistance [46, 61], considered effective against both intra- and extracellular *M. tuberculosis* bacteria [1].

Considering the limited information surrounding mechanisms by which drug resistance occurs, one of the few metabolomics investigations published regarding drug resistance to anti-TB drugs was done in 2012, by du Preez and Loots. Using GC-MS, they investigated an altered fatty acid metabolome of two *rpoB* mutant *M. tuberculosis* strains, by comparison to that of an isogenic drug sensitive wild-type parent strain, in order to characterize RIF-resistance, due to mutations in the *rpoB* gene. Metabolite markers best describing the variation were identified using multivariate statistical approaches, namely PCA and PLS-DA. This study was the first to identify a link between an altered fatty acid metabolism and RIF-resistance, *rpoB* mutations, and the β-subunit of RNA polymerase, present in *M. tuberculosis*. In short, they determined that the mutation in the *rpoB* gene of *M. tuberculosis*, resulting in RIF-resistance, also influences mycolic acid synthesis by disruption of the mRNA:NTP (ribonucleoside 5'-triphosphotases) equilibrium. 10-methyl branched chain fatty acids in the RIF. This was confirmed by the lack of detection of 10 methyl branched chain fatty acids in the RIF-resistant mutants, accompanied by the accumulation of their fatty acid substrates. Additionally, these mutants showed greater dependence on fatty acids and acetate, as altered energy sources, and subsequent up-regulation of the glyoxylic acid cycle. This study not only indicated the potential of pharmaco-metabolomics for the identification of metabolites associated with drug resistance, but also mapped its potential for global health and personalized medicine [13]. In view of the above, the methodology described by du Preez and Loots (2012), could furthermore be applied to investigating the mechanisms of drug resistance of *M. tuberculosis* to the other first-line, or even second-line drugs. The elucidation of xenobiotic related metabolites in these organisms, as opposed to that previously described for in the host, by comparison of drug-sensitive and drug-resistant strains, with low doses of these drugs in the growth media, could additionally provide further clues pertaining to the biotransformation of xenobiotics in these organisms, and shed light on additional metabolic pathways (and ultimately genes) involved in drug resistance.

In the following year, Chakraborty *et al.* (2013) conducted a metabolomics approach proving that para-aminosalicylic acid (PAS) does not inhibit *M. tuberculosis* growth by its inhibiting action on dihydropteroate synthase (DHPS). PAS was previously thought to act as a competitive inhibitor of DHPS by acting as an alternative substrate, in competition with *p*-aminobenzoate (PABA), in folate metabolism. They additionally indicated that other sulfonamides are ineffective against *M. tuberculosis* due to inactivation by bacterial metabolism, rather than inadequate uptake. As new approaches are needed to improve our understanding of drug action and its efficacy for drug development, this discovery highlights that catalysis by an enzyme, rather than the inhibition thereof, could possibly be the key to new drug development [63].

Furthermore, from a growth and pathogenesis perspective, bacteria are known to utilize catabolite repression in order to maximize growth, via consumption of host carbon substrates, in a preferred sequence. Considering this, a metabolomics study conducted by de Carvalho *et al.* (2010),

revealed that *M. tuberculosis* is capable of compartmentalized/simultaneous co-catabolism of multiple carbon sources/substrates, enhancing its growth. These carbon sources can be catabolized through the glycolytic, tricarboxylic acid, and/or pentose phosphate pathways, a phenomenon related to a previously undescribed bacterial metabolic network, and possibly linked to the pathogenicity of *M. tuberculosis*. Thus, using metabolomics, new knowledge was obtained, reshaping our basic understanding of *M. tuberculosis* and its ability to adapt to the host for survival [64].

Additionally, a GCxGC-TOFMS metabolomics study conducted by Meissner-Roloff *et al.* (2012), compared the metabolomes of a hyper- and hypo-virulent Beijing *M. tuberculosis* strain, attempting to characterise the varying degrees of virulence within the *M. tuberculosis* species, and consequently identified metabolite markers best differentiating between these two strains. This pilot study was the first of its kind to identify metabolites associated with an increased metabolic activity, cell wall synthesis, replication rates, and an altered anti-oxidant mechanism, which is thought to result in the increased pathogenicity and survival of this hyper-virulent strain of *M. tuberculosis* [65].

In the light of the above, a better understanding of the metabolome of the infectious organism, being it drug sensitive or drug-resistant, in response to a drug, may potentially lead to improved treatment approaches, and to either preventing drug resistance or to the development of alternative drugs for treating the disease [18]. With the emergence of multi-drug-resistant (MDR), extensively-drug-resistant (XDR), and more recently even totally-drug-resistant (TDR) TB, the robustness of our approaches to discover and develop new, more effective anti-TB drugs, are questionable. The discovery and development of new anti-TB drugs are largely dependent on our understanding of the complex host-pathogen-interactions [66], xenobiotic metabolism by both the host and bacteria, the specific drug action, their side-effects, etc.

Furthermore, another shortcoming, is translating basic research into drug discovery programs. To date, various genetic tools have been developed for manipulating *M. tuberculosis* in order to identify specific drug targets, but few candidates have emerged. Once validated drug targets are identified, lead compounds acting against these targets should be established, in order for medicinal chemistry to use these leads for developing new drugs. Additionally, mycobacterial persistence/resistance models need be characterized for rapid analyses of compounds with potent sterilizing activity against *M. tuberculosis*. Finally, the identification of metabolite markers providing an early indication or prediction of treatment outcome could facilitate clinical trials tremendously [66].

Unfortunately, drug discovery and development is a lengthy process, especially in the case of those developed against *M. tuberculosis*, and resources remain limiting. Although the estimated cost of drug development may vary, it is estimated that the full development of a new anti-TB drug, could cost up to approximately $100M. In 2000, The Global Alliance for TB Drug Development (GATB) was established, in order to address these issues. The aims of this alliance is to overcome natural barriers in the field of TB drug development, by forming partnerships between various diverse organizations, such as academic institutions, the pharmaceutical industry, government research laboratories, and non-governmental organizations [66]. The objectives of the GATB include (1) reducing the current duration and complexity of treatment and/or the doses needed to cure active, drug-sensitive TB, (2) improving the efficacy of MDR-TB treatment, (3) reducing the treatment period of latent TB infection, and (4) simplifying the treatment of TB-HIV/AIDS co-infected patients, by lowering the incidence of drug-drug interactions, especially those especially those associated with ARV's [66, 67]. These will not only result in reduced side-effects, better patient adherence and lower the costs of delivering DOTS, but will also ultimately have a major impact on treatment outcome. The challenges for reaching these goals include (1) elucidating the biological mechanisms of mycobacterial resistance and latency, (2) discovery and development of new anti-TB drugs with novel mechanisms of action, (3) developing and validating accurate animal models that can reliably be translated to humans, (4) developing and validating metabolite markers for predicting treatment outcome, subsequently reducing clinical trial duration, (5) identifying optimum drug combinations, and (6) the means to conduct clinical trials in high burden TB countries [67]. Thus, in order to reach these GATB goals, a significantly increased effort and resources is required at every stage of the drug discovery and development process. Furthermore, optimization of these programs will ensure that valuable research is not wasted and potential new anti-TB drugs missed, due to attrition in the development process [66].

3. NON-ADHERENCE, ESPECIALLY DUE TO ASSOCIATED SIDE-EFFECTS

According to numerous studies, reasons for non-adherence of patients to the anti-TB treatment program includes, amongst others, drug related side-effects (especially nausea, vomiting, giddiness), initial relief from TB-symptoms, alcohol consumption, work or domestic related problems, too ill to get to treatment points, or the stigma associated with having TB [68]. However, the anti-TB drug side-effects are most often associated with non-adherence, leading to treatment failure and drug resistance [66], and thus this aspect of non-adherence is of biological mechanistic relevance, topical for this review.

According to a 2008 World Health Organization (WHO) bulletin, the side-effects associated with anti-TB drug use include hepatotoxicity/hepatitis (liver inflammation), jaundice, exanthema (rash), dyspepsia (indigestion) and arthralgia (joint pain), and are considered the major factors leading to non-adherence. According to this bulletin, 23% of all TB patients terminated or defaulted treatment due to one or more of these side-effects. Furthermore, up to 86% of all patients may develop these side-effects, especially those co-infected with HIV/AIDS or those with a history of hepatitis [69]. TB treatment regimens require a high (>90%) compliance rate in order to achieve a successful treatment outcome [70]. Amuha *et al.* (2008) conducted a study, aimed at determining the prevalence and factors associated with non-adherence to anti-TB drugs. An interviewer administered questionnaire was used for qualitative data collection, and revealed that most patients tend to terminate treatment when they start to

feel better, or due to the severe side-effects and pill burden, as these drugs may be considered by the patient to be too many to take at once [70]. Although the aim of TB control is that trained health staff be able to recognize and manage these side-effects, it would serve best if these could be eliminated entirely. Further concern however is the additional expenses which may incur by supplying additional complementary medication to alleviate these associated side-effects [69].

To the best of our knowledge, only one such metabolomics study has been done to date. As described earlier, in 2005, Loots *et al.* conducted a metabolomics investigation into the side-effects associated with the oral administration of the combined anti-TB drug Rifater, and alleviating this using combination therapy with melatonin, in a Sprague-Dawley rat model. Rifater was seen to not only result in increased free radicals, but also induce increased organic acid levels characteristic of a MADD metabolic profile, which includes elevated levels of isovalerylglycine, ethylmalonic acid, butyrylglycine, 2-methylbutyrylglycine and suberic acid. On the basis of this, the similarity in symptoms, and the known mechanism by which this abnormal metabolite profile exists in MADD, a better understanding of how these anti-TB drugs cause these associated side-effects was proposed: by inhibition of the electron transport flavoproteins. Subsequently, melatonin, a well-known antioxidant previously proposed to increase the efficacy of INH, was investigated as a means of alleviating these symptoms, when administered in combination to these drugs. The result was alleviation of both the oxidative stress and MADD metabolite profile [22].

Despite the limited metabolomics data, a number of studies have been done investigating the general metabolites of first-line anti-TB drugs, formed in the host and their associated side-effects. These are summarized in (Table **2**).

Table 2. A summary of the general metabolites of first-line anti-tb drugs, and their associated side-effects.

Anti-TB Drug (Year Discovered)	Associated Metabolites	Associated Side-Effects
Isoniazid (INH) (1952): [72-75, 77]	Ammonia; Diacetylhydrazine (DiAcHZ), Hydrazine (HZ); Hydrazones with pyruvic (INH-PA) and ketoglutaric acids (INH-KA); Isonicotinic acid (INA); Monoacetylhydrazine (AcHZ); N1-Acetyl-N2-isonicotinylhydrazide (AcINH); Oxidizing free radicals	• Hepatotoxic: ATDH/hepatitis • Cutaneous: rash • Nephrology: increased urination • Abdominal: severe dizziness/nausea, vomiting • Flu-like symptoms: fever, headache • Other: drug-induced lupus erythematous, high anion gap metabolic acidosis, seizures (status epilepticus), mood changes (confusion, psychosis, depression), pyridoxine (vitamin B6) depletion (resulting in central nervous system (CNS), peripheral neuropathy, effects and sideroblastic anaemia)
Pyrazinamide (PZA) (1952): [1, 73, 75, 78]	Pyrazinoic acid (PA); 5-Hydroxypyrazinoic acid (5-OH-PA)	• Hepatotoxic: ATDH/hepatitis • Cutaneous: pruritus (skin irritation), rash and/or urticarial (hives) • Nephrology: dysuria (painful urination), interstitial nephritis (kidney inflammation) • Abdominal: nausea, vomiting • Flu-like symptoms: fever, arthralgia (due to decreased uric acid secretion), malaise • Other: sideroblastic anaemia
Ethambutol (EMB) (1966): [71, 75]	2,2'-(Ethylenediimino)-di-butyraldehyde; 2,2'-(Ethylenediimino)-di-butyric acid (EDA)	• Hepatotoxic: ATDH/hepatitis • Cutaneous: milk skin reaction • Optical: retrobulbar/optic neuritis, red-green color blindness, vertical nystagmus • Nephrology: hyperuricaemia • Flu-like symptoms: arthralgia • Other: peripheral neuropathy
Rifampicin (RIF) (1966): [72, 73, 75]	Desacetylrifampicin; 3-Formyl rifampicin	• Hepatotoxic: ATDH/hepatitis, upregulation of hormonal hepatic metabolism (decreasing hormone levels) • Cutaneous: flushing, pruritus, rash • Respiratory: breathlessness • Abdominal: cramps, diarrhea, nausea, vomiting • Flu-like symptoms: fever, headache, arthralgia, malaise, dysphoria • Other: immunological reactions, color change in bodily fluid (urine, tears, sweat)

From the literature however, very little information exists linking metabolism of drugs and the resulting metabolites to the above mentioned side-effects, except for that of the hepatotoxicity which is rather well described. All four first line drugs used as part of the DOTS program (EMB to a lesser extent), are associated with anti-TB drug induced hepatotoxicity (ATDH). ATDH generally occurs within the first two months of treatment, and it thought to be dose-dependent. Hence, decreasing the drug dosage may lighten the symptoms. Although the occurrence of ATDH varies in different populations, studies have shown the most common factors to determine a predisposition towards this is genetics, advanced age, gender, malnutrition, alcoholism, hepatitis B or C, and HIV/AIDS infection, which may in turn lead to increased morbidity, mortality, non-adherence, treatment failure/relapse and the consequential development of MDR-TB. Genetics is considered the most common cause of ATDH and polymorphisms in various drug metabolizing loci, such as N-acetyl transferase 2 (NAT 2), cytochrome P450 oxidase (CYP2E1) and glutathione S-transferase (GSTM1), contributes to the inter-individual variations and varying pharmacological responses to these drugs [75, 77]. Although the exact mechanism of ATDH remains unknown, the toxic metabolites associated with the first-line anti-TB drugs play a crucial role in ATDH development [73], as these drugs are metabolized by the liver. INH, RIF and PZA (especially when combined) are potentially hepatotoxic drugs and mostly associated with an increased risk of developing ATDH, if the patient is unable to clear the toxic metabolites formed in liver due to any of the aforementioned predispositions. Ammonia, diacetylhydrazine (DiAcHZ), hydrazine (HZ), hydrazones with pyruvic (INH-PA) and ketoglutaric acids (INH-KA), isonicotinic acid (INA), monoacetylhydrazine (AcHZ), N1-acetyl-N2-isonicotinylhydrazide (AcINH) and oxidizing free radicals, are the toxic metabolites produced when INH is metabolized, either via NAT2 or hydrolysis [72, 74, 77]. These reactions may further be facilitated with the co-administration of RIF, via its induction of the hepatic microsomal enzymes of CYP450. Furthermore, as RIF competes with bilirubin absorption and conjugation, for inhibition of glucuronosyltransferase, jaundice and increased serum bilirubin and/or hepatic enzyme levels may occur [75]. Additionally, pyrazinoic acid (PA) and 5-hydroxypyrazinoic acid (5-OH-PA), toxic metabolites associated with PZA metabolization, have also recently been implicated to be responsible for PZA-induced hepatotoxicity [76].

4. CONCLUSIONS

Olivier & Loots (2012) used a GC-MS metabolomics approach, in combination with multivariate statistical data analyses, to characterize and identify various *Mycobacterium* species. Based a combination of 12 metabolites, a discriminant classification model, based on Bayes' theorem, for distinguishing between various TB causing and non-TB *Mycobacterium* species, was developed [79]. Considering this, metabolomics shows potential for identifying surrogate markers which could be used to predict and/or monitor TB treatment outcome, and allow for personalized treatment approaches. These surrogate markers may be influenced by any of the factors described above, including environmental and genetic variance [53]. Predicting which TB patients would or would not respond to a specific treatment regimen early on, such that changes can be made to the regimen, is likely to result in a shortened therapy, improved patient adherence and a favorable clinical outcome. Selecting a personalized treatment regimen on the basis of individual patient characteristics, may ultimately improve drug efficacy and limit side-effects. Thus, implementing a monitoring system by measuring surrogate biomarkers before and during treatment, could have a significant impact on the global burden of this disease, especially in under-developed countries, were DOTS is already common practice [80].

Although we do not yet fully understand the exact mechanisms of TB disease and its full effect on the human host, the full mechanisms and side-effects of the current anti-TB drugs, and drug resistance by the organism, elucidating these from a metabolomics perspective will undoubtedly contribute to the existing knowledge pool, generated to date using genomics, proteomics and transcriptomics. Further studies aimed at elucidating TB disease and drug metabolite pathways, through the identification of new metabolite markers, especially those related to treatment monitoring or outcome, will certainly be of tremendous value. These markers may ultimately provide crucial information and clues towards the understanding of the exact varying effect of *M. tuberculosis* on the human host, and the subsequent host adaptions, in addition to individual variations to drug metabolism and side-effects, which may in turn lead to a better understanding of treatment failure, and subsequently measures for preventing this. Considering this, metabolomics is already showing promise for new and valuable scientific discoveries related to this.

CONFLICT OF INTEREST

The authors confirm that this article content has no conflict of interest.

ACKNOWLEDGEMENTS

Declared none.

ABBREVIATIONS

CE	=	Capillary electrophoresis
CNS	=	Central nervous system
CYP450	=	Cytochrome P450
DiAcHZ	=	Diacetylhydrazine
DHPS	=	dihydropteroate synthase
DOTS	=	Directly Observed Therapy, Short course
ESI	=	Electrospray ionization
EMB	=	Ethambutol
XDR	=	Extensively-drug-resistant
FDA	=	Food and Drug Administration
GC	=	Gas chromatography
HIV	=	Human immunodeficiency virus

HZ	=	Hydrazine
INH-KA	=	Hydrazones with ketoglutaric acids
INH-PA	=	Hydrazones with pyruvic acid
INH	=	Isoniazid
INA	=	Isonicotinic acid
LC	=	Liquid chromatography
MS	=	Mass spectrometry
AcHZ	=	Monoacetylhydrazine
MDR	=	Multi-drug-resistant
MADD	=	Multiple acyl-CoA dehydrogenase defect
M. tuberculosis	=	*Mycobacterium tuberculosis*
AcINH	=	N1-acetyl-N2-isonicotinylhydrazide
NMR	=	Nuclear magnetic resonance
OPLSDA	=	Orthogonal partial least-squares discriminant analysis
PABA	=	*p*-aminobenzoate
PAS	=	Para-aminosalicylic acid
PLS-DA	=	Partial least squares discriminant analysis
PCA	=	Principle component analysis
PZA	=	Pyrazinamide
PA	=	Pyrazinoic acid
RIF	=	Rifampicin
S/N	=	Signal to noise
GATB	=	The Global Alliance for TB Drug Development
TOFMS	=	Time of flight mass spectrometry
TDR	=	Totally-drug-resistant
TB	=	Tuberculosis
WHO	=	World Health Organization
XME	=	Xenobiotic-metabolizing enzyme

REFERENCES

[1] Olivier, I.; Loots, D. An overview of tuberculosis treatments and diagnostics. What role could metabolomics play. *J. Cell Tissue Res.,* **2011**, *11*(1), 2655-2671.

[2] Houshmand, M.; Banoei, M. M.; Tabarsi, P.; Panahi, M. S. S.; Kashani, B. H.; Ebrahimi, G.; Zargar, L.; Farnia, P.; Morris, M. W.; Mansouri, D. Do mitochondrial DNA haplogroups play a role in susceptibility to tuberculosis? *Respirology,* **2007**, *12*(6), 823-827.

[3] Sudre, P.; Ten Dam, G.; Kochi, A. Tuberculosis: a global overview of the situation today. *Bull. World Health Organ.,* **1992**, *70*(2), 149.

[4] World Health Organization. 2012 tuberculosis global facts *Geneva, Switzerland,* **2011**.

[5] Knechel, N.A. Tuberculosis: Pathophysiology, clinical features, and diagnosis. *Crit. Care Nurse,* **2009**, *29*(2), 34-43.

[6] Raviglione, M. C.; Narain, J. P.; Kochi, A. HIV-associated tuberculosis in developing countries: Clinical features, diagnosis, and treatment. *Bull. World Health Organ.,* **1992**, *70*(4), 515.

[7] Frieden, T. R.; Driver, C. R. Tuberculosis control: Past 10 years and future progress. *Tuberculosis,* **2003**, *83*(1), 82-85.

[8] Mukherjee, J. S.; Rich, M. L.; Socci, A. R.; Joseph, J. K.; Virú, F. A.; Shin, S. S.; Furin, J. J.; Becerra, M. C.; Barry, D.J.; Kim, J. Y. Programmes and principles in treatment of multidrug-resistant tuberculosis. *Lancet,* **2004**, *363*(9407), 474-481.

[9] Quy, H.; Lan, N.; Borgdorff, M.; Grosset, J.; Linh, P.; Tung, L.; Van Soolingen, D.; Raviglione, M.; Co, N.; Broekmans, J. Drug resistance among failure and relapse cases of tuberculosis: is the standard re-treatment regimen adequate? *Int. J. Tuberculosis Lung Dis.,* **2003**, *7*(7), 631-636.

[10] Wares, D.; Singh, S.; Acharya, A.; Dangi, R. Non-adherence to tuberculosis treatment in the eastern Tarai of Nepal. *Int. J. Tuberculosis Lung Dis.,* **2003**, *7*(4), 327-335.

[11] Adhvaryu, M.; Vakharia, B. Drug-resistant tuberculosis: emerging treatment options. *Clin. Pharmacol.: Adv. Applicat.,* **2011**, *3*, 51.

[12] Mukonzo, J. K. B. In *Pharmacogenetic aspects of HIV/AIDS, tuberculosis and malaria: Emphasis on Ugandan population;* Inst för laboratoriemedicin/Dept of Laboratory Medicine: **2011**.

[13] du Preez, I.; Loots, D. T. Altered fatty acid metabolism due to rifampicin-resistance conferring mutations in the rpoB gene of *Mycobacterium tuberculosis*: Mapping the potential of pharmacometabolomics for global health and personalized medicine. *OMICS: J. Integrat. Biol.,* **2012**, *16*(11), 596-603.

[14] Dunn, W. B.; Bailey, N. J.; Johnson, H. E. Measuring the metabolome: Current analytical technologies. *Analyst,* **2005**, *130*(5), 606-625.

[15] van der Werf, Mariët J; Overkamp, K. M.; Muilwijk, B.; Coulier, L.; Hankemeier, T. Microbial metabolomics: Toward a platform with full metabolome coverage. *Anal. Biochem.,* **2007**, *370*(1), 17-25.

[16] Orešič, M. Metabolomics, a novel tool for studies of nutrition, metabolism and lipid dysfunction. *Nutrit., Metabol. Cardiovasc. Dis.,* **2009**, *19*(11), 816-824.

[17] Kaddurah-Daouk, R.; Kristal, B. S.; Weinshilboum, R. M. Metabolomics: A global biochemical approach to drug response and disease. *Annu. Rev. Pharmacol. Toxicol.,* **2008**, *48*, 653-683.

[18] Chen, C.; Gonzalez, F. J.; Idle, J. R. LC-MS-based metabolomics in drug metabolism. *Drug Metab. Rev.*, **2007**, *39*(2-3), 581-597.

[19] Weckwerth, W.; Morgenthal, K. Metabolomics: From pattern recognition to biological interpretation. *Drug Discov. Today,* **2005**, *10*(22), 1551-1558.

[20] Parida, S. K.; Kaufmann, S. H. The quest for biomarkers in tuberculosis. *Drug Discov. Today,* **2010**, *15*(3), 148-157.

[21] Schoeman, J.; Loots, D. Improved disease characterisation and diagnostics using metabolomics: A review. *J. Cell Tissue Res.,* **2011**, *11*(1), 2673-2683.

[22] Loots, D. T.; Wiid, I. J.; Page, B. J.; Mienie, L. J.; Helden, P. D. Melatonin prevents the free radical and MADD metabolic profiles induced by antituberculosis drugs in an animal model. *J. Pineal Res.,* **2005**, *38*(2), 100-106.

[23] Schoeman, J. C.; du Preez, I.; Loots, D. T. A comparison of four sputum pre-extraction preparation methods for identifying and characterising *M. tuberculosis* using GCxGC-TOFMS metabolomics. *J. Microbiol. Methods,* **2012**, *91*(12), 301-311.

[24] du Preez, I.; Loots, D. New sputum metabolite markers implicating adaptations of the host to *Mycobacterium tuberculosis*, and vice versa. *Tuberculosis,* **2013**, *93*(3), 330-337.

[25] Halouska, S.; Fenton, R. J.; Barletta, R. G.; Powers, R. Predicting the *in vivo* mechanism of action for drug leads using NMR metabolomics. *ACS Chem. Biol.,* **2012**, *7*(1), 166-171.

[26] Powers, R. NMR metabolomics and drug discovery. *Magn. Reson. Chem.,* **2009**, *47*(S1), S2-S11.

[27] Shulaev, V. Metabolomics technology and bioinformatics. *Briefings Bioinformat.,* **2006**, *7*(2), 128-139.

[28] Aligent Technologies. Considerations for selecting GC/MS or LC/MS for metabolomics. O*nchnscohpcnshco Metabolom*. **2007**.

[29] Want, E.J.; Nordström, A.; Morita, H.; Siuzdak, G. From exogenous to endogenous: The inevitable imprint of mass spectrometry in metabolomics. *J. Proteome Res.,* **2007**, *6*(2), 459-468.

[30] Wallace, A.; Gibson, S.; De La Hunty, A.; Lamberg-Allardt, C.; Ashwell, M. Measurement of 25-hydroxyvitamin D in the clinical laboratory: Current procedures, performance characteristics and limitations. *Steroids,* **2010**, *75*(7), 477-488.

[31] Lei, Z.; Huhman, D. V.; Sumner, L. W. Mass spectrometry strategies in metabolomics. *J. Biol. Chem.,* **2011**, *286*(29), 25435-25442.

[32] Liska, D. The detoxification enzyme systems. *Altern. Med. Rev.,* **1998**, *3*(3), 187-198.
[33] Chasseaud, L. F.; Hawkins, D. Biotransformation of drugs *Encyclopedia of Pharmaceutical Technology, edited by J.Swarbrick and JC Boylan,(New York: Marcel Dekker),* **1990**, 129-157.
[34] Lu, A. Y. Drug-Metabolism research challenges in the new millennium individual variability in drug therapy and drug safety. *Drug Metab. Disposit.,* **1998**, *26*(12), 1217-1222.
[35] Matthews, H. W. Racial, ethnic and gender differences in response to medicines. *Drug Metabol. Drug Interact.,* **1995**, *12*(2), 77-92.
[36] Böttger, E. C. Drug resistance in *Mycobacterium tuberculosis*: molecular mechanisms and laboratory susceptibility testing **2011**, *40*,128-144.
[37] Loots, D. T.; Meissner-Roloff, R. J.; Newton-Foot, M.; Gey van Pittius, Nicolaas C. A metabolomics approach exploring the function of the ESX-3 type VII secretion system of *M. smegmatis*. *Metabolomics,* **2013**, *9*(3), 631-641.
[38] VandenHeuvel, W. J. Drug metabolite identification: Stable isotope methods. *J. Clin. Pharmacol.,* **1986**, *26*(6), 427-434.
[39] Evans, W. E.; Relling, M. V. Pharmacogenomics: translating functional genomics into rational therapeutics. *Science,* **1999**, *286*(5439), 487-491.
[40] Gurumurthy, P.; Ramachandran, G.; Kumar, A.H.; Rajasekaran, S.; Padmapriyadarsini, C.; Swaminathan, S.; Venkatesan, P.; Sekar, L.; Kumar, S.; Krishnarajasekhar, O. Malabsorption of rifampin and isoniazid in HIV-infected patients with and without tuberculosis *Clin. Infect. Dis.,* **2004**, *38*(2), 280-283.
[41] Bento, J.; Duarte, R.; Brito, M. C.; Leite, S.; Lobato, M. R.; do Carmo Caldeira, M.; Carvalho, A. Reminder of important clinical lesson: Malabsorption of antimycobacterial drugs as a cause of treatment failure in tuberculosis. *BMJ Case Reports,* **2010**, doi: 10.1136/bcr./2.2009.2554.
[42] Wells, C. D.; Cegielski, J. P.; Nelson, L. J.; Laserson, K. F.; Holtz, T. H.; Finlay, A.; Castro, K. G.; Weyer, K. HIV infection and multidrug-resistant tuberculosis-the perfect storm. *J. Infect. Dis.,* **2007**, *196*(Suppl 1), S86-S107.
[43] Dean, G. L.; Edwards, S. G.; Ives, N. J.; Matthews, G.; Fox, E. F.; Navaratne, L.; Fisher, M.; Taylor, G. P.; Miller, R.; Taylor, C. B. Treatment of tuberculosis in HIV-infected persons in the era of highly active antiretroviral therapy. *AIDS,* **2002**, *16*(1), 75-83.
[44] Wong, E. B.; Cohen, K. A.; Bishai, W. R. Rising to the challenge: new therapies for tuberculosis. *Trends Microbiol.,* **2013**, *21*, 493-501.
[45] Cheung, C.; Gonzalez, F. J. Humanized mouse lines and their application for prediction of human drug metabolism and toxicological risk assessment. *J. Pharmacol. Exp. Ther.,* **2008**, *327*(2), 288-299.
[46] King, T.; Brucker, M. In *Pharmacology for Women's Health;* Jones & Bartlett Publishers: Ontario, Canada, **2010**.
[47] Michalets, E. L. Update: Clinically significant cytochrome P-450 drug interactions. *Pharmacotherapy,* **1998**, *18*(1), 84-112.
[48] Ma, Z.; Lienhardt, C.; McIlleron, H.; Nunn, A. J.; Wang, X. Global tuberculosis drug development pipeline: The need and the reality. *Lancet,* **2010**, *375*(9731), 2100-2109.
[49] Gorski, J. C.; Vannaprasaht, S.; Hamman, M. A.; Ambrosius, W. T.; Bruce, M.A.; Haehner-Daniels, B.; Hall, S. D. The effect of age, sex, and rifampin administration on intestinal and hepatic cytochrome P450 3A activity&ast. *Clin. Pharmacol. Therapeut.,* **2003**, *74*(3), 275-287.
[50] Gupta, S.; Tyagi, S.; Almeida, D. V.; Maiga, M. C.; Ammerman, N. C.; Bishai, W. R. Acceleration of tuberculosis treatment by adjunctive therapy with verapamil as an efflux inhibitor. *Amer. J. Respirat. Crit. Care Med.,* **2013**, *188*(5), 600-607.
[51] Palomino, J. C.; Martin, A. TMC207 becomes bedaquiline, a new anti-TB drug. *Future Microbiol.,* **2013**, *8*(9), 1071-1080.
[52] Matteelli, A.; Carvalho, A. C.; Dooley, K. E.; Kritski, A. TMC207: the first compound of a new class of potent anti-tuberculosis drugs. *Future Microbiol,* **2010**, *5*(6), 849-858.
[53] Clayton, T. A.; Lindon, J. C.; Cloarec, O.; Antti, H.; Charuel, C.; Hanton, G.; Provost, J.; Le Net, J.; Baker, D.; Walley, R. J. Pharmaco-metabonomic phenotyping and personalized drug treatment. *Nature,* **2006**, *440*(7087), 1073-1077.
[54] Idle, J. R.; Gonzalez, F. J. Metabolomics. *Cell Metabol.,* **2007**, *6*(5), 348-351.
[55] Gonzalez, F. J.; Yu, A. Cytochrome P450 and xenobiotic receptor humanized mice. *Annu. Rev. Pharmacol. Toxicol.,* **2006**, *46*, 41.
[56] Tateno, C.; Yoshizane, Y.; Saito, N.; Kataoka, M.; Utoh, R.; Yamasaki, C.; Tachibana, A.; Soeno, Y.; Asahina, K.; Hino, H. Near completely humanized liver in mice shows human-type metabolic responses to drugs. *Amer. J. Pathol.,* **2004**, *165*(3), 901-912.
[57] Katoh, M.; Yokoi, T. Application of chimeric mice with humanized liver for predictive ADME. *Drug Metab. Rev.,* **2007**, *39*(1), 145-157.
[58] Driver, E. R.; Ryan, G. J.; Hoff, D. R.; Irwin, S. M.; Basaraba, R. J.; Kramnik, I.; Lenaerts, A. J. Evaluation of a mouse model of necrotic granuloma formation using C3HeB/FeJ mice for testing of drugs against *Mycobacterium tuberculosis*. *Antimicrob. Agents Chemother.,* **2012**, *56*(6), 3181-3195.
[59] Kaushal, D.; Mehra, S. Faithful Experimental Models of Human *Mycobacterium tuberculosis* Infection. *Mycobact. Dis.,* **2012**, *2*, e108.
[60] Carlton, W.; Kreutzberg, G. Isonicotinic acid hydrazide-induced spongy degeneration of the white matter in the brains of Pekin ducks. *Amer. J. Pathol.,* **1966**, *48*(1), 91.
[61] Ahmad, S.; Mokaddas, E. Recent advances in the diagnosis and treatment of multidrug-resistant tuberculosis. *Respir. Med.,* **2009**, *103*(12), 1777-1790.
[62] Gumbo, T.; Dona, C. S. S.; Meek, C.; Leff, R. Pharmacokinetics-pharmacodynamics of pyrazinamide in a novel *in vitro* model of tuberculosis for sterilizing effect: A paradigm for faster assessment of new antituberculosis drugs. *Antimicrob. Agents Chemother.,* **2009**, *53*(8), 3197-3204.
[63] Chakraborty, S.; Gruber, T.; Barry, C. E.; Boshoff, H. I.; Rhee, K. Y. Para-Aminosalicylic acid acts as an alternative substrate of folate metabolism in *Mycobacterium tuberculosis*. *Science,* **2013**, *339*(6115), 88-91.
[64] de Carvalho, L. S; Fischer, S. M.; Marrero, J.; Nathan, C.; Ehrt, S.; Rhee, K.Y. Metabolomics of *Mycobacterium tuberculosis* reveals compartmentalized co-catabolism of carbon substrates. *Chem. Biol.,* **2010**, *17*(10), 1122-1131.
[65] Meissner-Roloff, R. J.; Koekemoer, G.; Warren, R. M. A metabolomics investigation of a hyper-and hypo-virulent phenotype of Beijing lineage *M. tuberculosis. Metabolomics,* **2012**, *8*(6), 1194-1203.
[66] Duncan, K. Progress in TB drug development and what is still needed. *Tuberculosis,* **2003**, *83*(1), 201-207.
[67] Ginsberg, A.M.; Spigelman, M. Challenges in tuberculosis drug research and development. *Nature,* **2007**, *13*(3), 290-294.
[68] Jaggarajamma, K.; Sudha, G.; Chandrasekaran, V.; Nirupa, C.; Thomas, A.; Santha, T.; Muniyandi, M.; Narayanan, P. Reasons for non-compliance among patients treated under Revised National Tuberculosis Control Programme (RNTCP), Tiruvallur district, south India. *Indian J. Tuberculosis,* **2007**, *54*(3), 130-135.
[69] Awofeso, N. Anti-tuberculosis medication side-effects constitute major factor for poor adherence to tuberculosis treatment. *Bull. World Health Organ.,* **2008**, *86*(3), B-D.
[70] Amuha, M.G.; Kutyabami, P.; Kitutu, F.E.; Odoi-Adome, R.; Kalyango, J.N. Non-adherence to anti-TB drugs among TB/HIV co-infected patients in Mbarara Hospital Uganda: Prevalence and associated factors. *African Health Sci.,* **2009**, *9*(2), 8-15.
[71] Becker, C.; Dressman, J.; Amidon, G.; Junginger, H.; Kopp, S.; Midha, K.; Shah, V.; Stavchansky, S.; Barends, D. Biowaiver monographs for immediate release solid oral dosage forms: Ethambutol dihydrochloride. *J. Pharm. Sci.,* **2008**, *97*(4), 1350-1360.
[72] Ohno, M.; Yamaguchi, I.; Yamamoto, I.; Fukuda, T.; Yokota, S.; Maekura, R.; Ito, M.; Yamamoto, Y.; Ogura, T.; Maeda, K. Slow N-acetyltransferase 2 genotype affects the incidence of isoniazid and rifampicin-induced hepatotoxicity. *Int. J. Tuberculosis Lung Dis.,* **2000**, *4*(3), 256-261.
[73] Tostmann, A.; Boeree, M. J.; Aarnoutse, R. E.; De Lange, W.; Van Der Ven, J. A. M; Dekhuijzen, R. Antituberculosis drug-induced hepatotoxicity: Concise up-to-date review. *J. Gastroenterol. Hepatol.,* **2008**, *23*(2), 192-202.
[74] Tostmann, A.; Boeree, M. J.; Peters, W. H.; Roelofs, H. M.; Aarnoutse, R. E.; van der Ven, André JAM; Dekhuijzen, P. Isoniazid and its toxic metabolite hydrazine induce *in vitro* pyrazinamide toxicity. *Int. J. Antimicrob. Agents,* **2008**, *31*(6), 577-580.
[75] Arbex, M. A.; Varella, M. de C.L.; Siqueira, H. R. D.; Mello, F. A. F. Antituberculosis drugs: Drug interactions, adverse effects, and use in special situations-part 1: First-line drugs. *J. Brasileiro de Pneumologia,* **2010**, *36*(5), 626-640.

[76] Shih, T.; Pai, C.; Yang, P.; Chang, W.; Wang, N.; Hu, O. Y. A Novel Mechanism Underlies the Hepatotoxicity of Pyrazinamide. *Antimicrob. Agents Chemother.,* **2013**, *57*(4), 1685-1690.

[77] Roy, P.D.; Majumder, M.; Roy, B. Pharmacogenomics of anti-TB drugs-related hepatotoxicity. *Pharmacogenomics,* **2008**, *9*(3), 311-321.

[78] Heifets, L. B.; Flory, M. A.; Lindholm-Levy, P. J. Does pyrazinoic acid as an active moiety of pyrazinamide have specific activity against *Mycobacterium tuberculosis*? *Antimicrob. Agents Chemother.,* **1989**, *33*(8), 1252-1254.

[79] Olivier, I.; Loots, D.T. A metabolomics approach to characterise and identify various *Mycobacterium* species *J. Microbiol. Methods,* **2012**, *88*(3), 419-426.

[80] Betsou, F.; Parida, S. K.; Guillerm, M. Infectious diseases biobanking as a catalyst towards personalized medicine: *Mycobacterium tuberculosis* paradigm. *Tuberculosis,* **2011**, *91*(6), 524-532.

