Imperial College London

Investigation of Host Metabolic Interactions in Trypanosomatid Infections

Thesis submitted by

Sabrina Danielle Lamour

For the degree of Doctor of Philosophy of Imperial College London

Supervised by Dr Jasmina Saric and Prof. Elaine Holmes

Division of Computational Systems Medicine

Department of Surgery and Cancer

Faculty of Medicine

Imperial College London

United Kingdom

November 2015

Abstract

Infections with parasitic trypanosomatids cause serious neglected tropical diseases, which include cutaneous leishmaniasis (CL), causing skin ulceration, and human African trypanosomiasis (HAT), which causes flu-like symptoms followed by neurological features and is fatal if left untreated. Current disease management in both cases is inappropriate owing to suboptimal diagnosis and limited drugs with variable efficacy and severe adverse effects. Further research into their pathogeneses is critical for the discovery of novel diagnostic tools and therapeutic targets.

Experimental *Leishmania (L.) major* infection in two murine CL models with differing pathologies, and HAT caused by *Trypanosoma brucei (T. b.) rhodesiense* infection, were investigated from a systemic metabolic perspective, using ^{1}H nuclear magnetic resonance (NMR) spectroscopy. This was combined with cytokines and (for CL models) faecal microbial diversity measures, as additional markers of host response.

In addition to the discovery of a range of *Leishmania* infection-associated markers, including putative parasitic end-products alanine and succinate, significant divergences in measures were observed between the self-healing and non-healing CL mouse models. Such differences included variations in glucose and lipid metabolism, in inflammatory status, and in faecal *Clostridia* and *Gammaproteobacteria* abundance. Striking differences were also observed in the degree of statistical links between these measures, using newly developed correlation network analyses, implying that a rapid and inter-connected systems response was required for effective parasitic elimination.

This project also represents the first study to provide a comprehensive description of host metabolic changes in HAT. A total of 16 and 31 significant discriminatory metabolites were determined in plasma *via* ^{1}H NMR spectroscopy and ultra-performance liquid chromatography/mass spectrometry (UPLC-MS), respectively, including membrane phospholipids and amino acids, linked with inflammation. These findings provide a unique insight into underlying metabolic interactions that may contribute to pathology, and reveal a set of infection-related markers with possible diagnostic potential for CL and HAT.

Declarations

Statement of Originality

I certify that this thesis and the research to which it refers are the product of my own work, and that any ideas, contributions or quotations from the work of other people, published or otherwise, are fully acknowledged in accordance with the standard referencing practices of the discipline.

Copyright Declaration

The copyright of this thesis rests with the author and is made available under a Creative Commons Attribution Non-Commercial No Derivatives licence. Researchers are free to copy, distribute or transmit the thesis on the condition that they attribute it, that they do not use it for commercial purposes and that they do not alter, transform or build upon it. For any reuse or redistribution, researchers must make clear to others the licence terms of this work.

Acknowledgements

First and foremost, I would like to give a massive thank you to my supervisors Dr Jasmina Saric and Professor Elaine Holmes for their continuous support, guidance, mentoring, and advice throughout my time at Imperial, for giving me the opportunity to get involved in such fascinating research but also allowing and encouraging me to explore new ideas. I am deeply grateful to you both.

I would also like to sincerely thank my collaborators — including Dr Jeremy Sternberg, Aberdeen Univerisity (HAT study), Professor Simon Croft and Dr Karin Seifert, LSHTM (*Leishmania* murine study), Dr Julian Marchesi, Cardiff Univerisity (gut microbial profiling), Dr Matthew Robinson, LSHTM (*Leishmania* parasitic burden analyses) and Dr Michael Delves, Imperial College London (*Plasmodium* gametocyte study) — who, in addition to providing me with samples, technologies and/or lending their expertise for this project, have also given me extra guidance and helped shape our joint publications.

Much of the presented results heavily relied on statistical input from Dr Joram Posma and Dr Kirill Veselkov, who I would like to also give a big thank you to, for helping me focus my PhD research questions from a statistical point of view, providing me with customised MATLAB codes, and continuously providing statistical support. Thanks also to Dr Caroline Sands, who helped customise some MATLAB codes for me.

I would to thank the following people who have helped train me to use new techniques and/or aided in sample acquisition for my PhD: fellow colleagues Dr Maria Gomez-Romero and Dr Panos Vorkas (for UPLC-MS work and analyses), Dr Emilie Giraud at Dr M. Rogers' lab at LSHTM (for parasitic burden analyses), Dr James Evans at Dr J. Marchesi's lab at Cardiff University (for faecal DNA extraction and PCR), and lastly, fellow/former colleagues Dr Anthony Dona, Dr Richard Barton and Ms Samantha Lodge, for their guidance and trouble-shooting help with NMR.

Additionally, I would like to give thanks to the whole of the department for their support, friendship and banter, particularly Dr Mike Kyriakides, for his much appreciated nerdy advice and comments (even when not always asked for!), Nick Adesina-Georgiadis, for his greatly entertaining stories and stats advice, and who, together with Panos, always ensured interesting lunchtime

discussions. Also thank you to the fellow PhD students and/or former MRes classmates, including Julia Langer (especially for advice and motivational pushes), Robert Whiffin, Tsz "T" Law, Gabriel Valbuena, and Lea Maitre for keeping me sane and entertained throughout my Masters/PhD-life.

Lastly, I would like to sincerely thank my family and friends for always believing in me, pushing me and supporting me all the way, and for being a constant source of strength, happiness and love in my life. Also, to my mother Patricia, sister Akili, Sheryl and Steve, and April "Sephy", thank you also for helping to look after Elijah so that I could focus on thesis writing and rush off to meetings. To my nan (Nanny), thank you for always offering your advice (and constantly reminding me that I am in the "university of life"). Finally, to Jason... thank you for everything.

Unless specified otherwise, the work presented in this PhD Project was carried out at Imperial College London, within Computation Systems Medicine (CSM), Department of Surgery and Cancer. The work was funded by my Medical Research Council doctoral training award (MRC-DTA G1000390) and the Welcome Trust (Sir Henry Fellowship awarded to Dr Jasmina Saric, project grant 082786), with additional laboratory resources provided from the previously mentioned collaborators, for work carried out in their respective laboratories.

Abbreviations and Symbols

(d)	Doublet peak
(m)	Multiplet peak
(s)	Singlet peak
(t)	Triplet peak
1D	One dimensional
2D	Two dimensional
3HB	3-hydroxybutryate
A.U.	Arbitrary units
BC	Control BALB/c mice
BCAA	Branched-chain amino acids
BT	*L. major* infected BALB/c mice
CC	Control C57BL/6 mice
CL	Cutaneous Leishmaniasis
CE	Cholesterol ester
CE-MS	Capillary electrophoresis coupled to mass spectrometry
Cer	Ceramide
CPMG	Carr–Purcell Meiboom–Gill spin echo (NMR) pulse program
CSF	Cerebral spinal fluid
CT	*L. major* infected C57BL/6 mice
CV	Coefficient of variation
CXCL1	Chemokine [C-X-C motif] ligand 1
D-2	Day -2, prior to infection/injection
D_2O	Deuterium Oxide ("heavy" water)
D2	Day 2, post-infection/injection

D6	Day 6, post-infection/injection
D12	Day 12, post-injection (terminal day for control mice)
ds	Number of dummy scans
Dterm	Terminal day for infected mice
ESI	Electron spray ionisation
ESI(+)	ESI, positive mode
ESI(−)	ESI, negative mode
FA	Formic acid
FID	Free induction decay
g	Gram
g	Relative centrifugal force
GC	Gas chromatography
GPC	Glycerophosphocholine
^{1}**H**	Hydrogen nucleus/proton
H_2O	Distilled water
HAT	Human African Trypanomiasis
HILIC	Hydrophilic interaction chromatography
H-PCA	Hierarchical principal component analysis
HPLC	High performance liquid chromatography
ID	Identification
IFNγ	Interferon-gamma
IL	Interleukin
IPA	Isopropanol
K	Kelvin
kHz	Kilohertz
L.	*Leishmania*
LC	Liquid chromatography
LysoPC	Lysophosphatidylcholine
M	Molar
MAS-NMR	Magic-Angle Spinning NMR
mg	Milligram

mins	Minutes
MHz	Megahertz
mM	Millimolar
mm	Millimetre
MS	Mass spectrometry
μl	Microlitre
n	Number of samples/variables
ns	Number of acquired scans
NS	Not statistically signficant
NAG	*N*-acetylglycoprotein
NMR	Nuclear magnetic resonance
O-PLS-DA	Orthogonal projection to latent structures discriminatory analysis
OTU	Operational taxonomic unit
p	P-value, estimated probability of rejecting the null hypothesis
PBS	Phosphate buffered saline
PCA	Principal component analysis
PC	Principal Component (for PCA)/Phosphatidylcholine (for lipid analyses)
PCV	Packed cell volume
PE	Phosphatidylethanolamine
PLS-DA	Projection to latent structures- discriminatory analysis
ppm	Parts per million
q	Q-value, adjusted P-value following multiple testing correction
Q^2	Parameter of predictive ability of multivariate model for assigning samples correctly, based on x spectral data (Q^2X) or on y-matrix classifier data (Q^2Y)
QC	Quality control
Q-ToF	Quadrupole mass analyser coupled to a Time-of-flight mass analyser
R^2	Fraction of variation within data explained by multivariate model ("goodness of fit" parameter), based on x spectral data (R^2X) or y-matrix classifier data (R^2Y)
r	Correlation coefficient
ROS	Reactive oxygen species
rpm	Rotations per minute