Received: August 26, 2013 Revised: October 28, 2013 Accepted: November 12, 2013

Review

For reprint orders, please contact: reprints@futuremedicine.com

Biomarkers in Medicine

Metabolomics biomarkers for tuberculosis diagnostics: current status and future objectives

Ilse du Preez*[,1], Laneke Luies[1] & Du Toit Loots[1]

[1]School for Physical & Chemical Sciences, Human Metabolomics, North-West University (Potchefstroom Campus), Private Bag x6001, Box 269, Potchefstroom, South Africa, 2531

*Author for correspondence: ilse.dupreez@nwu.ac.za

Numerous studies have contributed to our current understanding of the complex biology of pulmonary tuberculosis and subsequently provided solutions to its control or eradication. Metabolomics, a newcomer to the Omics research domain, has significantly contributed to this understanding by identifying biomarkers originating from the disease-associated metabolome adaptations of both the microbe and host. These biomarkers have shed light on previously unknown disease mechanisms, many of which have been implemented toward the development of improved diagnostic strategies. In this review, we will discuss the role that metabolomics has played in tuberculosis research to date, with a specific focus on new biomarker identification, and how these have contributed to improved disease characterization and diagnostics, and their potential clinical applications.

First draft submitted: 13 October 2016; Accepted for publication: 18 November 2016; Published online: 18 January 2017

Keywords: biomarkers • blood and tissue • breath • culture • diagnostics • metabolites • metabolomics • sputum • tuberculosis • urine

Pulmonary tuberculosis (TB), a mostly curable disease caused by *Mycobacterium tuberculosis*, has reportedly infected approximately a third (1.9 billion) of the world's population, either in its active (symptomatic) or latent (asymptomatic) form. This disease results in the death of 1.5 million individuals per annum, ranking it the world's foremost cause of death from a single infectious bacterial agent [1]. The current diagnosis of latent TB relies primarily on the detection of the host immune response to the *M. tuberculosis* infection. Two methods based on this principle, the tuberculin skin test and the IFN-γ release assay, are currently the only methods recommended by the WHO for the detection of *M. tuberculosis* infection [2]. Although easy to perform, these tests have a number of limitations, since false-positive results can occur in individuals who were previously vaccinated or previously infected with *M. tuberculosis*, and false-negative results are common in patients with a compromised immune system, such as that caused by the HIV [3–5]. Smear microscopy is currently the most commonly used method for diagnosing active TB. Despite the fact that it is a low-cost, quick and simple method, it has a sensitivity of only 62% and cannot distinguish between various *Mycobacterium* species, nor can it detect drug resistance [6]. The current gold standard for diagnosing active TB is bacteriological culture. The latter is based on the observation of growth of *M. tuberculosis* harvested from patient sputum, and the method has a reported sensitivity and specificity of almost 100%, and can be used to detect drug resistance. This method is, however, time-consuming, considering the slow growth rates of mycobacteria, and even the most advanced culture systems such as MB/BacT, take approximately 2 weeks for a diagnosis, thereby delaying treatment onset [7]. More recently, newer technologies such as nucleic

Future Medicine part of fsg

acid amplification techniques, phage assays and serological tests, have also been implemented in the clinical environment. Although these tests have the capacity to outperform the gold standard with regards to diagnostic turnover time, they are far less sensitive [8–10]. Furthermore, the implementation of these tests, especially in low-income countries, is limited due to their high costs and the requirement for expensive infrastructure and highly trained personnel.

When considering the above argument in the light of the most recent TB prevalence statistics, it is clear that TB disease control needs to be re-evaluated and improved upon. One of the first steps toward achieving this goal is to address this problem from a different perspective and to identify new TB biomarkers, which will not only better characterize the disease, but also lead to the development of improved diagnostic strategies. In this review, we discuss the metabolomics-based TB biomarkers identified to date and critically evaluate the methodology used, and each biomarker in the context of their possible application toward TB diagnostics.

The metabolomics research approach

Although numerous single biomarkers have been implemented successfully for use in the diagnosis and prognosis of various disease states, the emergence of the more recent Omics technologies has unleashed the possibility of using biosignatures, which are profiles of combined biomarkers, toward this end. These biosignatures are especially useful when individual biomarkers do not exist or cannot be identified. The term Omics is used to describe research models aimed at acquiring large-scale data from each sample in a sample group, for the purpose of identifying disease biomarkers and/or elucidating novel functional or pathological mechanisms [11]. These datasets can comprise of a collection of genes (genomics), products of gene expression (transcriptomics), proteins (proteomics) or metabolites (metabolomics). Due to the largely untargeted means by which these datasets are usually generated, they may contain hundreds or even thousands of variables, and therefore, biomarkers and/or biosignatures are mined from the data using advanced biostatistical approaches. The validation and clinical performance of these biosignatures are evaluated by reporting the sensitivity, specificity and receiver operator characteristic curves, among others, according to the Standards for the Reporting of Diagnostic Accuracy Studies guidelines [12,13].

One of the newcomers to the Omics revolution, metabolomics, is defined as the nonbiased identification and quantification of the complete metabolome of a specific biological system, using an assortment of analytical techniques, in combination with various computational, statistical and mathematical analyses. The metabolome can be defined as the collection of all the metabolites, which are small molecular compounds (<1500 Da) participating in the metabolic reactions necessary for normal cell function, growth and maintenance [14]. Since the metabolome is the ultimate downstream product of the genome, transcriptome and proteome, a disturbance at any of these levels due to, for example, a pathological stimulus, will alter the metabolome with regards to the presence or absence of specific metabolites and/or their quantities [15]. The identification of these changes forms the basis for metabolomics biomarker discovery. Over the past few years, metabolomics has been used to identify potential diagnostic biomarkers for a variety of disease states, including TB, the latter of which is the focus of this review.

Existing metabolomics biomarkers with potential TB diagnostic capacity

To date, several metabolomics studies have been done with the aim of identifying specific single metabolite biomarkers and/or biosignatures, which can be used in either the initial screening or subsequent speciation phase of TB diagnostics. As indicated in Tables 1–5, these approaches include the use of various sample matrices, which were analyzed using a variety of analytical approaches. The identified biomarkers not only include the well-known TB biomarkers, tuberculostearic acid (TBSA), branched chain fatty acids and other cell wall components, but also novel compounds, never before associated with TB infection. Some of the compounds identified as potential biomarkers were detected in more than one study and in different sample matrices (indicated in bold text in Tables 1–5), the most significant of which will be discussed in the subsequent sections of this review.

Biomarkers detected using *Mycobacterium* cultures

Although the metabolite profiles of the various TB-causing mycobacteria grown *in vitro* differ to some extent from that grown *in vivo*, some researchers still prefer to use bacteriological cultures in the initial stages of diagnostic biomarker identification. Biomarkers identified in this manner are indicative of those organism-specific compounds that may be present in more complex patient-collected diagnostic samples, such as sputum or blood, and this pre-identification may assist in their detection in these patient samples where they may occur in diluted quantities and/or be masked by the matrix background.

Pavlou *et al.* (2004) used a metabolomics approach to distinguish between *M. tuberculosis*, *Mycobacterium*

Table 1. Tuberculosis biomarkers identified in culture, using a metabolomics research approach.

Sample cohort	Analytical apparatus	Biomarkers identified	Study (year)	Ref.
Mycobacterium tuberculosis	ATD-GC–MS	**1-methyl-naphthalene**; 3-heptanone; methylcyclododecane; 2,2,4,6,6-pentamethyl-heptane; trans-anti-1-methyldecahydronaphthalene; 2-hexene;1-methyl-4-(1-ethylethyl)-benzene; **1,4-dimethyl-cyclohexane**; 3,5-dimethylamphetamine; 3-methyl-butanal	Phillips *et al.* (2007)	[16]
M. tuberculosis *Mycobacterium bovis* *M. bovis* (BCG) *Mycobacterium avium complex* *Mycobacterium fortuitum* *Mycobacterium chelonae* *Mycobacterium abscessus*	SPME GC–MS	Methyl phenylacetate; methyl p-anisate; **methyl nicotinate**; o-phenylanisole	Syhre and Chambers (2008)	[17]
M. tuberculosis *M. avium* *M. bovis* *Mycobacterium kansasii*	GC–MS	Isopropyl-tetradecanoic acid; **9-hexadecenoic acid**; **TBSA**; **10-heptadecenoic acid**; **heptadecanoic** acid;**11-eicosanoic acid; 13-docosenoic acid**; tricosane; C32 mycocerosic acid; **15-tetracosenoic acid**	Olivier and Loots (2012)	[18]
M. tuberculosis *M. avium* *M. bovis* *M. kansasii* *Pseudomonas aeruginosa*	GC–MS	**TBSA**; **docosanoic acid**; 2-octyl-cyclopropaneoctanoic acid; **heptadecanoic** acid; **13-docosenoic acid**;**15-tetracosenoic acid**; 2,4-dimethyltetradecanoic acid;**9-hexadecenoic acid**; **eicosanoic acid**; tetradecanoic acid; octacosanoic acid; pentacosanoic acid; 2-hexyl-cyclopropaneoctanoic acid; purine; **9-octadecenoic acid**; indole-acetic acid; **11-eicosenoic acid**; citric acid; **cadaverine**; inositol; 5′-adenylic acid; decanoic acid;9-hexadecenoic acid (trans); **myo-inositol**; succinic acid; putrescine; erythritol; valeric acid	Olivier and Loots (2012)	[19]
M. tuberculosis *M. avium complex* *M. bovis* (BCG) *M. chelonae* *M. fortuitum* *M. kansasii*	UHPLC-ESI-QTOFMS	Dexpanthenol; Val-His-Glu-His; 1-tuberculosinyladenosine; 1-tuberculosinyl2′-deoxyadenosine; 1-tuberculosinyl derivative; 1-tuberculosinyl-O-acetyladenosine; phosphatidylglycerol	Lau *et al.* (2015)	[20]

Compounds in bold have been identified as potential TB biomarkers in more than one metabolomics study.
ATD: Automated thermal desorption; BCG: Bacillus Calmette–Guerin; ESI-QTOFMS: Electrospray ionisation quadrupole time-of-flight mass spectrometry; SPME: Solid-phase microextraction; TB: Tuberculosis; TBSA: Tuberculostearic acid; TOFMS: Time-of-flight mass spectrometry; UHPLC–MS/MS: Ultrahigh performance liquid chromatography–tandem mass spectrometry.

avium, *Pseudomonas aeruginosa* and a mixed infection *(M. tuberculosis* and *Mycobacterium scrofulaceum)*. They achieved a 96% prediction value when analyzing the headspace above the cultured samples of the aforementioned infectious *Mycobacterium* species, using an electronic nose (EN). The underlying principle of the method is as follows: the EN contains 14 conducting-polymer sensor arrays, which interact with the volatile compounds liberated in the headspace of the cultured *Mycobacterium* samples. This interaction results in a change in electrical resistance, producing characteristic signals of multiple sensor parameters. These variables collected for each culture were then used to construct a data matrix which was statistically analyzed via discriminant function analysis to differentiate the sample groups. When this methodology was applied to 46 patient sputum samples, it successfully differentiated sputum containing *M. tuberculosis*, *M. avium*, *P. aeruginosa*, a mixed infection and eight control samples (TB-negative patients admitted to the clinic) [37]. The

method was then further validated using 330 TB-positive and -negative patient-collected sputum samples, and correctly identified TB with a sensitivity of 89% and specificity of 91%, at a detection limit of 1×10^4 bacteria per ml sputum [38]. When investigating the potential clinical application of this EN methodology, two different off-the-shelf EN devices were tested in a real-life clinical setting. The first device could diagnose TB with a sensitivity of 68%, specificity of 69% and accuracy of 69%, while the second showed a sensitivity of 75%, specificity of 67% and accuracy of 69%. Although these results show promise, they do suggest that the EN technology is not yet sensitive, specific or accurate enough for clinical application toward TB diagnostics [39]. Furthermore, the chemical characteristics and identities of the signals detected cannot be characterized or quantified as yet, and therefore, the characteristic compounds causing these signals were not identified.

Mycobacterial cell wall components, specifically the characteristic mycolic acids, potentially serve as good indicators for the presence of this pathogen in a biological sample. Various studies have focused on the presence or absence and the ratios of these compounds, as a means to characterize, identify or differentiate the various infectious *Mycobacterium* species from one another, for application to TB diagnostics [40–42]. To date, the Sherlock™ Mycobacteria ID system (MYCO-LCS [MIDI Inc., DE, USA]) is the only commercially available HPLC-based, pattern recognition method for mycobacterial speciation, and functions by identifying variations in the infected mycobacteria's mycolic acid content. This method claims to have the ability to identify and differentiate 25 mycobacteria species and 35 nonmycobacteria species from pure culture, in <30 min, at a cost of approximately US$6 per analysis [43]. In 2012, Olivier and Loots applied a similar approach, identifying the characteristic lipid profiles of *M. tuberculosis*, *Mycobacterium bovis*, *Mycobacterium kansasii* and *M. avium,* using a gas chromatograpy–mass spectrometry (GC–MS) metabolomics research methodology, to fully automate this classification model. Twelve lipid biomarkers were identified (Table 1) and used to build a multivariate discriminant model, which could correctly assign unknown samples to their respective species groups with probabilities ranging from 72 to 100%. Although this method requires 16 h to achieve a diagnostic result, only 1×10^3 cultured cells were needed for correct classification, making it more sensitive than the previously described HPLC approach [19]. The research team, however, later improved on this method by implementing the use of a mixer mill to assist in the extraction of these biomarkers, resulting in a 5 h turnaround time and the need for smaller patient sample and solvent volumes. A further development was the application of an alternative extraction approach, allowing for the

Table 2. Tuberculosis biomarkers identified in human sputum, using a metabolomics research approach.

Sample cohort	Analytical apparatus	Biomarkers identified	Study (year)	Ref.
6 × TB^- spiked with *Mycobacterium tuberculosis* 6 × TB^- (not spiked)	GCxGC-TOFMS	**Nonadecanoic acid**; **TBSA**; hexacosanoic acid; acetohydroxamic acid; **myo-inositol**; **eicosanoic acid;á-D-glucopyranoside**; D-glycero-L-manno-heptanoic acid; palmitoleic acid; furan; tetracosanoic acid; arachidonic acid; octadecanoic acid; **D-glucose**; propane; **heptadecanoic acid**; **docosanoic acid**; **9-octadecenoic acid**; **D-mannose**; D-galactose; **D-glucosamine**; glycerol; **uridine**; D-fructose; 2-O-glycerol-à-D-galactopyranoside; **cadaverine**; á-D-galactopyranoside; á-D-xylopyranose; arabinofuranose; D-erythro-pentitol; **L-threonine**;á-D-galactofuranose; phenylethanolamine	Schoeman *et al*. (2012)	[21]
61 × TB^+ 34 × TB^-	GCxGC-TOFMS	à-D-glucopyranose-2-acetylamino-2-deoxy;**à-D-glucopyranoside**; á-L-mannopyranose; *N*-acetyl-glucosamine; à-D-mannopyranoside; D-galactose-6-deoxy; à-D-galactopyranose; **D-glucosamine**; D-citramalic acid; methyl-17-methyl-octadecanoic acid; D-gluconic acid;δ-lactone; **10-heptadecenoic acid**; **9-octadecenoic acid**; **nonadecanoic acid**; 2-deoxy-D-erythro-pentitol; glutaric acid; sebacic acid; ethane; butanal; γ-aminobutyric acid; 3,4-dihydroxybutanoic acid; normetanephrine	Du Preez and Loots (2013)	[22]

Compounds in bold have been identified as potential TB biomarkers in more than one metabolomics study.
TB: Tuberculosis; TB^+: Positively diagnosed for active TB via culture; TB^-: Patients portraying TB-related symptoms, but tested negative for TB via culture; TBSA: Tuberculostearic acid; TOFMS: Time-of-flight mass spectrometry.

Table 3. Tuberculosis biomarkers identified in blood and tissue, using a metabolomics research approach.

Sample matrix	Sample cohort	Analytical apparatus	Biomarkers identified	Study (year)	Ref.
Lung tissue (guinea pigs)	40 × TB^+ 10 × healthy controls	HRMAS NMR	**Aspartate**; **glutathione**; **betaine**; trimethylamine *N*-oxide; **lactate**; **alanine**; **acetate**; **creatine**; **phosphocholine**;**myo-inositol**; glycerophosphocholine	Somashekar *et al.* (2012)	[23]
Lung tissue (mice)	10 × TB^+ 10 × healthy controls	^{1}H NMR	**AMP**; **creatine**; **lactate**; **glucose**; NAD+; **NADP+**; **oxaloacetate**; **fumarate**; **succinate**; **alanine**; **aspartate**; **glutamate**; **leucine**; **lysine**; **isoleucine**; **phenylalanine**; **tyrosine**; **glutamine**; **betaine**; **phosphocholine**; **niacinamide**; phosphoethanolamine; acetaldehyde; **taurine**; **UDP-glucose**; **uracil**; **uridine**; xanthine; **glutathione**; itaconic acid	Shin *et al.* (2011)	[24]
Spleen tissue (mice)	10 × TB^+ 10 × healthy controls	^{1}H NMR	**AMP**; **lactate**; **NADP+**; **oxaloacetate**; **fumarate**; **succinate**; **alanine**; **leucine**; **isoleucine**; **tyrosine**; **glutamine**; **phosphocholine**; **niacinamide**; **uridine**; **glutathione**	Shin *et al.* (2011)	[24]
Liver tissue (mice)	10 × TB^+ 10 × healthy controls	^{1}H NMR	**AMP**; **creatine**; **lactate**; glycogen; **glucose**; **NADP+**; **oxaloacetate**; **succinate**; **aspartate**; **glutamate**; **leucine**; **lysine**; **isoleucine**; **uridine**; **phenylalanine**; **tyrosine**; **glutamine**; **phosphocholine**; **niacinamide**; **taurine**; **UDP-glucose**; **uracil**	Shin *et al.* (2011)	[24]
Serum (mice)	10 × TB^+ 10 × healthy controls	^{1}H NMR	**Glucose**; **leucine**; **isoleucine**; **phenylalanine**; **formate**	Shin *et al.* (2011)	[24]
Serum (guinea pigs)	40 × TB^+ 10 × healthy controls	HRMAS NMR	**Lactate**; **choline-containing compounds**; ethanolamine; phosphocreatine; nicotinamide; **glutamate**; **formate**; **acetate**	Somashekar *et al.* (2012)	[23]
Serum (human)	10 × TB^+ 6 × TB^+ (before and after treatment) 10 × TB^-	GC-TOFMS	5-oxoproline	Che *et al.* (2013)	[25]
Serum (human)	44 × TB^+ 46 × latent TB^+ (asymptomatic) 46 × healthy controls	UPLC–MS/MS (positive mode) UPLC–MS/MS (negative mode) GC–MS	Histidine; cysteine; **threonine**; citrulline; cysteine-glutathione disulfide; urea; *N*-acetylneuraminate; glycocholenate sulfate; inosine; **tryptophan**; **mannose**;3-carboxy-4-methyl-5-propyl-2-furanpropanoate; **phenylalanine**; pyroglutamine; taurocholenate sulfate; **glutamine**; octadecanedioate; γ-glutamylglutamine; glycylvaline; asparate	Weiner *et al.* (2012)	[26]

Compounds in bold have been identified as potential TB biomarkers in more than one metabolomics study.
Healthy control: Healthy with no TB-related symptoms and tested negative for TB via culture; HRMAS: High-resolution magic angle spinning; NAD+: Nicotinamide adenine dinucleotide; NADP+: Nicotinamide adenine dinucleotide phosphate; TB: Tuberculosis; TB^+: Positively diagnosed for active TB via culture; TB^-: Patients portraying TB-related symptoms, but tested negative for TB via culture; TOFMS: Time-of-flight mass spectrometry; UHPLC–MS/MS: Ultrahigh performance liquid chromatography-tandem mass spectrometry.

Table 3. Tuberculosis biomarkers identified in blood and tissue, using a metabolomics research approach (cont.).

Sample matrix	Sample cohort	Analytical apparatus	Biomarkers identified	Study (year)	Ref.
Serum (human)	38 × TB^+ 39 × healthy controls	^{1}H NMR	1-methylhistidine; acetoacetate; acetone; **glutamate**; **glutamine**; **isoleucine**; **lactate**; **lysine**; glycine; nicotinate; **phenylalanine**; pyruvate; **tyrosine**; **alanine**; **formate**; glycerolphosphocholine; low-density lipoproteins	Zhou *et al.* (2013)	[27]
Serum (human)	120 × TB^+ 105 × healthy controls 51 × lung cancer 45 × pneumonia 28 × COPD 22 × bronchiectasis	UPLC–MS	Hexadecanoic acid; lysoPC (C16:0); lysoPC (C18:0); 3D, 7D, 11D-phytanic acid; **docosanoic acid**; phytal; threoninyl-γ-glutamate; kynurenine; **quinolinic acid**; presqualene diphosphate; lysoPC (P-18:1 (9Z)); lysoPC (P-16:0)	Feng *et al.* (2015)	[28]
Serum (human)	20 × TB^+ 18 × healthy controls	LC–MS	99 biomarkers identified – related to glycerophospholipid and arachidonic acid metabolism	Zhong *et al.* (2016)	[29]
Plasma (human)	17 × TB^+ 17 × healthy controls	LC–MS	61 biomarkers identified, including anti-TB drugs, glutamate, choline derivatives, *Mycobacterium tuberculosis*-derived cell wall glycolipids (trehalose-6-mycolate and phosphatidylinositol), lipid mediators of inflammation	Frediani *et al.* (2014)	[30]

Compounds in bold have been identified as potential TB biomarkers in more than one metabolomics study.
Healthy control: Healthy with no TB-related symptoms and tested negative for TB via culture; HRMAS: High-resolution magic angle spinning; NAD+: Nicotinamide adenine dinucleotide; NADP+: Nicotinamide adenine dinucleotide phosphate; TB: Tuberculosis; TB^+: Positively diagnosed for active TB via culture; TB^-: Patients portraying TB-related symptoms, but tested negative for TB via culture; TOFMS: Time-of-flight mass spectrometry; UHPLC–MS/MS: Ultrahigh performance liquid chromatography-tandem mass spectrometry.

detection of metabolites belonging to not only the lipid markers described, but also other compounds including amino acids, alcohols, organic acids, monosaccharides, alkenes, alkanes, purines, pyrimidines, etc. Nineteen metabolites were identified as biomarkers for differentiating *M. tuberculosis*, *M. avium*, *M. bovis*, *M. kansasii* and *P. aeruginosa* (Table 1). Indole-acetic acid, cadaverine, purine, putrescine and two unknown compounds with masses of 343 and 373 were detected exclusively in *P. aeruginosa*, whereas inositol and myo-inositol were characteristic of the *Mycobacterium* species investigated. Succinic acid and an unknown compound with a mass of 268 were uniquely identified in *M. kansasii*, and another unknown compound with a mass of 541 was detected exclusively in *M. tuberculosis* [18]. In addition to the capacity of the extraction method to identify a greater variety of metabolites, it further reduced the solvent volumes required, and improved the detection limit to only 250 bacterial cells per analysis [44].

More recently, Lau *et al.* (2015) described an optimized ultrahigh performance liquid chromatography-ESI-quadruple time-of-flight mass spectrometry (Q-TOFMS) metabolomics approach to differentiate various *Mycobacterium* species from one another, by analyzing their excreted metabolite profiles. Although the main aim of this investigation was to differentiate *M. tuberculosis* and nontuberculous mycobacteria, clear separation between all species was achieved when the collected data were analyzed using principle component analysis and partial least squares-discriminant analysis. Of the 24 biomarkers identified (metabolites detected in significantly higher concentrations in *M. tuberculosis* as compared with the other nontuberculous mycobacteria strains), only seven could be annotated via MS/MS database matching (Table 1) [20].

Furthermore, in the context of the rising incidence of drug resistance, a successful TB diagnostic method should also be able to identify whether the infected *M. tuberculosis* strain is fully drug susceptible or drug resistant to either first- or second-line drugs. Although no metabolomics study to date has identified biomarkers for TB drug-susceptibility testing *per se*, differentiation of various drug-resistant strains from drug-susceptible parent strains has been demonstrated using the differences in the metabolite profiles of these organisms. In two related studies, GC–MS metabolomics

were used to compare the fatty acid metabolome [45] and total metabolome [46] profiles of two *rpoB* mutant *M. tuberculosis* strains (S522L and S531L) to that of a fully drug-susceptible wild-type parent strain, to better characterize rifampicin resistance. In each study, biomarkers for rifampicin resistance were identified and used to elucidate previously unknown biological mechanisms related to the survival and adaptation of these mutant strains. In a similar fashion, the group also investigated monoresistance to isoniazid (resulting from mutations in the *katG* gene) in *M. tuberculosis*. From the differentiating metabolites identified, it was shown that the isoniazid-resistant strains are more susceptible to oxidative stress, and a subsequent survival adaptation to this is to increase the uptake and utilization of alkanes and fatty acids as a carbon/energy source, and to synthesize antioxidant compounds, that is, ascorbic acid, and its oxidation via an ascorbate degradation pathway [47]. Although the diagnostic potentials of the identified biomarkers were not investigated, these studies show that metabolomics can be implemented toward the identification of characteristic drug-resistant biomarkers.

Although the methods described above all required a time-consuming culturing step, these attempts at improved TB diagnostics using bacterial cultures prove the capacity of metabolomics for detecting new diagnostic biomarkers. Ideally, a TB diagnostic method, including the initial screening and speciation steps, should be done directly from the collected clinical samples (preferably collected in a noninvasive manner) to shorten the delay between disease presentation and treatment onset. The true clinical application of these biomarkers can therefore only be determined when investigating their diagnostic value in patient-collected samples, as was done by Syhre *et al.* [34] and Phillips *et al.* [35] (see the 'Biomarkers detected using breath' subsection).

Biomarkers detected using sputum

The most commonly used sample matrix for diagnosing pulmonary TB to date is sputum. This mucus-like biofluid originates directly from the area of infection, that is, the airways of the lungs, and is highly populated with *M. tuberculosis* in TB-positive patients. A metabolic profile of a TB patient's sputum would subsequently contain *Mycobacterium*-specific metabolites, due to the physical presence of the organism in this sample, and also various altered, disease-induced host metabolites, thereby expanding the possibility of identifying diagnostic biomarkers. In addition to diagnostics, these biomarkers could also be implemented toward the elucidation of disease mechanisms, seeing that they are a true reflection of adaptations to the metabolome of *M. tuberculosis* due to *in vivo* growth and the host response to infection/disease. Despite this, however, very few metabolomics studies have been focused on biomarker identification for improved TB diagnostics using patient-collected sputum. The reason for this is most likely linked to the complexities in using this sample matrix, such as its viscosity and uneven consistency, and also the likelihood of possible TB infection when handling these samples in a standard analytical laboratory not equipped for these pur-

Table 4. Tuberculosis biomarkers identified in human urine, using a metabolomics research approach.

Sample cohort	Analytical apparatus	Biomarkers identified	Study (year)	Ref.
117 × TB^+ 37 × healthy controls 7 × COPD 5 × other pulmonary diseases	GC–MS	*o*-xylene; isopropyl acetate; 3-pentanol; dimethylstyrene; cymol	Banday *et al.* (2011)	[31]
21 × TB^+ 21 × TB^-	GC–MS	Lactic acid; norepinephrine; hydroquinone;4-hydroxybenzoic acid; **glucose**†	Das *et al.* (2015)	[32]
30 × healthy controls 46 × TB^+	GCxGC-TOFMS	2-C-methylglycerol; 2-octenoic acid; 5-hydroxyhexanoic acid; 5-hydroxyhydantoin; glycerol monostearate;indole-3-carboxylic acid; kynurenic acid; L-rhamnulose; oxalic acid; phenylacetic acid; **quinolinic acid**; ribitol; homovanillic acid; 5-hydroxyindoleacetic acid; **tryptophan**; phenyllactic acid; ***N*-acetyltyrosine (indicating elevated tyrosine)**	Luies and Loots (2016)	[33]

Compounds in bold have been identified as potential TB biomarkers in more than one metabolomics study.
†Of the 42 biomarkers, the identities of only these six compounds were confirmed.
COPD: Chronic obstructive pulmonary disease; Healthy control: Healthy with no TB-related symptoms and tested negative for TB via culture; TB: Tuberculosis; TB^+: Positively diagnosed for active TB via culture; TB^-: Patients portraying TB-related symptoms, but tested negative for TB via culture; TOFMS: Time-of-flight mass spectrometry.

Table 5. TB biomarkers identified in human breath using a metabolomics research approach.

Sample cohort	Analytical apparatus	Biomarkers identified	Study (year)	Ref.
10 × TB^+ 10 × healthy controls	SPME GC–MS	**Methyl nicotinate**	Syhre *et al.* (2009)	[34]
23 × TB^+ 19 × TB^- 59 × healthy controls	ATD-GC–MS	1,3-dimethyl-trans-cyclohexane;1,4-dichloro-benzene; **1,4-dimethyl-cyclohexane**;2-butyl-1-octanol; **1-methyl-naphthalene**; camphene; 4-methyl-decane; 3-ethyl-2-methyl-heptane; 2,6-dimethyl-octane; 1,2,3,4-tetramethyl-benzene; 3,6,6-trimethyl-bicyclo-3,1,1-hept-2-ene; 1-ethyl-4-methyl-trans-cyclohexane; l-β-pinene; ethyl-benzene; methyl-benzene; propyl-benzene; 3-methyl-heptane; 2-methoxy-2-me-propane; 1-octene; cyclohexane; heptanal;2,4-dimethyl-heptane; 4-methyl-heptane; nonanal;2-methyl-pentane; styrene; **tridecane**	Phillips *et al.* (2007)	[16]
30 × TB^+ 196 × TB^-	ATD-GC–MS	3-(1-methylethyl)-oxetane; 4-methyl-dodecane; hexyl-cyclohexane; Bis-(3,5,5-trimethylhexyl) phthalate; 1,3,5-trimethyl-benzene; 3,7-dimethyl-decane; **tridecane**; 4,6,8-trimethyl-1-nonene; 5-ethyl-2-methyl-heptane;4-methyl-1-hexene	Phillips *et al.* (2010)	[35]
50 × TB^+ 50 × TB^-	GC–MS	Dodecane; 3-heptafluorobutyroxypentadecane; octanol; 5-hexenoic acid; 2-ethyl-1-hexanol	Kolk *et al.* (2012)	[36]

Compounds in bold have been identified as potential TB biomarkers in more than one metabolomics study.
ATD: Automated thermal desorption; GC–MS: Gas chromatography–mass spectrometry; Healthy control: Healthy with no TB-related symptoms and tested negative for TB via culture; SPME: Solid-phase microextraction; TB: Tuberculosis; TB^+: Positively diagnosed for active TB via culture; TB^-: Patients portraying TB-related symptoms, but tested negative for TB via culture.

poses. The aforementioned viscosity and uneven consistency of sputum bring about the need for additional, time-consuming sample preprocessing steps before metabolomics analyses can commence. These methods were not always standardized, and even after successfully applying the necessary extraction procedures and analyses of the samples, the metabolic profiles obtained are complex and contain many compounds from a variety of compound classes. Subsequently comprehensive statistical analyses to sift through the masses of data, are required to identify potential biomarkers [22].

In 2012, specifically for metabolomics TB biomarker identification, Schoeman *et al.* developed a new sputum preprocessing method, in conjunction with a global metabolite extraction approach and GC×GC-TOFMS analysis, to differentiate sputum spiked with *M. tuberculosis* bacilli and control sputum (not spiked). The compounds identified as markers mainly represented those associated with the cell wall of *M. tuberculosis* (Table 2) [21]. This method was applied to differentiate culture-confirmed TB-positive (n = 34) and -negative (n = 61) patient sputum, and the 22 compounds (Table 2), best explaining the variation between the groups, were once again identified. Various new disease mechanisms were described, including the presence of a citramalate cycle in *M. tuberculosis*, and the interaction of this cycle with an upregulated glyoxylate cycle and increased fatty acid oxidation during *in vivo* growth in the human host. Furthermore, these biomarkers also shed light on an additional process by which the host produces hydrogen peroxide via glucose oxidation, as a means to eliminate the infected bacteria. The elevated concentrations of various neurotransmitters associated with the TB infection provided added information explaining many of the clinical symptoms that TB patients experience [22]. Although the objective of this study was to better characterize TB, the differentiation of the groups based on underlying metabolite differences suggests the potential for this type of metabolomics approach to be used for diagnostic purposes.

Although not using a traditional metabolomics approach, research groups have also investigated the possible use of individual sputum metabolites, such as the aforementioned TBSA, for diagnosing TB. In 2009, Cha *et al.* indicated that TBSA can be detected directly from patient sputum samples using solid-phase microextraction and postderivatization coupled to GC–MS. The group reported that this procedure is more sensitive than smear microscopy and requires only 5 h per analysis, thereby proving that the use of single metabolite biomarkers can be useful in the ini-

tial screening phase of TB diagnostics. However, similar to a positive smear microscopy result, the detection of TBSA in these samples is due to the presence of the mycolic acids, a common component in the cell walls of all mycobacteria, and can therefore not be used for speciation or the detection of drug resistance [48].

Biomarkers detected using blood & tissue

Blood samples (plasma and serum) are regarded as being comparatively more homogeneous in composition, less viscose and easier to process than sputum. Furthermore, although very low concentrations of potential metabolites originating directly from the infectious bacteria may be detected in the blood, it reflects those metabolites altered in the host due to the infection, and is subsequently better suited to identify host pathological and defense mechanisms [26], to examine disease progression [23], and/or for diagnostic and therapeutic monitoring purposes [27].

When analyzing serum samples collected from healthy controls, individuals with latent *M. tuberculosis* infection, and patients with active TB, using GC–MS and UPLC–MS/MS, Weiner *et al.* (2012) were able to detect alterations in amino acids, lipids and nucleotides, due to the anti-inflammatory changes that typically occur during TB disease progression. The altered metabolite profiles also revealed reduced phospholipase activity, increased indoleamine 2,3-dioxygenase 1 activity, an increased abundance of adenosine metabolic products and indicators of fibrotic lesions in those patients with active TB, comparatively. Additionally, 20 metabolites with potential diagnostic value were identified (Table 3) [26]. Similarly, using serum-based NMR metabolomics, Zhou *et al.* (2013) identified 17 metabolites with possible diagnostic value (Table 3), in addition to better describing the TB disease state in terms of an observed increased glycolysis, lipid degradation, nucleotide biosynthesis, energy consumption and a modified protein metabolism in TB patients. Expanding on this, they also investigated the metabolite changes induced by other disease states, including diabetes, malignancy and community-acquired pneumonia, and subsequently indicated that each of the changes to the metabolome induced by these perturbations is, in fact, disease specific [27]. With the main aim of identifying TB diagnostic biomarkers, Feng *et al.* (2015) also included patients with diseases other than TB in their patient cohort. The serum metabolome profiles of patients with active TB, chronic obstructive pulmonary disorder (COPD), pulmonitis, bronchiectasis, lung cancer and healthy controls, were compared using UPLC–MS analyses. By employing orthogonal partial least squares-discriminant analysis, they were able to differentiate TB patients from healthy controls and patients with other lung conditions, and subsequently identified a TB-specific serum biosignature. A set of 12 metabolites, mostly fatty acids, amino acids and lipids (Table 3), were identified as biomarkers for active TB, and a combination of four of these compounds (lysophosphatidylcholine (18:0), behenic acid, threoninyl-γ-glutamate and presqualene diphosphate) allowed for discrimination between active TB and control samples with an AUC value of 0.991 [28].

With the aim of explaining pathogen-induced changes in the host, in addition to using serum, Shin *et al.* (2011) compared the spleen, lung and liver tissue metabolome data, from *M. tuberculosis*-infected and healthy control mice, using ^{1}H NMR. Based on the metabolite profiles, clear differentiation between TB-positive and control mice was evident for all tissue and serum samples collected, and the most characteristic biomarkers were identified (Table 3). These compounds indicated that precursors of membrane phospholipids, that is, phosphocholine and phosphoethanolamine, as well as the oxidative stress response, glycolysis, amino acid metabolism and nucleotide metabolism, are altered due to the TB disease state [24]. Similarly, Somashekar *et al.* (2012) applied a ^{1}H NMR metabolomics approach using lung and serum samples collected from guinea pigs, 30 and 60 days after infection, with the aim of studying the effects of disease progression on the host metabolome. Sixteen lung metabolites were identified as TB biomarkers (Table 3), four of which (aspartate, glutathione, betaine and trimethylamine N-oxide) were uniquely detected in the active TB group as a result of increased oxidative stress and glutaminolysis in the lung lesions. Furthermore, several other metabolites, including lactate, choline-containing compounds, ethanolamine, phosphocreatine, nicotinamide and glutamate, were detected in reduced amounts in the serum of the infected animals comparatively, whereas formate and acetate were significantly increased after infection. The authors concluded that the metabolite variation detected in both lung tissue and serum indicates metabolic changes associated with necrotic disease processes, including anaerobic glycolysis, glutaminolysis, and gluconeogenesis [23].

In an attempt to find possible diagnostic markers for active TB, not influenced by anti-TB drug administration, Che *et al.* (2013) used a GC-TOFMS metabolomics approach to compare serum samples collected from healthy controls and TB patients, before and after anti-TB treatment. They indicated that cholesterol, galactose and malonic acid were significantly elevated in the TB patient samples, whereas phenylalanine, glycerol 3-phosphate, ornithine, inositol, lactic acid and 5-oxoproline were comparatively downregulated. Of these nine potential TB biomark-

ers, 5-oxoproline was determined to remain unaffected in the TB-positive patient sample group despite commencing treatment using first-line anti-TB drugs. Hence, this biomarker was validated using a bigger sample cohort. The importance of this study is that it indicates the capacity for metabolomics to identify useful biomarkers for early prediction of treatment outcome, and for possible use in new anti-TB drug development [25].

In some instances, due to the complexity of the blood metabolome and the increasing sensitivity of some of the analytical equipment used for the analysis thereof, an extremely large number of biomarkers can potentially be identified. In two such studies using LC–MS, Frediani *et al.* (2014) and Zhong *et al.* (2016) identified 99 and 61 biomarkers, respectively, when comparing the blood metabolomes of TB-positive patients and healthy controls [29,30]. Although the disease-induced alterations to the host's metabolic pathways were discussed in these studies (Table 3), neither investigated the potential diagnostic capacity of the identified biomarkers.

From a diagnostic perspective, although standardized protocols are available for the collection of blood samples, these techniques are considered invasive, and in many patients, the fear of needles can induce a stress response, altering the metabolic profile and inducing many of the markers previously described for the TB disease state [14].

Biomarkers detected using urine

Similar to blood, little or no mycobacteria are expected to be present in the urine of TB patients, and therefore, very low concentrations of those metabolites associated with or unique to *M. tuberculosis* would likely be detected. However, urine collected from TB patients would contain higher concentrations of the altered host-derived metabolites induced by the TB disease state, when compared with that of other patient sample matrixes. This may be considered an additional advantage for biomarker identification for better disease characterization and diagnostic applications, and therefore many metabolomics research investigations have been done using urine as the preferred sample matrix.

Banday *et al.* (2011) analyzed the volatile organic compounds (VOCs) in patient-collected urine for the purpose of discriminating TB patients from healthy controls, and from patients with other diseases such as COPD and lung cancer, using headspace GC–MS. They found significantly elevated levels of o-xylene and isopropyl acetate, accompanied by reduced levels of 3-pentanol, dimethylstyrene and cymol in the TB-positive patient group comparatively (Table 4) [31]. Similarly, Luies and Loots (2016) used a GC×GC-TOFMS approach to investigate the urinary TB metabolite markers induced by infection-related adaptations of the host metabolome and/or host–pathogen interactions. Clear principle component analysis differentiation was achieved between the TB-positive and healthy control group based on their respective metabolite profiles. The identified metabolite biomarkers (Table 4) indicated TB-induced changes resulting in abnormal host fatty acid and amino acid metabolism (phenylalanine in particular), mediated through changes to IFN-γ and possibly insulin [33]. These studies demonstrate the potential of metabolomics to identify TB biomarkers using patient-collected urine.

In addition to disease-induced metabolic changes, urine also directly reflects metabolites excreted by the liver and kidneys, and therefore provides additional information regarding changes to an individual's metabolome as a result of a toxic insult. To this end, Das *et al.* (2015) identified unique urinary TB biomarkers and determined the effects of TB treatment on the levels of these compounds. By comparing the GC–MS analyzed urine metabolic profiles of TB-positive patients to that of TB-negative individuals (with TB-like symptoms), they detected 42 features with potential diagnostic value, of which six could accurately be annotated compound names (Table 4). Most of the metabolites were related to the phenylalanine metabolic pathway (similar as to what was seen by Luies and Loots [2016]), and when using them to plot a receiver operator characteristic curve, the AUC was calculated as 0.85. The influence of first-line TB treatment on these metabolic profiles was investigated by analyzing samples obtained from TB patients at various time points during the treatment regimen. A treatment-dependent trend was observed, and those patients considered successfully cured after treatment completion, showed a metabolic profile similar to that of healthy controls [32]. This study indicates that in addition to pretreatment diagnostics, urinary biomarkers also have the capacity to monitor disease progression and to potentially monitor treatment response and predict treatment outcome. This information would be useful in new TB drug trials, since these markers could serve as surrogate end points for treatment outcome, long before the 6-month treatment regimen is completed.

Biomarkers detected using breath

Due to the recent advancements in breath collection techniques, this sample matrix has become an attractive option for diagnostic biomarker identification for TB. Breath originates directly from the infected lung,

and would be expected to contain the characteristic host and bacterial VOCs. Additionally, it can be collected in a noninvasive manner, even from children and HIV-positive individuals, where sputum collection is problematic.

Using an automated thermal desorption GC–MS, a set of 130 compounds (mostly alkanes, naphthalene and benzene derivatives) was uniquely detected in the headspace of *M. tuberculosis* cultures, compared with that produced in the headspace of sterile culture media. In a pilot study, the group applied this knowledge to patient-collected breath samples, and successfully differentiated healthy controls (n = 59) from hospitalized patients with a suspicion of pulmonary TB, with a 100% sensitivity and specificity. Using only the most abundant compounds detected in the headspace of the *M. tuberculosis* cultures (Table 5) as a diagnostic biosignature, they were also able to classify patients with negative (n = 19) and positive (n = 23) sputum culture results, with a specificity of 100% and sensitivity of 82.6% [16]. This approach was then applied to a larger sample cohort of 226 patients and successfully identified the TB-positive patients (those for whom the sputum culture, smear and chest radiograph were all positive) with a sensitivity of 84% and specificity of 64.7% [35].

Syhre and Chambers (2008) analyzed the compounds in the headspace of cultures of various *Mycobacterium* species (*M. tuberculosis*, *M. fortuitum*, *M. bovis*, *M. bovis* BCG, *M. abscessus* and *M. avium* complex) using solid-phase microextraction and GC–MS. The aim of their investigation was not to identify biomarkers for speciating infection, but rather to identify specific markers that can differentiate *M. tuberculosis* and *M. bovis* from all other mycobacteria, for later application to patient breath sample diagnostics. Two compounds, methyl p-anisate and methyl nicotinate, were identified exclusively in *M. tuberculosis* and *M. bovis*, whereas methyl phenylacetate was identified as a common biomarker in *M. tuberculosis*, *M. bovis* and *M. avium* complex (Table 5) [17]. An *in vivo* validation for methyl nicotinate as a TB-specific biomarker in patient-collected breath, showed significant differences in the concentration when comparing TB-positive patients to healthy controls, proving that it may potentially be valuable in a clinical environment [34]. However, a true validation of this would need to be done in breath collected from TB-positive patients and clinically sick controls (those with TB-like symptoms, but testing negative for TB via culture).

In a related study, Kolk *et al.* (2012), later also compared the GC–MS profiles of breath samples collected from 50 TB-positive patients and 50 hospitalized TB-negative patients. Seven VOCs (Table 5), including one unknown compound, were identified as potential biomarkers and were used to build a classification model that could predict TB with a sensitivity of 72%, specificity of 86% and accuracy of 79%. The model was further validated using a different set of samples (21 TB positive and 50 hospitalized TB negative), and correctly classified TB with a sensitivity of 62%, specificity of 84% and accuracy of 77% [36].

Considering the abovementioned evidence, breath biomarkers show huge potential for their application in a noninvasive TB diagnostic method, however, further validation is necessary in order to determine which of these compounds are present due to a general disease state, oxidative stress or other external confounders such as smoking or diet, and which of these are, in fact, truly TB specific.

Diagnostic validity of the TB biomarkers identified to date

A more comprehensive summary of the 49 metabolite biomarkers detected in more than one study (those markers indicated in bold text in Tables 1–5) is given in the supplementary section of this review (Supplementary Table 1). Considering the biomarkers identified from cultured *Mycobacterium*: 9-hexadecenoic acid, TBSA, heptadecanoic acid, 11-eicosenoic acid, 13-docosenoic acid, 15-tetracosenoic acid, eicosanoic acid and cadaverine, were all detected more than once in different studies conducted by the same research group [18–19,21], aimed at differentiating various infectious *Mycobacterium* species from one another. These eight compounds are associated with the unique cell wall structure of these organisms, which is already well-known to vary between *Mycobacterium* species, and subsequently these markers would be expected to have diagnostic value, especially for speciation [43]. Additionally, 10-heptadecenoic acid, 9-octadecenoic acid, nonadecanoic acid, á-D-glucopyranoside and D-glucosamine were detected more than once as biomarkers for cultured *M. tuberculosis* [19,21], and were confirmed when comparing TB-positive and -negative patient sputum [22]. These biomarkers are also associated with the unique cell wall of *M. tuberculosis* and therefore show promise for possible use in new TB diagnostic methods using patient-collected sputum. An additional 14 compounds characterizing TB were repeatedly detected in different tissue samples (liver, spleen and lung tissue) (Table 3). Interestingly, both myo-inositol and uridine were confirmed as characteristic compounds in TB-positive tissue from mice and guinea pigs [23,24], and were also identified as biomarkers in experiments using cultured *M. tuberculosis* [19,21]. The use of tissue samples in a clinical environment

is, however, impractical, and the presence of these biomarkers in human blood or urine, for instance, needs to be validated before considering their diagnostic capacity. Considering this, eight compounds identified in TB-positive host tissue were in fact also identified as TB biomarkers in serum and urine samples collected from TB-positive patients [26,27,33]. Of these, lysine, isoleucine, phenylalanine and tyrosine were detected in elevated concentrations in all of the TB-positive sample groups analyzed, irrespective of whether tissue, serum or urine was used, which serves as confirmation for their potential use as TB biomarkers. However, discrepancies in the concentrations of many other metabolite markers do leave room for debate, for their potential application for these purposes. Lactate and glutamate, for example, were seen to accumulate in various TB-positive host tissue [24] and TB-positive patient serum [27] samples, but were found to be reduced in serum collected from TB-infected guinea pigs [23]. Additionally, glutamine, was detected in reduced amounts in TB-positive patient serum by one study group [26], and then later in an elevated abundance by a different group, both using patient-collected serum [27]. Similarly, elevated levels of formate were identified in the serum collected from *M. tuberculosis*-infected mice [24] and guinea pigs [23], however, this metabolite was indicated to be reduced in serum collected from TB-positive patients when compared with healthy controls [27].

1-methyl-naphthalene, 1,4-dimethyl-cyclohexane [16] and methyl nicotinate [17,34] were identified as characteristic markers when using headspace analyses of *M. tuberculosis* cultures, and were also confirmed as biomarkers from TB-positive patient-collected breath samples. The use of methyl nicotinate as a possible TB diagnostic marker was further validated in a follow-up study using a larger sample cohort of collected breath samples from TB-positive and -negative patients [34]. Although this compound is used in a niacin test to distinguish cultured *M. tuberculosis* from other related species, its quantities in patient-collected breath may be influenced by other confounding factors such as smoking, chewing betel nuts and a suppressed immune system (for example in HIV-positive patients) [34]. Another TB breath biomarker, tridecane, was verified using two independent sample cohorts [35]. However, since this marker is thought to originate from increased oxidative stress in TB patients [35], it would most probably not be specific to a TB disease state, since elevated oxidative stress is a common occurrence in most diseases. Similarly, the biomarker quinolinic acid, identified in TB patient-collected serum [28] and urine [33], originates from the kynurenine pathway, induced by several proinflammatory cytokines, and chemokines involved immune reactions related to a number of disease states, and thus, probably also not specific enough for TB.

The challenges experienced to date for identifying a single, disease-specific metabolite for diagnostic application, could potentially be solved by using a combination of metabolites/biomarkers simultaneously. Docosanoic acid (behenic acid), for example, was detected as one of the most abundant fatty acids in *M. tuberculosis* [19], but its diagnostic capacity using TB-positive patient samples was rather poor. However, when used in combination with three other compounds (lysophosphatidylcholine (18:0) threoninyl-γ-glutamate and presqualene diphosphate), it could differentiate TB patients from healthy controls as well as patients with other lung diseases with an AUC of 0.991 [28].

Regardless of the aim of the study, all biomarkers or biosignatures identified should be validated both statistically and biologically [49]. Although all of the abovementioned metabolomics investigations identified potentially valuable biomarkers, the majority of these were used to better characterize the disease and the pathogen-induced metabolic changes in the host. For this reason, only 4 of the 21 studies described in this review, and summarized in Tables 1–5, actually validated the identified biomarkers for their potential diagnostic applications, using either an independent test or validation sample set. Another point to note is that many of the TB characterization and diagnostic studies described above, used rather strict inclusion and exclusion criteria for sample selection. Although this serves well to eliminate variation in the sample cohort to better characterize the disease mechanisms, it could potentially bias the diagnostic outcome and would not necessarily be attained in a true clinical setting. Additionally, only two of the diagnostic studies described [28,31] compared the identified TB biomarkers to those identified in patients with related diseases, such as COPD and lung cancer, while only one other research group [26] included asymptomatic individuals with latent TB infection. Considering this, in the immediate future, it would be beneficial to see increased efforts toward validating these newly identified biomarkers for use in a clinical context, especially with regards to improved TB diagnostics.

Future perspective & challenges: from benchtop to clinical application

Considering the above, over the last 10 years, metabolomics has contributed exponentially to the number of new TB biomarkers identified, which have not

only shed light on new disease mechanisms but also improved TB diagnostics. Ideally, a TB diagnostic method should be sensitive, specific, fast, inexpensive and easy to use without extensive training or clinical facilities. Such a method should be able to accurately diagnose TB, and be capable of identifying the species of *Mycobacterium* causing the disease, while also indicating whether or not it is drug resistant. The method should additionally be able to predict and/or monitor treatment outcome since this would determine the most appropriate treatment approach [50]. Although these idealistic diagnostic requirements seem impossible considering the limitations of current TB diagnostic approaches, with the recent advancements in metabolomics and the new TB biomarkers identified to date, it may be achievable. Improvements to the sensitivities and specificities of these approaches can easily be attained using other detection technologies or targeted analytical approaches on the same metabolomics MS-based equipment. The analyses times can also be significantly reduced to only a few minutes, or even seconds, when searching for the identified biomarkers in a targeted manner. Furthermore, metabolomics has identified characteristic biomarkers of active TB using patient-collected sputum, blood, urine and breath (hence the application of these to less invasive diagnostics), and also for speciation, characterizing drug resistance and predicting/monitoring treatment outcome. The latter is not only important in the sense that one could potentially predict a patient's response to first-line treatment at time of diagnosis (i.e., before the medication is administrated) or well before the 6-month treatment regimen is completed, but these markers could accelerate the time-consuming process of new TB medication efficacy testing, and elucidate their mechanisms of action.

The major limitations considering the idealistic diagnostic requirements we propose above regarding the current metabolomics strategies, would be that a metabolomics approach typically uses expensive analytical equipment, operated by highly trained staff, in a laboratory environment. One should, however, bear in mind that the general idea behind metabolomics is to use these high-end analytical apparatus only for the initial biomarker(s) identification and validation phase. Once these markers have been identified, other technologies (like nanotechnology, or other detection technologies) can be used to develop strategies to increase the sensitivity and specificity of detection and to develop an inexpensive kit for rapid point-of-care diagnostics, with minimal training to the user. Alternatively, kits can also be developed for high-throughput laboratory-based TB diagnostics on 96-well plates, using inexpensive laboratory equipment, such as a basic spectrophotometer. Another limitation is that no research group to date has reported a single, unique biomarker present in only TB-positive patients (with the exception of the well-known TBSA). That is not to say that such biomarkers have not in fact been identified, and/or are being developed for diagnostic applications. Using a biosignature rather than a single characteristic marker to diagnose a disease, does complicate the matter to a certain extent, however, with the aid of statistical or mathematical prediction models, or pattern recognition software, this can easily be overcome and the diagnosis automated.

Conclusion

The recent application of metabolomics for TB research has led to an exponential increased number of new TB biomarkers identified over the last decade. These biomarkers have shed light on new mechanisms related to the biology of the causative pathogen, *M. tuberculosis*, as well as to various underlying disease mechanisms and adaptations, such as drug resistance, host–microbe interactions and treatment failure. Additionally, these compounds have also contributed toward the development of new diagnostic procedures for active and latent TB, pathogen speciation and predicting treatment outcome, using patient samples collected via less invasive sampling techniques. Considering this, the new TB biomarkers identified using metabolomics have not only elucidated previously unknown disease mechanisms, but also have the potential to contribute to all aspects of TB clinical care. However, increased efforts toward validating these newly identified biomarkers for use in a clinical context, especially toward improved diagnostics and TB treatment strategies, are of the utmost importance.

Disclosure

The opinions expressed and conclusions derived are those of the authors and are not necessarily those of the National Research Foundation.

Financial & competing interests disclosure

The financial assistance of the National Research Foundation of South Africa for this research is gratefully acknowledged (UID: 95269 and 94217). The authors have no other relevant affiliations or financial involvement with any organization or entity with a financial interest in or financial conflict with the subject matter or materials discussed in the manuscript apart from those disclosed.

No writing assistance was utilized in the production of this manuscript.

Executive summary

- Current tuberculosis (TB) diagnostic strategies fail since there is still no method with the necessary sensitivity, specificity, speed, affordability, accuracy and ease to use without extensive training or clinical facilities, which can determine which species of *Mycobacterium* is causing the disease, whether or not it is drug resistant and with the ability to predict treatment outcome.
- Over the past 10 years, metabolomics has contributed exponentially to the number of new TB biomarkers, which have shed light on new disease mechanisms and also improved TB diagnostics using patient-collected sputum, blood, urine and breath (hence the application of these to less invasive diagnostics).
- These biomarkers can be used together with other technologies (like nanotechnology or other detection technologies) to not only increase the sensitivity and specificity of detection, but also to develop inexpensive kits for application to rapid point-of-care diagnostics, or alternatively high-throughput spectrophotometer-based TB diagnostic methods using 96-well plates.
- In the immediate future, however, it would be beneficial to see increased efforts toward validating these newly identified biomarkers for use in a clinical context.

References

Papers of special note have been highlighted as:
• of interest; •• of considerable interest

1 World Health Organization. Tuberculosis: WHO global tuberculosis report 2014. (Report), 1–2 (2014). http://apps.who.int/iris/bitstream/10665

2 World Health Organization. Diagnostics for tuberculosis: global demand and market potential. (Report), 33–49 (2006). www.who.int/tdr/publications/documents/

3 Ferrara G, Losi M, D'Amico R *et al.* Use in routine clinical practice of two commercial blood tests for diagnosis of infection with *Mycobacterium tuberculosis*: a prospective study. *Lancet* 367(9519), 1328–1334 (2006).

4 Pai M, Kalantri S, Dheda K. New tools and emerging technologies for the diagnosis of tuberculosis: Part I. Latent tuberculosis. *Expert Rev. Mol. Diagn.* 6(3), 413–422 (2006).

5 Pouchot J, Grasland A, Collet C, Coste J, Esdaile JM, Vinceneux P. Reliability of tuberculin skin test measurement. *Ann. Intern. Med.* 126(3), 210–214 (1997).

6 Dhingra VK, Aggarwal N, Rajpal S, Aggarwal JK, Gaur SN. Validity and reliability of sputum smear examination as diagnostic and screening test for tuberculosis. *Indian J. Allergy Asthma Immunol.* 17(2), 67–69 (2003).

7 Moore DF, Curry JI. Detection and identification of *Mycobacterium tuberculosis* directly from sputum sediments by Amplicor PCR. *J. Clin. Microbiol.* 33(10), 2686–2691 (1995).

8 Albert H, Heydenrych A, Brookes R *et al.* Performance of a rapid phage-based test, FASTPlaqueTB, to diagnose pulmonary tuberculosis from sputum specimens in South Africa. *Int. J. Tuberc. Lung Dis.* 6(6), 529–537 (2002).

9 Kim JH, Kim YJ, Ki CS, Kim JY, Lee NY. Evaluation of Cobas TaqMan MTB PCR for detection of *Mycobacterium tuberculosis*. *J. Clin. Microbiol.* 49(1), 173–176 (2011).

10 Lee J-C, Yu F-L, Lin M-H *et al.* Utility of immunochromatographic assay for detecting *Mycobacterium tuberculosis* from positive BACTEC MGIT 960 cultures. *J. Biomed. Lab. Sci.* 22(64), e9 (2010).

11 Wheelock CE, Goss VM, Balgoma D *et al.* Application of 'omics technologies to biomarker discovery in inflammatory lung diseases. *Eur. Respir. J.* 42(3), 802–825 (2013).

12 Bossuyt PM, Reitsma JB, Bruns DE *et al.* The STARD statement for reporting studies of diagnostic accuracy: explanation and elaboration. *Ann. Intern. Med.* 138(1), W1–W12 (2003).

13 Xia J, Broadhurst DI, Wilson M, Wishart DS. Translational biomarker discovery in clinical metabolomics: an introductory tutorial. *Metabolomics* 9(2), 280–299 (2013).

14 Dunn WB, Bailey NJC, Johnson HE. Measuring the metabolome: current analytical technologies. *Analyst* 130(5), 606–625 (2005).

15 Parida SK, Kaufmann SHE. The quest for biomarkers in tuberculosis. *Drug Discov. Today* 15(3), 148–157 (2010).

16 Phillips M, Cataneo RN, Condos R *et al.* Volatile biomarkers of pulmonary tuberculosis in the breath. *Tuberculosis* 87(1), 44–52 (2007).

17 Syhre M, Chambers ST. The scent of *Mycobacterium tuberculosis*. *Tuberculosis* 88(4), 317–323 (2008).

18 Olivier I, Loots DT. A comparison of two extraction methods for differentiating and characterising various *Mycobacterium* species and *Pseudomonas aeruginosa* using GC-MS metabolomics. *Afr. J. Microbiol. Res.* 6(13), 3159–3172 (2012).

19 Olivier I, Loots DT. A metabolomics approach to characterise and identify various *Mycobacterium* species. *J. Microbiol. Methods* 88(3), 419–426 (2012).

• **Describes a multivariate discriminant model for speciation, with probabilities ranging from 72 to 100%.**

20 Lau SKP, Lam C-W, Curreem SOT *et al.* Identification of specific metabolites in culture supernatant of *Mycobacterium tuberculosis* using metabolomics: exploration of potential biomarkers. *Emerg. Microbes Infect.* 4(1), e6 (2015).

21 Schoeman JC, Du Preez I, Loots DT. A comparison of four sputum pre-extraction preparation methods for identifying and characterising *M. tuberculosis* using GCxGC-TOFMS metabolomics. *J. Microbiol. Methods* 91, 301–311 (2012).

22 Du Preez I, Loots DT. New sputum metabolite markers implicating adaptations of the host to *Mycobacterium tuberculosis*, and vice versa. *Tuberculosis* 93(3), 330–337 (2013).

23 Somashekar BS, Amin AG, Tripathi P *et al.* Metabolomic signatures in guinea pigs infected with epidemic-associated

W-Beijing strains of *Mycobacterium tuberculosis*. *J. Proteome Res.* 11(10), 4873–4884 (2012).

24 Shin J-H, Yang J-Y, Jeon B-Y *et al.* ^{1}H NMR-based metabolomic profiling in mice infected with *Mycobacterium tuberculosis*. *J. Proteome Res.* 10(5), 2238–2247 (2011).

- **Compared various sample matrices (serum, as well as spleen, lung, liver tissue) from *Mycobacterium tuberculosis*-infected and healthy control mice, with the aim of explaining pathogen-induced changes in the host.**

25 Che N, Cheng J, Li H *et al.* Decreased serum 5-oxoproline in TB patients is associated with pathological damage of the lung. *Clin. Chim. Acta* 423, 5–9 (2013).

26 Weiner J, Parida SK, Maertzdorf J *et al.* Biomarkers of inflammation, immunosuppression and stress with active disease are revealed by metabolomic profiling of tuberculosis patients. *PLoS ONE* 7(7), e40221 (2012).

- **The only research group who compared asymptomatic individuals with latent TB infection to TB-positive patients.**

27 Zhou A, Ni J, Xu Z *et al.* Application of 1H NMR spectroscopy-based metabolomics to sera of tuberculosis patients. *J. Proteome Res.* 12(10), 4642–4649 (2013).

- **Not only better describes the TB disease state, but also investigated the metabolite changes induced by other disease states, including diabetes, malignancy and community-acquired pneumonia.**

28 Feng S, Du YQ, Zhang L, Zhang L, Feng RR, Liu SY. Analysis of serum metabolic profile by ultra-performance liquid chromatography-mass spectrometry for biomarkers discovery: application in a pilot study to discriminate patients with tuberculosis. *Chin. Med. J.* 128(2), 159–168 (2015).

- **Aimed at identifying TB diagnostic biomarkers, and included patients with chronic obstructive pulmonary disorder, pulmonitis, bronchiectasis, lung cancer and healthy controls in their study cohort.**

29 Zhong L, Zhou J, Chen X, Yin Y. Serum metabolomic study for the detection of candidate biomarkers of tuberculosis. *Int. J. Clin. Exp. Pathol.* 9(3), 3256–3266 (2016).

30 Frediani JK, Jones DP, Tukvadze N *et al.* Plasma metabolomics in human pulmonary tuberculosis disease: a pilot study. *PLoS ONE* 9(10), e108854 (2014).

31 Banday KM, Pasikanti KK, Chan ECY *et al.* Use of urine volatile organic compounds to discriminate tuberculosis patients from healthy subjects. *Anal. Chem.* 83(14), 5526–5534 (2011).

- **Volatile organic compounds were analyzed to discriminate TB patients from healthy controls, and from patients with other diseases such as chronic obstructive pulmonary disorder and lung cancer.**

32 Das MK, Bishwal SC, Das A *et al.* Deregulated tyrosine-phenylalanine metabolism in pulmonary tuberculosis patients. *J. Proteome Res.* 14(4), 1947–1956 (2015).

33 Luies L, Loots D. Tuberculosis metabolomics reveals adaptations of man and microbe in order to outcompete and survive. *Metabolomics* 12(3), 1–9 (2016).

34 Syhre M, Manning L, Phuanukoonnon S, Harino P, Chambers ST. The scent of *Mycobacterium tuberculosis* – Part II breath. *Tuberculosis* 89(4), 263–266 (2009).

35 Phillips M, Basa-Dalay V, Bothamley G *et al.* Breath biomarkers of active pulmonary tuberculosis. *Tuberculosis* 90(2), 145–151 (2010).

36 Kolk A, Van Berkel J, Claassens M *et al.* Breath analysis as a potential diagnostic tool for tuberculosis. *Int. J. Tuberc. Lung Dis.* 16(6), 777–782 (2012).

37 Pavlou AK, Magan N, Jones JM, Brown J, Klatser P, Turner APF. Detection of *Mycobacterium tuberculosis* (TB) *in vitro* and *in situ* using an electronic nose in combination with a neural network system. *Biosensors Bioelectron.* 20(3), 538–544 (2004).

38 Fend R, Kolk AH, Bessant C, Buijtels P, Klatser PR, Woodman AC. Prospects for clinical application of electronic-nose technology to early detection of *Mycobacterium tuberculosis* in culture and sputum. *J. Clin. Microbiol.* 44(6), 2039–2045 (2006).

39 Kolk A, Hoelscher M, Maboko L *et al.* Electronic-nose technology using sputum samples in diagnosis of patients with tuberculosis. *J. Clin. Microbiol.* 48(11), 4235–4238 (2010).

40 Butler WR, Jost KC Jr, Kilburn JO. Identification of mycobacteria by high-performance liquid chromatography. *J. Clin. Microbiol.* 29(11), 2468–2472 (1991).

41 Freeman R, Goodacre R, Sisson PR, Magee JG, Ward AC, Lightfoot NF. Rapid identification of species within the *Mycobacterium tuberculosis* complex by artificial neural network analysis of pyrolysis mass spectra. *J. Med. Microbiol.* 40(3), 170–173 (1994).

42 Glickman SE, Kilburn JO, Butler WR, Ramos LS. Rapid identification of mycolic acid patterns of mycobacteria by high-performance liquid chromatography using pattern recognition software and a *Mycobacterium* library. *J. Clin. Microbiol.* 32(3), 740–745 (1994).

43 Midi Inc. Sherlock™ Mycobacteria Identification System. (2015). www.midi-inc.com/pages/microbial_id.html

44 Du Preez I, Loots DT. Detection limit for differentiating between various *Mycobacterium* species and *Pseudomonas aeruginosa* using gas chromatography–mass spectrometry (GC-MS) metabolomics: a comparison of two extraction methods. *Afr. J. Microbiol. Res.* 7(9), 797–801 (2013).

45 Du Preez I, Loots DT. Altered fatty acid metabolism due to rifampicin-resistance conferring mutations in the rpoB gene of *Mycobacterium tuberculosis*: mapping the potential of pharmaco-metabolomics for global health and personalized medicine. *OMICS* 16(11), 596–603 (2012).

46 Loots DT. New insights into the survival mechanisms of rifampicin-resistant *Mycobacterium tuberculosis*. *J. Antimicrob. Chemother.* 71(3), 655–660 (2016).

47 Loots DT. An altered *Mycobacterium tuberculosis* metabolome induced by katG mutations resulting in isoniazid resistance. *Antimicrob. Agents Chemother.* 58(4), 2144–2149 (2014).

48 Cha D, Cheng DE, Liu M, Zeng Z, Hu X, Guan W.

The role of metabolomics in tuberculosis treatment research

Journal:	*Biomarkers in Medicine*
Manuscript ID	BMM-2017-0141.R1
Manuscript Type:	Review
Keywords:	Metabolomics, Treatment biomarkers, Tuberculosis

The role of metabolomics in tuberculosis treatment research

Abstract

Despite the fact that tuberculosis (TB) is a curable disease, it still results in approximately 1.8 million deaths annually. Various inadequacies in the current TB treatment strategies are major contributors to this high disease prevalence, including the long duration of therapy, the severe side-effects associated with TB drugs, treatment failure due to drug-resistance, post-treatment disease relapse, and HIV co-infection. In this review, we describe how metabolomics has contributed towards better explaining/elucidating the mechanisms of drug action/metabolism, drug toxicity, and microbial drug-resistance, and how metabolite biomarkers may serve as prognostic indicators for predicting treatment outcome as well as for the development of new TB drugs. We also discuss possible future contributions that metabolomics can make toward more efficient, less toxic TB treatment strategies.

Keywords: Biomarkers; Metabolomics; Treatment; Tuberculosis

1. Introduction

Despite the fact that tuberculosis (TB) has been studied for decades, using various research approaches, including genomics, transcriptomics and proteomics, this disease is still regarded as the foremost cause of death from a single infectious bacterial agent, with a global annual mortality rate of 1.8 million individuals [1].

Currently, the World Health Organization (WHO) recommends the Directly Observed Therapy Short-course (DOTS) strategy for treating fully drug-susceptible TB. This program consists of a six month therapeutic regimen, starting with an initial two month (intensive) treatment phase, using a combination of isoniazid (INH), rifampicin (RIF), pyrazinamide (PZA) and ethambutol (EMB), followed by a four month (maintenance) phase, using only INH and RIF. Additionally, each dose of TB drugs is consumed under the supervision of a trained medical professional (direct observation), in an attempt to improve therapeutic adherence [2]. When the DOTS program is fully adhered to, it has a high cure rate in patients with drug-susceptible TB [3], however, almost 15% of the 10.4 million new TB cases reported annually, are still treated unsuccessfully [1]. Several factors, including varying individual metabolism, human immunodeficiency virus (HIV) co-infection, disruptions and/or premature termination of this first-line treatment approach (which are mostly attributed to the drug-related side-effects), and/or limited access to the medical facilities where these drugs are administered, have been associated with an unsuccessful treatment outcome and drug-resistance [4].

For treating multidrug-resistant (MDR)-TB (defined as resistance to both RIF and INH), the WHO recommends the addition of expensive and more toxic second-line drugs to the DOTS program, thereby extending the total therapeutic duration to 24 months (known as DOTS plus) [3]. However, despite these treatment recommendations, globally only 125 000 (20%) of the 580 000 patients with confirmed MDR-TB, actually received treatment in 2015, and of these individuals, only 52% were reportedly treated successfully [5].

An essential process towards TB eradication, would be to reduce the number of failed treatment outcome cases, and also to prevent the development and spread of drug-resistant TB. The first step towards achieving these objectives would be to get a better understanding of the underlying mechanisms of current TB drugs and their related side-effects, in addition to elucidating the mechanisms of drug-resistance and the varying individual drug responses. These new insights could potentially contribute to the development of new, less toxic and more efficient TB drugs, requiring a shorter treatment duration, and resulting in fewer drug-resistant cases. In this review, we will discuss the contributions that metabolomics, and the identification of new metabolite biomarkers using this research approach, have made towards achieving these goals.

2. Using metabolomics towards improving tuberculosis treatment strategies

Metabolomics can be defined as the comprehensive identification and quantification of all small intra- and extra-cellular metabolites (<1500 Da) present in a biological system at a specific point in time, using highly selective and sensitive analytical methods, in conjunction with bioinformatics. The application of metabolomics for biomarker discovery is based on the identification of alterations to the specific metabolome being investigated, due to an external stimulus, such as a disease state or xenobiotics, or as a result of changes to the genome originating from mutations or other genetic variation [6, 7]. However, it is important to consider both the advantages and limitations of the analytical instrumentation/techniques used for metabolomics applications, which have been summarized by De Villiers and Loots (2013). Briefly, gas chromatography mass spectrometry (GC-MS) is relatively affordable with high sensitivity and a good dynamic range, but requires lengthy sample preparation, especially when analyzing non-volatile compounds (requiring derivatization). Liquid chromatography mass spectrometry (LC-MS) is ideal for highly polar and ionic compounds and offers a comprehensive metabolome coverage when being applied in both positive and negative ionization modes, however, it has comparatively lower chromatographic resolution and higher signal-to-noise ratios when compared to that of GC-MS. Capillary electrophoresis mass spectrometry (CE-MS), on the other hand, is fast, affordable, and highly efficient for analyzing polar and ionic

compounds, but has poor a sensitivity and reproducibility when compared to the other techniques described. Lastly, although nuclear magnetic resonance (NMR) is considered highly selective and does not result in sample destruction during analysis, it is has a far lower sensitivity when compared to the aforementioned techniques [4].

With regards to its applications to TB treatment strategies, metabolomics has successfully identified new metabolite biomarkers contributing to (a) an improved understanding of the mechanisms of TB drug action [8], (b) their side-effects [9], and (c) mechanisms by which TB drug-resistance occur [10]. Furthermore, metabolomics has also contributed to the identification of biomarkers which may be used as prognostic indicators for predicting treatment outcome [11], and/or towards the development or clinical testing of new TB drugs [12]. These metabolomics contributions, and the identified biomarkers, will be discussed in greater detail in the remainder of this review.

2.1 Metabolome variations related to tuberculosis drug action and drug metabolism

Although the DOTS and DOTS plus TB treatment regimens are considered effective when fully adhered to by the patient, the exact mechanisms of action and xenobiotic metabolism of each these drugs, in addition to the molecular mechanisms responsible for mycobacterial dormancy (suspending or temporarily stopping all physiological functions, i.e. latent infection), persistence (inability rid the host of metabolically active bacteria despite treatment) and drug-resistance, remain poorly understood [13, 14]. By using a global metabolomics approach, hundreds of different compounds can be analyzed and identified simultaneously, allowing for the discovery of new enzyme activities and metabolic pathways [15]. As summarized in Table 1, a number of research groups to date have applied this approach towards a better understanding of first/second-line TB drug action and metabolism, as well as to investigate the therapeutic characteristics of new TB drug candidates, with the hope that such insights may eventually lead to the development of more effective, less toxic, shortened treatment regimes.

Li *et al.* (2011) applied an ultra-performance liquid chromatography time-of-flight mass spectrometry (UPLC-TOFMS) metabolomics approach, to characterize INH metabolism in humans. When comparing the urinary metabolite profiles obtained from healthy controls to those from INH-administered healthy volunteers, eight known and seven new INH-derived metabolites, the latter of which comprised of two new hydroxylated INH metabolites (2-oxo-1,2-dihydro-pyridine-4-carbohydrazine and isoniazid N-oxide) and five new hydrazones (INH conjugated to 4-methyl-2-oxopentanoic acid and/or methyl-2-oxopentanoic acid; INH conjugated to 2-oxohexanedioic acid; INH conjugated to 3-(4-hydroxyphenyl)-2-oxopropanoic acid; INH conjugated to 2-oxo-3-phenylpropanoic acid; and INH conjugated to 3-(1H-indol-3-yl)-2-oxopropanoic acid). The hydrazones are intermediates of the leucine and/or isoleucine, lysine, tyrosine, tryptophan and phenylalanine metabolic pathways, suggesting a disruption in the essential amino acid metabolism as an additional mechanism by which INH inhibits *M. tuberculosis* growth [16]. Similarly, Nandakumar *et al.* (2014) used LC-TOFMS metabolic profiling to identify metabolic changes in replicating *M. tuberculosis* populations induced by a sub-lethal dose of INH, RIF and streptomycin, respectively. By applying unsupervised hierarchical clustering analysis, a subset of drug-induced metabolome changes common to all three antibiotics was identified. These changes represented the tricarboxylic acid (TCA) cycle, glyoxylate metabolism and biosynthetic pathways of various amino acids. With regards to the TCA intermediates, pyruvate, succinate, fumarate and malate levels were seen to be elevated subsequent to drug exposure, whereas a-ketoglutarate levels reduced, and citrate/isocitrate levels remained unchanged. This result suggests, that despite the previously known diverse bacterial targets of these TB drugs, all three also result in the activation of isocitrate lyases (ICLs). These findings were further substantiated by fact that the ICL-deficient *M. tuberculosis* strains are significantly more sensitive to these TB drugs. Additionally, this increased susceptibility could be chemically restored using antioxidants, confirming that the observed TCA changes are indicative of an adaptive respiratory slowdown caused by an antibiotic-triggered increase in the production of respiratory-derived reactive oxygen intermediates. This metabolomics approach therefore identified a new mechanism by which first-line TB drugs may act on *M. tuberculosis*, thereby suggesting an alternative strategy which may be exploited in the development of more effective TB drugs [17].

With application of metabolomics to investigating second-line TB drug action, Halouska *et al.*, (2007) used a NMR metabolomics approach to identify biomarkers which may explain the mechanisms by which D-cycloserine (DCS), a MDR-TB drug, functions. Although DCS was previously reported to inhibit both D-alanine racemase (Alr) and D-alanine-D-alanine ligase (Ddl) activity *in vitro*, its *in vivo* activity against *M. tuberculosis* was not yet confirmed. Therefore, they compared the NMR metabolite profiles from a wild-type *M. smegmatis* strain to that of an *alr* null mutant, its complemented derivatives, and two DCS-resistant mutant strains. The metabolite changes identified when comparing the complemented derivatives, indicated that a second D-alanine pathway exists in the *alr* null mutants, which appears to persist after exposure to DCS. Principle component analysis (PCA) differentiation of the *alr* null mutant and its complemented derivatives was, however, lost after DCS exposure, which suggests that the multiple DCS targets (Alr and Ddl) determined *in vitro,* are also present *in vivo*. Clustering of the DCS-resistant strains with the complemented derivative strains suggests that DCS treatment also results in inhibition of the cell wall enzymes, with minimal anti-microbial impact however, and that Ddl appears to be the primary *in vivo* lethal target of DCS [8]. In a follow-up metabolomics study, the same group confirmed these *in vivo* drug targets by indicating cell growth is inhibited when D-alanyl-D-alanine production is halted [18]. Prosser and De Carvalho (2013) later confirmed these results using a more direct *ex vivo* stable isotope metabolomics approach, identifying Ddl as a possible target for future drug development initiatives against *M. tuberculosis* [19].