RT	Room temperature
s	Seconds
SCFA	Short-chain fatty acids
σ	Chemical shift
SEM	Standard error of the mean
SM	Sphingomyelin
sp	species (singular)
spp	species (plural)
ssp	sub-species (plural)
STOCSY	Statistical total correlation spectroscopy
T.	*Trypanosoma*
T. b.	*Trypanosoma brucei*
TCA	Tricarboxylic acid
TG	Triglyceride
Th	T helper (cell)
TNFα	Tumour-necrosis-factor-alpha
ToF	Time-of-flight mass analyser
UK	Unknown/unassigned metabolite
UV	Unit variance
UPLC	Ultra performance liquid chromatography
WHO	World Health Organization

Contents

Abstract **2**

Declarations **3**

Acknowledgements **4**

Abbreviations and Symbols **6**

List of Figures **16**

List of Tables **17**

1 Introduction **18**

1.1 Trypanosomatid Parasites Cause Neglected Tropical Diseases 18

1.2 Cutaneous Leishmaniasis 21

1.2.1 Impact and Clinical Presentation of the Leishmaniases 21

1.2.2 Diagnosis and Treatment of CL 22

1.2.3 The Life-cycle and Biology of Leishmania Parasites 24

1.2.4 Immune Response to Leishmania 29

1.3 Human African Trypanosomiasis 30

1.3.1 Impact and Clinical Presentation of HAT 30

1.3.2 Diagnosis and Treatment of HAT 33

1.3.3 The Life-cycle and Biology of Trypanosoma brucei ssp 36

1.3.4 Host Immune Factors in HAT 38

1.4 Metabolic Profiling of Parasitic Infections 39

1.5 Project Aims and Thesis Outline 41

1.6 Additional Projects 42

2 General Materials and Methods **43**

2.1 Methodology Overview . . . 43

2.2 Sample Preparation for ^{1}H NMR Analysis . . . 44

2.2.1 Basic Principles of NMR Spectroscopy . . . 44

2.2.2 NMR Buffer Reagents . . . 46

2.2.3 Preparation of Mouse Urine and Plasma Samples . . . 46

2.2.4 Preparation of Mouse Tissue and Faecal Extracts . . . 47

2.2.5 Mouse Sample Preparation for MAS-NMR . . . 48

2.2.6 Clinical Plasma and CSF Preparation for NMR . . . 49

2.3 NMR Sample Acquisition and Data Processing . . . 49

2.3.1 Mouse Sample Acquisition and Processing for NMR . . . 49

2.3.2 Clinical Sample Acquisition and Processing . . . 50

2.4 Clinical Sample Preparation for MS . . . 51

2.4.1 Principle of UPLC-MS . . . 51

2.4.2 Human Plasma Preparation for Lipid-MS . . . 55

2.4.3 Lipid-MS Acquisition and Pre-processing . . . 57

2.5 Metabolic Analysis and Metabolite Identification . . . 58

2.5.1 Multivariate Analyses for NMR . . . 58

2.5.2 NMR Metabolite Assignment . . . 61

2.5.3 MS Data Processing and Multivariate Analysis . . . 62

2.5.4 UPLC-MS Metabolic Feature Assignment and Nomenclature . . . 63

2.6 Immune and Gut Microbial Profiling . . . 64

2.6.1 Cytokine Measurements . . . 64

2.6.2 Faecal Microbiota Diversity Analysis . . . 67

2.6.3 Univariate Statistical Analyses of Cytokine and Microbial Data . . . 74

2.6.4 Multiple Test Corrections . . . 74

2.7 Correlation Analyses between Datasets . . . 75

3 Results: Systemic Characterisation of Murine Strains **77**

3.1 Background . . . 77

3.2 Objectives . . . 79

3.3 Experimental Methods . . . 79

3.3.1 Animal Study Design and Sample Collection . . . 79

3.3.2 Data Acquisition and Analyses . . . 82

3.4 Results . . . 84
3.4.1 Experimental Physiological Measurements . . . 84
3.4.2 Batch Adjustment of Metabolic Datasets . . . 84
3.4.3 Metabolic Characterisation of Biofluids . . . 88
3.4.4 Metabolic Characterisation of Tissues . . . 91
3.4.5 Plasma Cytokine Profiles . . . 95
3.4.6 Faecal Microbial Results . . . 96
3.4.7 Hierarchical PCA of Metabolic, Cytokine and Microbial Profiles . . . 100
3.5 Discussion . . . 103
3.5.1 Inherent Differences in Haematocrit between Strains . . . 103
3.5.2 Metabolic Characterisation Reveals Strain-Differences on a Systemic and Local Level . . . 103
3.5.3 Strain-related Variation in Faecal Microbial Families . . . 105
3.5.4 Cytokine Measurements Demonstrates Intrinsic Pro-inflammatory Profile in BALB/c mice . . . 106
3.5.5 Study Limitations . . . 107
3.5.6 Conclusions . . . 107

4 Results: Systemic Characterisation of Contrasting Murine Models of Cutaneous Leishmaniasis 109

4.1 Background . . . 109
4.2 Objectives . . . 111
4.3 Experimental Methods . . . 111
4.3.1 Parasite Culture Methods . . . 111
4.3.2 Murine Leishmaniasis Study Overview . . . 112
4.3.3 Data Acquisition and Analyses . . . 114
4.4 Results . . . 116
4.4.1 CL Model Physiological and Pathological Measures . . . 116
4.4.2 Effect of L. major Infection on Host Metabolism in Non-healing CL Model . 116
4.4.3 Strain Adjustment of Data to Determine Infection-related Responses . . . 119
4.4.4 Comparison of Individual Infection-Associated Host Responses between Non-healing and Self-healing CL Models . . . 120
4.4.5 Interactions between Two Sets of Infection-Related Host Responses via 2D Matrix Correlation Analyses . . . 124

4.4.6 Discriminatory Co-analyses of Metabolic, Cytokine and Microbial Datasets that Differentiate CL Models . . . 131
4.4.7 Parasitic Burden Associated with a Set of Metabolic, Immune, and Microbial Markers . . . 133
4.4.8 Correlation Network Analyses for Interactive Systems Host Response towards L. major Infection . . . 133
4.5 Discussion . . . 139
4.5.1 Pathological Measures Indicate Contrasting Phenotype between Two CL Models already in Early Infection . . . 139
4.5.2 L. major Infection Associated with Changes in Glucose and Lipid Turnover and in Gut-Microbial Metabolism . . . 139
4.5.3 Infection-related Differences between CL Models Observed in Every Assessed Response System . . . 140
4.5.4 Links between Metabolic, Immune and Microbial Markers Highlight a Systematic Response to L. major Infection . . . 143
4.5.5 Study Limitations . . . 144
4.5.6 Conclusions and Future Work . . . 145

5 Results: Metabolic and Immune Characterisation of Human African Trypanomiasis (*rhodesiense*) 147

5.1 Background . . . 147
5.2 Objectives . . . 148
5.3 Experimental Methods . . . 149
5.3.1 Ethics . . . 149
5.3.2 Subject Information and Sample Collection . . . 149
5.3.3 Data Acquisition and Analyses . . . 150
5.4 Results: Plasma Analyses . . . 154
5.4.1 Metabolic Profiling via ^{1}H NMR . . . 154
5.4.2 Lipid Profiling via UPLC-MS . . . 159
5.4.3 Plasma Cytokine Analyses . . . 167
5.5 Results: CSF Analyses . . . 168
5.5.1 Metabolic Profiling via ^{1}H NMR . . . 168
5.5.2 CSF Cytokine Analyses . . . 172
5.6 Discussion . . . 174

5.6.1 Strong HAT-associated Differences in Plasma Linked to Changes in Lipid and Amino Acid Metabolism . . . 175
5.6.2 Disease-related Changes in CSF Provide Insights into Neuropathology of HAT 178
5.6.3 Inflammation Observed in HAT Linked with Altered Metabolism in Host . . 179
5.6.4 Study Limitations . . . 180
5.6.5 Conclusions and Future Work . . . 181

6 General Discussion and Concluding Remarks 183

6.1 Summary of Findings . . . 183
6.1.1 Systemic Characterisation of Murine Strains . . . 183
6.1.2 Sytemic Characterisation of Contrasting Murine Models of CL . . . 184
6.1.3 Metabolic and Immune Characterisation of HAT . . . 186
6.2 General Considerations and Challenges . . . 188
6.3 Recommendations for Future Work . . . 190
6.4 Conclusion . . . 191