Although there are currently a number of new potential TB drugs undergoing phase II and III preclinical trials, delamanid and bedaquiline are the only newly approved TB drugs in over 50 years. These drugs, however, are currently only used for treating adults with MDR-TB when no other alternatives show effect [20]. Considering the urgent need for new TB drugs and alternative TB treatment approaches, De Carvalho *et al.* (2011) used a targeted LC-MS metabolic profiling method to investigate the potential mechanisms of drug action of nitazoxanide (NTZ) and tizoxanide (TIZ) (two broad spectrum anti-parasitic drugs) against *M. tuberculosis*. The results of their metabolomics investigation proved that both compounds are capable of penetrating and accumulating in

M. tuberculosis cells, and act against replicating and non-replicating populations by disrupting the membrane potential and pH homeostasis. This mechanism of action is identical to that of an anti-mycobacterial drug with a similar chemical structure, niclosamide (NCS), which is an oxidative phosphorylation uncoupler. This study shows how metabolomics can be used in the development of more potent and bioavailable TB drug analogues of NTZ and/or TIZ [21], and similarly to both Halouska *et al.*, (2007) [8] and Prosser and De Carvalho (2013) [19], the capability for using metabolomics to successfully elucidate previously unknown mechanisms of drug action.

2.2 Drug-induced changes to the host metabolome suggesting mechanisms of toxicity

One of the major obstacles in the control of pulmonary TB is the rising incidence of drug-resistant *M. tuberculosis* strains, an occurrence that can largely be attributed to patient non-adherence as a result of the side-effects associated with these drugs. These adverse drug reactions (ADRs) including, but not limited to, hepatotoxicity/hepatitis, jaundice, hypoglycemia, rashes, myocardial damage, neurological complications, respiratory complications, indigestion, joint pain and acidosis [4, 9], have been shown to contribute to approximately 30% of all TB treatment failures [26]. The ability to predict and/or prevent ADRs would undoubtedly improve patient adherence.

To date, several metabolomics studies have focused on characterizing the altered TB drug-induced metabolome, for the purpose of elucidating mechanisms related to TB drug toxicity (Table 2). With the aim of elucidating mechanisms related to INH-induced hepatotoxicity, Bando *et al*. (2011) applied a GC-MS based metabolomics approach to investigate changes in the plasma and urine metabolite profiles of rats subsequent to oral hydrazine (an INH metabolite) administration (with a dosage of 120 and 240 mg/kg). Various metabolite alterations associated with histopathological changes, including fatty acid degradation and glycogen accumulation, were identified. These alterations included a dose-dependent elevation in the amino acid precursors of glutathione (cysteine, glutamate and glycine), 5-oxoproline (a product of glutathione metabolism) and ascorbate. TCA cycle intermediates were reduced in the hydrazine treated rats, whereas urea cycle metabolites and other amino acids were found to be elevated. Collectively, the metabolite changes detected in this study proved the

importance of oxidative stress and glutathione consumption in the etiology of hydrazine-induced hepatotoxicity [27], something which can potentially be utilized for pharmacology and toxicology screening purposes. Using a similar approach, Rawat *et al*. (2017) investigated the toxicity index for PZA and its metabolic products, pyrazinoic acid (PA) and 5-hydroxy pyrazinoic acid (5-OHPA), in rat serum samples, using a ^{1}H NMR metabolomics approach. A list of 19 metabolites were identified as being perturbed in the samples collected from the treated rats. Comparatively, elevated concentrations of the inflammatory associated metabolites: choline/glycerylphosphorylcholine, N-acetyl glycoproteins and O-acetyl glycoproteins, were detected in the treated rat samples, accompanied by reduced lipid levels due to pronounced drug induced oxidative stress and consequent lipid peroxidation. Furthermore, the elevated levels of glucose and lactate in the treated group indicated a disturbance in glucose metabolism and a reduced aerobic glycolytic activity, characteristic of liver toxicity. Furthermore, reduced levels of glucogenic and ketogenic amino acids (glutamate, alanine, valine and leucine) were detected in the treated rats, substantiating their role as key energy metabolites during conditions of oxidative stress. These perturbations were found to be more significant in the serum of those rats dosed with 5-OHPA and PA, indicating that PA and 5-OHPA are more hepatotoxic than the PZA drug itself, by inducing liver injury via oxidative stress and inflammation [28]. The above mentioned studies, prove the capacity of metabolomics to elucidate the mechanisms of toxicity induced by TB drugs individually administered. However, since TB drugs are rarely administrated individually, Loots *et al.* (2005) used a GC-MS metabolomics approach to investigate the effects of the combined TB drug, Rifater, on the metabolome of Sprague-Dawley rats. Their results indicated that Rifater administration increased oxidative stress and induced a metabolic profile resembling that of a multiple acyl-CoA dehydrogenase deficiency (MADD) in humans (Table 2). The metabolic biomarkers identified suggest that Rifater treatment inhibits the functioning of the electron transport chain flavoproteins, and could be associated with many of ADRs reported for Rifater. Furthermore, the observed MADD metabolite profile and the associated oxidative stress could be alleviated by the co-administration of melatonin [9], a potent antioxidant. Considering this, Ozkan *et al*. (2012) measured the plasma melatonin and urinary 6-hydroxymelatonin sulphate levels in 31 patients with pulmonary TB as compared to that of 31 healthy controls. They also proved that

supplementation with melatonin during the course of TB treatment, reduces the severity of the side effects [29]. However, to date, current treatment protocols have not been adapted to include the use of melatonin, possibly since its effective dosage and possible side-effects, particularly during long-term use, have as yet not been adequately evaluated. Additionally, melatonin manufacturing protocols are as yet not standardized, hence the dosage per tablet may vary and/or may also contain harmful substances [30].

2.3 Metabolome variations related to TB drug-resistance

Drug-resistance in *M. tuberculosis* can be classified as either originating from a genetic or phenotypic occurrence. Genetic derived drug-resistance occurs as a result of mutations in the chromosomal genes of the pathogen, resulting in MDR- or extensively drug-resistant (XDR) *M. tuberculosis* strains. Phenotypic derived drug-resistance, on the other hand, occurs as a result of epigenetic alterations in gene expression and protein modification, resulting in a *M. tuberculosis* strain with increased tolerance to TB drugs, especially in non-replicating bacteria (requiring prolonged treatment) [39]. Considering the rising incidence of drug-resistance, especially in many third world countries, it is important to better understand the molecular mechanisms leading to the associated phenotype, in order to successfully diagnose and treat individuals infected with these drug-resistant *M. tuberculosis* strains, or to prevent the occurrence thereof. In 2012, Du Preez and Loots used GC-MS metabolomics to compare the fatty acid metabolomes of two *rpoB* mutant *M. tuberculosis* strains (S522L and S531L) to that of a fully drug-susceptible wild-type parent strain, in order to better characterize RIF-resistance. Their results indicated a decreased synthesis of 10-methyl branched-chain fatty acids and various cell wall lipids, and an increased dependency for short-chain fatty acids as an alternative carbon source in the drug-resistant strains. The more clinically prevalent *rpoB* S531L mutants additionally showed an increased capacity for using these alternative energy sources, compared to the less frequently occurring *rpoB* S522L mutants. This study was the first of its kind to link an altered fatty acid metabolism with RIF-resistance in *M. tuberculosis* [10] (Table 3). In a follow-up investigation, Loots (2016) indicated a total depletion of aconitic acid, suggesting a shift in aconitase

functionality towards mRNA binding and stability (as a survival response to the primary mutation causing drug-resistance) and away from energy production and growth, explaining the reduced viability of these RIF-resistant strains. This explained and substantiated the previously observed increased dependency on alternative energy sources as a survival mechanism. A number of other metabolic changes observed confirmed an additional survival response to the mutation via maintaining/remodeling the cell wall [40]. Contributing to this, Lahiri *et al*. (2016) used a LC-MS based organism-wide screen to compare approximately 10 000 cell wall lipids occurring in RIF-resistant mutants, in order to identify the lipid alterations induced by various *rpoB* mutations. Their results indicated altered mycobactin siderophores and acylated sulfoglycolipid concentrations, both of which are essential for bacterial growth. This study provided further evidence of characteristic cell wall lipid remodeling in RIF-resistant *M. tuberculosis* strains first proposed by Loots (2015), and additionally suggests new diagnostic and therapeutic targets against drug-resistant strains of *M. tuberculosis* [41]. Loots (2014) also investigated mono-resistance to INH (resulting from mutations in the *katG* gene) in *M. tuberculosis*, using GCxGC-TOFMS metabolomics, and compared the metabolite profiles of two INH-resistant strains (H15 and H71) to that of a wild-type parent strain. Results indicated that the INH-resistant strains were more susceptible to oxidative stress, and adapted to this via the upregulation of those mechanisms involved in the uptake and utilization of alkanes and fatty acids as an additional carbon/energy source. Additionally, the synthesis of compounds associated with reducing oxidative stress — which included an ascorbic acid degradation pathway, never before associated with this infectious pathogen — was also upregulated [42].

As previously mentioned, *M. tuberculosis* drug-resistance develops largely from patient non-adherence, however, new evidence suggests that the causal genetic mutations in *M. tuberculosis* may also result from pharmacokinetic-pharmacodynamic variability associated with the induction of the mycobacterial drug efflux pump [43]. Therefore, although significant progress has been made in the understanding of drug-resistance in *M. tuberculosis* to date, many contributing factors are still largely unknown, with many new discoveries still to be made. In future, this will undoubtedly contribute towards the development of more efficient treatment strategies against this organism.

2.4 Metabolomics investigations aimed towards better monitoring of TB treatment efficiency and predicting treatment outcome

Currently, smear microscopy, is the most widely used method for monitoring TB treatment efficacy. This method for detecting the presence of *M. tuberculosis* in patient collected sputum, functions on the basis of the ability of these acid-fast mycobacteria for retaining a dye subsequent to exposure with an acid-alcohol solution [44]. During a successful TB treatment protocol, a visible decrease in the bacterial load is expected when applying this method to subsequent patient collected sputum samples, and a positive sputum smear during the second month of treatment is indicative of possible treatment failure. However, a period of two months before the first signs of treatment failure is still considered too timely, since the bacteria may have developed drug resistance during this time [45, 46]. Additionally, although this method is fast (requiring less than two hours) and easy to perform [44], the color reaction is based on the characteristic mycolic acids present in the cell walls of mycobacteria, and since these fatty acids persist even when the bacteria die, false positive results are a frequent anomaly [47]. Furthermore, numerous quantitative studies have indicated that this approach is rather insensitive for the detection of acid fast mycobacteria, (requiring 5 000–10 000 bacteria/mL to be present in the patient collected sputum) [48], the latter resulting in a false negative result. Alternative approaches for monitoring or predicting treatment outcome includes monitoring patient mass over the duration of the treatment regimen, and a less than 5% body weight gain after the first month of treatment may also be indicative of a possible negative treatment outcome [45, 49]. Although patient mass monitoring is a simple and inexpensive approach for predicting treatment outcome, it is rather unspecific and influenced by numerous confounders including socioeconomic status, HIV co-infection, nutritional habits and stress, to name but a few. Walzl *et al.* (2008) reviewed a number of novel immunological biomarkers in serum for monitoring TB treatment response and predicting treatment outcome [50]. Although these biomarkers showed promise for monitoring treatment response, none of these have been validated as yet, or even considered for clinical trial evaluation, due to the costs, complexed logistics and special field site characteristics that would be required.

Despite these efforts, a fast, simple and accurate approach for the early detection/prediction of treatment failure, especially in high-risk patients, does as yet not exist, and this is one of the major contributors to the resulting extended therapeutic duration on inefficient TB medication and the development of drug-resistant TB strains [51].

An effective method for accurately monitoring treatment progression and/or accurately predicting treatment outcome or potential relapse, well before the end of the treatment regimen or even before therapy with standard TB drugs commences, would be regarded as a major breakthrough, since alternative treatment approaches can be considered at an earlier stage. Considering this, various metabolomics studies have been done for the purpose of identifying the metabolome variation and associated biomarkers which could be accurately used for this purpose.

Mahapatra *et al.* (2014) used LC-MS metabolomics to identify metabolic variations characterizing treatment outcome using patient-collected urine. By comparing the metabolite profiles of urine collected at time of diagnosis (before treatment) to those collected one month after treatment had commenced, they identified 45 metabolites differentiating these sample groups. After reducing these metabolites to a set of six statistically significant molecular features, they were able to classify the urine as either baseline or one month treatment samples, with an error rate of 11.8% [12]. In future, these metabolites may serve as biomarkers for evaluating drug response, which could potentially accelerate TB drug clinical trials. In the following year, Das *et al.* (2015) investigated the altered urinary metabolite profiles of active TB-positive patients on first-line TB drugs by comparing samples collected at various time points during the treatment regimen, with those collected from healthy controls. A time-dependent treatment trend was observed, and the successfully cured patients were shown to have a metabolic profile similar to that of asymptomatic healthy controls. Although this study indicates the potential use of metabolomics towards broadening our understanding of TB treatment progression, the main aim of the study was to identify diagnostic biomarkers and, therefore, no metabolites explaining this trend were actually identified or discussed [52]. The same group also investigated the individual variation of TB drug metabolism, which can be associated with treatment outcome, drug-related side-effects and the development of drug-resistance. They applied an

untargeted GC-TOFMS metabolomics approach and analyzed urine samples collected from newly diagnosed TB patients at various intervals (i.e. 2h, 6h, 12h, 24h, 36h and 48h) after first-line oral TB drug administration. The majority of the first-line drugs (INH, RIF and PZA) administered were detected, as well their known metabolites. This metabolomics study not only allowed for a comprehensive description of the drug metabolism phenotype associated with first-line TB drugs (by interpretation of the variation in drug abundance and their known metabolites), but also identified a previously unknown drug metabolite of EMB [25]. Using a GCxGC-TOFMS metabolomics approach, Luies and Loots (2017, unpublished results) were able to differentiate the urinary metabolite profiles of active TB-positive patients with a successful and unsuccessful treatment outcome at time of diagnosis (i.e. before treatment onset). The major metabolite variations identified suggested that a gut microbiome imbalance predicts and explains why certain individuals have an unsuccessful treatment outcome [53]. Additionally, these patients presented with elevated levels of those metabolites associated with abnormalities in the long-chain fatty acid β-oxidation pathway, accompanied by reduced L-carnitine and short-chain fatty acids, which is indicative of a possible mitochondrial trifunctional protein (MTP) defect. Abnormalities to the amino acid profile of the unsuccessful treatment outcome group also indicated compromised insulin secretion [54]. In a similar study, Tientcheu *et al.* (2015) compared the transcriptomic and metabolic profiles of patients with either a *M. africanum* or *M. tuberculosis*-infection after receiving the same TB treatment, in order to identify host variations associated with lineage-specific pathogenesis and treatment response. The serum metabolic profiles showed a decline in pro-inflammatory metabolites after treatment, which was more pronounced in the *M. tuberculosis*-infected patients, comparatively. They concluded that differences in the transcriptomic and serum metabolic profiles associated with these two lineages may be an indication of intrinsic host factors associated with susceptibility to TB and/or efficacy of standard TB treatment [55].

Over the past few years, progress has also been made in the development and application of electronic nose (sensor array) technologies toward metabolomics research. An electronic nose contains conducting-polymer sensor arrays that interact with volatile compounds, resulting in electrical

resistance producing characteristic signals of multiple sensor parameters [56, 57]. In 2016, Zetola *et al.* applied this technology towards the development of improved TB diagnostics and monitoring TB treatment efficacy. Breath samples were collected from healthy, HIV-uninfected controls and pulmonary TB-positive patients, of which 61% were co-infected with HIV, at time of diagnosis and at various intervals during the treatment regimen. The electronic nose signals obtained from the volatile organic compounds in these breath samples indicated that the device had a sensitivity of 94.1% and a specificity of 90.0% for differentiating TB-positive patients (before treatment onset) and healthy controls. Changes in these breath signals were also observed longitudinally (at 0, 2, 7, 14, and 30 days following TB treatment onset), indicating its potential application for monitoring treatment progression [58].

Considering the above, metabolomics has proved to be a useful tool for identify metabolite variations characterizing treatment efficiency and outcome. These variations could, in future, be implemented towards the development of a targeted diagnostic/prognostic approach, by specifically searching for characteristic metabolite signatures using MS based analytical equipment. The application of such a method to patient collected urine for example, would be considered advantageous since large volumes are readily available via non-invasive sampling, and minimal sample preparation is required, reducing the analysis time to merely a few minutes. Furthermore, the sensitivity/or limits of detection using such approaches is as little as 250 bacterial cells per sample [59], which is a vast improvement on the detection limits and poor sensitivities of currently used methods as describe above. A possible disadvantage of a metabolomics based approach however, would be the high cost of the analytical equipment used. However, the idea behind biomarker identification using these specialized analytical techniques are for the purpose of developing inexpensive, easy-to-use, color reaction kits, using nanotechnology, for example, for later clinical applications.

2.5 The application of metabolomics in new anti-tuberculosis drug clinical trials

The identification of new drug related biomarkers is an important facet of new drug discovery, since these can potentially explain drug targets and mechanisms, drug-drug interactions, drug-related

immune reactions, dose-response relationships, and toxicity, well before the resulting changes to the phenotype occurs [60]. When considering clinical drug trials, these biomarkers (including those identified via metabolomics studies) can be applied to (a) identify and monitor drug delivery to targets, (b) elucidate and predict the associated pathophysiological mechanisms of disease and drug interactions, and (c) determine and monitor clinical effects. The prioritization of rational drug discovery resources can be facilitated by evaluating biomarkers in these three ways early in the drug development process, which enables early proof-of-concept studies for novel therapeutic targets [60].

To this end, metabolic profiling (metabolomics) is an efficient tool to identify biomarkers for elucidating both the *in vitro* and *in vivo* activities of metabolic drug targets, and also to identify the metabolic pathways influenced and their associated enzymes, revealing additional drug targets [61]. As reviewed by Prosser *et al*. (2014) and Jansen and Rhee (2017), metabolomics can also be used towards the discovery of novel metabolic functions/pathways, and additionally the identification of previously unknown drug enzyme targets and the influenced metabolic pathways [15, 61]. Such metabolomics studies typically compare the metabolomes of a patient group receiving the drug from those receiving a placebo. The detected altered metabolite profile, when interpreted in the context of known metabolism and metabolic defects, can be used to predict drug action and possible toxicities. Alternatively, one can use metabolomics-based methods to track drug-specific metabolites and their associated metabolism by using a stable isotope-labeled xenobiotic treatment group in comparison to an unlabeled group [4, 6].

The ideal TB drug, or combinations thereof, should be fast acting, effective against drug-susceptible and drug-resistant strains, and have anti-mycobacterial activity against both replicating and non-replicating mycobacteria populations, with little or no side-effects. Since no such drug has been discovered or developed for TB to date, new compounds for this purpose are constantly being investigated. Currently, the efficacy of new candidate TB drugs are evaluated by determining the degree of growth inhibition of *M. tuberculosis* in lung homogenates, a process requiring approximately three weeks of incubation [11]. Although qualitative analysis of tuberculostearic acid (TBSA) in sputum samples has been investigated for diagnosing pulmonary TB, no attempts for using

TBSA as a quantitative biomarker for bacterial growth have been reported. To this end, Cai *et al.* (2013) used a GC-MS/MS approach to quantify the derivatized methyl ester of TBSA produced during *M. tuberculosis* growth in (a) an axenic medium, (b) macrophage cell (J774) cultures, and (c) the lungs of interferon-gamma knockout mice, with and without exposure to various TB drugs. In both the macrophage and knockout mice experiments, TBSA could be positively associated with the number of colony forming units in the *M. tuberculosis* cultures. It was also evident that TBSA persists once synthesized and does not accumulate further after the inhibition of bacterial replication, thereby validating the use of TBSA as a quantitative biomarker for *M. tuberculosis* growth. However, the poor stability of TBSA in non-viable bacteria limits its use to drug studies measuring growth inhibition. Thus, the authors suggested using TBSA as a bacteriostatic biomarker, and not as a bactericidal biomarker. This group also investigated the use of hexacosanoic acid as a secondary biomarker to TBSA, for detecting *M. tuberculosis* in sputum samples. They reported a similar response to TB drugs as that seen with TBSA; however, hexacosanoic acid concentrations were approximately 100-fold lower and provided no additional advantages [11]. This study proved the potential of using a single *M. tuberculosis* metabolite for monitoring the efficacy of TB drugs. However, complete metabolite biosignatures, from data obtained using one or more analytical platform(s), are typically preferred for such applications due to the comprehensive amounts of data generated, which in turn can be used to not only evaluate drug efficacy, but also drug mechanisms [62].

3. Future perspectives of metabolomics for the development of improved anti-tuberculosis treatment strategies

As discussed above, metabolomics can be considered a powerful tool for identifying previously unknown metabolite variations characterizing a xenobiotic perturbation, which in turn can be used for evaluating drug mechanisms, drug-related side-effects, drug-resistance, treatment progression and disease relapse, and also prompt assessment of the efficacy and safety of new TB drug candidates. Although the rapid development of new TB drug candidates will undoubtedly bring about numerous

advantages to the general TB population and society as a whole, it is important to realize that individual variation in drug response, as is the case for many other xenobiotics for that matter, also exists. Many studies have indicated that inter-individual variation in the healthy human metabolome can directly be associated with an individual's genetic makeup, age, gender and race, as well as to various environmental factors such as an individual's social habits, diet, stress levels, and lifestyle [63]. By including these known confounders into a metabolomics experiment, the metabolite variations characterizing and explaining individual variation and the associated drug-related side-effects or a lack of treatment response, can be identified and better explained, allowing for personalized treatment approaches [64]. This in turn could also lead to a shorter treatment duration for first-line TB drugs (currently six months), further lowering the risk for developing late onset drug side-effects. These metabolite variations could additionally be used diagnostically for predicting treatment outcome/response, perhaps even before treatment onset [7, 65], and also identify those individuals who would potentially relapse. In practice, however, the identification of such metabolite variations would require very large sample cohorts, including patient and control groups, representing all possible environmental and genetic variation. Achieving this however, would most likely require a global initiative to firstly standardize all metabolomics procedures used for biomarker identification, including all analytical approaches (such as extraction methods, quantification procedures and the quality control samples used), and statistical analysis, which may affect results. The metabolomics data generated by all platforms globally should then be made available publically, enabling researchers to validate biomarkers identified by different research groups using diverse sample cohorts. This is something the international metabolomics community is slowly adopting, through the initiatives of the "International Metabolomics Society".

4. Concluding remarks

Although the use of metabolomics towards the identification of biomarkers aimed at improved TB treatment strategies could be considered a relatively new research domain, significant progress has been made towards the discovery of numerous metabolites better explaining previously unknown

mechanisms of drug action, toxicity, treatment progression, relapse and drug-resistance. The future identification and validation of these biomarkers may pave the way toward more effective, less toxic and less expensive TB drugs, which could be administered for shorter treatment periods, with a reduced incidence of an unsuccessful treatment outcome and/or relapse.

5. Executive summary

- Abstract: Current TB treatment strategies need to be improved due to the many limitations associated with these, including the long duration of therapy, severe side-effects associated with TB drugs, occurrence of treatment failure (due to patient genetics or drug-resistance), post-treatment disease relapse, and human immunodeficiency virus (HIV) co-infection.
- Conclusion: Using a metabolomics research approach, numerous new metabolite biomarkers have been identified for better explaining previously unknown mechanism of drug action, toxicity, treatment progression, relapse and drug-resistance.
- Conclusion: The validation of these biomarkers could potentially contribute to the development of new, more effective, less toxic and less expensive TB drugs, leading to shorter treatment duration and fewer drug-resistant cases.

6. Financial disclosure/Acknowledgements

The authors declare that there are no conflicts of interest, and that this manuscript, and the work described therein, is unpublished and has not been submitted for publication elsewhere.

The financial assistance of the National Research Foundation (NRF) of South Africa for this research is gratefully acknowledged (UID: 95269). The opinions expressed and conclusions derived are those of the authors and are not necessarily those of the NRF.

7. References

1. World Health Organization. Global tuberculosis report 2015. (2015). http://apps.who.int/iris/bitstream/10665/191102/1/9789241565059_eng.pdf

2. Olivier I, Loots DT. An overview of tuberculosis treatments and diagnostics. What role could metabolomics play. *Cell Tissue Res.* 11(1), 2655-2671 (2011).

3. World Health Organization. The global plan to stop TB 2011-2015: Transforming the fight towards elimination of tuberculosis. (2010). http://www.stoptb.org/assets/documents/global/plan/tb_globalplantostoptb2011-2015.pdf

4. De Villiers L, Loots DT. Using metabolomics for elucidating the mechanisms related to tuberculosis treatment failure. *Curr. Metabolomics* 1(4), 306-317 (2013).

5. World Health Organization. Global tuberculosis report 2016. (2016). http://apps.who.int/iris/bitstream/10665/250441/1/9789241565394-eng.pdf?ua=1

6. Chen C, Gonzalez FJ, Idle JR. LC-MS-based metabolomics in drug metabolism. *Drug Metab. Rev.* 39(2-3), 581-597 (2007).

7. Kaddurah-Daouk R, Kristal BS, Weinshilboum RM. Metabolomics: A global biochemical approach to drug response and disease. *Annu.Rev.Pharmacol.Toxicol.* 48 653-683 (2008).

8. Halouska S, Chacon O, Fenton RJ, Zinniel DK, Barletta RG, Powers R. Use of NMR metabolomics to analyze the targets of D-cycloserine in mycobacteria: Role of D-alanine racemase. *J. Proteome Res.* 6(12), 4608-4614 (2007).

9. Loots DT, Wiid IJ, Page BJ, Mienie LJ, Helden PD. Melatonin prevents the free radical and MADD metabolic profiles induced by antituberculosis drugs in an animal model. *J. Pineal Res.* 38(2), 100-106 (2005).

10. Du Preez I, Loots DT. Altered fatty acid metabolism due to rifampicin-resistance conferring mutations in the rpoB gene of *Mycobacterium tuberculosis*: Mapping the potential of pharmaco-metabolomics for global health and personalized medicine. *OMICS* 16(11), 596-603 (2012).

11. Cai G, Pauli GF, Wang Y, Jaki BU, Franzblau SG. Rapid determination of growth inhibition of *Mycobacterium tuberculosis* by GC–MS/MS quantitation of tuberculostearic acid. *Tuberculosis* 93(3), 322-329 (2013).

12. Mahapatra S, Hess AM, Johnson JL *et al*. A metabolic biosignature of early response to anti-tuberculosis treatment. *BMC Infect. Dis.* 14(1), 53 (2014).

13. Van Den Boogaard J, Kibiki GS, Kisanga ER, Boeree MJ, Aarnoutse RE. New drugs against tuberculosis: Problems, progress, and evaluation of agents in clinical development. *Antimicrob. Agents Chemother.* 53(3), 849-862 (2009).

14. Lipworth S, Hammond R, Baron V, Hu Y, Coates A, Gillespie S. Defining dormancy in mycobacterial disease. *Tuberculosis* 99 131-142 (2016).

15. Prosser GA, Larrouy-Maumus G, De Carvalho LPS. Metabolomic strategies for the identification of new enzyme functions and metabolic pathways. *EMBO reports* 15(6), 657-669 (2014).

16. Li F, Miao Y, Zhang L, Neuenswander SA, Douglas JT. Metabolomic analysis reveals novel isoniazid metabolites and hydrazones in human urine. *Drug Metab. Pharmacokinet.* 26(6), 569-576 (2011).

17. Nandakumar M, Nathan C, Rhee KY. Isocitrate lyase mediates broad antibiotic tolerance in *Mycobacterium tuberculosis. Nature communications* 5 (2014).

18. Halouska S, Fenton RJ, Zinniel DK, Marshall DD, Barletta RLG, Powers R. Metabolomics analysis identifies d-Alanine-d-Alanine ligase as the primary lethal target of d-Cycloserine in mycobacteria. *J. Proteome Res.* 13(2), 1065-1076 (2013).

19. Prosser GA, De Carvalho LP. Metabolomics reveal d-alanine: d-alanine ligase as the target of d-cycloserine in *Mycobacterium tuberculosis. ACS Med. Chem. Lett.* 4(12), 1233-1237 (2013).

20. Zumla A, Nahid P, Cole ST. Advances in the development of new tuberculosis drugs and treatment regimens. *Nat. Rev. Drug Discov.* 12(5), 388-404 (2013).

21. De Carvalho LPS, Darby CM, Rhee KY, Nathan C. Nitazoxanide disrupts membrane potential and intrabacterial pH homeostasis of *Mycobacterium tuberculosis. ACS Med. Chem. Lett.* 2(11), 849-854 (2011).

22. Halouska S, Fenton RJ, Barletta RLG, Powers R. Predicting the in vivo mechanism of action for drug leads using NMR metabolomics. *ACS Chem. Biol.* 7(1), 166-171 (2012).

23. Mahapatra S, Woolhiser LK, Lenaerts AJ *et al*. A novel metabolite of antituberculosis therapy demonstrates host activation of isoniazid and formation of the isoniazid-NAD+ adduct. *Antimicrob. Agents Chemother.* 56(1), 28-35 (2012).

24. Kim B, Moon J-Y, Choi MH *et al*. Global metabolomics and targeted steroid profiling reveal that rifampin, a strong human PXR activator, alters endogenous urinary steroid markers. *J. Proteome Res.* 12(3), 1359-1368 (2013).

25. Das MK, Arya R, Debnath S *et al*. Global urine metabolomics in patients treated with first-line tuberculosis drugs and identification of a novel metabolite of ethambutol. *Antimicrob. Agents Chemother.* 60(4), 2257-2264 (2016).

26. Hopkins AL. Network pharmacology: The next paradigm in drug discovery. *Nat. Chem. Biol.* 4(11), 682-690 (2008).

27. Bando K, Kunimatsu T, Sakai J *et al*. GC-MS-based metabolomics reveals mechanism of action for hydrazine induced hepatotoxicity in rats. *Journal of Applied Toxicology* 31(6), 524-535 (2011).

28. Rawat A, Chaturvedi S, Singh A *et al*. Metabolomics approach discriminates toxicity index of pyrazinamide and its metabolic products, pyrazinoic acid and 5-hydroxy pyrazinoic acid. *Hum. Exp. Toxicol.* 0960327117705426 (2017).

29. Ozkan E, Yaman H, Cakir E *et al*. Plasma melatonin and urinary 6-hydroxymelatonin levels in patients with pulmonary tuberculosis. *Inflammation* 35(4), 1429-1434 (2012).

30. World Health Organization. Mode of travel: Health considerations. *International Travel and Health* (2010).

31. Sumner SJ, Burgess JP, Snyder RW, Popp JA, Fennell TR. Metabolomics of urine for the assessment of microvesicular lipid accumulation in the liver following isoniazid exposure. *Metabolomics* 6(2), 238-249 (2010).

32. Cunningham K, Claus SP, Lindon JC *et al*. Pharmacometabonomic characterization of xenobiotic and endogenous metabolic phenotypes that account for inter-individual variation in isoniazid-induced toxicological response. *J. Proteome Res.* 11(9), 4630-4642 (2012).

33. Cheng J, Krausz KW, Li F, Ma X, Gonzalez FJ. CYP2E1-dependent elevation of serum cholesterol, triglycerides, and hepatic bile acids by isoniazid. *Toxicol. Appl. Pharmacol.* 266(2), 245-253 (2013).

34. Church RJ, Wu H, Mosedale M *et al*. A systems biology approach utilizing a mouse diversity panel identifies genetic differences influencing isoniazid-induced microvesicular steatosis. *Toxicol. Sci.* 140(2), 481-492 (2014).

35. Nicholls AW, Holmes E, Lindon JC *et al*. Metabonomic investigations into hydrazine toxicity in the rat. *Chem. Res. Toxicol.* 14(8), 975-987 (2001).

36. Bollard ME, Keun HC, Beckonert O *et al*. Comparative metabonomics of differential hydrazine toxicity in the rat and mouse. *Toxicol. Appl. Pharmacol.* 204(2), 135-151 (2005).

37. Garrod S, Bollard ME, Nicholls AW *et al*. Integrated metabonomic analysis of the multiorgan effects of hydrazine toxicity in the rat. *Chem. Res. Toxicol.* 18(2), 115-122 (2005).

38. Zhao H, Si Z-H, Li M-H *et al*. Pyrazinamide-induced hepatotoxicity and gender differences in rats as revealed by a 1 H NMR based metabolomics approach. *Toxicology Research* 6(1), 17-29 (2017).

39. Zhang Y, Yew W. Mechanisms of drug resistance in *Mycobacterium tuberculosis*: Update 2015. *The International Journal of Tuberculosis and Lung Disease* 19(11), 1276-1289 (2015).

40. Loots DT. New insights into the survival mechanisms of rifampicin-resistant *Mycobacterium tuberculosis*. *J. Antimicrob. Chemother.* 71(3), 655-660 (2015).

41. Lahiri N, Shah RR, Layre E *et al*. Rifampin Resistance Mutations Are Associated with Broad Chemical Remodeling of Mycobacterium tuberculosis. *J. Biol. Chem.* 291(27), 14248-14256 (2016).

42. Loots DT. An altered *Mycobacterium tuberculosis* metabolome induced by katG mutations resulting in isoniazid resistance. *Antimicrob. Agents Chemother.* 58(4), 2144-2149 (2014).

43. Pasipanodya JG, Gumbo T. A new evolutionary and pharmacokinetic–pharmacodynamic scenario for rapid emergence of resistance to single and multiple anti-tuberculosis drugs. *Curr. Opin. Pharmacol.* 11(5), 457-463 (2011).

44. American Thoracic Society and the Centers for Disease Control and Prevention. Diagnostic Standards and Classification of Tuberculosis in Adults and Children. *Am. J. Respir. Crit. Care Med.* 161(4), 1376-1395 (1999).

45. Bernabe-Ortiz A, Carcamo CP, Sanchez JF, Rios J. Weight variation over time and its association with tuberculosis treatment outcome: a longitudinal analysis. *PLoS One* 6(4), e18474 (2011).

46. González-Martín J, García-García JM, Anibarro L *et al*. Consensus document on the diagnosis, treatment and prevention of tuberculosis. *Archivos de Bronconeumología (English Edition)* 46(5), 255-274 (2010).

47. Ruiz-Manzano J, Blanquer R, Calpe JL *et al*. Diagnosis and treatment of tuberculosis. *Archivos de Bronconeumología ((English Edition))* 44(10), 551-566 (2008).

48. Kox L. Tests for detection and identification of mycobacteria. How should they be used? *Respir. Med.* 89(6), 399-408 (1995).

49. Krapp F, Veliz J, Cornejo E, Gotuzzo E, Seas C. Bodyweight gain to predict treatment outcome in patients with pulmonary tuberculosis in Peru. *The international journal of tuberculosis and lung disease* 12(10), 1153-1159 (2008).

50. Walzl G, Ronacher K, Siawaya JFD, Dockrell HM. Biomarkers for TB treatment response: Challenges and future strategies. *J. Infect.* 57(2), 103-109 (2008).

51. Menzies D, Benedetti A, Paydar A *et al*. Effect of duration and intermittency of rifampin on tuberculosis treatment outcomes: A systematic review and meta-analysis. *PLoS Med.* 6(9), e1000146 (2009).

52. Das MK, Bishwal SC, Das A *et al*. Deregulated tyrosine–phenylalanine metabolism in pulmonary tuberculosis patients. *J. Proteome Res.* 14(4), 1947-1956 (2015).

53. Luies L, Ronacher K, Walzl G, Loots D. Predicting tuberculosis treatment outcome using metabolomics (Unpublished manuscript). (2017).

54. Luies L, Mienie J, Motshwane C, Ronacher K, Walzl G, Loots D. Urinary metabolite markers characterizing tuberculosis treatment failure (Unpublished manuscript). (2017).

55. Tientcheu LD, Maertzdorf J, Weiner J *et al*. Differential transcriptomic and metabolic profiles of *M. africanum*-and *M. tuberculosis*-infected patients after, but not before, drug treatment. *Genes Immun.* 16(5), 347-355 (2015).

56. Pavlou AK, Magan N, Jones JM, Brown J, Klatser P, Turner APF. Detection of *Mycobacterium tuberculosis* (TB) in vitro and in situ using an electronic nose in combination with a neural network system. *Biosens. Bioelectron.* 20(3), 538-544 (2004).

57. Wilson AD. Biomarker metabolite signatures pave the way for electronic-nose applications in early clinical disease diagnoses. *Curr. Metabolomics* 4(2), 12 (2016).

58. Zetola NM, Modongo C, Matsiri O *et al*. Diagnosis of pulmonary tuberculosis and assessment of treatment response through analyses of volatile compound patterns in exhaled breath samples. *J. Infect.* (2016).

59. Du Preez I, Loots DT. Detection limit for differentiating between various *Mycobacterium* species and *Pseudomonas aeruginosa* using gas chromatography-mass spectrometry (GC-MS) metabolomics: A comparison of two extraction methods. *Afr. J. Microbiol. Res.* 7(9), 797-801 (2013).

60. Parida SK, Kaufmann SHE. The quest for biomarkers in tuberculosis. *Drug Discov. Today* 15(3), 148-157 (2010).

61. Jansen RS, Rhee KY. Emerging Approaches to Tuberculosis Drug Development: At Home in the Metabolome. *Trends Pharmacol. Sci.* (2017).

62. Jacobsen M, Mattow J, Repsilber D, Kaufmann SH. Novel strategies to identify biomarkers in tuberculosis. *Biol. Chem.* 389(5), 487-495 (2008).

63. Koen N, Du Preez I. Chapter Three: Metabolomics and personalized medicine. *Adv. Protein Chem. Struct. Biol.* 102 53-78 (2016).

64. Kumar B, Prakash A, Ruhela RK, Medhi B. Potential of metabolomics in preclinical and clinical drug development. *Pharmacol. Rep.* 66(6), 956-963 (2014).

65. Kaddurah-Daouk R, Weinshilboum RM. Pharmacometabolomics: implications for clinical pharmacology and systems pharmacology. *Clin. Pharmacol. Ther.* 95(2), 154-167 (2014).

Papers of special note have been highlighted as: * of interest or ** of considerable interest.

*Halouska et al., 2007

In this study, nuclear magnetic resonance (NMR) metabolomics was used to investigate the mechanism of action of D-cycloserine (DCS), a drug currently used to treat multidrug-resistant TB (MDR-TB). The results showed that that D-alanine-D-alanine ligase (Ddl) is the *in vivo* lethal target of DCS.

*Loots et al., 2005

Here, metabolomics was used to indicate that a combined TB drug, Rifater, induced a metabolic profile resembling that of multiple acyl-CoA dehydrogenase deficiency (MADD) in Sprague-Dawley rats, thereby explaining a number of drug-related side-effects, which could be alleviated by the co-administration of melatonin.

*Du Preez and Loots, 2012

The groups applied GC-MS metabolomics to link rifampicin resistance in *M. tuberculosis* to an altered fatty acid metabolism.

**Luies and Loots, 2017a

In this study, GCxGC-TOFMS acquired urinary metabolome profiles were used to describe novel mechanisms related to TB treatment failure, even before treatment onset, unlocking the possibility of developing an early prognostic method for TB treatment outcome.

* Tientcheu et al., 2015

The group compared the serum transcriptomic and metabolic profiles of *M. africanum*- to that of *M. tuberculosis*-infected patients after receiving the same TB treatment, and identified host biomarkers associated with lineages-specific pathogenesis and treatment response, which may be an indication of intrinsic host factors associated with susceptibility to TB and/or efficacy of standard TB treatment.

Table 1: A summary of various metabolomics applications investigating mechanisms of TB drug action and drug metabolism.

Drug/drug metabolite	Research model	Number of samples	Sample matrix	Analytical apparatus	Reference
DCS	*M. smegmatis*:	-	Cultures	^{1}H NMR	[8]
	Wild-type	12			
	TAM23	12			
	TAM23 pTAMU3	12			
	GPM16	12			
	GPM14	12			
	Wild-type with DCS	12			
	TAM23 with DCS	12			
	TAM23 pTAMU3 with DCS	12			
	GPM16 with DCS	12			
	GPM14 with DCS	12			
Controls	*M. smegmatis*	40	Cultures	^{1}H NMR	[22]
Amiodarone		10			
Amiodarone		10			
Ampicillin		10			
Chloramphenicol		10			
Chlorprothixene		10			
Ciproflaxacin		10			
DCS		10			
EMB		10			
Ethionamide		10			
Gentamicin		10			
INH		10			
Kanamycin		10			
RIF		10			
Streptomycin		10			
Vancomycin		10			
DCS	*M. smegmatis*	6	Cultures	^{1}H NMR	[18]
	M. tuberculosis	6			
DCS	*M. tuberculosis*	-	Cultures	LC-MS	[19]
INH	TB+ human patients	8	Urine	LC-MS	[23]
	TB+ mice	8	Urine		
	M. tuberculosis	-	Cultures		
Controls	*M. tuberculosis*	-	Cultures	LC-MS	[17]
INH					
RIF					
Streptomycin					
INH	Healthy humans	6	Urine	UPLC-TOFMS	[16]
RIF	Healthy humans	12	Urine	UPLC-QTOF	[24]
Controls	Healthy humans	10	Urine	GC-TOFMS	[25]
DOTS	TB+ humans	20			
EMB	Healthy humans	2			

Abbreviations: DCS, D-cycloserine; TAM23, D-alanine racemase null mutant; TAM23 pTAMU3, TAM23 complemented with wild-type alr gene; GPM16, DCS-resistant mutant; GPM14, D-alanine racemase overproducing mutant; LCS, L-Cycloserine; ^{1}H NMR, nuclear magnetic resonance; LC-MS, liquid chromatography mass spectrometry; UPLC, ultra-performance liquid chromatography; TOFMS, time-of-flight mass spectrometry; QTOF, quadrupole time-of-flight; GC, gas chromatography; TB+, tuberculosis positive; INH, isoniazid; RIF, rifampicin; EMB, ethambutol; DOTS, directly observed therapy short-course; TB+, tuberculosis positive.

Table 2: A summary of various metabolomics applications investigating TB drug side-effects.

Drug/drug metabolite	Side-effect	Research model	Number of samples	Sample matrix	Analytical apparatus	Reference
INH	Hepatotoxicity (Hepatic MVLA)	Rats	12	Urine	^{1}H NMR	[31]
INH	Neurotoxicity	Rats	30	Urine	^{1}H NMR	[32]
INH	Hepatotoxicity	Wild-type mice Cyp2e1-null mice	7 7	Urine	LC-MS	[33]
INH	Hepatotoxicity (Steatosis)	Mice (34 strains)	272	Liver	^{1}H NMR	[34]
Hydrazine (INH metabolite)	Hepatotoxicity (overall metabolic effect)	Rats	20	Urine Plasma	^{1}H NMR	[35]
Hydrazine (INH metabolite)	Hepatotoxicity (overall metabolic effect)	Rats Mice	30 24	Urine	^{1}H NMR	[36]
Hydrazine (INH metabolite)	Hepatotoxicity Neurotoxicity	Rats	18	Kidney Liver Brain	^{1}H NMR	[37]
Hydrazine (INH metabolite)	Hepatotoxicity	Rats	24	Urine Plasma Liver	GC-MS	[27]
PZA	Hepatotoxicity	Rats	24	Serum	^{1}H NMR	[28]
PZA	Hepatotoxicity	Rats (Female) Rats (Male)	30 30	Liver Serum	^{1}H NMR	[38]
Rifater (combination therapy)	Oxidative stress MADD profile	Rats	6	Urine	GC-MS	[9]

Abbreviations: INH, isoniazid; PZA, pyrazinamide; MVLA, microvesicular lipid accumulation; LC-MS, liquid chromatography mass spectrometry; ^{1}H NMR, nuclear magnetic resonance; GC-MS, gas chromatography mass spectrometry; MADD, multiple acyl-CoA dehydrogenase deficiency; TB+, tuberculosis positive; HPLC, high performance liquid chromatography; ELISA, enzyme-linked immunosorbent assay.

Table 3: A summary of various metabolomics applications investigating *M. tuberculosis* drug resistance mechanism to TB drugs.

Drug/drug metabolite	Resistance conferring mutation	*M. tuberculosis* mutation	Analytical apparatus	Ref.
INH	*katG*	Deletion of codon 315 W321F	GCxGC-TOFMS	[42]
RIF	*rpoB*	S531L S522L	GC-MS	[10]
RIF	*rpoB*	S531L S522L	GCxGC-TOFMS	[40]
RIF	*rpoB*	H526Y S531L Q513E	UHPLC-ESI-Q-TOFMS	[41]

Abbreviations: INH, isoniazid; RIF, rifampicin; GCxGC-TOFMS, two dimensional gas chromatography time-of-flight mass spectrometry; GC-MS, *gas chromatography mass spectrometry; UHPLC-ESI-Q-TOFMS, ultra-high performance liquid chromatography–electrospray ionization–quadruple time-of-flight mass spectrometry.*

Metabolomics (2016)12:40
DOI 10.1007/s11306-016-0969-x

ORIGINAL ARTICLE

Tuberculosis metabolomics reveals adaptations of man and microbe in order to outcompete and survive

Laneke Luier[1] · **Du Toit Loots**[1]

Received: 7 September 2015 / Accepted: 10 December 2015

Abstract Despite numerous research efforts to control tuberculosis, it is still regarded as a global pandemic. It is clear that the infectious agent responsible and its associated disease mechanisms remain poorly understood. Alternative research strategies are therefore urgently needed to better characterize active-TB, especially the adaptations of the host and microbe as they compete to survive. Using a GCxGC-TOFMS metabolomics approach, we identified new urinary TB metabolite markers induced by adaptations of the host metabolome and/or host-pathogen interactions. The most significant of these were the TB-induced changes resulting in abnormal host fatty acid and amino acid metabolism, in particular to tryptophan, phenylalanine and tyrosine, inducing a metabolite profile similar to that of patients suffering from phenylketonuria, mediated through changes to INF-γ and possibly insulin. This subsequently also explains some of the symptoms associated with TB and provides clues to better treatment approaches.

Keywords Tuberculosis · Urinary metabolomics · Metabolite markers · Host-pathogen · Adaptations · Interactions

Electronic supplementary material The online version of this article (doi:10.1007/s11306-016-0969-x) contains supplementary material, which is available to authorized users.

✉ Du Toit Loots
dutoit.loots@nwu.ac.za

Laneke Luier
laneke.devilliers@gmail.com

[1] Human Metabolomics, School for Physical and Chemical Sciences, North-West University (Potchefstroom Campus), Private Bag X6001, Box 269, Potchefstroom 2531, South Africa

1 Introduction

Tuberculosis (TB), a highly contagious bacterial disease caused by *Mycobacterium tuberculosis* (*M. tuberculosis*), is considered the leading cause of death globally from a single bacterial pathogen. This disease affects approximately one-third of the world's population, either in its active or latent form. The latest reports indicate 9 million new TB cases per annum, resulting in an estimated 1.5 million deaths (3800 deaths a day) (World Health Organization, 2014).

TB transmission occurs through small infectious aerosol droplets that contain live *M. tuberculosis* bacilli. When inhaled into the lungs, they undergo rapid replication, before being engulfed by host macrophages. Following macrophage internalization, *M. tuberculosis* avoids death by blocking phagolysosomal fusion, allowing the bacilli to persist in a non- or slowly-replicating state within the relatively hospitable niche of these phagosomes, where they can survive for decades (Philips and Ernst 2012, Warner 2014). Thus, although the host immune response is unable to eradicate *M. tuberculosis*, a fully immune-competent host can suppress the infection. This asymptomatic and non-infectious state is referred to as latent TB. However, when the immune system becomes compromized, latent TB develops into active TB—the symptomatic and highly infectious state of the disease. Common symptoms associated with active pulmonary TB include a bad cough with mucus, pleurisy, haemoptysis, dyspnoea, wheezing, weakness/fatigue, weight-loss, loss of appetite, chills/fever, and night sweats (Long et al. 2002).

Despite numerous research efforts to date, since the discovery of *M. tuberculosis* in 1882, it is clear that this pathogen and its underlying disease mechanisms remain poorly understood. Thus, alternative research approaches,

Published online: 03 February 2016

giving a different perspective of the disease, are urgently needed to better characterize TB, especially those investigating host-microbe interactions and each of their adaptations, as they compete with each other in order to survive. To this end, metabolomics has the potential to identify new disease markers, describing previously unidentified disease mechanisms, adaptations of the pathogen to environmental changes and how these react to antimicrobial treatments (Drapal et al. 2014). 'Metabolomics' refers to the unbiased identification and quantification of all metabolites present in a biological system, utilizing highly sensitive and selective analytical methods, in conjunction with biostatistics for metabolite marker selection. This 'omics' approach allows for a holistic view of a patient's metabolic state/metabolome, detecting alterations due to a specific perturbation, such as TB (De Villiers and Loots 2013). Using this approach on sputum collected from TB patients, Du Preez and Loots (2013) were able to identify new metabolites from both *M. tuberculosis* and the infected host, thereby revealing previously unknown metabolic pathways and adaptations of each. These included the: (a) presence of a citramalate cycle in *M. tuberculosis*; (b) interaction of this cycle with an upregulated glyoxylate cycle; and (c) increased utilization of fatty acids and glutamate as additional carbon sources by *M. tuberculosis* during pulmonary infection. Adaptations of the human host included: (a) an alternative mechanism for hydrogen peroxide (H_2O_2) synthesis via glucose oxidation, in order to counteract the bacterial infection more efficiently; (b) subsequent inhibition of oxidative phosphorylation due to pronounced oxidative stress; and (c) elevated concentrations of various neurotransmitters and other metabolites related to some of the symptoms associated with TB (Du Preez and Loots 2013).

In the current investigation, we applied a two-dimensional gas chromatography time-of-flight mass spectrometry (GCxGC-TOFMS) metabolomics approach, to compare and differentiate culture-confirmed active TB-positive ($n = 46$) and TB-negative healthy control ($n = 30$) groups, based on the comparative detected changes to their urinary metabolomes. Apart from the fact that urine from TB patients is readily available, easily stored and can be collected by non-invasive techniques, it serves as an ideal matrix for identifying metabolites related to an altered TB host metabolome, since it contains metabolic signatures of many pathways that may change as a result of infection or disease.

2 Materials and methods

A summary of the experimental design is given in Fig. 1, and described in detail below.

2.1 Urine sample collection and storage

Urine samples were collected in standard vials, from 30 TB-negative healthy controls and 46 culture-confirmed active TB-positive patients from the same geographical area, prior to any treatment, after which they were placed on ice, immediately transported to the lab, and frozen at −20 °C until analysis.

2.2 Reagents and chemicals

Ethyl acetate and diethyl ether were from Honeywell International Inc. (Muskegon, MI, USA). These organic solvents were ultra-pure Burdick and Jackson brands. Hydrochloric acid (HCl), pyridine, anhydrous sodium sulphate (Na_2SO_4) and trimethylchlorosilane (TMCS) were from Merck (Darmstadt, Germany). 3-Phenylbutyric acid, hexane and bis(trimethylsilyl)-trifluoroacetamide (BSTFA) were from Sigma-Aldrich (St. Louis, MO, USA).

2.3 Organic acid extraction procedure and derivatization

Creatinine values for all urine samples were determined prior to organic acid extraction, using a creatinine enzyme kit (Thermo Scientific; reference number 981845) and subsequent analysis on an Indiko Clinical Analyser, Type 863 (Thermo Scientific). These creatinine values are used to normalize metabolite concentrations, as well as to determine the volume of urine, internal standard, BSTFA, TMCS and pyridine used for each extraction. 5 N HCl (six drops) was added to the determined amount of urine, followed by the addition of the internal standard, 3-phenylbutyric acid (25 µmol/mg creatinine). Thereafter, ethyl acetate (6 mL) was added, followed by mixing and centrifugation (3000 rpm for 3 min) to achieve solvent phase separation, after which the organic phase was aspirated into a clean tube. Diethyl ether (3 mL) was then added to the remaining aqueous phase, again followed by mixing, centrifugation and aspirating the organic solvent into the tube containing the previously collected phase. Hereafter, Na_2SO_4 (approximately 3 g) was added to remove any water still present in the sample, followed by a brief vortex and centrifugation. The organic phase was then decanted from the pellet and evaporated to complete dryness under a stream of nitrogen at 37 °C. This was followed by a 60 min derivatization period at 60 °C with BSTFA (22.6 µL/µmol creatinine), TMCS (4.5 µL/µmol creatinine) and pyridine (4.5 µL/µmol creatinine) (Loots et al. 2004).

2.4 GCxGC-TOFMS analyses

Each derivatized sample was transferred to a 1.5 mL GC–MS sample vial, capped and placed into the auto-sampler

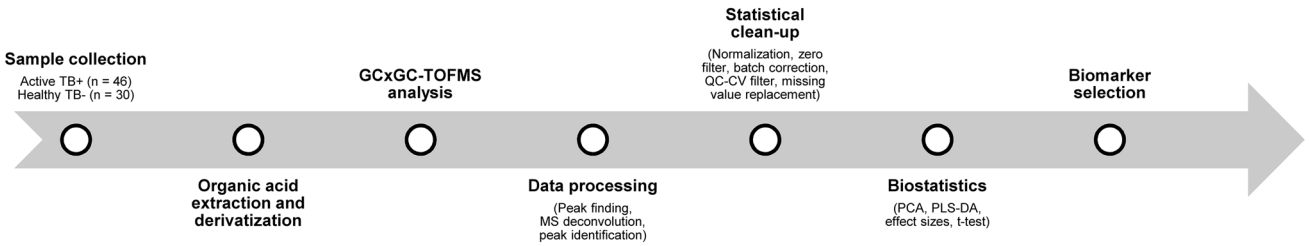


Fig. 1 Summary of the experimental design. Subsequent to sample collection, an organic acid extraction and derivatization of each sample extract was performed, followed by GCxGC-TOFMS analysis, data processing, statistical data clean-up and biostatistics. Hereafter biomarkers were selected on the basis of various predefined statistical "cut offs"

tray of a Gerstel Multi-Purpose Sampler (Gerstel GmbH and Co. KG, Mülheim an der Ruhr, Germany), coupled to a Pegasus 4D GCxGC-TOFMS (LECO Africa (Pty) Ltd, Johannesburg, South Africa), equipped with an Agilent 7890A gas chromatograph (Agilent, Atlanta, USA), and a TOFMS (LECO Africa). All samples (1 µL) were injected in random sequence, along with quality control (QC) samples, at regular intervals throughout the analytical run, using a 1:12 split ratio. Compound separation was achieved using a Restek Rxi-5Sil MS primary capillary column (30 m, 0.25 µm film thickness and 250 µm internal diameter), and a Rxi-17 secondary capillary column (0.9 m, 0.1 µm film thickness and 100 µm internal diameter). Cryomodulation was achieved with a hot pulse of nitrogen gas for 0.7 s, every 3 s. The injector temperature was maintained at a constant 280 °C for the entire run, and ultra-high-purity helium was used as the carrier gas at a constant flow of 1 mL/min. The temperature program for the primary column was 55 °C for 1 min, after which it was increased at a rate of 7 °C/min to a temperature of 285 °C, at which it was maintained for a further 4 min, followed by a temperature increase of 20 °C/min, to a final temperature of 305 °C, where it was maintained for a further 1 min. The secondary oven's temperature ramp was programmed identically to that of the primary column, with a +5 °C deviation at all of the time points. The filament was switched off for the first 8 min of each sample run, as this period was considered to correspond to a solvent delay during which no mass spectra were recorded. However, this 8 min interval was included on the first column's time axis to reflect accurate retention times. The transfer line temperature was held at a constant 270 °C and the ion source temperature at a constant 200 °C for the entire run. The detector voltage was 1600 V and the filament bias −70 eV. Mass spectra were collected over the range 50–600 *m*/*z* at an acquisition rate of 200 spectra/s.

2.5 Data processing

Mass spectral deconvolution, peak alignment and peak identification/annotation was conducted using ChromaTOF software (version 4.32). Mass spectral deconvolution was performed at a signal to noise ratio of 300, with a minimum of 3 apexing peaks, and any possible retention time shifts were corrected across all samples by aligning identical mass spectra displaying similar retention times. Peaks were identified by comparison of their characteristic mass fragmentation patterns and retention times with libraries compiled from previously injected standards.

2.6 Statistical data analyses

Prior to statistical data analysis, a standard metabolomics data clean-up procedure was applied. All compounds were normalized relative to the internal standard by calculating the relative concentration of each. Following this, variables showing no variation between the groups were removed, and a data filter was applied to each variable to eliminate those with more than 50 % zero values in each group (Smuts et al. 2013). Using the QC samples, quantile equating was applied to correct for any batch effects (Wang et al. 2012). Hereafter, a 50 % QC coefficient of variation (CV) filter was applied. Lastly, as most zero values present in the dataset were due to low abundance rather than being absent, they were replaced with a value calculated as half of the smallest value (detection limit) present in the dataset (Schoeman et al. 2012).

Various multivariate (principal components analysis (PCA); partial least squares–discriminant analysis (PLS-DA)) and univariate (effect sizes; unpaired *t* test) biostatistical analyses were applied using MetaboAnalyst, a metabolomics web-based server based on the statistical package "R" (version 2.10.0). PCA is an unsupervised method used to determine whether or not a natural grouping or differentiation exists between sample groups. The variables/metabolites best describing this differentiation are ranked according to their respective modelling powers. PLS-DA, on the other hand, is a supervised method used to determine group membership of an individual sample, and identifies those variables/metabolites that best characterize the differentiated sample groups. Each metabolite is ranked according to the variables' influence on the projection (VIP) parameter, which directly

corresponds to its importance (Du Preez and Loots 2013). Effect sizes are used to compare each variable individually between the groups, without taking sample size into account, and indicates practical/clinical significance. For this parametric dataset, an effect size >0.5 indicates a moderate effect. The unpaired *t*-test establishes statistical significance (*P*-value) by determining whether or not the averages of two groups differ. Conventionally, a *P*-value <0.05 is considered significant (Ellis and Steyn 2003).

3 Results and discussion

Figure 2 shows clear PCA differentiation of the TB-positive and TB-negative healthy control sample groups on the basis of the GCxGC-TOFMS metabolomics data collected. The total amount of variance explained by the first two principal components (PCs) (R_2X cum) was 71.0 %, of which PC1 accounted for 62.5 %, and PC2 accounted for 8.5 %. A PLS-DA model was subsequently built and showed a modelling parameter R^2Y (cum) of 93.6 %, indicative of the total explained variation of the response Y, and a Q^2 (cum), corresponding to the cross-validated variation accounted for by the response Y, of 81.0 %.

As indicated in Fig. 3, 12 compounds with a PCA modelling power >0.5 (Brereton 2003), a PLS-DA VIP > 1.0 (Smuts et al. 2013), an effect size >0.5, and a *t* test with a *P* value <0.05 (Ellis and Steyn 2003), were selected as the metabolite markers that best explained the variation between the analysed groups (listed in Table 1). Box-and-whisker plots of these 12 compounds are provided in the supplementary information. This multi-statistical approach is based on the assumption that the different statistical tests correct/compensate for each other's flaws with the elimination of false-positive compounds (Venter et al. 2015).

Most of the identified metabolites can be explained by changes in the host metabolome in response to TB, supporting previously speculated mechanisms and shedding light on hitherto unknown mechanisms related to *M. tuberculosis* pathogenesis, and host-microbe interactions and adaptations, as will be described below.

M. tuberculosis infection results in a host immune response, signalling macrophages to engulf the invading bacteria. Accompanying this is an inflammatory response resulting in the release of tumor necrosis factor alpha (TNF-α) and interferon gamma (IFN-γ) (Philips and Ernst 2012), the latter of which has been shown to elevate

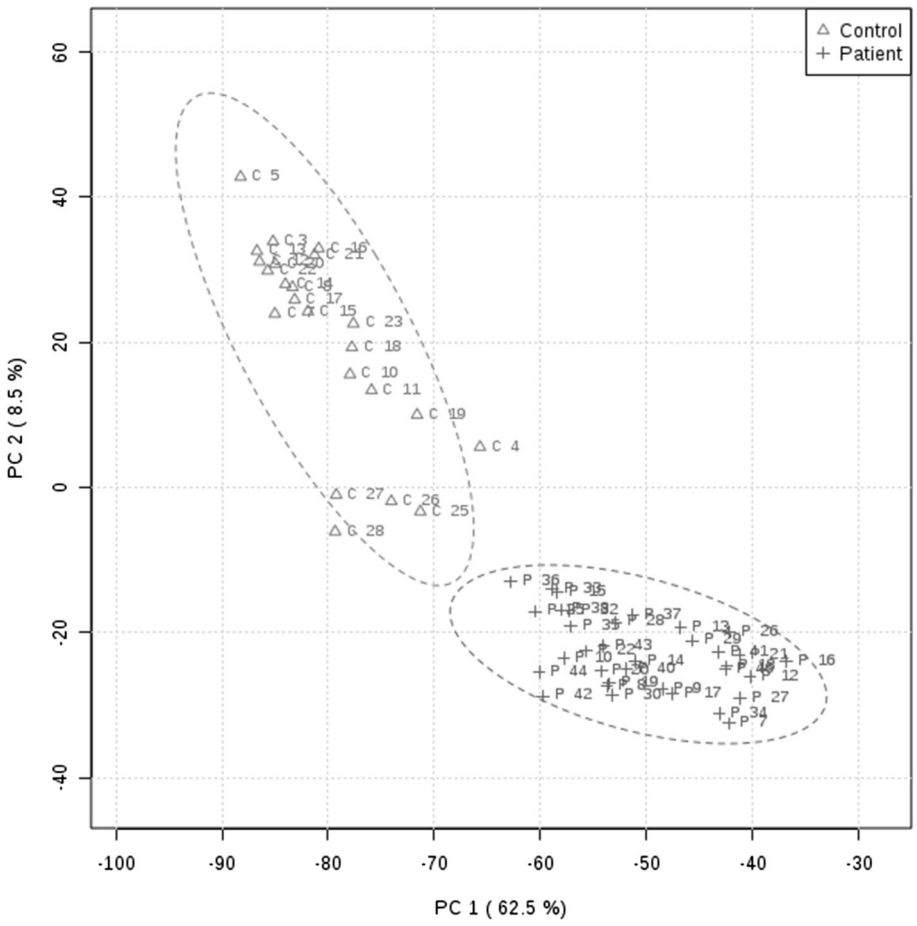


Fig. 2 PCA differentiation of GCxGC-TOFMS organic acid analyzed data, of urine samples collected from active TB-positive and TB-negative patients. A plot of PC1 versus PC2 of the TB-negative healthy control (denoted by *C*) and active TB-positive patient (denoted by *P*) groups, showing clear differentiation of these two groups as a result of variation in their metabolite profiles. The variances accounted for are indicated in *parenthesis*

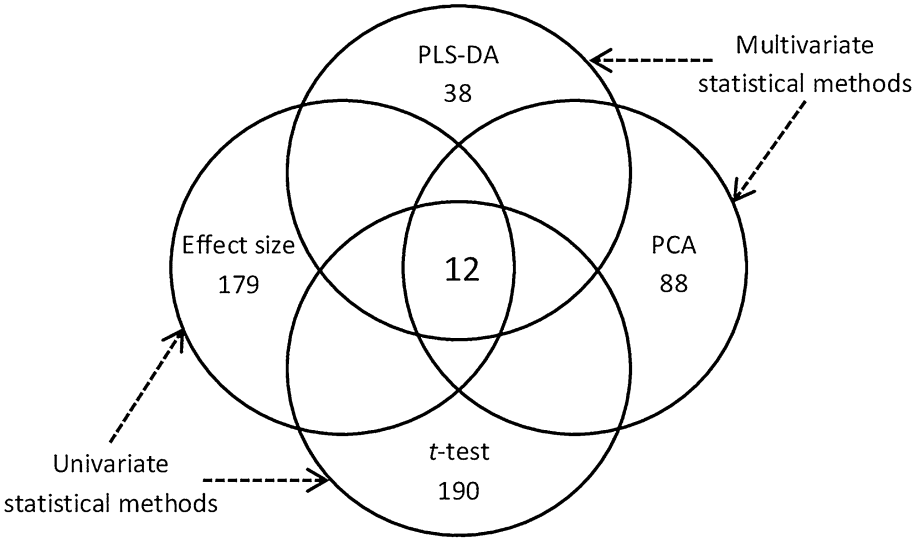


Fig. 3 Venn diagram indicating compound selection using a multi-statistical approach. A total of 12 compounds were identified as most significant for describing the variation between the TB-negative healthy control and TB-positive patient groups

Table 1 The 12 urinary metabolite markers that best explain the variation detected in the urine samples of the TB-negative healthy control and TB-positive patients, ranked according to the retention times

Metabolite name (Chemspider ID)	Retention time (RT) (s)	TB-negative healthy controls: Average concentration (mg/g creatinine) (standard error of the mean)	Active TB-positive: Average concentration (mg/g creatinine) (standard error of the mean)	PCA (Power)	PLS-DA (VIP)	Effect sizes (*d*-value)	*t*-test (*P*-value)
5-Hydroxyhexanoic acid (21230835)	780,0.17	0.045 (0.021)	0.594 (0.127)	0.534	2.312	0.730	<0.001
Phenylacetic acid (10181341)	1107,2.31	0.214 (0.034)	1.137 (0.114)	0.719	1.028	1.372	<0.001
2-Octenoic acid (4445840)	1134,2.35	0.011 (0.007)	0.114 (0.015)	0.568	2.366	1.150	<0.001
2-C-Methylglycerol (19304)	1300,0.23	0.071 (0.020)	0.727 (0.084)	0.666	2.432	1.319	<0.001
5-Hydroxyhydantoin (3369395)	1373,6.24	0.049 (0.019)	0.287 (0.045)	0.522	1.961	0.902	<0.001
Oxalic acid (946)	1416,1.34	0.030 (0.009)	0.143 (0.024)	0.821	2.147	0.813	<0.001
L-Rhamnulose (388459)	1575,0.29	0.028 (0.021)	0.302 (0.041)	0.718	2.838	1.146	<0.001
Quinolinic acid (1037)	1590,1.92	0.615 (0.046)	6.422 (1.042)	0.743	1.151	0.942	<0.001
Ribitol (10254628)	1743,1.28	1.089 (0.276)	4.768 (0.745)	0.663	2.694	0.838	<0.001
Indole-3-carboxylic acid (13985315)	1854,6.66	0.137 (0.026)	0.929 (0.130)	0.642	1.048	1.032	<0.001
Kynurenic acid (3712)	1889,6.12	46.055 (0.812)	23.312 (3.872)	0.701	1.017	0.993	<0.001
Glycerol monostearate (71407)	2392,8.16	0 (0)	1.079 (0.247)	0.688	3.804	0.737	<0.001

indoleamine 2,3-dioxygenase 1 (IDO1), which in turn is also associated with the prognosis of various pathological conditions, including TB (Suzuki et al. 2012). This IDO1-upregulating host response is thought to deplete the bacterial environment of tryptophan, an amino acid essential to *M. tuberculosis* growth (Warner 2014), and also expected to result in alterations in the downstream metabolites in this pathway (see Fig. 4). Quinolinic acid, for instance, a neurotoxic metabolite formed via tryptophan catabolism, has been linked to inflammatory diseases, and is known to be

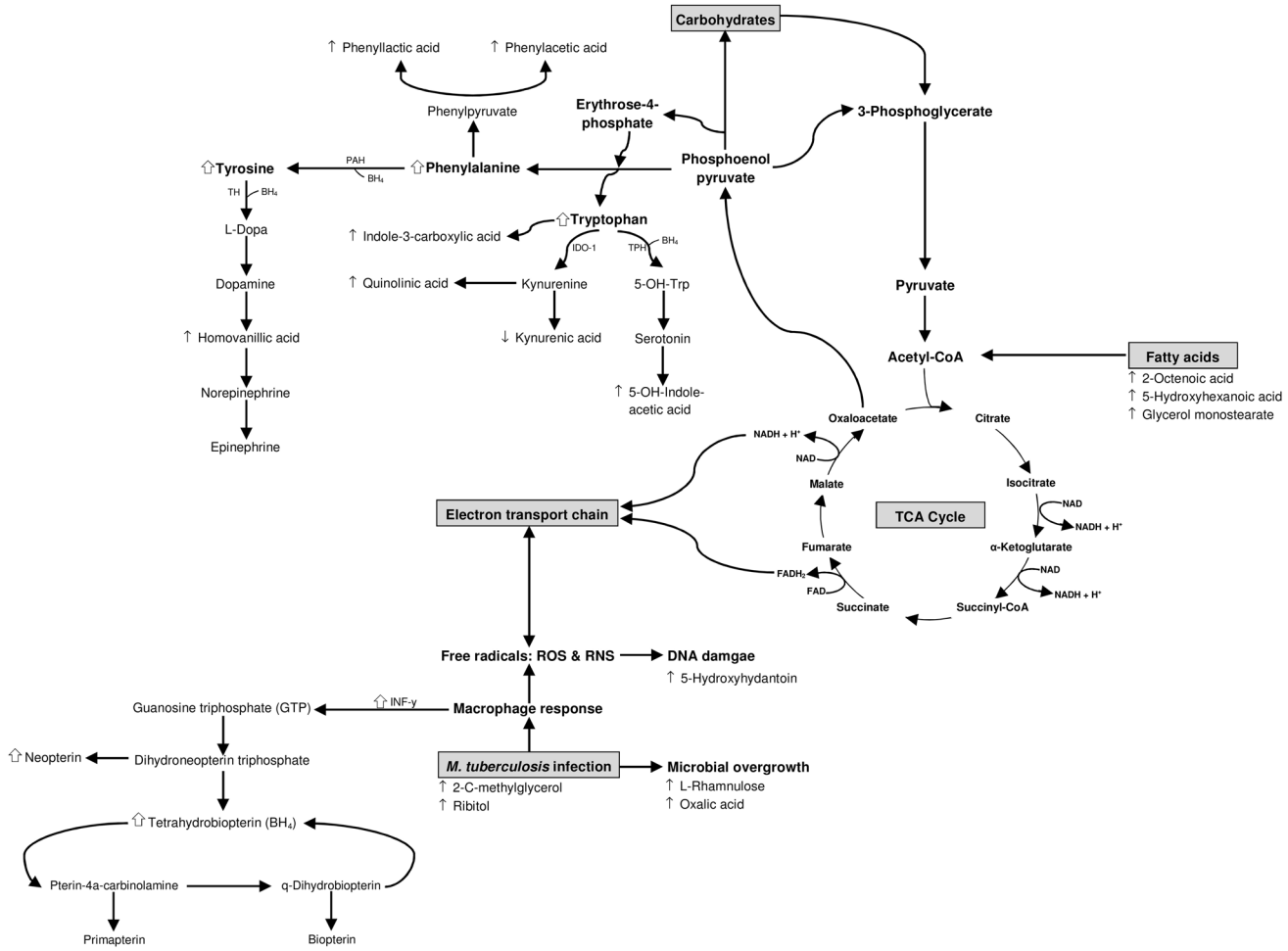


Fig. 4 Altered host metabolome induced by *M. tuberculosis*. A schematic representation of the metabolite markers (and their associated metabolism), identified in the urine of TB-positive patients, which are either increased or decreased relative to the healthy TB-negative group, indicated by ↑↓ respectively. Metabolites hypothesized/previously indicated in the literature to be elevated, are indicated with open arrows

elevated through IDO1 activation (Heyes et al. 1992). A study by Asp et al. (2011) indicated that TNF-α and IFN-γ reduced the levels of kynurenic acid (a neuroprotective agent), by inhibiting kynurenine aminotransferases (KATs) synthesis (Asp et al. 2011). For these reasons, one would expect elevations in quinolinic acid and reductions in kynurenic acid in the TB-infected host, which is what we see from our metabolomics data. Additionally, indole-3-carboxylic acid, the dehydrogenation product of tryptophan (Lübbe et al. 1983), was elevated approximately nine times in the TB-positive group, which may serve as an additional means by which the host attempts to reduce the much-needed tryptophan for *M. tuberculosis* survival. Interestingly, although not detected as one of the 12 metabolite markers using the multi-statistical method described, tryptophan was found to be elevated in the TB-positive group (0.667 vs. 0.096 mmol/L; $P = 0.041$), either because *M. tuberculosis* can synthesize tryptophan de novo (Warner 2014), or from another phenomenon explained below.

Another important observation is the presence of more than five times the normal levels of phenylacetic acid detected in the urine of TB-positive patients. Phenyllactic acid, although not identified as a metabolite marker, was also significantly elevated in the TB patients (4.107 vs. 0.009 mmol/L; $P < 0.0001$). Elevations in these two metabolites are indicative of phenylalanine accumulation in the TB-positive patients, similarly to what is seen in phenylketonuria (PKU) patients. Normally, phenylalanine is metabolized to tyrosine via phenylalanine hydroxylase (PAH), using tetrahydrobiopterin (BH_4) as a cofactor. If PAH or BH_4 are deficient for whatever reason, phenylalanine will accumulate and undergo transamination to produce phenylpyruvic acid, which is then either oxidized to phenylacetic acid or reduced to phenyllactic acid (Puri, 2006). Since the usual tryptophan [5-hydroxyindole-acetic acid (0.155 vs. 0.917 mmol/L; $P > 0.05$)], and tyrosine [N-acetyltyrosine (0.014 vs. 0.117 mmol/L; $P < 0.0001$); homovanillic acid (7.356 vs. 19.642 mmol/L; $P < 0.0001$)]

catabolism metabolites were elevated in the TB-positive group—and considering that both of these amino acid catabolic pathways require BH_4—a respective catabolic enzyme defect or BH_4 deficiency can be ruled out. Elevated concentrations of these amino acids are more likely due to compromised insulin production, which further supports earlier evidence for the association between TB and diabetes (Dooley and Chaisson 2009). This is further substantiated by the recently described detection of D-gluconic acid δ-lactone in the sputum of TB-positive patients, thought to be induced by elevated glucose levels, in addition to elevated levels of normetanephine, a metabolite of norepinephrine, associated with this response and also formed from tyrosine (Du Preez and Loots 2013). Insulin not only influences carbohydrate but also amino acid and protein metabolism, as a means to preserve amino acids for other essential anabolic processes. Since tryptophan and phenylalanine, as well as tyrosine to a certain extent, are essential amino acids, their concentrations in the human systemic circulation is dependent on protein consumption and uptake from the portal circulation system. These amino acids can also be released from tissue reservoirs or circulating proteins, which is typically what happens during wasting, a well-known occurrence in patients suffering from active TB, and also linked to reduced insulin secretion. The quantities of dietary phenylalanine, tyrosine and tryptophan entering systemic circulation are regulated by their catabolic hepatic enzymes, namely PAH, tyrosine aminotransferase (TAT), and tryptophan dioxygenase (TDO), respectively, which are induced by various factors, including reduced insulin. Although recent developments have shed light on the association between compromized insulin secretion and altered amino acid metabolism, the exact mechanism by which this occurs remains unknown (Adams 2011; Cansev and Wurtman 2007). In a previous metabolomics investigation, in which insulin-sensitive and insulin-resistant subjects were compared, both elevated phenylalanine and tyrosine concentrations were factors differentiating these groups (Tai et al. 2010), and serves as further evidence for the role of reduced insulin in the elevation of these amino acids. Moreover, reduced insulin secretion also results in free fatty acid oxidation in vivo (Adams 2011). When the body uses fat as the primary energy source, it is broken down and released into the bloodstream, resulting in increased urinary fatty acids and their breakdown products, as was detected in our study. These fatty acids and breakdown products included glycerol monostearate, 2-octenoic acid and 5-hydroxyhexanoic acid. This could potentially explain the weight-loss associated with TB if these fatty acids are sourced from the body's fat stores (Kairamkonda et al. 2003; Niu et al. 2012). This host response may additionally be driven by various bacterial mechanisms. Upon infection, *M. tuberculosis* initially utilizes host glucose and triacylglycerides as primary carbon sources, under aerobic conditions. However, with increased IFN-γ production, the glucose-deficient macrophages become hypoxic, and access to iron and its usual carbon sources become limited, forcing *M. tuberculosis* to induce an iron-scavenging reaction and utilize host cholesterol and fatty acids exclusively as the primary carbon sources. Hence, *M. tuberculosis* is able to adapt its metabolic activity in order to survive and persist within its human host, as well as to maximize energy production from alternative nutrient sources (Philips and Ernst 2012; Rhee et al. 2011; Singh et al. 2012). *M. tuberculosis* can also co-catabolize many carbon sources simultaneously via compartmentalization using multiple pathways (De Carvalho et al. 2010).