7 Bibliography 193

A Appendix 219

List of Appendix Figures . . . 219
List of Appendix Tables . . . 220
Appendix Figures and Tables . . . 221
Publications . . . 245
Permissions for Third Party Copyright Works . . . 309

List of Figures

1.1 The Main Morphological Classes of Trypanosomatidae 20
1.2 Diverse Dermal Manifestations of Cutaneous Leishmaniasis 22
1.3 Life-cycle of Leishmania Parasites 25
1.4 Central Carbon Metabolism of Leishmania Parasites 28
1.5 Distribution of Human African Trypanosomiasis 32
1.6 Life-cycle of Trypanosoma brucei Parasites 37

2.1 A Representative ^{1}H NMR Spectrum of Mouse Urine 45
2.2 Main Instrument Components of UPLC-MS used for Lipid Profiling 52
2.3 A Representative Chromatogram and MS Spectrum via UPLC-MS 56
2.4 Geometric Representation of PCA 59
2.5 Comparison between PLS-DA and O-PLS-DA 61
2.6 Solid-Phase "Sandwich" ELISA 65
2.7 Schematic of Polymerase Chain Reaction Cycles 71

3.1 Overview of Experimental Murine Studies 81
3.2 Batch and Strain Variation in Plasma Metabolic Profiles of Plasma across the Two Murine Studies 87
3.3 O-PLS-DA Plot of Strain Differences in Plasma 89
3.4 Urinary Metabolic Profiles from Day 12 Different from Other Time-points 91
3.5 Metabolic Variation across Tissue Extracts 93
3.6 Elevated Pro-inflammatory Cytokines in Plasma of BALB/c Mice 96
3.7 Erysipelotrichia Class Found Significantly Higher in C57BL/6 than in BALB/c Mice 98
3.8 Variations in Faecal Bacterial Family Composition Show Partial Strain Separation . 99
3.9 Differentially Expressed Faecal Microbial Families in the Two Mouse Strains 100
3.10 Hierarchical PCA Demonstrates Urinary Metabolites as Most Discriminatory Markers between Two Mouse Strains 102

4.1 Parasitic Levels Measured in the Lesion, Liver Spleen of CL Models . . . 117
4.2 Metabolic Variation Linked with *L. major* Infection in Plasma and Urine of BALB/c Mice . . . 118
4.3 Elevated Pro-inflammatory Cytokines in Plasma of BALB/c Mice . . . 123
4.4 Differences in Dominant Faecal Bacterial Composition between CL Models . . . 124
4.5 Changes in Relative Abundance of Less Prevalent Faecal Microbiota Observed in Self-Healing Model . . . 125
4.6 Variation in Immune-Metabolic Correlation Patterns Associated with Infection in two CL Models . . . 126
4.7 Differences in Metabolic and Microbial Responses to L. major Infection Dominate Separation between Self-Healing and Non-healing Models . . . 132
4.8 Correlation Analyses between Parasitic Burden and Host Response Markers in Non-healing and Self-healing CL models . . . 134
4.9 Correlation Network Wheels of Host Responses to Infection Show Differences in Inter-connectivity between CL Models and Time-points . . . 135
4.10 Correlation Network Maps of Host Responses Highlight Hubs in CL Models . . . 138

5.1 Overview of HAT Patient Sample Numbers Analysed Across the Three Platforms . . 153
5.2 Representative ^{1}H NMR Spectra of Human Plasma . . . 155
5.3 Strong Metabolic Response Associated with HAT . . . 155
5.4 Permutation Analyses of PLS-DA Model Comparing HAT Patients with Controls . . 156
5.5 Marked Disparities in Plasma Lipid Profiles between HAT Patients and Controls . . 160
5.6 Selection of HAT Discriminatory Lipid Markers . . . 161
5.7 Weak Correlations Observed between Plasma ^{1}H NMR and UPLC-MS Markers . . . 166
5.8 Elevated Concentrations of Selected Cytokines in Plasma HAT Patients . . . 167
5.9 No Dominant Metabolic Variation in CSF ^{1}H NMR Phenotypes Associated with HAT Disease or Diagnostic Stage . . . 169
5.10 Specific Set of CSF Metabolites Differentially Expressed between HAT Patients and Controls . . . 170
5.11 Metabolic Changes in CSF Associated with Drowsiness . . . 173
5.12 Elevated Concentrations of Selected Cytokines in Plasma HAT Patients . . . 174
5.13 IL-10 and IL-6 Display Similar Correlation Patterns with CSF Metabolites . . . 175

List of Tables

1.1 Standard Anti-trypanosomal Drug Treatments for Ealy and Late Stage HAT 34

2.1 Buffer Preparations used for NMR Spectroscopy 47
2.2 Sample Weight and Solvent Volumes used for Metabolite Extraction 48

3.1 Weight and PCV Measures from Control Animals in Mouse Study 1 84
3.2 Weight and PCV Measures from Mouse Study 2 85
3.3 Batch Differences Observed Before and After Batch-Adjustment of Data 85
3.4 Strain Discriminatory Metabolites in Biofluids 90
3.5 Evaluation of Time-related Differences Metabolic Profiles of Biofluids 90
3.6 Strain Discriminatory Metabolites in Tissues 94

4.1 Metabolic Changes Associated with *L. major* Infection in the Non-Healing Model . . 119
4.2 Discriminatory Infection-Associated Metabolites in Biofluids 121
4.3 Discriminatory Infection-Associated Metabolites in Tissues 122
4.4 Immune–Metabolic Correlations in Plasma, Urine and Faeces of CL Models 127
4.5 Immune–Metabolic Correlations in Tissues of CL Models 128
4.6 Microbial–Metabolic Correlations in Plasma and Faeces of CL Models 129
4.7 Microbial–Metabolic Correlations in Urine of CL Models 130
4.8 Immune–Microbial Correlations in CL Models 131

5.1 Patient and Control Cohort Demographic and Clinical Information 151
5.2 Plasma Metabolite Integrals Measured via ^{1}H NMR Spectroscopy 157
5.3 Accuracy of the Collection of Plasma NMR Discriminatory Markers for HAT 158
5.4 Disease Discriminatory Lipid Species Measured in Plasma via UPLC-MS 163
5.5 Accuracy of the Collection of Plasma UPLC-MS Discriminatory Markers for HAT . 164
5.6 CSF Metabolites Measured via ^{1}H NMR Spectroscopy 171

Chapter 1

Introduction

1.1 Trypanosomatid Parasites Cause Neglected Tropical Diseases 18
1.2 Cutaneous Leishmaniasis 21
1.3 Human African Trypanosomiasis 30
1.4 Metabolic Profiling of Parasitic Infections 39
1.5 Project Aims and Thesis Outline 41
1.6 Additional Projects 42

1.1 Trypanosomatid Parasites Cause Neglected Tropical Diseases

The neglected tropical diseases (NTD's) represent a diverse collection of disabling infectious diseases that are poorly managed and afflict particularly the poorest populations across Asia, Africa, and the Americas.[1,2] Collectively, they affect over 1.4 billion people globally[1,2], and have a disease burden in sub-Saharan Africa that is comparable to malaria and tuberculosis.[3] The vast majority of NTD's are caused by parasites, which include a range of intestinal and soil-transmitted helminth (worm/fluke) infections, and protozoan (single-celled eukaryotes) infections. Of these pathogenic protozoa, the most common belong to the family of *Trypanosomatidae*, within the order of *Kinetoplastida*.

Trypanosomatidae are exclusively parasitic, and reside within the gut and associated organs of insects for at least a portion of their life-cycle. Many *Trypanosomatidae* species live only in insects, such as genus *Crithidia*, whereas a few also infect plants, e.g. genus *Phytomonas*.[4] Of those which are additionally capable of infecting vertebrate hosts — typically mammals — the ones of most clinical interest belong to the genera *Leishmania (L.)* and *Trypanosoma (T.)*. *Leishmania* parasites

are causative agents of the leishmaniases and represent the most wide-spread of the protozoan NTD's. Pathogenic trypanosomes include *T. cruzi*, responsible for Chagas disease, primarily within South America, and *Trypanosoma brucei (T. b.) gambiense* and *T. b. rhodesiense*, which both cause human African Trypanosomiasis.

As is common with parasitic protozoa, *Trypanosomatids* display extensive capabilities at adapting in response to their local environment, in order to survive, invade, evade and proliferate within their host. Despite such morphological adaptations, certain key structures remain common to all *Trypanosomatidae*, including the presence of a single flagellum, which apart from the intracellular morphology (i.e. amastigote), is long and mobile, positioned near the apex or attached laterally to the cell body (Figure 1.1). They also all contain a kinetoplast near the basal body of the flagellum, which consists of a network of intertwined circular DNA, encoding ribosomal RNA's and mitochondrial proteins.[5] Unlike most other eukaryotic cells where glycolysis takes place in the cytosol, the *Trypanosomatidae* compartmentalise the majority of this pathway within specialised membrane-bound vesicles, known as the glycosomes, which contain mostly glycolytic enzymes.[6] The six main morphological variations of the *Trypanosomatidae* are shown in Figure 1.1, highlighting the differences in sizes and positions of selected cellular structures.