The elevated 5-hydroxyhydantoin detected in the TB-positive group is associated with DNA damage induced by oxidative stress, a common occurrence in TB patients, primarily due to H_2O_2 production by the macrophages in an attempt to overcome the bacterial invasion (Du Preez and Loots 2013). The ascorbic acid (vitamin C) oxidation product, oxalic acid (Robitaille et al. 2009), was observed in elevated concentrations in the TB-positive group. This is most likely due to ascorbic acid intake by the TB patients as this vitamin is not naturally synthesized by humans. Apart from this, foodborne moulds of *Aspergillus* species proliferate in the gastrointestinal tract of the host, as a secondary condition to the primary TB disease state (Singh and Toskes 2004) and are also known to produce oxalic acid (Alam et al. 2002).

The ribitol, L-rhamnulose and 2-C-methylglycerol (2MG) detected in elevated concentrations in the TB patients, are most likely derived directly from *M. tuberculosis,* as these are components of its cell walls (Mikusova et al. 1996; Silhavy et al. 2010; Velagapudi et al. 2010). Furthermore, 2MG is also an intermediate of the 2-C-methyl-D-erythritol-4-phosphate (MEP) pathway, unique to *M. tuberculosis*, enabling this pathogen to synthesize essential metabolites derived from isoprenoid compounds, and subsequently considered a target for the development of new anti-TB drugs, since homologous enzymes are absent in humans (Eoh et al. 2007; Kholodar and Murkin 2013).

Finally, previous studies have proposed that the nutritional imbalance (resulting in wasting) caused by *M. tuberculosis* trigger autophagy, a homeostatic intracellular degradation process targeting intracellular *M. tuberculosis*. During this process, macromolecules and organelles are fused with lysosomes and degraded, in order to sustain cellular anabolic processes (Choi et al. 2013; Deretic and Levine 2009; Goletti et al. 2013; Gutierrez et al. 2004). Once these macromolecules have been degraded, monomeric units, such as amino acids, fatty acids and DNA components, are either exported, excreted or reused

(Goletti et al. 2013; Mizushima 2007). Various studies have also shown that reduced insulin secretion (Mao et al. 2011) or increased glucagon (Deter and De Duve 1967) may induce autophagy as an additional host response to TB, contributing to the elevated amino acids, fatty acids, DNA breakdown products, and bacterial components detected in the TB patients.

Apart from the novel contribution of this discovery towards better understanding host-microbe interactions/adaptations, these markers may help explain some of the symptoms associated with TB, including neurological abnormalities (since quinolinic acid is a N-methyl-D-aspartate receptor agonist) (Heyes et al. 1992), vomiting, nausea, diarrhea, drowsiness, fatigue, loss of appetite and weight-loss (associated with elevated levels of phenylacetic acid) (Sherwin and Kennard 1919). Co-administration of anti-TB drugs with melatonin may assist in reducing elevated quinolinic acid (Cabrera et al. 2000), since melatonin has also been shown to lower a number of toxic metabolites associated with the side-effects of anti-TB drugs and also increase its efficacy for eliminating *M. tuberculosis* (Loots et al. 2005). Furthermore, BH_4 is given as a treatment for PKU patients in an attempt to reduce phenylacetic acid (McInnes et al. 1984), hence co-administration of anti-TB drugs with this merits consideration. Lastly, considering the association between TB and diabetes, co-administration of anti-TB drugs with antidiabetic drugs, such as metformin, could also be investigated. Singhal et al. (2014) showed how treatment of TB patients with metformin reduced intracellular *M. tuberculosis* growth by inducing reactive oxygen species production; facilitated phagolysosomal fusion by enhancing the host immune response; reduced the TB-incuded tissue pathology and inflammation; decreased TB severity; increased conventional anti-TB drug efficacy; improved clinical outcome; and reduced the incidence of latent TB (Singhal et al. 2014).

4 Concluding remarks

Metabolomics is considered a multistage process, often aimed at collecting data for the purpose of formulating new hypotheses and elucidating previously unknown disease mechanisms and/or host-microbe interactions and adaptations. This metabolomics study compared a healthy TB-negative group to that of an active TB-positive patient group, to identify the host and microbe interactions and adaptations, in response to TB infection. The 12 identified urinary metabolites, best explaining the differences occurring between these groups, could directly be linked to an abnormal host fatty acid and amino acid metabolism, especially those affecting tryptophan, phenylalanine and tyrosine, and inducing a metabolite profile similar to that of a PKU patient. This is due to changes to INF-γ and possibly insulin, and may also explain some of the TB-associated symptoms, providing clues to better treatment approaches.

Author contribution Both authors contributed equally to this work: DTL conceptionalized the study design, and LDV did the data acquisition/analysis. LDV worked on data interpretation, and drafted the article, and DTL revised it critically for important intellectual content. Both authors approved the final version to be submitted.

Compliance with Ethical standards

Conflict of interest The authors declare that there are no conflicts of interest, and that this manuscript, and the work described therein, is unpublished and has not been submitted for publication elsewhere.

Ethical approval Ethical approval for this investigation, carried out according to the Declaration of Helsinki and International Conference of Harmonization guidelines, was obtained from the Ethics Committee of the North-West University, South Africa (Number NWU-00127-11-A1). All participants gave written informed consent for collection of urine and its use as described.

References

Adams, S. H. (2011). Emerging perspectives on essential amino acid metabolism in obesity and the insulin-resistant state. *Advances in Nutrition, 2*, 445–456.

Alam, S., Khalil, S., Ayub, N., & Rashid, M. (2002). *In vitro* solubilization of inorganic phosphate by phosphate solubilizing microorganisms (PSM) from maize rhizosphere. *International Journal of Agriculture and Biology., 4*, 454–458.

Asp, L., Johansson, A. S., Mann, A., et al. (2011). Effects of pro-inflammatory cytokines on expression of kynurenine pathway enzymes in human dermal fibroblasts. *Journal of Inflammation (London), 8*, 25.

Brereton, R. G. (2003). *Chemometrics: Data analysis for the laboratory and chemical plant*. Chichester: Wiley.

Cabrera, J., Reiter, R. J., Tan, D., et al. (2000). Melatonin reduces oxidative neurotoxicity due to quinolinic acid: In vitro and in vivo findings. *Neuropharmacology, 39*, 507–514.

Cansev, M., & Wurtman, R. (2007). Aromatic amino acids in the brain. *Handbook of neurochemistry and molecular neurobiology* (pp. 59–97). Berlin: Springer.

Choi, A. M., Ryter, S. W., & Levine, B. (2013). Autophagy in human health and disease. *New England Journal of Medicine, 368*, 651–662.

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

De Villiers, L., & Loots, D. T. (2013). Using metabolomics for elucidating the mechanisms related to tuberculosis treatment failure. *Current Metabolomics, 1*, 306–317.

Deretic, V., & Levine, B. (2009). Autophagy, immunity, and microbial adaptations. *Cell Host & Microbe, 5*, 527–549.

Deter, R. L., & De Duve, C. (1967). Influence of glucagon, an inducer of cellular autophagy, on some physical properties of rat liver lysosomes. *Journal of Cell Biology, 33*, 437–449.

Dooley, K. E., & Chaisson, R. E. (2009). Tuberculosis and diabetes mellitus: Convergence of two epidemics. *The Lancet Infectious Diseases, 9*, 737–746.

Drapal, M., Perez-Fons, L., Wheeler, P. R., & Fraser, P. D. (2014). The application of metabolite profiling to *Mycobacterium* spp.: Determination of metabolite changes associated with growth. *Journal of Microbiol Methods, 106*, 23–32.

Du Preez, I., & Loots, D. T. (2013). New sputum metabolite markers implicating adaptations of the host to *Mycobacterium tuberculosis*, and vice versa. *Tuberculosis, 93*, 330–337.

Ellis, S., & Steyn, H. (2003). Practical significance (effect sizes) versus or in combination with statistical significance (p-values): Research note. *Management Dynamics, 12*, 51–53.

Eoh, H., Brown, A. C., Buetow, L., et al. (2007). Characterization of the *Mycobacterium tuberculosis* 4-diphosphocytidyl-2-C-methyl-D-erythritol synthase: Potential for drug development. *Journal of Bacteriology, 189*, 8922–8927.

Goletti, D., Petruccioli, E., Romagnoli, A., Piacentini, M., & Fimia, G. M. (2013). Autophagy in *Mycobacterium tuberculosis* infection: A passepartout to flush the intruder out? *Cytokine & Growth Factor Reviews, 24*, 335–343.

Gutierrez, M. G., Master, S. S., Singh, S. B., Taylor, G. A., Colombo, M. I., & Deretic, V. (2004). Autophagy is a defense mechanism inhibiting BCG and *Mycobacterium tuberculosis* survival in infected macrophages. *Cell, 119*, 753–766.

Heyes, M. P., Saito, K., & Markey, S. P. (1992). Human macrophages convert L-tryptophan into the neurotoxin quinolinic acid. *Biochemical Journal, 283*, 633–635.

Kairamkonda, V., Dalzell, M., Losty, P. D., & Davidson, C. (2003). Perforated duodenal ulcer disclosing medium chain acyl-CoA dehydrogenase deficiency. *Archives of Disease in Childhood, 88*, 88–89.

Kholodar, S. A., & Murkin, A. S. (2013). DXP reductoisomerase: Reaction of the substrate in pieces reveals a catalytic role for the nonreacting phosphodianion group. *Biochemistry, 52*, 2302–2308.

Long, N. H., Diwan, V. K., & Winkvist, A. (2002). Difference in symptoms suggesting pulmonary tuberculosis among men and women. *Journal of Clinical Epidemiology, 55*, 115–120.

Loots, D. T., Mienie, L. J., Bergh, J. J., der Schyf, Van, & Cornelis, J. (2004). Acetyl-L-carnitine prevents total body hydroxyl free radical and uric acid production induced by 1-methyl-4-phenyl-1,2,3,6-tetrahydropyridine (MPTP) in the rat. *Life Sciences, 75*, 1243–1253.

Loots, D. T., Wiid, I. J., Page, B. J., Mienie, L. J., & Helden, P. D. (2005). Melatonin prevents the free radical and MADD metabolic profiles induced by antituberculosis drugs in an animal model. *Journal of Pineal Research, 38*, 100–106.

Lübbe, C., Van Pée, K., Salcher, O., & Lingens, F. (1983). The metabolism of tryptophan and 7-chlorotryptophan in Pseudomonas pyrrocinia and Pseudomonas aureofaciens. *Hoppe-Seyler's Zeitschrift für physiologische Chemie, 364*, 447–454.

Mao, F., Chen, T., Zhao, Y., et al. (2011). Insulin resistance: A potential marker and risk factor for active tuberculosis? *Medical Hypotheses, 77*, 66–68.

McInnes, R. R., Kaufman, S., Warsh, J. J., et al. (1984). Biopterin synthesis defect: Treatment with L-dopa and 5-hydroxytryptophan compared with therapy with a tetrahydropterin. *Journal of Clinical Investigation, 73*, 458–469.

Mikusova, K., Mikus, M., Besra, G. S., Hancock, I., & Brennan, P. J. (1996). Biosynthesis of the linkage region of the mycobacterial cell wall. *Journal of Biological Chemistry, 271*, 7820–7828.

Mizushima, N. (2007). Autophagy: Process and function. *Genes & Development, 21*, 2861–2873.

Niu, J., Pi, Z., Yue, H., Wang, Y., Yu, Q., & Liu, S. (2012). Effect of ginseng polysaccharide on the urinary excretion of type 2 diabetic rats studied by liquid chromatography–mass spectrometry. *Journal of Chromatography B, 907*, 7–12.

Philips, J. A., & Ernst, J. D. (2012). Tuberculosis pathogenesis and immunity. *Annual Review of Pathology: Mechanisms of Disease, 7*, 353–384.

Puri, D. (2006). Textbook of medical biochemistry: Clinical cases Chapter 13.2. *Textbook of medical biochemistry* (p. 376). New-Delhi: Elsevier.

Rhee, K. Y., Carvalho, De, Sorio, Luiz Pedro, Bryk, R., et al. (2011). Central carbon metabolism in *Mycobacterium tuberculosis*: An unexpected frontier. *Trends in Microbiology, 19*, 307–314.

Robitaille, L., Mamer, O. A., Miller, W. H, Jr, et al. (2009). Oxalic acid excretion after intravenous ascorbic acid administration. *Metabolism, Clinical and Experimental, 58*, 263–269.

Schoeman, J. C., Du Preez, I., & Loots, D. T. (2012). A comparison of four sputum pre-extraction preparation methods for identifying and characterising *M. tuberculosis* using GCxGC-TOFMS metabolomics. *Journal of Microbiol Methods, 91*, 301–311.

Sherwin, C. P., & Kennard, K. S. (1919). Toxicity of phenylacetic acid. *Journal of Biological Chemistry, 40*, 259–264.

Silhavy, T. J., Kahne, D., & Walker, S. (2010). The bacterial cell envelope. *Cold Spring Harbor Perspectives in Biology, 2*, a000414.

Singh, V., Jamwal, S., Jain, R., Verma, P., Gokhale, R., & Rao, K. V. (2012). *Mycobacterium tuberculosis*-driven targeted recalibration of macrophage lipid homeostasis promotes the foamy phenotype. *Cell Host & Microbe, 12*, 669–681.

Singh, V. V., & Toskes, P. P. (2004). Small bowel bacterial overgrowth: Presentation, diagnosis, and treatment. *Current Treatment Options in Gastroenterology, 7*, 19–28.

Singhal, A., Jie, L., Kumar, P., et al. (2014). Metformin as adjunct antituberculosis therapy. *Science Translational Medicine, 6*, 263ra159.

Smuts, I., Der Westhuizen, Van, Francois, H., Louw, R., et al. (2013). Disclosure of a putative biosignature for respiratory chain disorders through a metabolomics approach. *Metabolomics, 9*, 379–391.

Suzuki, Y., Suda, T., Asada, K., et al. (2012). Serum indoleamine 2,3-dioxygenase activity predicts prognosis of pulmonary tuberculosis. *Clinical and Vaccine Immunology, 19*, 436–442.

Tai, E., Tan, M., Stevens, R., et al. (2010). Insulin resistance is associated with a metabolic profile of altered protein metabolism in Chinese and Asian-Indian men. *Diabetologia, 53*, 757–767.

Velagapudi, V. R., Hezaveh, R., Reigstad, C. S., et al. (2010). The gut microbiota modulates host energy and lipid metabolism in mice. *Journal of Lipid Research, 51*, 1101–1112.

Venter, L., Lindeque, Z., Van Rensburg, P. J., Van Der Westhuizen, F., Smuts, I., & Louw, R. (2015). Untargeted urine metabolomics reveals a biosignature for muscle respiratory chain deficiencies. *Metabolomics, 11*, 111–121.

Wang, S., Kuo, C., & Tseng, Y. J. (2012). Batch normalizer: A fast total abundance regression calibration method to simultaneously adjust batch and injection order effects in liquid chromatography/time-of-flight mass spectrometry-based metabolomics data and comparison with current calibration methods. *Analytical Chemistry, 85*, 1037–1046.

Warner, D. F. (2014). Mycobacterium tuberculosis metabolism. *Cold Spring Harbor Perspectives in Medicine,*. doi:10.1101/cshperspect.a021121.

World Health Organization. (2014). Tuberculosis: WHO global tuberculosis report 2014, 1–2.

Metabolomics

URINARY METABOLITE MARKERS CHARACTERIZING TUBERCULOSIS TREATMENT FAILURE

--Manuscript Draft--

Manuscript Number:	MEBO-D-17-00069
Full Title:	URINARY METABOLITE MARKERS CHARACTERIZING TUBERCULOSIS TREATMENT FAILURE
Article Type:	Original Article
Keywords:	Biomarkers; M. tuberculosis; Metabolomics; Treatment failure; Tuberculosis; Urine
Corresponding Author:	Laneke de Villiers, B.Sc North-West University Potchefstroom, North-West SOUTH AFRICA
Corresponding Author Secondary Information:	
Corresponding Author's Institution:	North-West University
Corresponding Author's Secondary Institution:	
First Author:	Laneke de Villiers, B.Sc
First Author Secondary Information:	
Order of Authors:	Laneke de Villiers, B.Sc
	Japie Mienie, Ph.D.
	Christinah Motshwane, B.Sc.
	Katharina Ronacher, Ph.D.
	Gerhard Walzl, Ph.D.
	Du Toit Loots, Ph.D.
Order of Authors Secondary Information:	
Funding Information:	
Abstract:	Background: Considering that approximately 15% of the nine million new tuberculosis (TB) cases reported per annum are not treated successfully, new, distinctive and specific biomarkers are needed to better characterize the biological basis of a poor treatment outcome. Methods: Urine samples from 41 active pulmonary TB patients were collected at baseline (time of diagnosis), during treatment (weeks 1, 2 and 4) and two weeks after treatment completion (week 26). These samples were divided into successful (cured) and unsuccessful (failed) treatment outcome groups and analyzed using a GCxGC-TOFMS metabolomics research approach. Results: The metabolite data collected showed clear differentiation of the cured and failed treatment outcome groups using the samples collected at the time of diagnosis, i.e. before any treatment was administered. Conclusions: The treatment failure group was characterized by an imbalanced gut microbiome, in addition to elevated levels of metabolites associated with abnormalities in the long-chain fatty acid β-oxidation pathway, accompanied by reduced L-carnitine and short-chain fatty acids, indicative of a mitochondrial trifunctional protein defect in particular. Furthermore, an altered amino acid metabolism was also observed in these patients, which confirms previous findings and associations to increased interferon gamma due to the host's immune response to M. tuberculosis and a compromised insulin secretion.

URINARY METABOLITE MARKERS CHARACTERIZING TUBERCULOSIS TREATMENT FAILURE

RUNNING TITLE: Characterizing TB treatment failure

Laneke Luies[1], Japie Mienie[1], Christinah Motshwane[1], Katharina Ronacher[2,3], Gerhard Walzl[2] and Du Toit Loots[1*]

[1]School for Physical and Chemical Sciences, Human Metabolomics, North-West University (Potchefstroom Campus), Private Bag x6001, Box 269, Potchefstroom, South Africa, 2531. [2]DST/NRF Centre of Excellence for Biomedical Tuberculosis Research/MRC Centre for Molecular and Cellular Biology, Division of Molecular Biology and Human Genetics, Faculty of Medicine and Health Sciences, Stellenbosch University, Tygerberg, South Africa, 7505. [3]Mater Medical Research, The University of Queensland, Brisbane, Australia

CORRESPONDING AUTHOR

Du Toit Loots; E-mail: dutoit.loots@nwu.ac.za; Address: North-West University, Potchefstroom Campus, Building F3: Biochemistry, 11 Hoffman Street, Potchefstroom, South Africa, 2531.

To avoid confusion, please note that the author's first name is 'Du Toit' and last name is 'Loots'

WORD COUNT: Abstract: 199

Text: 2 746

ABSTRACT

Background: Considering that approximately 15% of the nine million new tuberculosis (TB) cases reported per annum are not treated successfully, new, distinctive and specific biomarkers are needed to better characterize the biological basis of a poor treatment outcome.

Methods: Urine samples from 41 active pulmonary TB patients were collected at baseline (time of diagnosis), during treatment (weeks 1, 2 and 4) and two weeks after treatment completion (week 26). These samples were divided into successful (cured) and unsuccessful (failed) treatment outcome groups and analyzed using a GCxGC-TOFMS metabolomics research approach.

Results: The metabolite data collected showed clear differentiation of the cured and failed treatment outcome groups using the samples collected at the time of diagnosis, i.e. before any treatment was administered.

Conclusions: The treatment failure group was characterized by an imbalanced gut microbiome, in addition to elevated levels of metabolites associated with abnormalities in the long-chain fatty acid β-oxidation pathway, accompanied by reduced L-carnitine and short-chain fatty acids, indicative of a mitochondrial trifunctional protein defect in particular. Furthermore, an altered amino acid metabolism was also observed in these patients, which confirms previous findings and associations to increased interferon gamma due to the host's immune response to *M. tuberculosis* and a compromised insulin secretion.

KEY WORDS: Biomarkers; *M. tuberculosis*; Metabolomics; Treatment failure; Tuberculosis; Urine.

1. INTRODUCTION

Tuberculosis (TB), caused by *Mycobacterium tuberculosis*, remains the world's foremost cause of death from a single bacterial agent, which is alarming as it is considered curable. Nearly 10.4 million new TB cases are reported per annum, of which almost 15% are not treated successfully, resulting in approximately 1.5 million deaths globally (3 800 deaths a day) (World Health Organization 2016). TB treatment failure may be attributed to, amongst others: (a) irregular or inadequate anti-TB drug supplies to rural areas and third-world countries, (b) poor patient TB-education, (c) poor socio-economic circumstances, (d) the prolonged treatment duration, (e) treatment non-adherence, (f) anti-TB drug resistance, as well as (g) various biological/biochemical factors (De Villiers and Loots 2013).

The World Health Organization (WHO) recommends a six-month multi-drug treatment regimen, in which a combination of four drugs (isoniazid (INH), rifampicin (RIF), pyrazinamide (PZA) and ethambutol (EMB)) are used for treating active TB. In patients with drug-susceptible TB, this regimen reportedly has a 1–4% failure rate, and only 7% of those patients with a successful treatment outcome relapse within 24 months (Dye et al. 2005). When considering that every active TB patient can potentially infect an additional 10–15 individuals per annum (World Health Organization 2016), it becomes evident that TB treatment failure and relapse are important considerations in achieving the millennium goals pertaining to the eradication of TB. Despite the many TB research efforts to date, the biological mechanisms associated with anti-TB drug response remains poorly understood. It is also unclear if certain patients have more efficient mechanisms for eliminating this disease and responding to the treatment thereof, and to what extent lifestyle and environmental factors may contribute to this (De Villiers and Loots 2013).

Early disease diagnosis and effective treatment protocols are the two primary objectives of TB control, aimed at reducing mortality and morbidity while also preventing the development of drug resistance (Antoine et al. 2007). Thus, monitoring treatment progression and having the means to accurately predict treatment outcome, well before the six-month treatment regimen is completed, would be regarded as a major breakthrough, as poor treatment outcomes could be identified early, and

alternative treatment approaches implemented, ultimately reducing non-adherence, drug resistance and treatment failure (Walzl et al. 2008; Horne et al. 2010). Considering this, there is a need for new, sensitive and specific biomarkers, not only for use in the early prediction of treatment failure, but also for better characterizing and explaining the underlying mechanisms related to this occurrence, so that alternative treatment approaches can be developed. Once identified, these biomarkers need to be sufficiently validated for use as surrogate endpoints of treatment failure, early in the treatment regimen, ideally even before treatment begins. To date, no such urinary biomarkers have been identified with absolute certainty, hence the aim of this study was to use a two-dimensional gas chromatography time-of-flight mass spectrometry (GCxGC-TOFMS) metabolomics approach to identify biomarkers differentiating individuals with a successful (n = 26) and unsuccessful (n = 15) treatment outcome, as early as possible during the treatment regimen, which would also better characterize and explain the biological mechanisms related to TB treatment failure. Urine was selected as the preferred sample for addressing the above mentioned aim, since large quantities can easily be obtained and less complex sample preparation is required for analyses, as compared to sputum for instance. Furthermore, in recent years, the need for a holistic approach to metabolism has led to the development of urinary metabolomics for biomarker discovery in various diseases (Ryan et al. 2011; Mahapatra et al. 2014).

2. METHODS

2.1 Clinical samples

Anonymised archived urine samples were procured from the Faculty of Medicine and Health Sciences, NRF/DST Centre of Excellence for Biomedical Research, at the Stellenbosch University/MRC Centre for TB Research, from where they were transported to the North-West University (NWU), Human Metabolomics: Infectious Disease Laboratory, for metabolomic analysis. These samples were selected from a prospective observational cohort study of individuals with active pulmonary TB, diagnosed using smear microscopy and bacteriological cultures (Hesseling et al. 2010). From these original

samples, all treatment failure patients were included and matched by age, gender and extent of disease on chest X-rays to cured patients. Hence, the samples included were from 41 culture-confirmed active TB-positive South African patients (22 males and 19 females, between the age of 17 and 58) at baseline (time of diagnosis, thus prior to initiation of treatment), during the course of treatment with the directly observed treatment short-course (DOTS) strategy (weeks 1, 2 and 4) and two weeks after treatment completion (week 26). These patients had drug-susceptible TB, were human immunodeficiency virus (HIV)-seronegative, not pregnant, and with no other diseases (including diabetes, malignancy, lung cancer, chronic bronchitis and sarcoidosis). The samples were divided into successful (n=26) and unsuccessful (n=15) treatment outcome groups. According to the WHO, treatment failure can be defined as a patient whose sputum smear or culture is positive at month five or later during treatment (World Health Organization 2014).

2.2 Sample analysis

Creatinine values for all urine samples, including quality control (QC) samples, were determined using a creatinine enzyme kit (Thermo Scientific; reference number 981845) and analyzed on an Indiko Clinical Analyzer, Type 863 (Thermo Scientific). These creatinine values are used to normalize metabolite concentrations, and to determine the volume of urine, internal standard, bis-(trimethylsilyl)-trifluoroacetamide (BSTFA), trimethylchlorosilane (TMCS) and pyridine needed for each extraction and derivatization. An organic acid extraction of the collected patient urine samples were performed, analyzed on a Pegasus 4D GCxGC-TOFMS (LECO Africa (Pty) Ltd, Johannesburg, South Africa) along with the necessary QC samples, and processed to identify all compounds, following the methods described in Luies and Loots (2016) (Luies and Loots 2016).

2.3 Statistical data analysis

A standard metabolomics data clean-up procedure was applied before statistical data analysis. All metabolites were normalized relative to the internal standard by calculating the relative concentration

of each, and a 50% zero filter was applied to each variable (Smuts et al. 2013) to eliminate those compounds with more than 50% zero values within both groups. Quantile equating was applied to correct for any batch effects (Wang et al. 2012), followed by a 50% quality control–coefficient of variation (QC–CV) filter. Lastly, all zero values were replaced with a value calculated as half of the lowest detected value present in the entire dataset, as these may be due to low abundance rather than being absent (Schoeman et al. 2012). GCxGC-TOFMS analysis yielded 782 compounds of which only 241 remained for further statistical analysis after the above mentioned data clean-up were completed. The data were analyzed via a variety of multi- and univariate statistical methods, using MetaboAnalyst (based on the statistical package "R"; version 2.10.0), which included principal components analysis (PCA) (Wold et al. 1987), partial least squares–discriminant analysis (PLS–DA) (Westerhuis et al. 2008), fold change (Du Preez and Loots 2013) and Mann-Whitney test (Pallant 2001).

3. RESULTS AND DISCUSSION

Using the GCxGC-TOFMS data generated, a PCA was done to determine at which time point individuals who responded to the TB treatment can be differentiated from those who were not cured from TB. The PCA scores plots show clear differentiation of the successful and unsuccessful treatment outcome groups, using the urine collected at time of diagnosis and again at week 26 (Figure 1, a and b respectively), the latter of which can be expected since the cured individuals are TB culture negative, and those with a poor treatment outcome are still TB culture positive at this point in time (Luies and Loots 2016). No PCA differentiation was achieved for the other time intervals investigated (see supplementary material), most likely due to the effects of the anti-TB medication on the human metabolome, masking any underlying differences which occur in the metabolism initially differentiating the successful and unsuccessful treatment outcome groups at time of diagnosis. For the differentiation achieved at time of diagnosis (Figure 1a), the total amount of variance explained by the first two principal components (PCs) (R^2X cum) was 93.4%, of which PC1 accounted for 91.8%, and PC2 accounted for 1.6%. For the differentiation achieved at week 26 (Figure 1b), the total amount of

variance explained by the first two PCs (R^2X cum) was 91.9%, of which PC1 accounted for 90.2% and PC2 accounted for 1.7%.

Since the aim of this study was to identify biomarkers for the early prediction of treatment response and explain the mechanisms associated with treatment failure, only the PCA differentiation achieved at time of diagnosis was of further interest in this investigation. The separation of the groups at week 26 are mostly due to metabolic changes since one group is TB-positive and the other TB-negative, which was already investigated on a previous occasion (Luies and Loots 2016). Considering the above, a PLS–DA model was built for the time of diagnosis data and showed a modeling parameter R^2Y (cum) of 93.48%, indicative of the total explained variation of the response Y.

Clear separation of the successful and unsuccessful treatment outcome groups is remarkable considering the possible predictive value that these may have for identifying individuals who will or will not respond to conventional treatment, even before treatment commences. However, before this biosignature or the individual metabolite markers can be utilized for this purpose, further validation is needed using new subjects to ensure these findings can be generalized. In the context of this study, the biosignature/metabolite markers were used to better describe the biological mechanisms as to why certain individuals with drug-sensitive TB are not successfully treated. Subsequently, those compounds with a PLS–DA variable influence on the projection (VIP) >1.0 (Smuts et al. 2013), or a fold change >|2|, or a Mann-Whitney with a *P*-value <0.05 (Du Preez and Loots 2013), were selected as the metabolite markers that best explained the variation between the analyzed groups. Of these 72 characteristic metabolite markers, 50 could be annotated using libraries compiled from previously injected standards, and are listed in Table 1.

Table 1: *The 50 urinary metabolite markers identified at time of diagnosis that best explain the variation detected between the successful and unsuccessful treatment outcome groups.*

Metabolite name	Successful treatment outcome: Average concentration (mg/g creatinine) (Standard error of the mean)	Unsuccessful treatment outcome: Average concentration (mg/g creatinine) (Standard error of the mean)	PLS–DA (VIP)	Fold change	Mann-Whitney test (*P*-value)
Amino acid metabolites:					
2-Ketovaleric acid	2.093 (0.428)	3.799 (1.302)	1.30	1.82	0.1318
2-Methyl-3-hydroxybutyric acid	3.777 (0.489)	5.487 (0.708)	2.30	1.45	0.0273
2-Piperidinecarboxylic acid (Pipecolic acid)	0.893 (0.230)	0.428 (0.104)	1.24	2.09	0.0197
3-Hydroxy-3-methylglutaric acid	9.437 (0.456)	12.179 (0.945)	1.08	1.29	0.0077
3-Hydroxyvaleric acid	7.691 (1.341)	8.896 (1.861)	1.05	1.16	0.3301
4-Methylcatechol	6.916 (1.150)	10.299 (1.854)	1.36	1.49	0.1022
Glyceric acid	0.268 (0.098)	0.356 (0.091)	1.14	1.33	0.0062
Methylsuccinic acid	3.912 (0.448)	6.581 (1.020)	2.60	1.68	0.0030
N-Tiglylglycine	0.834 (0.247)	1.567 (0.371)	1.91	1.88	0.0094
o-Hydroxyphenylacetic acid	3.093 (0.384)	5.293 (0.602)	2.50	1.71	0.0030
Oxalic acid	14.250 (1.308)	19.933 (2.116)	2.29	1.40	0.0191
Phenylacetic acid	0.305 (0.157)	0.903 (0.273)	2.14	2.97	0.0306
Quinolinic acid	5.311 (0.854)	9.448 (1.403)	5.50	1.78	0.0040
Vanillylmandelic acid	17.088 (0.878)	22.269 (1.830)	1.27	1.30	0.0040
***M. tuberculosis* cell wall components:**					
Ribitol	1.910 (0.627)	2.455 (0.523)	1.84	1.29	0.0708
Gut microbiota imbalance metabolites:					
2,3-Butanediol	17.831 (4.351)	16.301 (5.137)	3.65	1.09	0.4450
3-(3-Hydroxyphenyl)-3-hydroxypropionic acid (HPHPA)	20.030 (3.029)	28.087 (4.093)	5.40	1.40	0.0553
3-Hydroxy-3-(4-hydroxy-3-methoxyphenyl)propionic acid	2.912 (0.575)	8.366 (2.431)	6.45	2.87	0.0474
3-Hydroxyhippuric acid	17.580 (1.716)	23.662 (2.525)	2.37	1.35	0.0289
4-Hydroxybenzoic acid	4.642 (0.663)	9.966 (1.754)	4.92	2.15	0.0027
4-Hydroxyphenyllactic acid	5.695 (1.526)	6.724 (1.720)	1.21	1.18	0.1485
5-Hydroxymethyl-2-furoic acid	17.563 (4.279)	21.994 (6.053)	4.46	1.25	0.1795
Benzoic acid	0.936 (0.275)	1.367 (0.251)	1.56	1.46	0.0229
cis-4-Hydroxycyclohexanecarboxylic acid	1.428 (0.610)	0.666 (0.190)	0.87	2.15	0.0306
Citramalic acid	8.700 (1.608)	10.742 (3.124)	1.66	1.23	0.4550
Furan-2,5-dicarboxylic acid	8.755 (2.540)	10.156 (3.446)	1.14	1.16	0.3032
Furoylglycine	7.466 (1.443)	10.421 (2.241)	3.39	1.40	0.1115
Vanillic acid	7.731 (0.876)	11.249 (1.230)	2.23	1.46	0.0082
DNA damage and oxidative stress markers:					
2-Deoxyribonic acid	0.477 (0.113)	1.080 (0.245)	1.85	2.27	0.0169
Parabanic acid	2.984 (0.799)	3.584 (0.865)	2.68	1.20	0.0474
TCA and GABA metabolites:					

3,4-Dihydroxybutyric acid	9.160 (0.825)	11.730 (0.89)	1.03	1.28	0.0108
4-Hydroxybutyric acid	0.305 (0.097)	0.544 (0.158)	1.68	1.79	0.0088
Citric acid	2.730 (0.838)	3.526 (1.619)	1.34	1.29	0.2691
Malic acid	4.579 (2.712)	5.085 (2.668)	3.13	1.11	0.4351
Dicarboxylic acids and 3-hydroxy fatty acids:					
2,3-Dihydroxypentonic acid	4.632 (1.255)	6.118 (2.265)	2.56	1.32	0.3485
3,5-Dihydroxypentonic acid	3.956 (0.469)	5.698 (0.790)	1.29	1.44	0.0743
3-Hydroxydodecanedioic acid	4.946 (1.659)	2.313 (1.075)	2.27	2.14	0.2000
3-Hydroxysebacic acid	1.499 (0.386)	3.716 (1.140)	2.36	2.48	0.0582
3-Ketosebacic acid	2.419 (0.535)	6.451 (1.560)	6.96	2.67	0.0324
3-Methylhexanoic acid (3-Methyladipic acid)	7.723 (0.800)	12.631 (1.637)	3.30	1.64	0.0058
cis,*cis*-4,7-Decadiene-1,10-dioic acid	9.523 (1.259)	12.113 (1.359)	1.19	1.27	0.0743
cis-4-Decene-1,10-dioic acid	1.403 (0.316)	2.561 (0.649)	1.74	1.83	0.0499
Heptanedioic acid (Pimelic acid)	2.895 (0.434)	4.967 (0.684)	2.32	1.72	0.0024
Hexanedioic acid (Adipic acid)	0.713 (0.192)	1.369 (0.443)	1.38	1.92	0.0816
Nonanedioic acid (Azelaic acid)	4.596 (0.918)	6.465 (1.271)	2.01	1.41	0.0894
Octanedioic acid (Suberic acid)	3.452 (0.595)	4.577 (0.394)	1.68	1.33	0.0169
Other:					
1,2,3-Trihydroxybutane	1.962 (0.540)	2.842 (1.415)	1.22	1.45	0.4950
2-Methyl-1,2-dihydroxypropane	5.360 (1.768)	6.456 (2.985)	2.78	1.20	0.3672
4-Pentenoic acid	5.860 (0.898)	8.201 (1.289)	1.86	1.40	0.0611
Glucuronic acid	5.923 (0.442)	8.307 (0.757)	1.22	1.40	0.0062

*Abbreviations: mg/g, milligram per gram; PLS–DA, partial least squares–discriminant analysis; VIP, variable influence on the projection.

When considering the metabolite markers best differentiating the successful and unsuccessful treatment outcome patient groups, the first important observation is the elevated amounts of metabolites associated with tryptophan (quinolinic acid (Heyes et al. 1992)), phenylalanine (phenylacetic acid (Puri 2006), o-hydroxyphenylacetic acid (Taniguchi and Armstrong 1963)), and tyrosine (4-methylcatechol (Li et al. 2007), vanillylmandelic acid (Eisenhofer et al. 2004)) metabolism (see Figure 2). The same metabolic flux and the resulting accumulation of these amino acids, were also previously reported by Luies and Loots (2016) in TB-positive patient urine, and explained by elevations in interferon gamma (IFN-γ) due to the host's immune response to *M. tuberculosis* and a compromised insulin secretion by the host (Luies and Loots 2016). Furthermore, markers directly associated with the presence of *M. tuberculosis* (ribitol (Silhavy et al. 2010)), a gut microbiota imbalance (2,3-butanediol (Hong 2011), 3-(3-hydroxyphenyl)-3-hydroxypropionic acid (Shaw 2013), 3-hydroxy-3-(4-hydroxy-3-methoxyphenyl)propionic acid, 3-hydroxyhippuric acid (Gonthier et al.

2003), 4-hydroxybenzoic acid (Russell et al. 2013; Tomás-Barberán and Clifford 2000), 4-hydroxyphenyllactic acid (Russell et al. 2013), 5-hydroxymethyl-2-furoic acid (Jellum et al. 1973), benzoic acid (Russell et al. 2013), *cis*-4-hydroxycyclohexanecarboxylic acid (Kronick et al. 1983), citramalic acid (Du Preez and Loots 2013), furan-2,5-dicarboxylic acid (Pettersen and Jellum 1972; Jellum et al. 1973), furoylglycine (Pettersen and Jellum 1972), vanillic acid (Tomás-Barberán and Clifford 2000)), and DNA damage (2-deoxyribonic acid (Zhou and Greenberg 2012)) due to oxidative stress (parabanic acid (Marklund et al. 2000)), were also detected and correlate with those markers previously characterizing TB, as described by Luies and Loots (2016). Considering the fact that these markers are not only indicative of a general TB disease state (Luies and Loots 2016), but also more pronounced in the treatment failure group in this study, may indicate that these individuals who are unsuccessfully treated have an increased disease severity due to a microbiota imbalance and/or underlying host abnormality, which will be discussed in greater detail below.

The highly significant evidence from the almost three times elevated 3-hydroxy-3-(4-hydroxy-3methoxyphenyl)propionic acid and more than two times lower *cis*-4-hydroxycyclohexanecarboxylic acid in the treatment failure group, strongly testifies towards an imbalance in gut microbiota in the unsuccessfully treated patients, which is associated with a weaker immune system and inability to fight disease (Clemente et al. 2012). Both the innate and adaptive immune response systems have evolved to rely on microbiota interactions, which not only promote immune cell maturation but also influence the normal development of immune functions (Clemente et al. 2012). Additionally, gut microbiota is also well known to interact with various drugs, influencing many factors relating to their absorption and subsequent plasma concentrations, and hence is considered an important contributor to xenobiotic/drug bioavailability and toxicity (Gonzalez et al. 2011).

Furthermore, altered levels of various dicarboxylic acids, acylcarnitines and 3-hydroxy fatty acids were detected in the urine of the treatment failure group. Lipid molecules are increasingly recognized as having the potential to shape the immune response to infectious pathogens (Kaushal 2012) since these are important bioactive mediators of cellular activity during pathophysiological processes, and affects various activities including cell apoptosis, monocyte adhesion, platelet aggregation, and

regulation of immune responses (Hasanally et al. 2014). Interestingly, however, the presence of these metabolites in urine are also used to distinguish between individuals with inherited defects of the following long-chain fatty acid (LCFA) transport-associated enzymes: (a) carnitine palmitoyltransferase 1 (CPT-1), (b) carnitine palmitoyltransferase 2 (CPT-2), (c) translocase (TL), (d) very-long-chain acyl-coenzyme A (CoA) dehydrogenase (VLCAD) or (e) any of the three enzymes making up the mitochondrial trifunctional protein (MTP); long chain 2-enoyl-Co A hydratase, long-chain 3-hydroxyacyl-CoA dehydrogenase and 3-ketoacyl-CoA thiolase (Sim et al. 2002; Garg and Dasouki 2006). A CPT-1 deficiency results in reduced C_{16}–C_{18} acylcarnitines and increased free carnitine, a CPT-2, TL or VLCAD deficiency results in both elevated long chain acylcarnitines and dicarboxylic acids with reduced free carnitine, and a MTP deficiency (on any of its three associated enzymes) results in elevated long chain acylcarnitines, dicarboxylic acids and 3-hydroxy fatty acids with reduced free carnitine (Garg and Dasouki 2006). Since the organic acid extraction method used in this study was not suited for the extraction or detection of acylcarnitines, hence their absence in Table 1, we did a further ultra-performance liquid chromatography (UPLC) (Van Aardt et al. 2016) urinary acylcarnitine analysis (Venter et al. 2015), and found significantly elevated palmitoylcarnitine (C_{16}) in the urine of the treatment failure group comparatively (136.19 vs. 58.22 mmol/L; P=0.0095). Additionally, although not statistically significant, the total short chain acylcarnitines were comparatively decreased in the treatment failure group (8.77 vs. 2.75 mmol/L; P=0.2298). Free carnitine (C_0) was also decreased (0.64 vs. 2.58 mmol/L; P=0.4329), as expected in such deficiencies, because these remain bound to the LCFAs and are unable to cross the mitochondrial membrane. Considering this diagnostic metabolite profile, a MTP deficiency or abnormality is suggested to occur in the treatment failure group (see Figure 3). MTP mutations are estimated at a prevalence of approximately 1:75 000 (Garg and Dasouki 2006), and according to unpublished results of the Potchefstroom Laboratory for Inborn Errors of Metabolism (PLIEM), may be as high as 1 in 100 for a mutation in any of the seven enzyme/protein systems mentioned above. Additional evidence for this, is the serum carnitine deficiencies previously reported in 47.7% of all TB-positive patients (Hatamkhani et al. 2014), which are also associated with these carnitine transporters (Flanagan et al. 2010). Furthermore, L-carnitine has been shown to play a significant role in T-cell-dependent

antibacterial activity in the host, and enhance the immune response via reduced macrophage and lymphocyte malfunction in TB patients (Emilio Jirillo et al. 1993; E Jirillo et al. 1991), hence a deficiency in L-carnitine, due to an underlying enzyme/protein system deficiency, may explain why these individuals did not respond to treatment. Confirmation of a reduced capacity for mitochondrial β-oxidation due to such a deficiency, is the elevated concentrations of fatty acids with odd numbered carbons, such as azelaic acid and heptanedioic (pimelic) acid in the treatment failure group, due to the above mentioned accumulating fatty acids undergoing peroxisomal α-oxidation prior to peroxisomal β-oxidation (Figure 3) (Van Veldhoven 2010).

Accumulation of fatty acyl-CoA derivatives, due to the defective LCFA transport, can also contribute to a compromised insulin secretion (Luies and Loots 2016). Long-chain fatty acyl-CoA's are strong inhibitors of glutamate dehydrogenase (GDH) (Lai et al. 1994), and since this enzyme is responsible for the conversion of glutamate to α-ketoglutarate, and an elevated flux of the tricarboxylic acid (TCA) cycle, eventually resulting in insulin secretion (see Figure 3) (Newsholme et al. 2006), suggests an additional means by which insulin secretion may be compromised in the treatment failure group. Further confirmation of this is the elevated levels of 3,4-dihydroxybutyric acid and 4-hydroxybutyric acid, indicating an increased metabolic flux in the gamma-aminobutyric acid (GABA) shunt pathway (Shinka et al. 2002), in the opposite direction of the glutamate to α-ketoglutarate reaction previously mentioned (see Figure 2). Compromised insulin secretion has also been shown to result in the accumulation of various branched chain amino acids (BCAAs) (Lu et al. 2013), which explains the elevated leucine (3-hydroxyvaleric acid, 3-hydroxy-3-methylglutaric acid) and isoleucine (2-methyl-3-hydroxybutyric acid, methylsuccinic acid, N-tiglylglycine) metabolite intermediates observed in the treatment failure group comparatively.

Whatever the proposed mechanism, the patients in the treatment failure group show a reduced capacity to transport LCFA into the mitochondria for metabolism into short-chain fatty acids (SCFAs), which are well known for their antimicrobial effects (i.e. increase the fluidity of the bacterial cell wall and subsequently negatively influence cell wall integrity (Royce et al. 2013), and induce host defense peptide LL-37 (Mily et al. 2015)). Considering this, apart from the reduced L-carnitine detected in the

treatment failure group, the reduced SCFAs may be an additional explanation as to why these individuals may not be responding to the treatment and why the metabolite markers suggest elevated disease severity.

4. CONCLUSION

This study highlights the capacity of metabolomics to identify markers which predict a poor response to treatment, and also better characterize or propose previously unknown mechanisms resulting in TB treatment failure. The most significant observations in this metabolomics study were the elevated levels of those metabolites associated with an imbalance in the gut microbiome. Since this influence xenobiotic uptake and toxicity, the synchronous use of probiotics for optimizing the microbiome during first-line anti-TB treatment, may improve treatment outcome, and could be a topic of further investigation. Another interesting observation was those metabolites associated with abnormalities in any of the three enzymes of the MTP complex in the treatment failure group. Since L-carnitine and SCFAs are also reduced in these individuals, and well known for their anti-mycobacterial properties, this metabolic pathway may explain why these individuals have an increased disease severity and/or a poor response to TB treatment.

5. REFERENCES

Antoine, D., French, C. E., Jones, J., & Watson, J. M. (2007). Tuberculosis treatment outcome monitoring in England, Wales and Northern Ireland for cases reported in 2001. *Journal of epidemiology and community health, 61*(4), 302-307.

Clemente, J. C., Ursell, L. K., Parfrey, L. W., & Knight, R. (2012). The impact of the gut microbiota on human health: an integrative view. *Cell, 148*(6), 1258-1270.

De Villiers, L., & Loots, D. T. (2013). Using metabolomics for elucidating the mechanisms related to tuberculosis treatment failure. *Current Metabolomics, 1*(4), 306-317.

Du Preez, I., & Loots, D. T. (2013). New sputum metabolite markers implicating adaptations of the host to *Mycobacterium tuberculosis*, and vice versa. *Tuberculosis, 93*(3), 330-337.

Dye, C., Watt, C. J., Bleed, D. M., Hosseini, S. M., & Raviglione, M. C. (2005). Evolution of tuberculosis control and prospects for reducing tuberculosis incidence, prevalence, and deaths globally. *Jama, 293*(22), 2767-2775.

Eisenhofer, G., Kopin, I. J., & Goldstein, D. S. (2004). Catecholamine metabolism: A contemporary view with implications for physiology and medicine. *Pharmacological reviews, 56*(3), 331-349.

Flanagan, J. L., Simmons, P. A., Vehige, J., Willcox, M. D., & Garrett, Q. (2010). Role of carnitine in disease. *Nutrition & metabolism, 7*(1), 1.

Garg, U., & Dasouki, M. (2006). Expanded newborn screening of inherited metabolic disorders by tandem mass spectrometry: Clinical and laboratory aspects. *Clinical biochemistry, 39*(4), 315-332.

Gonthier, M.-P., Cheynier, V., Donovan, J. L., Manach, C., Morand, C., Mila, I., et al. (2003). Microbial aromatic acid metabolites formed in the gut account for a major fraction of the polyphenols excreted in urine of rats fed red wine polyphenols. *The Journal of nutrition, 133*(2), 461-467.

Gonzalez, A., Stombaugh, J., Lozupone, C., Turnbaugh, P. J., Gordon, J. I., & Knight, R. (2011). The mind-body-microbial continuum. *Dialogues in clinical neuroscience, 13*(1), 55.

Hasanally, D., Chaudhary, R., & Ravandi, A. (2014). Role of phospholipases and oxidized phospholipids in inflammation. In *Phospholipases in Health and Disease* (pp. 55-72): Springer.

Hatamkhani, S., Khalili, H., Karimzadeh, I., Dashti-Khavidaki, S., Abdollahi, A., & Jafari, S. (2014). Carnitine for prevention of antituberculosis drug-induced hepatotoxicity: A randomized, clinical trial. *Journal of gastroenterology and hepatology, 29*(5), 997-1004.

Hesseling, A., Walzl, G., Enarson, D., Carroll, N., Duncan, K., Lukey, P., et al. (2010). Baseline sputum time to detection predicts month two culture conversion and relapse in non-HIV-infected patients. *The International Journal of Tuberculosis and Lung Disease, 14*(5), 560-570.

Heyes, M. P., Saito, K., & Markey, S. P. (1992). Human macrophages convert L-tryptophan into the neurotoxin quinolinic acid. *Biochem.J, 283*, 633-635.

Hong, Y. S. (2011). NMR-based metabolomics in wine science. *Magnetic Resonance in Chemistry, 49*(S1), S13-S21.

Horne, D. J., Royce, S. E., Gooze, L., Narita, M., Hopewell, P. C., Nahid, P., et al. (2010). Sputum monitoring during tuberculosis treatment for predicting outcome: systematic review and meta-analysis. *The Lancet Infectious Diseases, 10*(6), 387-394.

Jellum, E., Børresen, H. C., & Eldjarn, L. (1973). The presence of furan derivatives in patients receiving fructose-containing solutions intravenously. *Clinica Chimica Acta, 47*(2), 191-201.

Jirillo, E., Altamura, M., Marcuccio, C., Tortorella, C., De Simone, C., & Antonaci, S. (1993). Immunological responses in patients with tuberculosis and in vivo effects of acetyl-L-carnitine oral administration. *Mediators of inflammation, 2*(7), S17-S20.

Jirillo, E., Altamura, M., Munno, I., Pellegrino, N., Sabato, R., Fabio, S. D., et al. (1991). Effects of acetyl-L-carnitine oral administration on lymphocyte antibacterial activity and TNF-α levels in patients with active pulmonary tuberculosis. A randomized double blind versus placebo study. *Immunopharmacology and immunotoxicology, 13*(1-2), 135-146.

Kaushal, D. (2012). Eicosanoids, prostaglandins, and the progression of tuberculosis. *Journal of Infectious Diseases*, jis611.

Kronick, J., Mamer, O., Montgomery, J., & Scriver, C. (1983). Apparent gastrointestinal origin of cis-4-hydroxycyclohexanecarboxylic acid. *Clinica Chimica Acta, 132*(2), 205-208.

Lai, J. C., Liang, B. B., Zhai, S., Jarvi, E. J., & Lu, D. R. (1994). Brain mitochondrial citrate synthase and glutamate dehydrogenase: differential inhibition by fatty acyl coenzyme A derivatives. *Metabolic brain disease, 9*(2), 143-152.

Li, G., Zhang, H., Sader, F., Vadhavkar, N., & Njus, D. (2007). Oxidation of 4-methylcatechol: implications for the oxidation of catecholamines. *Biochemistry, 46*(23), 6978-6983.

Lu, J., Xie, G., Jia, W., & Jia, W. (2013). Insulin resistance and the metabolism of branched-chain amino acids. *Frontiers of medicine, 7*(1), 53-59.

Luies, L., & Loots, D. (2016). Tuberculosis metabolomics reveals adaptations of man and microbe in order to outcompete and survive. *Metabolomics, 12*(3), 1-9.

Mahapatra, S., Hess, A. M., Johnson, J. L., Eisenach, K. D., DeGroote, M. A., Gitta, P., et al. (2014). A metabolic biosignature of early response to anti-tuberculosis treatment. *BMC infectious diseases, 14*(1), 53.

Marklund, N., Östman, B., Nalmo, L., Persson, L., & Hillered, L. (2000). Hypoxanthine, uric acid and allantoin as indicators of in vivo free radical reactions. Description of a HPLC method and human brain microdialysis data. *Acta neurochirurgica, 142*(10), 1135-1142.

Mily, A., Rekha, R. S., Kamal, S. M., Arifuzzaman, A. S. M., Rahim, Z., Khan, L., et al. (2015). Significant effects of oral phenylbutyrate and vitamin D3 adjunctive therapy in pulmonary tuberculosis: A randomized controlled trial. *PLoS One, 10*(9), e0138340.

Newsholme, P., Brennan, L., & Bender, K. (2006). Amino acid metabolism, β-cell function, and diabetes. *Diabetes, 55*(Supplement 2), S39-S47.

Pallant, J. (2001). Manual SPSS Survival: A step by step guide to data analysis using SPSS. (pp. 302): Open University Press, McGraw-Hill Education, Philadelphia, USA.

Pettersen, J. E., & Jellum, E. (1972). The identification and metabolic origin of 2-furoylglycine and 2, 5-furandicarboxylic acid in human urine. *Clinica Chimica Acta, 41*, 199-207.

Puri, D. (2006). Textbook of Medical Biochemistry: Clinical cases, Chapter 13.2. In *Textbook of Medical Biochemistry* (Vol. 2nd ed., pp. 376). New-Delhi, India: Elsevier.

Royce, L. A., Liu, P., Stebbins, M. J., Hanson, B. C., & Jarboe, L. R. (2013). The damaging effects of short chain fatty acids on Escherichia coli membranes. *Applied Microbiology and Biotechnology, 97*(18), 8317-8327.

Russell, W. R., Hoyles, L., Flint, H. J., & Dumas, M.-E. (2013). Colonic bacterial metabolites and human health. *Current opinion in microbiology, 16*(3), 246-254.

Ryan, D., Robards, K., Prenzler, P. D., & Kendall, M. (2011). Recent and potential developments in the analysis of urine: A review. *Analytica Chimica Acta, 684*(1), 17-29.

Schoeman, J. C., Du Preez, I., & Loots, D. T. (2012). A comparison of four sputum pre-extraction preparation methods for identifying and characterising M. tuberculosis using GCxGC-TOFMS metabolomics. *Journal of microbiological methods, 91*, 301-311.

Shaw, W. (2013). Increased urinary excretion of a 3-(3-hydroxyphenyl)-3-hydroxypropionic acid (HPHPA), an abnormal phenylalanine metabolite of Clostridia spp. in the gastrointestinal tract, in urine samples from patients with autism and schizophrenia. *Nutritional neuroscience*.

Shinka, T., Inoue, Y., Ohse, M., Ito, A., Ohfu, M., Hirose, S., et al. (2002). Rapid and sensitive detection of urinary 4-hydroxybutyric acid and its related compounds by gas chromatography–mass spectrometry in a patient with succinic semialdehyde dehydrogenase deficiency. *Journal of Chromatography B, 776*(1), 57-63.

Silhavy, T. J., Kahne, D., & Walker, S. (2010). The bacterial cell envelope. *Cold Spring Harbor perspectives in biology, 2*(5), a000414.

Sim, K. G., Hammond, J., & Wilcken, B. (2002). Strategies for the diagnosis of mitochondrial fatty acid β-oxidation disorders. *Clinica Chimica Acta, 323*(1), 37-58.

Smuts, I., Van Der Westhuizen, F. H., Louw, R., Mienie, L. J., Engelke, U. F. H., Wevers, R. A., et al. (2013). Disclosure of a putative biosignature for respiratory chain disorders through a metabolomics approach. *Metabolomics, 9*(2), 379-391.

Taniguchi, K., & Armstrong, M. D. (1963). The enzymatic formation of o-hydroxyphenylacetic acid. *The Journal of biological chemistry, 238*, 4091-4097.

Tomás-Barberán, F. A., & Clifford, M. N. (2000). Dietary hydroxybenzoic acid derivatives–nature, occurrence and dietary burden. *Journal of the science of food and agriculture, 80*(7), 1024-1032.

Van Aardt, W. J., le Roux, J. M., Lindeque, J. Z., Mason, S., & Louw, R. (2016). The effect of temperature on the respiration and metabolism of the African burrowing scorpion (Opistophthalmus latimanus). *Comparative Biochemistry and Physiology Part D: Genomics and Proteomics, 20*, 50-56.

Van Veldhoven, P. P. (2010). Biochemistry and genetics of inherited disorders of peroxisomal fatty acid metabolism. *Journal of lipid research, 51*(10), 2863-2895.

Venter, L., Lindeque, Z., Van Rensburg, P. J., Van Der Westhuizen, F., Smuts, I., & Louw, R. (2015). Untargeted urine metabolomics reveals a biosignature for muscle respiratory chain deficiencies. *Metabolomics, 11*(1), 111-121.

Walzl, G., Ronacher, K., Siawaya, J. F. D., & Dockrell, H. M. (2008). Biomarkers for TB treatment response: Challenges and future strategies. *Journal of Infection, 57*(2), 103-109.

Wang, S.-Y., Kuo, C.-H., & Tseng, Y. J. (2012). Batch normalizer: A fast total abundance regression calibration method to simultaneously adjust batch and injection order effects in liquid chromatography/time-of-flight mass spectrometry-based metabolomics data and comparison with current calibration methods. *Analytical Chemistry, 85*(2), 1037-1046.

Westerhuis, J. A., Hoefsloot, H. C. J., Smit, S., Vis, D. J., Smilde, A. K., van Velzen, E. J. J., et al. (2008). Assessment of PLSDA cross validation. *Metabolomics, 4*(1), 81-89.

Wold, S., Esbensen, K., & Geladi, P. (1987). Principal component analysis. *Chemometrics and Intelligent Laboratory Systems, 2*(1), 37-52.

World Health Organization (2014). Definitions and reporting framework for tuberculosis – 2014 revision. (pp. 47). Geneva, Switzerland (WHO Press).

World Health Organization (2016). Global tuberculosis report 2016. (pp. 214). Geneva, Switzerland (WHO Press): World Health Organization.

Zhou, C., & Greenberg, M. M. (2012). Histone-catalyzed cleavage of nucleosomal DNA containing 2-deoxyribonolactone. *Journal of the American Chemical Society, 134*(19), 8090-8093.

FOOTNOTE PAGE

6 AUTHOR CONTRIBUTION

DTL conceptualized the study design; KR and GW provided all urine samples. LL performed the GC data acquisition/analysis and CM the UPLC data acquisition/analysis. LL, JM and DTL interpreted the data and drafted the article. All authors approved the final version to be submitted.

7 COMPLIANCE WITH ETHICAL REQUIREMENTS

7.1 Conflict of interest

The authors declare that there are no conflicts of interest, and that this manuscript, and the work described therein, is unpublished and has not been submitted for publication elsewhere.

7.2 Funding statement

The authors have no specific funding to report.

7.3 Ethical approval

Ethical approval for this investigation, conducted according to the Declaration of Helsinki and International Conference on Harmonization Guidelines, was obtained from the Ethics Committee of the North-West University, South Africa (reference number NWU-00127-11-A1), as well as from Stellenbosch University Health Research Ethics Committee (reference number 99/039) and Cape Town City Health. All participants gave written informed consent for study participation and HIV testing.

7. TABLE AND FIGURE LEGENDS

Table 1: The 50 urinary metabolite markers identified at time of diagnosis that best explain the variation detected between the successful and unsuccessful treatment outcome groups.

Figure 1: Principal components analysis (PCA) scores plots of principal component 1 versus principal component 2 of the successful and unsuccessful treatment outcome groups, showing clear differentiation of these two groups at (a) time of diagnosis and (b) week 26, due to variation in their underlying metabolite profiles. The variances accounted for are indicated in parenthesis.

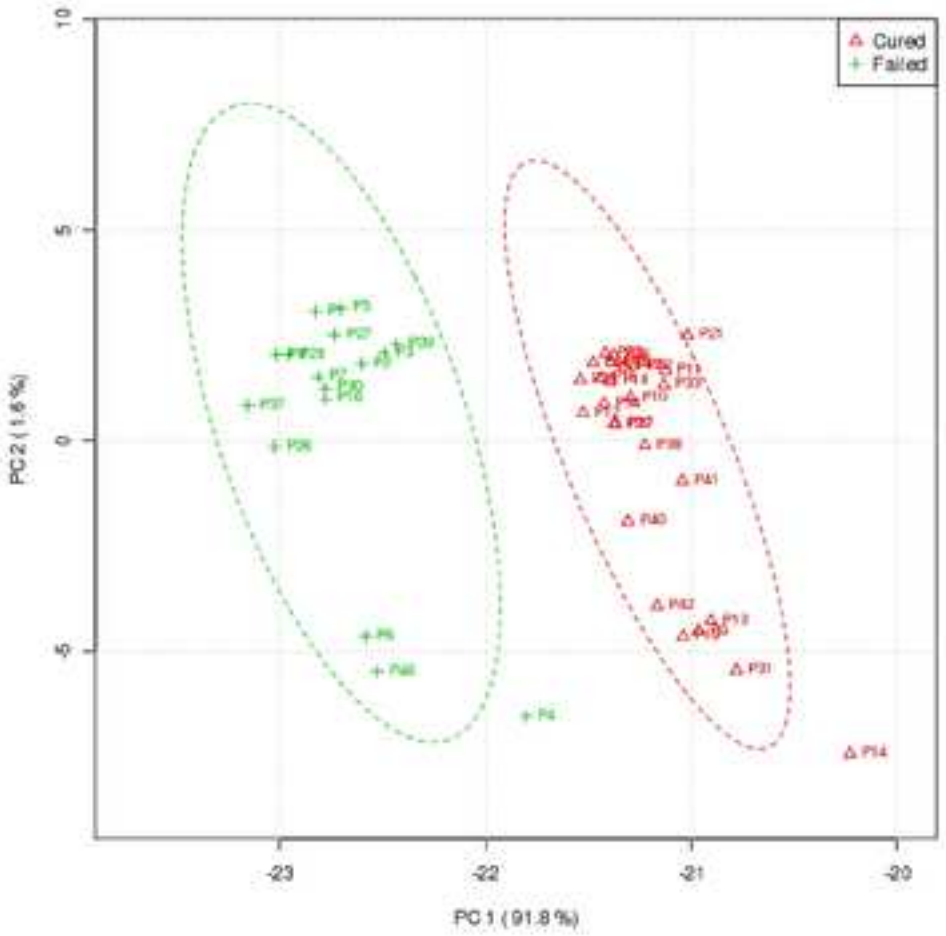
Figure 2: Amino acid metabolism of the identified urinary metabolite markers, which are either increased (↑) or decreased (↓) in the treatment failure group comparetively.

Figure 3: Fatty acid oxidation of the identified urinary metabolite markers, which are either increased (↑) or decreased (↓) in the treatment failure group comparetively. Abbreviations: CPT I, carnitine palmitoyltranferase 1; CPT II, carnitine palmitoyltransferase 2; CACT, carnitine-acylcarnitine translocase; VLCAD, very long chain acyl-CoA dehydrogenase; MCAD, medium chain acyl-CoA dehydrogenase; SCAD, short chain acyl-CoA dehydrogenase; M/SCHAD, medium/short chain hydroxyacyl-CoA dehydrogenase; MTP, mitochondrial trifunctional protein; SCYD, short chain enoyl-CoA hydratase; MCT, medium chain 3-ketoacyl-CoA thiolase; SCT, medium chain 3-ketoacyl-CoA thiolase; ACOX 1, acyl-CoA oxidase 1; ACOX 2, acyl-CoA oxidase 2; D-PBE, peroxisomal bifunctional enzyme; ACAA1, 3-ketoacyl-CoA thiolase; PHYH, Phytanoyl-CoA hydroxylase; HPLC2, Unknown; pistanalDH, pristanaldehyde dehydrogenase.

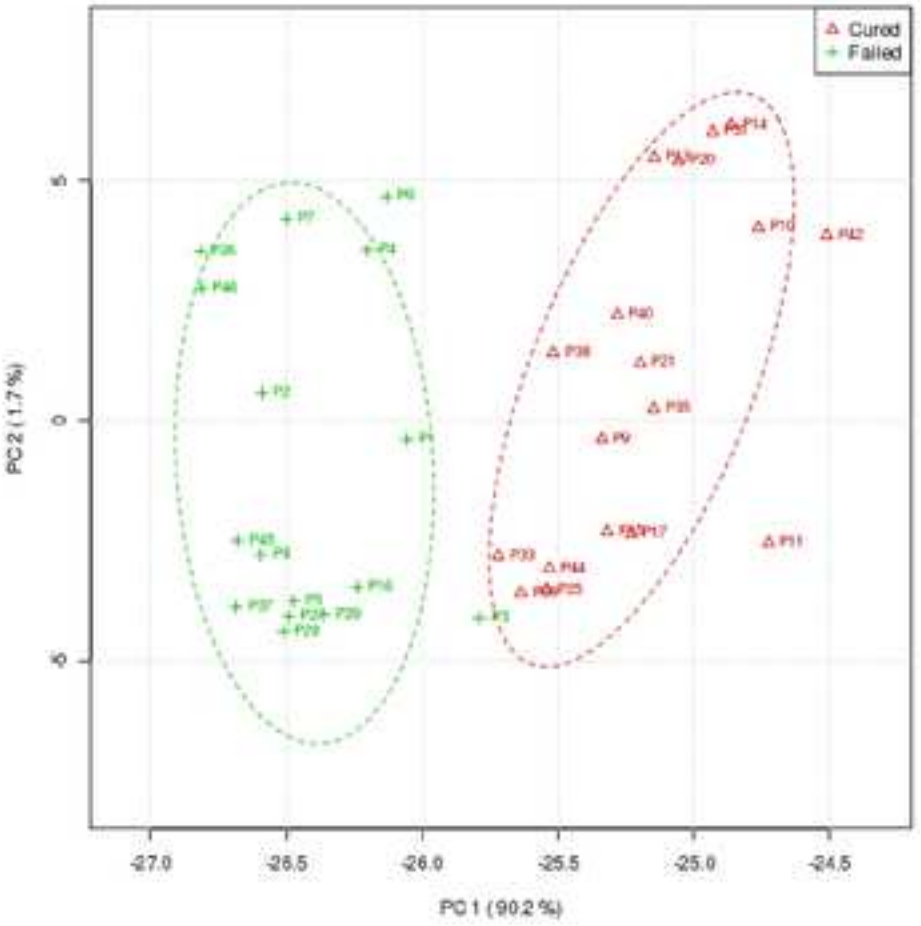
Supplementary figure: Principal components analysis (PCA) scores plots of principal component 1 versus principal component 2 of the successful and unsuccessful treatment outcome groups, at (a) time of diagnosis, (b) week 1, (c) week 2, (d) week 4 of treatment and (e) two weeks after treatment completion (week 26), subsequent to an organic acid extraction and GCxGC-TOFMS analyses. The explained variances are indicated in parenthesis.

Figure1

Click here to download Figure Figure_1.jpg



(a)



(b)

Figure2

Click here to download Figure Figure_2.png

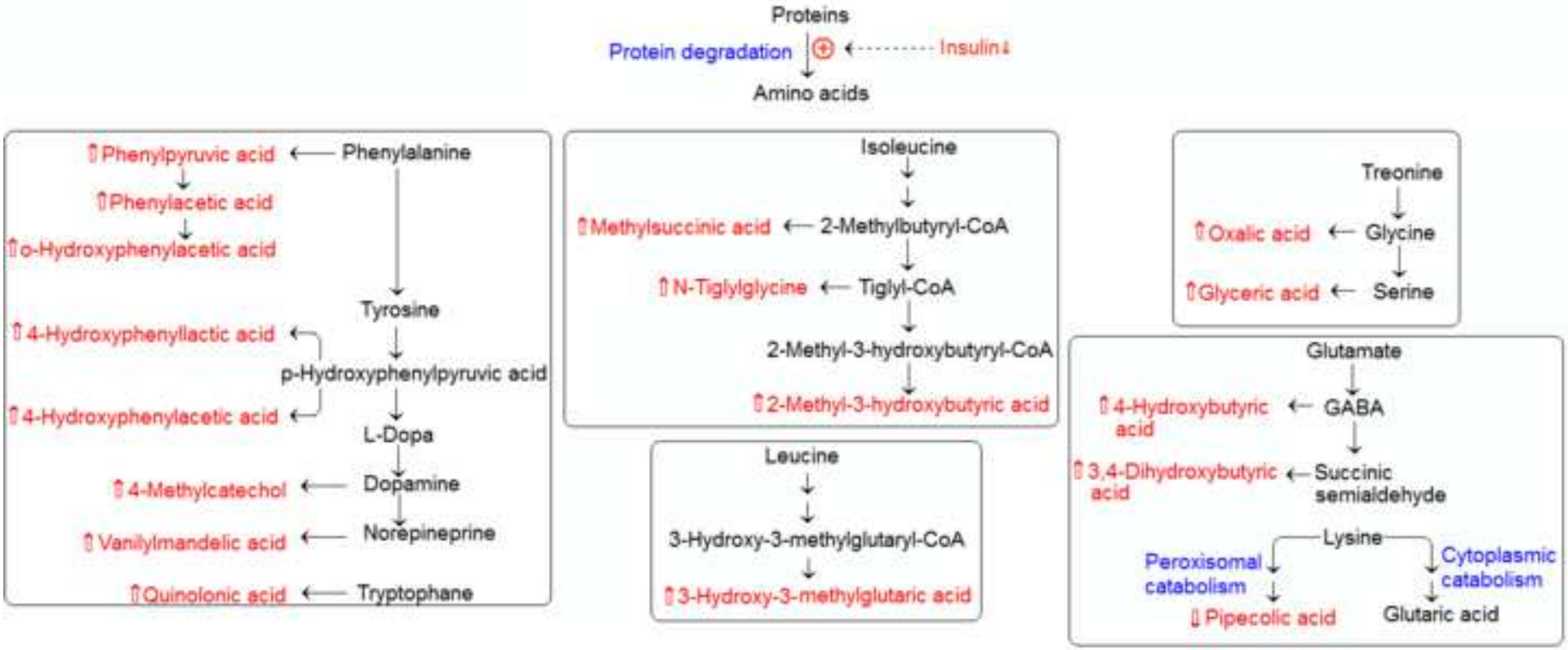
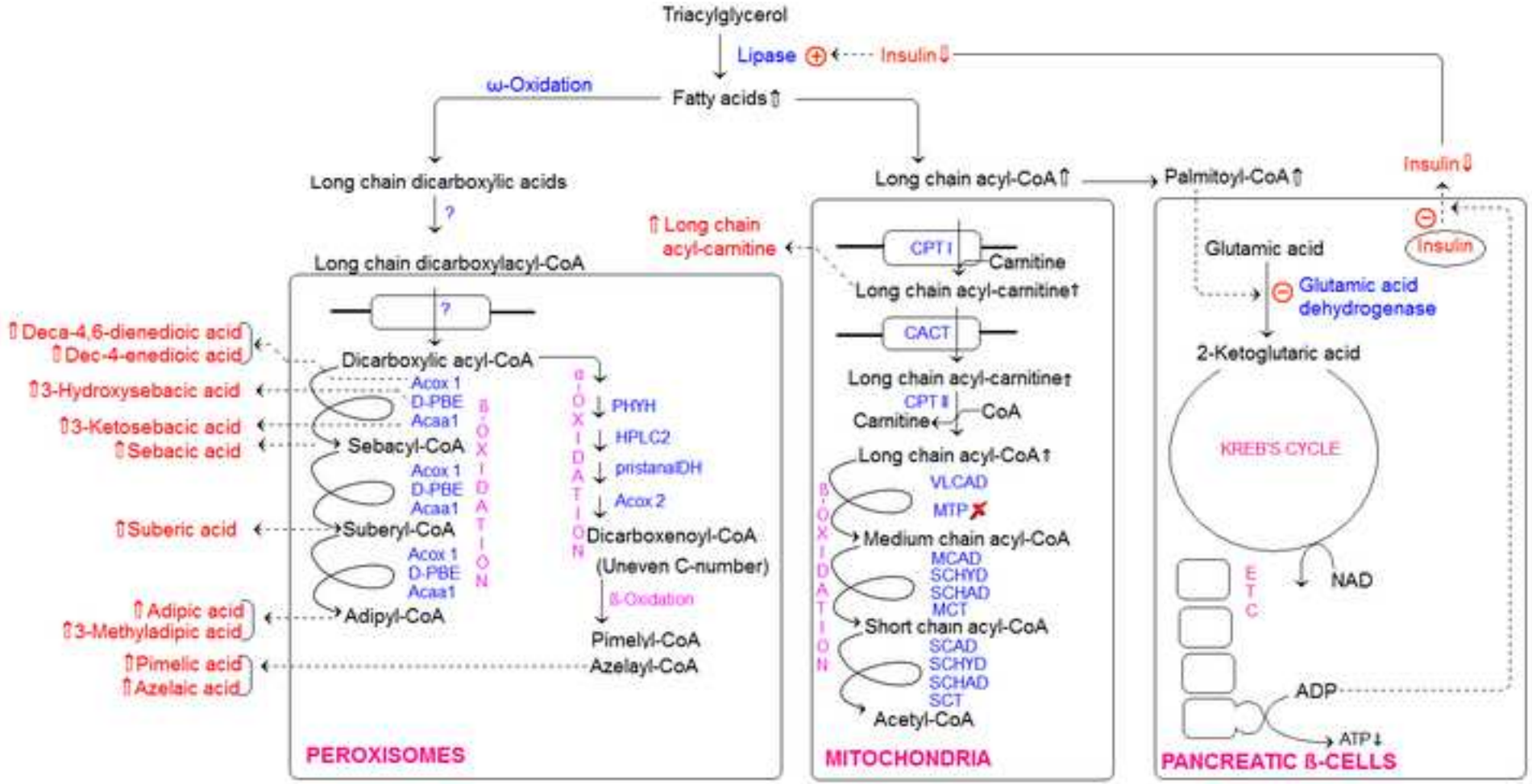


Figure3

Click here to download Figure Figure_3.png



PREDICTING TUBERCULOSIS TREATMENT OUTCOME USING METABOLOMICS

Journal:	*Biomarkers in Medicine*
Manuscript ID	BMM-2017-0133.R1
Manuscript Type:	Preliminary Communication
Keywords:	tuberculosis, Metabolomics, predicting treatment outcome

ABSTRACT

Background: Predicting a poor treatment outcome would offer significant benefits for patient care and for new drug development.

Methods/Results: Urine samples from TB-positive patients with a successful and unsuccessful treatment outcome were collected at baseline and analyzed. The identified metabolites were used in a forward logistic regression model, which achieved a ROC AUC of 0.94 (95% CI 0.84–1) and cross-validated well in a leave-one-out context, with an AUC of 0.89 (95% CI 0.7–1). Two possible predictors were identified, which are associated with a gut microbiota imbalance.

Discussion/Conclusion: Our findings show the capacity of metabolomics to predict treatment failure at time of diagnosis, which potentially offers significant benefits for use in new drug development clinical trials and individualized patient care.

Key words: *M. tuberculosis*; tuberculosis; metabolomics; predicting treatment outcome; treatment failure

INTRODUCTION

The latest reports indicate 10.4 million individuals are infected with *M. tuberculosis* globally, resulting in an estimated 1.8 million deaths from tuberculosis (TB), annually. Clearly, current strategies to control this pandemic are inadequate, especially in high-burden countries with limited resources, where the treatment target goals, as recommended by the World Health Organization (WHO), have not yet been met [1].

Currently, the most effective anti-TB treatment strategy, Directly Observed Treatment Short-course (DOTS), consists of an initial two month (intensive) phase using a combination of isoniazid (INH), rifampicin (RIF), pyrazinamide (PZA) and ethambutol (EMB), followed by a four month (maintenance) phase, consisting of only INH and RIF [2]. The DOTS strategy, when fully adhered to, has a relatively high success rate in patients with drug-susceptible TB [1]. TB treatment failure using this approach, however, still occurs with drug-sensitive TB, and is defined as the occurrence of persistently positive sputum smears or cultures at the end of treatment [3]. An additional concern is that these individuals, who are not treated successfully, remain infectious for prolonged periods of time, and continue to transmit the disease to others. Improved treatment outcomes and timely intervention is a priority for TB control programs and requires the identification of those factors associated with treatment failure [4]. Previous studies suggest the following risk factors for drug-susceptible TB treatment failure and relapse: (a) a less than 5% gain in bodyweight [5]; (b) age; (c) human immunodeficiency virus (HIV) co-infection; (d) diabetes type II, (e) illiteracy; (f) alcoholism and (g) prior TB treatment [6]. The most prominent contributor to this occurrence, however, is a slow patient response to treatment. It is also important to note that a number of novel immunological biomarkers (i.e cytokines and chemokines) have been proposed for monitoring TB treatment response and predicting treatment outcome [7], including those directly associated with *M. tuberculosis*, such as lipoarabinomannan [8, 9] or those associated with the overall inflammatory response, such as induced protein (IP)-10 [10-12].

Two diagnostic approaches are frequently used to detect active TB and monitor treatment outcome, namely sputum smear microscopy and bacteriological cultures, the latter of which is considered the gold standard. Response to treatment is usually determined based on sputum culture conversion after two months of treatment [13]. However, a period of two months or more before the first indications of treatment efficacy

can be determined, is considered too long, since the bacteria have had the opportunity to adapt and create drug-resistant strains [14, 15], and also considering that these patients may still transfer the disease to others during this time. Additionally, these treatment failure predictive methods require frequent laboratory monitoring, which is not only time consuming, but also costly. Considering this, there is an urgent need to search for alternative approaches or biomarkers which can be used for the early detection of treatment failure [16].