Within the scope of this PhD thesis, I will be presenting work with *L. major* and *T. b. rhodesiense*, as two representatives species of pathogenic *Trypanosomatids*. This chapter will give an overview of their associated pathologies, immunology and parasitology. Additional emphasis shall be given on the use of a metabolic profiling approach, to explore potential host-parasitic metabolic interactions and reveal candidate infection-associated biomarkers, that may contribute towards pathogenesis.

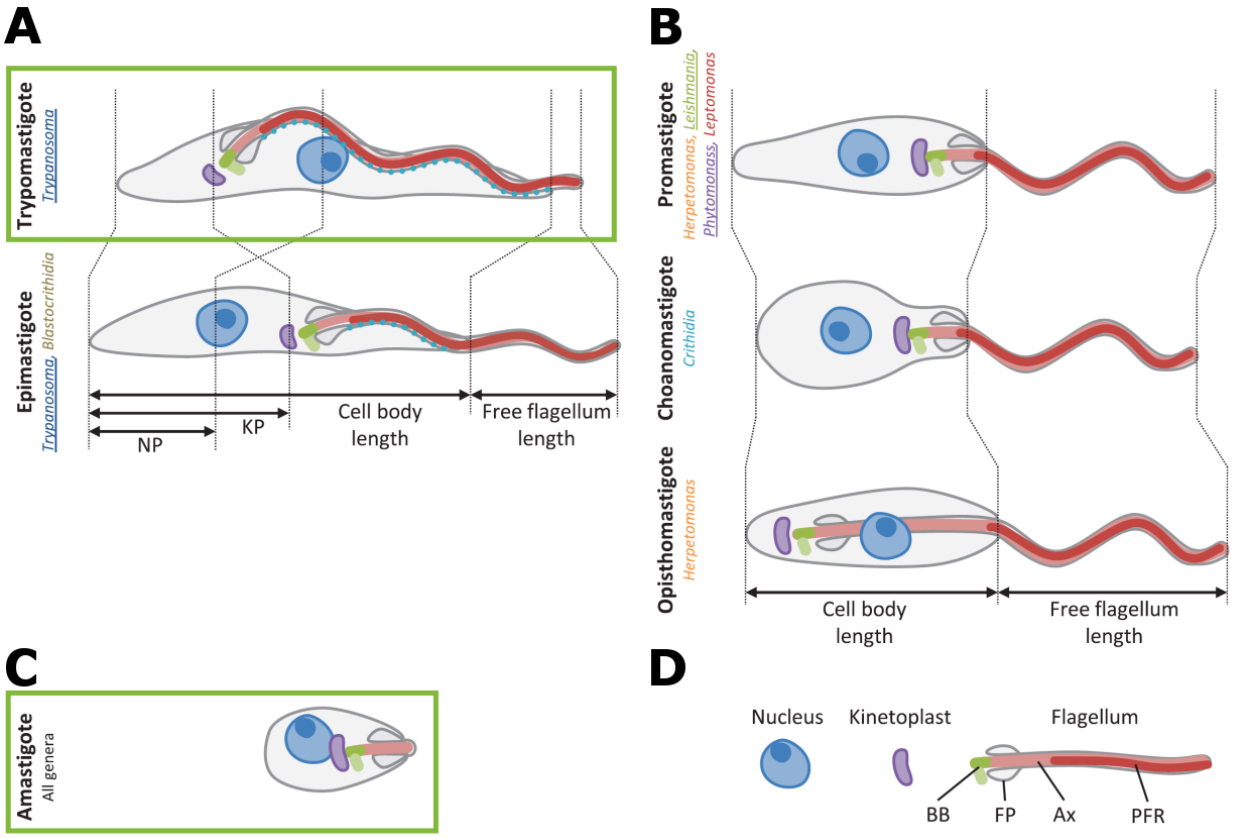


Figure 1.1: The Main Morphological Classes of Trypanosomatidae

Schematic diagrams of six major morphological forms of *Trypanosomatids*. Parasitic forms observed in infected human hosts have been highlighted in green boxes. (A) Morphologies with a flagellum laterally attached to the cell body. (B) Morphologies with free flagellum, extending beyond the flagellar pocket neck. Metrics indicate cell body length, free flagellum length, kinetoplast–posterior distance (KP) and nucleus–posterior distance (NP). Genera in which each morphology occurs are listed, with monophyletic genera underlined. (C) The amastigote represents the intracellular morphology, which lacks a long motile flagellum. (D) Key. Structures associated with the flagellum: BB, basal body/pro-basal body pair; FP, flagellar pocket; Ax, axoneme; PFR, paraflagellar rod. *Figure modified with permission from Wheeler, R.J. et al. (2013)*[7] *under the Creative Commons Attribution license.*

1.2 Cutaneous Leishmaniasis

1.2.1 Impact and Clinical Presentation of the Leishmaniases

The leishmaniases describe a group of vector-borne diseases that primarily affect the poorest communities across four continents, mainly found in tropical, sub-tropical regions and the Mediterranean basin. The most common and mildest form of the leishmaniases is known as cutaneous leishmaniasis (CL), which will be investigated in this PhD project. Latest estimates by the Global Burden of Disease Study 2010 claim that 8–13 million cases of human CL currently exist worldwide. CL also has the most negative impact on people's well-being of all NTD's, as measured by disability-adjusted life years (DALY's)[1], with the highest DALY score of approximately 3.32 million.[2,8] In addition to being a serious public health concern, the leishmaniases pose important economical costs by relying on relatively expensive drug treatments. Moreover, as is the case for all NTD's, they cause morbidity, leading to disabilities to school attendance and/or temporal depletion of labour, which in turn generate further losses to agricultural/industrial development programmes.[9]

CL typically manifests as an ulcerative skin lesion near the bite location of the sandfly vector, known as localised CL, commonly caused by infections with *L. major* or *L. tropica*. Another common manifestation is the presence of many non-ulcerative papules or nodules across the skin (termed diffuse CL), caused by *L. mexicana*, *L. amazonensis*, and *L. aethiopica* infections (though all of these are also capable of only presenting as localised CL).[10] Photographic examples of these are shown in Figure 1.2. The ulcers are usually self-healing in immune-competent individuals, yet the healing process is slow, taking several weeks to months, depending on the causative *Leishmania* species *(spp)*. Furthermore, lesions may become more severe due to bacterial co-infections, resulting in disfiguring scars.[11,12] A variant of CL is muco-cutaneous leishmaniasis (MCL), which presents as destructive ulcerations from the nose/mouth to the throat, largely due to *L. braziliensis* infections, found in Bolivia, Brazil, and Peru.[10,13]

Another form of the leishmaniases is visceral leishmaniasis (VL), also known as "kala-azar" or "black fever". This represents the most severe form which results in wide-spread parasitic dissemination within the host, affecting multiple internal organs, and is potentially fatal if untreated. VL is usually the result of infection with *L. donovani*, *L. infantum* or *L. chagasi spp.*[11,13] It is not uncommon for individuals who have recovered from VL to relapse weeks to years after their initial infection, to present with secondary dermal manifestations.[14] Additionally, many *Leishmania* infections are considered to be symptomatic in some individuals, who inadvertently, may act as

[1] Calculated as the sum of years of life lost owing to premature mortality and number of years lived with disability.

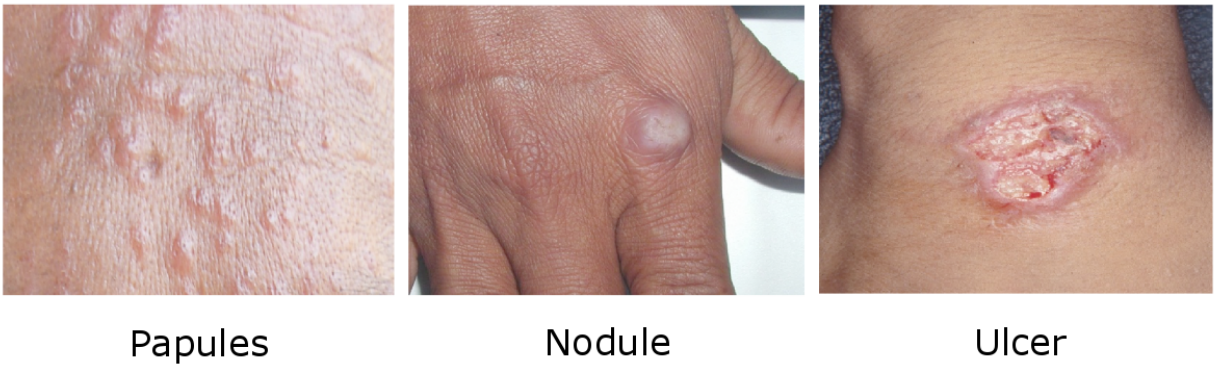


Figure 1.2: Diverse Dermal Manifestations of Cutaneous Leishmaniasis

Figure modified with permission from Olliaro, P. et al. (2013)[12] *under the Creative Commons Attribution license.*

additional reservoirs and maintain parasitic transmission within a population.[10]

1.2.2 Diagnosis and Treatment of CL

Despite efforts, no vaccine is currently available for any of the leishmaniases.[15] Current disease management relies mainly on vector control through insecticide spraying, and identifying and treating cases.[9] Yet the latter proves particularly difficult, owing to the wide variability of clinical presentations and disease severity of CL (as well as the other leishmaniases). Eliminating differential diseases is important for diagnosis as many share similar dermal characteristics with CL, e.g. leprosy, skin cancers and cutaneous tuberculosis.[10] Parasitological assessment through microscopic examination of lesion biopsies continues to be the gold standard for CL diagnosis, due to its high specificity. However, the sensitivity of this method is low, generating variable results, largely dependent on the technical capabilities and experience of the laboratory personnel.[16] Nevertheless, its wide use is down to the relatively low cost of this technique compared with more sophisticated technologies outlined below.