Over the last 10 years, metabolomics has proved to be a very effective tool to search for and identify new diagnostic and prognostic biomarkers and/or disease risk factors [17]. Metabolomics is defined as the unbiased identification and quantification of all intermediates of metabolism in a biological system, using specialized analytical techniques, in conjunction with biostatistical and mathematical analyses [18]. In 2012, Olivier and Loots applied a gas chromatography mass spectrometry (GC-MS) metabolomics approach to characterize and differentiate various infectious *Mycobacterium* species (i.e. *M. tuberculosis, M. bovis, M. kansasii and M. avium*) and *Pseudomonas aeruginosa* based on their characteristic lipid profiles. Twelve lipid biomarkers were identified and used to build a multivariate discriminant model, which could correctly assign unknown samples to their respective species groups with probabilities ranging from 72 to 100% [19]. Che *et al.* (2013) used a similar approach to compare serum samples collected from healthy controls and TB patients, before and after TB therapy, to find diagnostic markers for active TB, not influenced by anti-TB therapy. They demonstrated nine potential diagnostic TB biomarkers, one of which (i.e. 5-oxoproline) remained unaffected by first-line TB therapy [20]. These studies indicate the capacity of metabolomics to not only differentiate *M. tuberculosis* from other *Mycobacterium* species, but also to identify useful biomarkers that can be used for the early prediction of treatment outcome, and for possible use in new anti-TB drug development. Considering this, we used an untargeted two-dimensional gas chromatography time-of-flight mass spectrometry (GCxGC-TOFMS) urinary metabolomics approach to differentiate TB-positive patients with a successful (n=21) and unsuccessful (n=10) treatment outcome at time of diagnosis, and identified those metabolite markers which best differentiate the groups, in order to build a logistic regression model to possibly predict treatment failure before first-line anti-TB drug administration commences.

MATERIALS AND METHODS

Clinical samples

Anonymized archived urine samples were procured from the Faculty of Medicine and Health Sciences, NRF/DST Centre of Excellence for Biomedical Research, at the Stellenbosch University/MRC Centre for TB Research, from where they were transported (frozen at -80ºC) to the North-West University (NWU), Human Metabolomics: Infectious Disease Laboratory, for metabolomics analysis. These samples were part of a prospective observational cohort study of individuals with active pulmonary TB, diagnosed with *M. tuberculosis* infection using smear microscopy, bacteriological cultures and strain typing [21]. Representation of different *M. tuberculosis* strains was not significantly different between cured and failed participants. The samples were not collected at a specific time of day in order to ensure robustness of any identified metabolites. From the original samples, all treatment failure cases were included and matched by age, gender and extent of disease on chest X-rays to cured cases. Thus, the samples included were from 31 culture-confirmed active TB-positive South African patients (16 males and 15 females, aged between 17 and 58) at baseline (i.e. time of diagnosis, thus prior to initiation of treatment). These patients had drug-susceptible TB, were HIV-seronegative, not pregnant, and had no other diseases (i.e. diabetes, malignancy, lung cancer, chronic bronchitis and sarcoidosis), nor were they receiving any other medication, including antibiotics. The samples were divided into successful (n=21) and unsuccessful (n=10) treatment outcome groups.

Sample analysis

A creatinine value was determined for each urine sample collected using a creatinine enzyme kit (Thermo Scientific; reference number 981845) and analyzed using an Indiko Clinical Analyzer, Type 863 (Thermo Scientific). This value is used to determine the amount of urine, internal standard, bis-(trimethylsilyl)-trifluoroacetamide (BSTFA), trimethylchlorosilane (TMCS) and pyridine needed for each extraction and derivatization. An organic acid extraction was performed on all patient-collected urine samples, after which they were analyzed by randomly injecting each sample, along with the necessary quality control (QC) samples, into a Pegasus 4D GCxGC-TOFMS (LECO Africa (Pty) Ltd, Johannesburg, South Africa) and

processed in order to identify compounds by comparison of their mass spectra to libraries generated from previously injected standards, as described by Luies and Loots (2016) [22].

Statistical data analysis

It is standard practice in metabolomics investigations to perform various data clean-up procedures, prior to statistical data analysis [23]. The concentrations of all the identified metabolites were calculated relative to the internal standard and expressed in relation to the creatinine value. A 50% zero filter was applied to each variable [24] to eliminate compounds with more than 50% zero values in both groups. Any batch effects were corrected for using quantile equating [25], followed by a 50% quality control–coefficient of variation (QC–CV) filter. Finally, in order to select the most significant metabolite markers, various univariate statistical analyses (i.e. fold change, Mann-Whitney test and effect sizes) were applied to the untransformed data using MATLAB with Statistics and PLS Toolbox Release (2012; The MathWorks Inc., Natick, MA, USA) and SAS (2015; The SAS System for Windows Release 9.3 TS Level 1M0 Copyright© by SAS Institute Inc., Cary, NC, USA). Fold change compares the absolute value changes of specific variables across two groups, and fold changes greater than two are considered significant [26]. The Mann-Whitney test is considered the non-parametric alternative to the *t*-test, when assessing independent samples. This test determines significance by comparing the medians of variables across two groups, where a p-value below 0.05 is considered significant [27]. Effect sizes indicate practical significance by comparing variables individually between groups. For this non-parametric dataset, effect sizes were based on the associated Mann-Whitney z-values and an effect size above 0.3 (which indicates a moderate effect) were deemed practically relevant [28].

RESULTS

As indicated in Figure 1, compounds with an absolute fold change exceeding two, a Mann-Whitney p-value of 0.05 or lower and an effect size larger than 0.3 were selected as the metabolite markers most capable of predicting treatment failure, and are listed in Table 1 (n=18). This multi-selection approach is based on the assumption that these different statistical methods compensate for each other's limitations and should reduce the rate of false discovery of unimportant compounds [22]. A volcano plot was used to compare the size of

the fold change to the statistical significance level by plotting the $\log_{10}$ scaled Mann-Whitney p-values against the $\log_2$ scaled fold change values (see Figure 2).

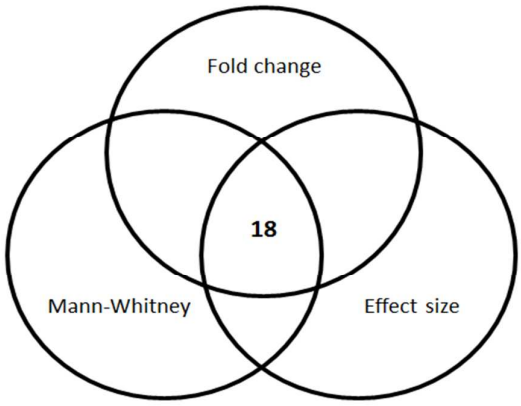


Figure 1: *Venn-diagram indicating the multi-selection approach used to select the most significant metabolite markers for predicting treatment failure.*

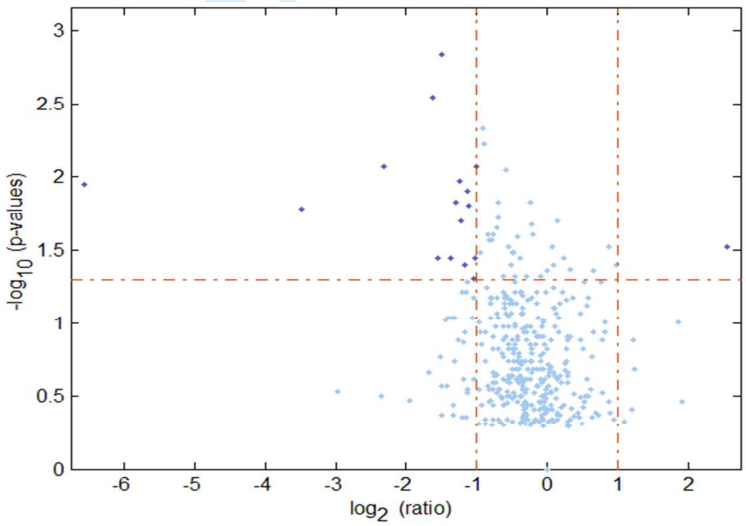


Figure 2: *Volcano plot of the $\log_{10}$ scaled Mann-Whitney p-values against the $\log_2$ scaled fold change values.*

Table 1: *The 18 univariate urine metabolite markers considered for potentially predicting treatment outcome at time of diagnosis.*

Metabolite name	Retention time (RT) (s)	Successful treatment outcome: Average concentration (mg/g creatinine) (Standard deviation)	Samples detected (n=21)	Unsuccessful treatment outcome: Average concentration (mg/g creatinine) (Standard deviation)	Samples detected (n=10)	Fold change (absolute value)	Mann-Whitney test (*p*-value)	Mann-Whitney Effect Size (ω value)
2,3,4,5-Tetrahydroxypentanoic acid-1,4-lactone	1396,1.59	17.509 (12.067)	95.2%	35.154 (35.4)	100%	2.008	0.008	0.429
2-Hydroxy-1,2,3-propanetricarboxylic acid	1573,1.52	13.388 (6.72)	100%	31.249 (36.752)	100%	2.334	0.020	0.368
2-Methoxy-5-methylphenol	1018,2.12	0.489 (0.35)	95.2%	5.458 (13.715)	100%	11.171	0.016	0.383
2-Octenoic acid	1021,1.69	0.576 (0.6)	95.2%	1.364 (1.018)	100%	2.367	0.011	0.414
3,5-Dihydroxybenzoic acid	1573,1.64	1.054 (0.49)	100%	2.374 (1.655)	100%	2.251	0.001	0.573
3-(4-Hydroxy-3-methoxyphenyl)propionic acid	1720,1.78	7.285 (4.019)	100%	22.472 (21.333)	100%	3.085	0.003	0.497
3-Hydroxy-hexanedioic acid	1444,1.48	1.737 (0.779)	100%	3.554 (3.813)	100%	2.046	0.050	0.296
3-Methoxy-4-hydroxy-benzenepropanoic acid	1642,1.97	1.077 (1.799)	100%	2.649 (2.383)	100%	2.460	0.015	0.391
cis-4-Decene-1,10-dioic acid	1612,1.71	0.679 (0.706)	100%	1.759 (1.714)	100%	2.591	0.036	0.323
Phthalic acid mono-2-ethylhexyl ester	2128,2.53	11.231 (39.786)	100%	1.926 (1.077)	100%	5.832	0.030	0.338
tert-Butylhydroquinone	1363,1.58	0.45 (0.725)	90.5%	2.234 (2.357)	100%	4.969	0.008	0.429
Unknown 1	1489,1.7	0.43 (0.364)	100%	0.967 (0.858)	100%	2.251	0.040	0.315
Unknown 2	1648,1.77	0.158 (0.179)	95.2%	0.347 (0.232)	100%	2.201	0.013	0.402
Unknown 3	1753,1.59	0.087 (0.096)	100%	0.248 (0.21)	100%	2.838	0.001	0.535
Unknown 4	2050,1.84	98.212 (168.565)	100%	289.043 (416.387)	100%	2.943	0.036	0.323
Unknown 5	2101,1.85	8.294 (7.618)	100%	16.9 (13.627)	100%	2.038	0.036	0.323
Unknown 6	2116,2.05	0.129 (0.112)	95.2%	12.211 (36.959)	100%	94.334	0.011	0.410
Unknown 7	520,1.65	1.324 (0.511)	95.2%	2.874 (2.568)	100%	2.172	0.016	0.387

DISCUSSION

The predictive ability of these compounds (Table 1) were compared and combined by performing forward, backward and stepwise selection logistic regression. Therefore, all metabolites in Table 1 were given an equal opportunity to form part of the final model since all were provided to the logistic regression procedure. The procedure then considers various models, but penalizes models with higher numbers of predictors. Metabolites are added (forward), removed (backward) or both (stepwise) until a model is found which performs the best in terms of possible prediction and number of predictors. The final model was based on forward selection and was chosen as it showed the most predictive potential (i.e. had the highest area under the receiver operating characteristic) — it should be noted however that not all compounds with a high fold change or low Mann-Whitney *p*-value are necessarily good predictors (see Figure 3). The model contained two metabolite markers as potential predictors of treatment outcome, namely 3,5-dihydroxybenzoic acid and 3-(4-hydroxy-3-methoxyphenyl)propionic acid. The model showed acceptable fit (goodness-of-fit) with a non-significant Hosmer-Lemeshow statistic (p=0.444), as well as a reasonable relationship between treatment outcome and the above predictors with a maximum rescaled R-squared statistics of 0.75. Both predictors made significant contributions to the prediction, as indicated by the Wald criterion reported in Table 2.

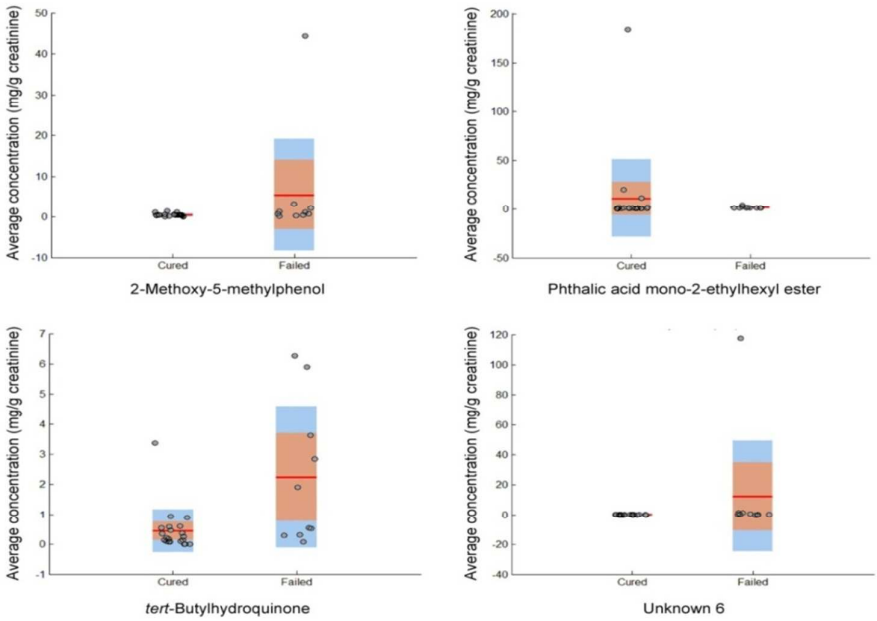


Figure 3: *Box-like plots of the top four fold change metabolites, demonstrating that not all compounds are necessarily good predictors when judged on their fold change value only.*

Table 2: *The maximum likelihood parameter estimates and odds ratios.*

Parameter	Estimate	Wald Chi-square	p-value	95% CI for Estimates	Odds Ratio (unit=1)	95% CI for Odds Ratio
3,5-Dihydroxybenzoic acid	3.24	5.71	0.017	1.2 – 6.8	25.6	3.2 – 886.1
3-(4-Hydroxy-3-methoxyphenyl) propionic acid	0.29	3.53	0.06	0.1 – 0.7	1.3	1.1 – 2.0

The odds ratio, reported for each of the possible predictors in Table 2, is a measure of the association between a specific feature (i.e. metabolite) and a treatment outcome (i.e. failure to respond to treatment) [29]. Thus, the odds ratio represents the increased or decreased probability that a patient will not be cured successfully (treatment failure) after anti-TB drug treatment, for every unit increase or decrease in the metabolite measured. An odds ratio of greater than 20 is considered significant for a diagnostic biomarker [17]. An increase of one unit in the levels of 3,5-dihydroxybenzoic acid and 3-(4-hydroxy-3-methoxyphenyl)propionic acid are associated with an increased risk of treatment failure, i.e. the odds of treatment failure increase by 25.6 and 1.3, respectively, for each one unit increase. Although the odds ratio associated with 3-(4-hydroxy-3-methoxyphenyl)propionic acid are not indicative of a biomarker, the combination of multiple "weak" metabolite markers into a single model may also provide a high level of discrimination [17]. The large confidence intervals (CIs) are due, as least in part, to the small groups used here and the relatively large variation within these groups. As with all newly developed models, further validation of these possible predictors is an absolute requirement prior to application.

The predictive ability of the model was assessed based on the receiver operating characteristic (ROC) curve, as shown in Figure 4. The model achieved an area under the ROC (AUC) value of 0.94 (95% CI 0.84 – 1) and cross-validated well in a leave-one-out context, with an AUC equal to 0.89 (95% CI 0.7 – 1). The predictive ability of a model is established if it achieves an AUC of 0.9 or higher [17], however, to truly validate the model will require a larger, independent sample cohort. Therefore, we focused our attention on interpreting the two possible predictors in the model.

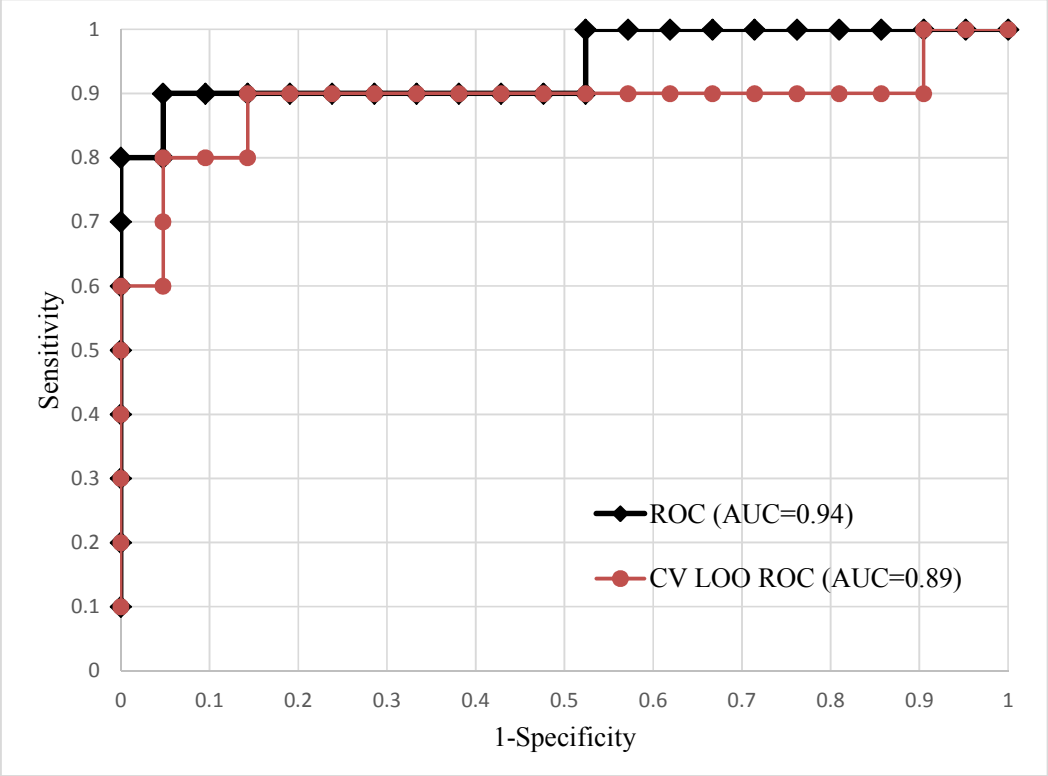


Figure 4: *Receiver operating characteristic (ROC) curve of the final logistic regression model. The black line represents the ROC when all samples are used to access sensitivity and specificity of the model, while the red line represents the ROC when cross-validated using a leave-one-out approach.*

The box-like plots of these two possible predictors (and their chemical structures), as illustrated in Figure 5, indicate the individual variation in the data. These plots were generated using notBoxPlot.m, a MATLAB function developed by Rob Campbell (http://www.mathworks.com/matlabcentral/fileexchange/26508-raacampbell13-notboxplot). The raw data is scattered over a 95% confidence interval (red box) and one standard deviation (blue box) of the mean (red line). The individual observations (indicated by the grey dots) show a much larger variation in the unsuccessful treatment outcome group comparatively, as well as some overlap. It is therefore evident from these graphs that neither predictor will perform well in isolation. We also include a bivariate scatter plot (see Figure 6) to graphically illustrate the combined predictive ability of 3,5-dihydroxybenzoic acid and 3-(4-hydroxy-3-methoxyphenyl)propionic acid to discriminate between individuals with successful and unsuccessful treatment outcome.

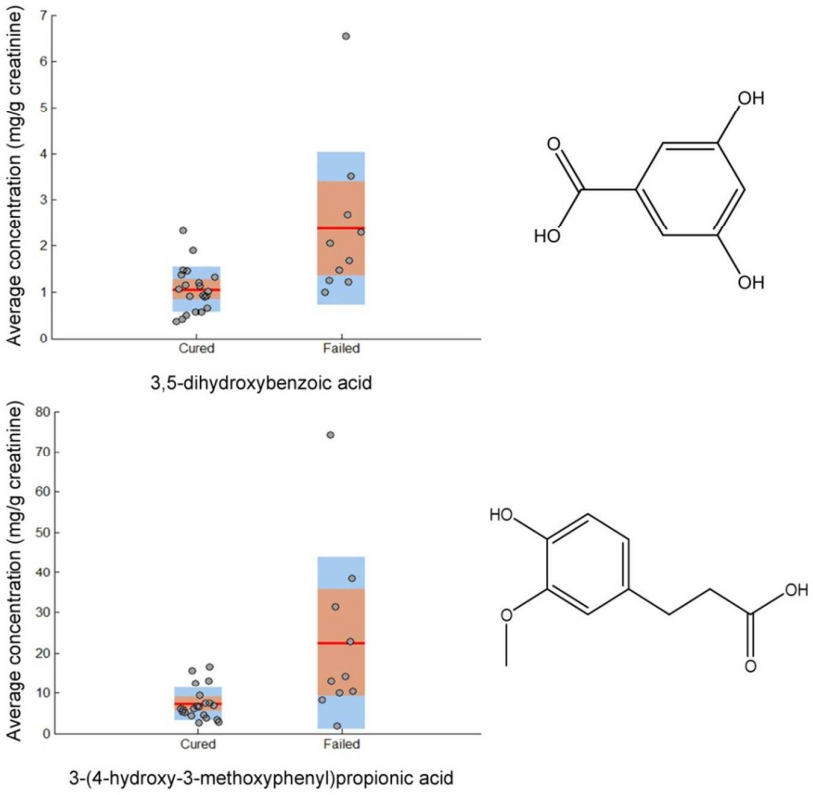


Figure 5: *Box-like plots of the two identified possible predictors, with the accompanying chemical structures.*

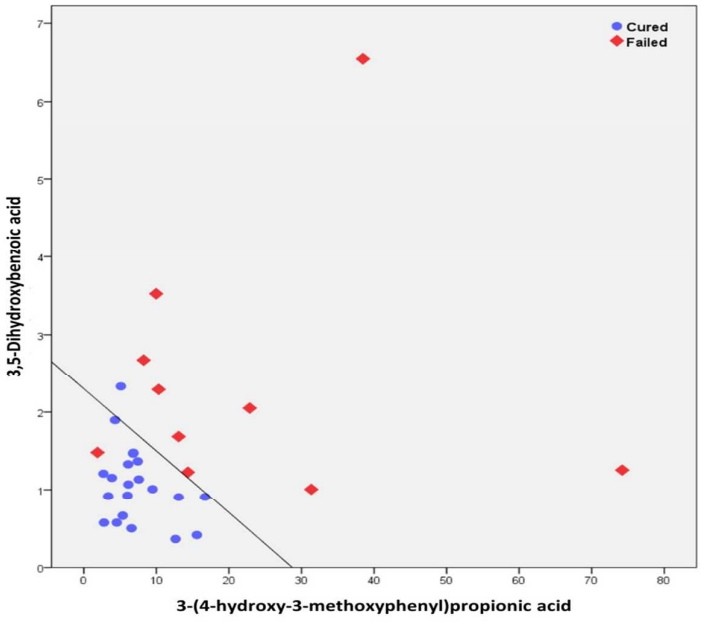


Figure 6: *Bivariate scatter plot of the raw data generated for 3,5-dihydroxybenzoic acid versus 3-(4-hydroxy-3-methoxyphenyl)propionic acid. The line indicates clear separation between the two comparative groups based on the two metabolites selected as predictors.*

When considering these two possible predictors from a biological mechanistic perspective, both 3,5-dihydroxybenzoic acid (α-resorcylic acid) [30] and 3-(4-hydroxy-3-methoxyphenyl)propionic acid (dihydroferulic acid) [31] are compounds with aromatic rings associated with a microbial-derived microbial imbalance [30, 31], and also with a TB disease state [22, 32]. Hence, an altered microbiome in the treatment failure group may be contributing to the lacking efficacy of the anti-TB drugs [33]. These results also confirm previous findings by Luies et al., (2017), who identified a number of metabolites related to a microbial imbalance, using a combination of multivariate and univariate statistical approaches, which characterized treatment failure to first-line drugs [34]. Since the innate and adaptive immune response systems rely on various microbiota interactions to promote immune cell maturation and function, a microbial imbalance results in a weakened immune system and an inability to fight disease [35]. Additionally, the gut microbiota is considered an important contributor to xenobiotic/drug bioavailability and toxicity since it interacts with various drugs, influencing factors relating to their absorption, distribution and secretion, and hence plasma concentrations [36].

Compounds with aromatic rings [33, 37, 38], including the majority of the anti-TB drugs used in this investigation (INH, RIF and PZA), and their phase II metabolites, are transported across cell membranes and eliminated via the ATP-binding cassette (ABC)-type multidrug transporters, which are highly expressed in human enterocytes of the liver, intestine, blood-brain barrier, blood-testis barrier, placenta and kidneys [38, 39] as well as on the cell surface of various intersinal microbiota [40]. If for whatever reason (such as a microbial imbalance for instance), there is an overexpression/activity of these proteins on the host's cell surfaces, it would lead to an elevated excretion of these aromatics, explaining the elevated concentrations of the 3,5-dihydroxybenzoic acid and 3-(4-hydroxy-3-methoxyphenyl)propionic acid in the urine of the treatment failure patient group, suggesting that these aromatic anti-TB drugs would have suffered the same fate, possibly explaining treatment failure.

The genome of *M. tuberculosis* also contains 2.5% ABC transporters, which may be classified as either importers or exporters, determined by the direction of translocation of substrates [39]. The known mechanisms of drug-resistance to first-line anti-TB drugs, can be summarized as follows: INH-resistance is due to mutations in the *katG* and *inhA* genes; RIF-resistance results from alterations or point mutations in the *rpoB* gene of RNA polymerase; PZA-resistance may result from various mutations in the *pncA*, *rpsA* and

panD genes; and EMB-resistance is mainly due to mutations on the *embB* gene [41]. However, these mutations do not explain the mechanism of drug-resistance in all clinical isolates, suggesting that additional, unknown mechanisms may also be at play, such as the production of enzymes/proteins which modify/degrade drugs resulting in inactivation, cell wall impermeability and efflux mechanisms. Although such individual mechanisms may not result in clinical resistance, the combinations of these may confer high-level resistance, which may help explain the variation in resistance seen in *M. tuberculosis* isolates [39].

Although, the scope of this investigation was only to identify metabolite markers and test their capacity for predicting treatment failure, using patient-collected urine samples, and not to elucidate the possible mechanisms for treatment failure, the above mentioned hypotheses are interesting topics for future investigation, and also serve as further validation for their possible use for predicting a failed treatment outcome to first-line anti-TB drugs. Furthermore, it may be interesting to compare these patients to healthy humans with a normal intestinal biosis in order to firmly allocate the microbiota dysbiosis to the treatment failure cohort. Additionally, the issue of whether the metabolites analyzed in the urine of these patients are due to human host metabolism or whether they originate from the commensal intestinal bacteria can be addressed in future by experiments with germ-free and conventionalized *M. tuberculosis*-infected laboratory animals.

CONCLUSION

We were able to identify two urinary metabolite markers which possibly predict treatment failure of TB patients to first-line anti-TB drugs, before treatment onset. When these possible predictors were combined into a logistic regression model, they performed well with regards to sensitivity and specificity using leave-one-out cross-validation. It is important to note that the purpose of this model was to identify a combination of variables with high discriminatory ability containing as few variables as needed to achieve this, for potential diagnostic purposes. We propose that these two possible predictors warrant further investigation and validation using a bigger sample cohort. Once validated, the development of a prediction rule or predictive model can commence. Nonetheless, our findings show the capacity of metabolomics to identify patients at greater risk of a poor treatment outcome, at time of diagnosis, which potentially offers significant benefits for use in new drug development clinical trials and an individualized approach to patient care.

SUMMARY POINTS

- Several independent studies have suggested that an unsuccessful sputum culture conversion after two months of treatment is a risk factor for treatment failure and can possibly be used for predicting this outcome. However, predicting a poor treatment outcome well before two months would offer significant benefits for patient care and also in the development of new anti-TB agents.
- We identified metabolite markers in urine samples collected from TB-positive patients at time of diagnosis (before treatment commences) using GCxGC-TOFMS metabolomics, which when used in a logistic regression model, serve as possible predictors for TB treatment failure.
- The predictive ability of the model was assessed based on the ROC curve, and achieved an AUC of 0.94 (95% CI 0.84 – 1) and cross-validated well in a leave-one-out context, with an AUC of 0.89 (95% CI 0.7 – 1).
- These two identified possible predictors are also associated with a gut microbiota imbalance, confirming previously proposed mechanisms related to treatment failure in these individuals.

REFERENCES

1. World Health Organization. Global tuberculosis report 2015. (2015). http://apps.who.int/iris/bitstream/10665/191102/1/9789241565059_eng.pdf
2. Olivier I, Loots DT. An overview of tuberculosis treatments and diagnostics. What role could metabolomics play. J. Cell Tiss. Res. 11(1), 2655-2671 (2011).
3. Mukherjee JS, Rich ML, Socci AR et al. Programmes and principles in treatment of multidrug-resistant tuberculosis. Lancet 363(9407), 474-481 (2004).
4. Alobu I, Oshi DC, Oshi SN, Ukwaja KN. Profile and determinants of treatment failure among smear-positive pulmonary tuberculosis patients in Ebonyi, Southeastern Nigeria. Int. J. Mycobacteriol. 3(2), 127-131 (2014).
5. Krapp F, Veliz JC, Cornejo E, Gotuzzo E, Seas C. Bodyweight gain to predict treatment outcome in patients with pulmonary tuberculosis in Peru. Int. J. Tuberc. Lung Dis. 12(10), 1153-1159 (2008).
6. Albuquerque MDFPM, Ximenes RaDA, Lucena-Silva N et al. Factors associated with treatment failure, dropout, and death in a cohort of tuberculosis patients in Recife, Pernambuco State, Brazil. Cad. Saude Publica 23(7), 1573-1582 (2007).
7. Walzl G, Ronacher K, Djoba Siawaya JF, Dockrell HM. Biomarkers for TB treatment response: challenges and future strategies. J. Infect. 57(2), 103-109 (2008).
8. Lawn SD, Kerkhoff AD, Vogt M, Wood R. Diagnostic accuracy of a low-cost, urine antigen, point-of-care screening assay for HIV-associated pulmonary tuberculosis before antiretroviral therapy: a descriptive study. Lancet Infect. Dis. 12(3), 201-209 (2012).
9. Lawn SD, Kerkhoff AD, Burton R et al. Diagnostic accuracy, incremental yield and prognostic value of Determine TB-LAM for routine diagnostic testing for tuberculosis in HIV-infected patients requiring acute hospital admission in South Africa: a prospective cohort. BMC Med. 15(1), 67 (2017).
10. Cannas A, Calvo L, Chiacchio T et al. IP-10 detection in urine is associated with lung diseases. BMC Infect. Dis. 10(1), 333 (2010).
11. Petrone L, Cannas A, Aloi F et al. Blood or urine IP-10 cannot discriminate between active tuberculosis and respiratory diseases different from tuberculosis in children. BioMed Res. Int. 2015 (2015).
12. Petrone L, Cannas A, Vanini V et al. Blood and urine inducible protein 10 as potential markers of disease activity. Int. J. Tuberc. Lung Dis. 20(11), 1554-1561 (2016).
13. Horne DJ, Royce SE, Gooze L et al. Sputum monitoring during tuberculosis treatment for predicting outcome: systematic review and meta-analysis. Lancet Infect. Dis. 10(6), 387-394 (2010).
14. Baumann R, Kaempfer S, Chegou NN et al. Serodiagnostic markers for the prediction of the outcome of intensive phase tuberculosis therapy. Tuberculosis (Edinb) 93(2), 239-245 (2013).
15. Siawaya JFD, Bapela NB, Ronacher K et al. Immune parameters as markers of tuberculosis extent of disease and early prediction of anti-tuberculosis chemotherapy response. J. Infect. 56(5), 340-347 (2008).
16. Namukwaya E, Nakwagala FN, Mulekya F, Mayanja-Kizza H, Mugerwa R. Predictors of treatment failure among pulmonary tuberculosis patients in Mulago hospital, Uganda. Afr. Health Sci. 11(3), S105-S111 (2011).
17. Dunn WB, Wilson ID, Nicholls AW, Broadhurst D. The importance of experimental design and QC samples in large-scale and MS-driven untargeted metabolomic studies of humans. Bioanalysis 4(18), 2249-2264 (2012).

18. De Villiers L, Loots DT. Using metabolomics for elucidating the mechanisms related to tuberculosis treatment failure. Current metabolomics 1(4), 306-317 (2013).
19. Olivier I, Loots DT. A comparison of two extraction methods for differentiating and characterising various Mycobacterium species and Pseudomonas aeruginosa using GC-MS metabolomics. Afr. J. Microbiol. Res. 6(13), 3159-3172 (2012).
20. Che NY, Cheng JH, Li HJ et al. Decreased serum 5-oxoproline in TB patients is associated with pathological damage of the lung. Clin. Chim. Acta 423 5-9 (2013).
21. Hesseling AC, Walzl G, Enarson DA et al. Baseline sputum time to detection predicts month two culture conversion and relapse in non-HIV-infected patients. Int. J. Tuberc. Lung Dis. 14(5), 560-570 (2010).
22. Luies L, Loots D. Tuberculosis metabolomics reveals adaptations of man and microbe in order to outcompete and survive (vol 12, 40, 2016). Metabolomics 12(3), 1-9 (2016).
23. Van Den Berg RA, Hoefsloot HC, Westerhuis JA, Smilde AK, Van Der Werf MJ. Centering, scaling, and transformations: improving the biological information content of metabolomics data. BMC Genomics 7(1), 142 (2006).
24. Smuts I, Van Der Westhuizen FH, Louw R et al. Disclosure of a putative biosignature for respiratory chain disorders through a metabolomics approach. Metabolomics 9(2), 379-391 (2013).
25. Wang X, Zhang A, Sun H. Urine metabolomics. Clin. Chim. Acta 414 65-69 (2012).
26. Du Preez I, Loots DT. New sputum metabolite markers implicating adaptations of the host to *Mycobacterium tuberculosis*, and vice versa. Tuberculosis (Edinb) 93(3), 330-337 (2013).
27. Pallant J. Manual SPSS Survival: A step by step guide to data analysis using SPSS. 302 (2001).
28. Ellis SM, Steyn HS. Practical significance (effect sizes) versus or in combination with statistical significance (p-values): Research note. Management dynamics: Journal of the southern african institute for management scientists 12(4), p. 51-53 (2003).
29. Field A. Discovering statistics using SPSS statistics. (3). Sage Publications Ltd, London. (2013).
30. Liu C, Kuei C, Zhu J et al. 3,5-Dihydroxybenzoic acid, a specific agonist for hydroxycarboxylic acid 1, inhibits lipolysis in adipocytes. J. Pharmacol. Exp. Ther. 341(3), 794-801 (2012).
31. Rechner AR, Spencer JP, Kuhnle G, Hahn U, Rice-Evans CA. Novel biomarkers of the metabolism of caffeic acid derivatives in vivo. Free Radic. Biol. Med. 30(11), 1213-1222 (2001).
32. Singh VV, Toskes PP. Small Bowel Bacterial Overgrowth: Presentation, Diagnosis, and Treatment. Curr. Treat. Options Gastroenterol. 7(1), 19-28 (2004).
33. Estudante M, Morais JG, Soveral G, Benet LZ. Intestinal drug transporters: an overview. Adv. Drug Del. Rev. 65(10), 1340-1356 (2013).
34. Luies L, Mienie J, Motshwane C, Ronacher K, Walzl G, Loots D. Urinary metabolite markers characterizing tuberculosis treatment failure (Unpublished manuscript). (2017).
35. Clemente JC, Ursell LK, Parfrey LW, Knight R. The impact of the gut microbiota on human health: an integrative view. Cell 148(6), 1258-1270 (2012).
36. Gonzalez A, Stombaugh J, Lozupone C, Turnbaugh PJ, Gordon JI, Knight R. The mind-body-microbial continuum. Dialogues Clin. Neurosci. 13(1), 55-62 (2011).
37. Warren G, Houslay M, Metcalfe J. Cholesterol is excluded from the phospholipid annulus surrounding an active calcium transport protein. Nature 255 684-687 (1975).
38. Vasiliou V, Vasiliou K, Nebert DW. Human ATP-binding cassette (ABC) transporter family. Hum. Genomics 3(3), 281-290 (2009).

39. Louw GE, Warren RM, Gey Van Pittius NC, Mcevoy CR, Van Helden PD, Victor TC. A balancing act: efflux/influx in mycobacterial drug resistance. Antimicrob. Agents Chemother. 53(8), 3181-3189 (2009).
40. Mercado-Lubo R, Mccormick BA. The interaction of gut microbes with host ABC transporters. Gut Microbes 1(5), 301-306 (2010).
41. Zhang Y, Yew WW. Mechanisms of drug resistance in *Mycobacterium tuberculosis*: update 2015. Int. J. Tuberc. Lung Dis. 19(11), 1276-1289 (2015).

Reference annotations:

* Lawn *et al*. 2017

This study determined the diagnostic yield, accuracy and prognostic value of urine-lipoarabinomannan (LAM) testing. They reported that the diagnostic yield of urine-LAM was unrelated to respiratory symptoms, and that a positive urine-LAM status was strongly associated with a poor prognosis.

* Dunn *et al*. 2012

This review provides an understanding/importance of ROC (AUC) and odds ratios, as well as how to interpret these, which forms an important part of this manuscript.

* Louw *et al*. 2009

This minireview describes ATP-binding cassette (ABC)-type multidrug transporters in the context of *M. tuberculosis* and drug delivery.