Current serologically-based tools, such as immunochromatographic dipstick tests or antigen-based direct agglutination tests, have been successfully implemented for VL diagnosis, yet have proven too variable for accurate results for CL.[16] Diagnosis of CL can also be performed *via* the Montenegro skin test (also termed Leishmanin skin test) with high sensitivity and specificity, which rapidly tests the antibody response towards *Leishmania* antigens.[17] However, culturing facilities are thus required to generate these antigens (usually crude parasite extracts), which can greatly vary in quality across batches and alter test sensitivity. Another limitation of this test is that it cannot differentiate between current and past infections.[16] More recently, polymerase chain reaction (PCR)-based methods for the detection of *Leishmania* parasite DNA from lesions have

been implemented for CL diagnosis, with higher accuracy than microscopic techniques. PCR also allows species-specific identification, though their use is usually limited in resource-restrained settings.[18]

Whilst lesions may heal themselves, treatment is usually sought to accelerate the healing process (particularly if on the face or in/near joints), prevent progression to more aggressive forms, if the infected individual is immune-comprised (e.g. HIV co-infection), and/or depending on the parasite species.[19] A plethora of different treatment options are available for CL but their potencies are highly variable, especially between *Leishmania spp.* Pentavalent antimonials have been used for several decades and continue to be recommended as first-line drugs in most *Leishmania* endemic countries for both CL and VL, providing drug resistance is absent — a factor which is becoming an increasingly important problem.[14] These antimonials include the parental drugs sodium stibogluconate (Pentostam^{TM}, GlaxoSmithKline), and meglumine antimoniate (Glucantime, Aventis). Evidence suggests their main mechanism of action is through interference of trypanothione metabolism, a thiol molecule with anti-oxidative properties in *Trypanosomatids.*[20] However, these drugs are frequently associated with multiple side-effects, including muscle, joint and abdominal pain, nausea, and cardiotoxicity. Administration generally involves multiple intravenous or intramuscular injections (intra-lesion injections are less common), over a number of days, with varying degrees of efficacy.[14]

Second-line drugs against *Leishmania* include Amphotericin B and and pentamidine, though as with the antimonials, they too require a long course of parenteral administration.[19] Amphotericin B (Fungizone, Bristol-Myers Squibb) is a polyene antifungal drug that targets the *Leishmania* membrane sterols (known as ergosterols), resulting in membrane pore formation and parasitic cell lysis.[14] It is given intravenously to treat CL, as well as MCL and VL. However, the drug is highly toxic, particularly damaging to renal cells, and interferes with renal filtration.[14,21] Various lipid formulations of this drug have been developed that have reduced toxicity, yet their use and/or availability is restricted in endemic countries, owing to their high cost.[21,22] Pentamidine on the other hand is only used to treat specific forms of CL in South America.[14]

There has been some evidence supporting the use of other antibiotics and antifungals against CL but these have also been very species dependent, e.g. antifungal fluconazole, for *L. major* infection. Miltefosine, a phosphocholine analogue which was originally developed as a cancer drug but was found to have leishmanicidal effect is accompanied with multiple significant adverse effects (including being teratogenic and abortifacient) and is thus generally used to treat the more serious VL rather than the milder forms of CL.[14,21]

Topical formulations that are currently used to treat CL are usually comprised of Paromomycin (typically at 15% with 10% urea and/or methyl benzylmethonium chloride), an aminoglycoside antibiotic thought to interrupt *Leishmania* mitochondrial function.[21,23] In theory a topical formulation may sound appealing to treat a dermal disease, yet inconsistencies in formulation have resulted in varying degrees of skin penetration and have led to incomplete parasitic elimination, as well as presence of unwanted adverse events, such as blisters and bruising.[19] Efficacy is also very species dependent. Combined with restricted availability, Paramomycin usage has been largely limited in endemic areas.[21] Other treatments that have been used for CL include cryotherapy, local heat therapy *via* radio-frequency, and a wide range herbal remedies, again with varying degrees of success.[19,22]

Given the lack of consistency and/or complications of current treatments, and high financial costs to affected individuals in low-income settings, there remains consequently an urgent need for the discovery of new efficacious and cost-effective treatments for the leishmaniases. Furthermore, as the aim of CL therapy is largely to accelerate self-healing, immunomodulatory approaches aimed at enhancing the host's own response to eliminate the parasite may provide important alternatives treatment strategies, alongside the long-term goal of ultimately developing a successful vaccine.[14]

1.2.3 The Life-cycle and Biology of Leishmania Parasites

The leishmaniases are transmitted by infected female sandflies (*Phlebotomus* and *Lutzomyia spp*)[9,10] either directly between humans, or through zoonotic transmissions, which includes multiple animal reservoirs (dogs, rodents, and marsupials).[24] Approximately 21 out of the 30 *Leishmania* species are known to be infective to humans.[11] They exist in two different forms, depending on the host stage during their life-cycle, as summarised in Figure 1.3.

Within the sandfly vector, the parasite exists as a motile flagellated promastigote (Figures 1.3 and 1.1). Once deposited in the midgut of the sandfly, promastigotes attach themselves to the epithelial cell lining and proliferate. At this stage they are avirulent and are termed procyclic promastigotes. This is followed by detachment from the epithelial lining and migration from the midgut to the proboscis (i.e. the elongated mouth-part of the fly). During this process, they evolve from the non-infective procyclics into non-dividing infectious metacyclic promastigotes. The metacyclic promastigotes are deposited from the proboscis into the mammalian host dermis and/or bloodstream during a blood meal, where they are engulfed in non-lethal vacuoles of local neutrophils and phagocytosed by resident dendritic cells and macrophages, the latter representing the main host cell targeted by *Leishmania* (Figure 1.3).[13,25]

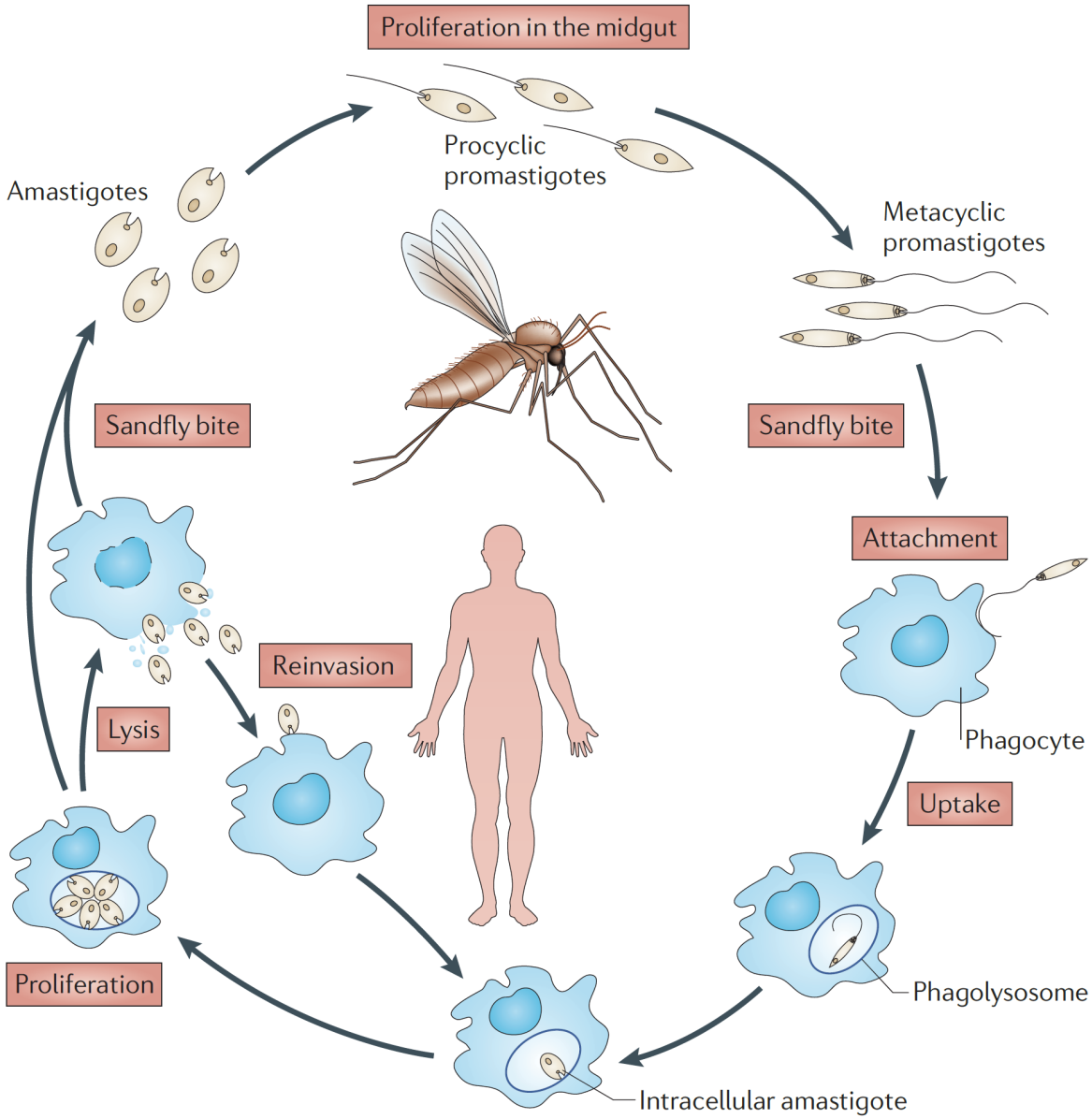


Figure 1.3: Life-cycle of Leishmania Parasites

Leishmania procyclic promastigotes differentiate in sandflies into infective, non-dividing metacyclic promastigotes, which are located ready for transmission at the stomodeal valve (an invagination of the foregut into the midgut). During blood feeding, the sandfly regurgitates metacyclic promastigotes, together with immunomodulatory parasite-derived proteophosphoglycans and various salivary components. The metacyclic promastigotes are then engulfed by neutrophils in non-leishmanicidal vacuoles and/or phagocytosed by local macrophages and dendritic cells. After establishing an intracellular residence, metacyclic promastigotes transform into aflagellate amastigotes. Amastigotes undergo replication within host cells, which rupture when too many amastigotes are present, allowing reinfection of local phagocytes. The transmission cycle is complete when infected phagocytes are taken up by another sandfly with the blood meal, and amastigotes then convert into promastigotes in the sandfly midgut.
Figure reproduced and legend modified with permission from Kaye, P. and Scott, P. (2011)[25]
Copyright 2011 Nature Publishing Group.

Within the host cells, the promastigotes differentiate into spheroid-shaped, non-motile, aflagellated amastigotes that rapidly divide *via* binary fission until they eventually rupture the cell. Released amastigotes then re-infect surrounding macrophages and other immune and stromal cells. These parasitised host cells become ingested by sandflies in subsequent bloodmeals, where the amastigotes are released into the midgut. Here, the amastigotes transform back into promastigotes, to restart the infectious cycle (Figure 1.3).[25,26]

In addition to the striking morphological and motility changes observed between promastigotes and amastigotes, the entire surface (including flagellum) of all *Leishmania* promastigotes are covered by a thick glycocalyx coat, composed largely of lipophosphoglycan (LPG), whilst amastigotes express little LPG, if at all.[27] LPG is a lipid-containing polysaccharide, composed of a polymer of dissacharide-phosphate repeats, that is linked to the parasite membrane *via* a glycosylphosphatidylinositol (GPI) anchor.[27,28] LPG has been show to be a major virulence factor for *Leishmania* parasites, and has been linked to a number of processes critical for survival and establishing infection within the host, including attachment to macrophage cell surface, resisting complement-mediated lysis, acting as a physical barrier against anti-leishmanial antibodies, and modulating macrophage signal transduction.[28] Within the sandfly, LPG has also been implicated in promastigote attachment and detachment to epithelial cells, and resistance to hydrolytic activity within the gut.[28]

In contrast to most intra-cellular parasites that tend to avoid lysosomal vesicles, *Leishmania* represent some of the few species that have adapted to specifically reside and actively proliferate within phagosomes/endosomes, even after fusion with the acidic lysosomes (i.e. within phagolysomes).[26] In fact, *in vitro* studies exploit this characteristic by artificially dropping the pH of promastigotes cultures in a host cell free environment, in addition to increasing the temperature from 26°C (mimicking temperatures within sandflies) to 37°C (inside mammalian hosts), to induce their differentiation into amastigotes.[29]

In macrophages that have been activated with certain cytokines, e.g. by Interferon-γ (IFNγ) and Tumour-necrosis-factor-α (TNFα), lysosomes also contain enzymes that generate a range of other molecules toxic to the parasite, such as reactive oxygen species (ROS), including superoxide radical (O_2^-), nitric oxide (NO), hydrogen peroxide (H_2O_2) and peroxynitrite ($ONOO^-$).[13] One of the defence mechanisms of *Trypanosomatids* (including *Leishmania* and *Trypanosoma* species) is that they contain trypanothione, an unusual dithiol molecule used as an antioxidant, which consists of two glutathione molecules, linked by spermidine (a polyamine).[13] Trypanothione has been shown to reduce $ONOO^-$, H_2O_2, and NO, as well as be able to sequester NO (combined

with iron) into stable non-toxic complexes[2], and thus minimise oxidative damage towards the parasite.[13,30] Consequently, the survival of *Leishmania spp* within macrophages is determined by the balance between an appropriate host response that results in sufficient activation of *Leishmania* -infected macrophages, and the abilities of the parasite to subvert host immunity and resist the cytotoxic mechanisms of the activated macrophages.[25,31]

Leishmania Metabolism

The nutritional differences in local environment between the two stages of *Leishmania* life-cycle is reflected in the disparities in their metabolism, each adapted to suit their needs. Both amastigotes and promastigotes can use carbohydrates (preferentially glucose)[3] for energy, which are broken down into pyruvate *via* glycolysis, a process that takes place within their glycosomes.[32,33] The first steps of glycolysis have a high requirement for co-factors nicotinamide adenine dinucleotide (NAD^+) and high energy molecule adenosine triphosphate (ATP). These molecules are replenished through the production of succinate and pyruvate from phosphoenolpyruvate, as shown in Figure 1.4.[33,34]

In most mammalian systems, pyruvate is further converted into acetyl coenzyme A (acetyl-coA), *via* pyruvate dehydrogenase, in the presence of oxygen, resulting in the high production of ATP and the emission of CO_2 as a waste product. Alternatively, when oxygen supply is restricted (and thus the respiratory chain function is inhibited), pyruvate is reduced to lactate by lactate dehydrogenase, in order to maintain the redox balance.[32] *Trypanosomatids* lack lactate dehydrogenase and thus use alternate pathways: pyruvate is converted instead into either alanine in the cytosol (*via* alanine aminotransferase), or enters the mitochondrion to be converted into acetyl-coA.[32,34] Whilst the presence of genes encoding all TCA enzymes have been detected, new evidence suggests that *Leishmania* parasites favour the conversion of acetyl-coA into acetate, transferring the coA group onto a succinate molecule, to form succinyl-coA instead, *via* acetate–succinate-coA transferase. Succinyl-coA is subsequently converted back to succinate, generating ATP in the process.[33,34] Consequently the major end products from *Leishmania* parasites are acetate, succinate and alanine, highlighted in Figure 1.4.[33,34]

Evidence has shown that the availability of glucose in gut of the sandlfy is relatively limited, yet the gut contains relatively high concentrations of amino acids. Unlike other trypanosomatids, *Leishmania* promastigotes have shown to be able to use amino acids as alternative energy sources, e.g.

[2] with much higher affinity than glutathione, the main mammalian antioxidant molecule

[3] Though *Leishmania spp* contain several kinases that can break down other sugars (e.g. ribose and ribulose), a well as amylase, presumably for the degradation of plant starch and disaccharides that are ingested by sandflies that feed on nectar as well as bloodmeals.[32]

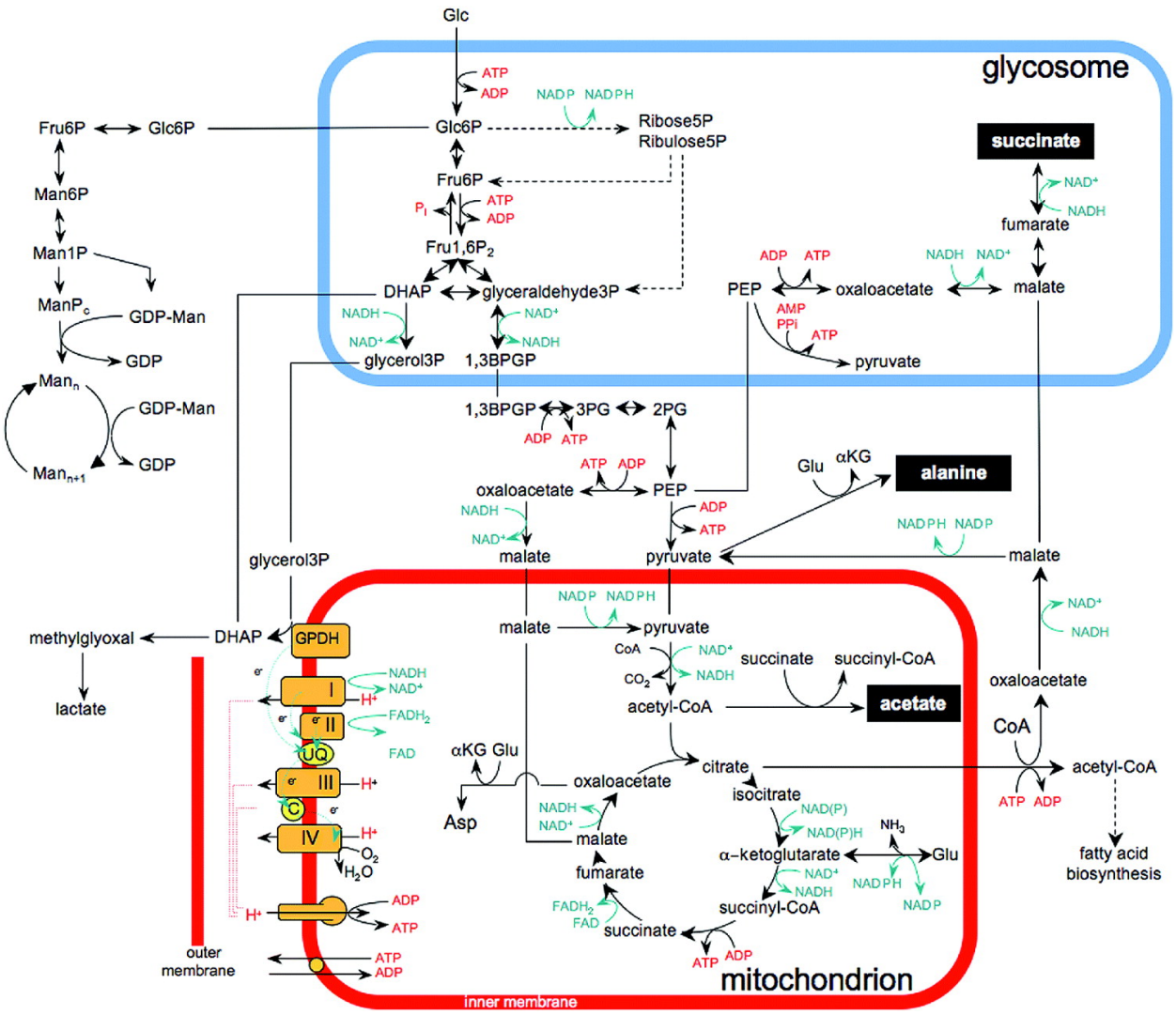


Figure 1.4: Central Carbon Metabolism of Leishmania Parasites

Schematic representation of central carbon metabolism in *Leishmania* promastigotes, cultured in glucose-rich medium. The major secreted end-products (succinate, alanine and acetate) are shown on a black background. A number of enzymes in the glycolytic, gluconeogenic, pentose phosphate and succinate fermentation pathways are thought to be partially (with dual localisation in the cytosol) or exclusively localised to the glycosome. Dotted arrows refer to multi-enzyme steps that are not shown. Abbreviations: I–IV, complexes of the respiratory chain; αKG, α-ketoglutarate; 1,3BPGP, 1,3-bisphosphoglycerate; ADP, adenosine diphosphate; ATP, adenosine triphosphate; DHAP, dihydroxyacetone phosphate; e^-, electrons; FAD, flavin adenine dinucleotide; $FADH_2$, FAD (reduced); $Fru1,6P_2$, fructose-1,6-bisphosphatase; GDP, guanosine diphosphate; Glu, glutamate; GPDH, FAD-dependent glycerol 3-phosphate dehydrogenase; Glc6P, glucose-6-phosphate; Man6P, mannose-6-phosphate; $ManP_c$, Man1,4-cyclic-phosphate; Man_n, mannogen oligomers (Manβ1-2Man)n; NAD^+, nicotinamide adenine dinucleotide (oxidised); NADH; NAD (reduced); NADP, nicotinamide adenine dinucleotide phosphate; NADPH, NADP (reduced); PEP, phosphoenolpyruvate; 2PG, 2-phosphoglycerate; 3PG, 3 phosphoglycerate.
Figure and legend reproduced with permission from Saunders, E.C. et al. (2010)[33]
Copyright 2010 Cambridge University Press.

proline, glutamine and threonine, which are metabolised into TCA-intermediates.[32,34,35] Amino acids are also used as carbon skeletons for gluconeogenesis.[36] In contrast, amastigotes live within host cell phagolysosomes, which have been shown to be rich in multiple nutrients such as lipids, nucleobases, vitamins, cations, polyamines and amino acids, yet poor in sugars such as hexose and mannose, all of which are largely scavenged by the *Leishmania* species for parasitic growth.[26] Studies have shown that *Leishmania* amastigotes heavily rely on lipid catabolism, using fatty-acid derived acetyl-coenzyme A and glycerol as major energy sources.[33–35]

1.2.4 Immune Response to Leishmania

In addition to the nature of *Leishmania* species, the way the host responds to the pathogen, notably the host's immune response, is critical in determining parasitic clearance and disease progression. Host immunology in *Leishmania* infection has been investigated in a range of murine experimental models, which have been instrumental in gaining an insight into the host-parasite relationship and thus, a better understanding of the pathophysiology of this complex set of diseases.[31] In particular, the type of dominant CD4 T helper (Th) response, that directs the type of macrophage activation towards the beginning of infection, plays a key role in mediating pathology.

Evidence has shown that infection with *L. major* in most laboratory strains, including C57BL/6 mice, result in a slow and self-healing skin lesion at the site of injection – similar to uncomplicated cutaneous leishmaniasis in humans. In contrast, the susceptible BALB/c mouse strain develops wide-spread non-healing lesions, followed by internal dissemination of the parasite that eventually prove fatal, resembling the visceral form of leishmaniases. Whilst both models have been shown to develop an initial Th 2 type immune response, characterised by increased IL-4 production shortly after infection, self-healing strains subsequently develop a dominant Th 1 response. These mice increase their production of IL-12 and INFγ and subsequently eradicate the parasite.[37–41] However, the susceptible non-healing BALB/c mice sustain their IL-4 response and thus fail to eliminate infection, resulting in exacerbated disease.[37–41]

The critical roles of Th 1 vs. Th 2 cytokines have been reflected in several *in vitro* studies, with the discovery of "classical activation" of macrophages by Th 1 cytokines IL-12, IFNγ and TNFα, which induce the intracellular production of NO from arginine *via* nitric oxide synthase (iNOS), as well as other ROS, and promote parasitic elimination.[42,43] In contrast, macrophages that are "alternatively activated", stimulated by Th 2 cytokines IL-4 and IL-13, allow the survival of the parasite through indirect downregulation of ROS by favouring the conversion of arginine to ornithine. Ornithine is an essential metabolite for polyamine synthesis (e.g. spermidine), which

in turn are required for parasitic growth[42–44] and, as recently discussed, for resistance against oxidative damage, as constituents of trypanothione.[13] Furthermore, studies have shown that one of the mechanisms that *Leishmania* parasites use to avoid the induction of Th 1 response is by depleting cholesterol levels from the membranes of infected macrophages and dendritic cells, interefering with the assembly of receptors and co-factors involved in IL-12 signalling, whilst also inducing regulatory anti-inflammatory cytokine IL-10.[45]

Our current understanding of the adaptive Th response has since evolved, with the appearance of novel Th subsets including T regulatory cells (T regs), Th 17 cells, follicular T helper cells (Tfh), and the concept of plasticity between these subsets.[46,47] Furthermore, an increasing number of studies have highlighted a multifactorial host response to *Leishmania* infection, including the involvement of innate cells[48–50] and protective anti-leishmanial antibodies.[51,52] Evidence implies that *Leishmania* infection actively induces secretion of transforming growth factor beta (TGFβ), another regulatory cytokine which suppresses iNOS activity[45], and recruits T regs, which suppress effector T cells and have been linked to persistent *Leishmania* infection.[53]

Furthermore, results from an *in vivo L. major* infection models have demonstrated that mice lacking natural gastrointestinal microbiota have increased pathological symptoms compared with their conventional counterparts, despite generating an appropriate Th cytokine response to the parasite, highlighting the role of gut microbes in regulating host responses.[54] Nevertheless the oversimplification of the Th 1 vs. Th 2 responses generated towards *Leishmania spp* has still proven useful in the partial explanation of disease pathology, not only in murine models but also in many clinical cases. Studies have shown that dominant Th 1 cytokines have been associated with milder localised forms of CL in humans, whilst Th 2 and/or regulatory responses have been linked to the more aggressive MCL[55] and VL.[56,57]

1.3 Human African Trypanosomiasis

1.3.1 Impact and Clinical Presentation of HAT

Human African Trypanosomiasis (HAT), also known as African sleeping sickness, is endemic in sub-saharan Africa, as shown in Figure 1.5. The geographical spread of the disease is defined by the distribution of its vector, the tseste fly. It is caused by two morphologically identical subspecies of *T. brucei*: *T. b. gambiense*, which infects mainly humans in west and central Africa, and *T. b. rhodesiense*, which is primarily zoonotic with a vast animal reservoir, found largely in eastern and southern Africa. As yet, the two species are still geographically distinct but there is

potential for these two to overlap, as highlighted in Figure 1.5.[58] There also other species that cause trypanosomiasis in animals (also known as *nagana*) but do not generally infect humans, including *T. congolense*, *T. vivax* and *T. b. brucei*.[59]

Historically, HAT has caused a huge burden on endemic countries, with several devastating epidemics across Africa in the late 19th century (killing ∼0.5 million people) and early to mid 20th century.[60] These epidemics spurred on the investment in vector control measures and active surveillance by colonial powers, which still represent the two main strategies to battle HAT today.[60] However, following almost complete removal of HAT, the disease re-emerged in the late 1990's, partly due to civil conflicts and break-down of control measures.[58,60]

Since then, proactive measures agreed by the WHO and partners have helped to dramatically reduce the number of new reported cases of HAT, from 40,000 in 1998 to under 10,000 in 2009.[61] Latest estimates indicate a disease burden of 560,000 DALY's, with approximately 37,000 total number of reported cases.[2,8] Whilst these numbers may appear relatively insignificant in a world-wide context, HAT has a serious socio-economic effect in the affected areas, owing to its focal distribution and detrimental health consequences. Combined with the animal form of African trypanosomiasis, HAT is thought to have represented a key factor for curbing rural development in Sub-Saharan Africa.[58] Additionally, the number of reported cases is considered to be a gross under-estimation as a large number of HAT infections are estimated to be under-reported, particularly in remote locations that are inaccessible to conventional health care systems, where the disease is most prevalent. In order to achieve the global target of disease elimination by 2020, the WHO has recommended that continued efforts in diagnostic research and vector control are required to successfully identify, monitor and eradicate HAT.[61–63]

The disease has two clinical stages. In the first "early" stage, the parasite replicates within the blood and lymphatic system (this stage is thus also referred to as the haemo-lymphatic stage), which may present in individuals as non-specific inflammatory symptoms (e.g. fever, swollen lymph nodes, hepatomegaly, splenomegaly) and, in some cases, some neurological symptoms.[64,65] If left to progress without appropriate treatment, the disease enters the "late" stage of HAT (also known as meningo-encephalitic stage), where the parasites have disseminated to other organs and have crossed the blood-brain barrier, to infect the central nervous system (CNS). At this stage, individuals characteristically display neurological symptoms (e.g. headaches, tremors, rigidities, and impaired gait) and/or psychological features (e.g. changes in personalities and psychosis), which, if left untreated, progress to increased drowsiness, cachexia (body wasting), coma, and ultimately, death.[64,65]

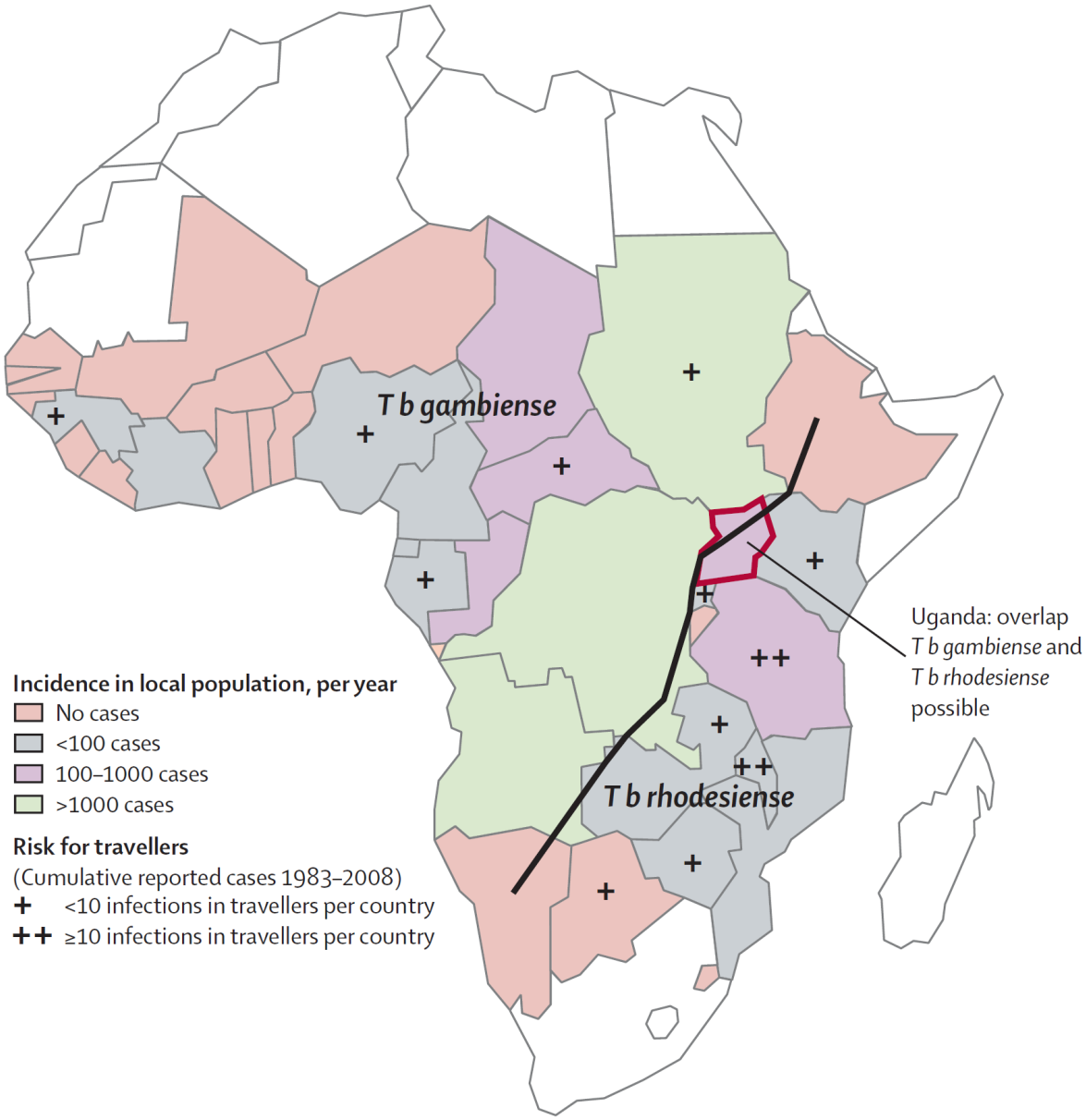


Figure 1.5: Distribution of Human African Trypanosomiasis

Border between HAT caused by *T. b. gambiense* and *T. b. rhodesiense* is shown by black line. *Figure reproduced with permission from Brun, R. et al. (2010)*[58] *Copyright 2010 Elsevier.*

The severity and time period of disease differ widely between the two *T. brucei* subs-species: infection with *T. b. gambiense* is frequently asymptomatic during the early stage and leads to low-grade chronic disease that can last months or years. On the other hand, *T. b. rhodesiense* infection usually induces an acute febrile illness, lasting 7–10 days, and is difficult to distinguish from other endemic illnesses, such as malaria or enteric fever.[65,66] A localised skin lesion (chancre) may develop at the site of the fly bite after a few days yet this is often only observed in infected travellers and typically heals spontaneously after a few weeks.[65] Furthermore, patients suffering from HAT often have other concurrent endemic infections that can confound diagnosis and hinder them from being treated effectively.

1.3.2 Diagnosis and Treatment of HAT

Diagnosis of HAT consists of an initial screening, followed by parasitological confirmation, and finally diagnostic staging. For *T. b. gambiense*, screening is performed largely using serological tests, such as the Card Agglutination Test for Trypanosomiasis (CATT), a rapid simple (and cheap) test with high sensitivity, that is often used in mass prevention programmes.[60,65] Unfortunately there is still no test equivalent to CATT available for trypanosomiasis *rhodesiense*. Other serological screening tests include enzyme-linked immunosorbent assay (ELISA) and immunofluorescent assays, which have yielded promising results particularly for *T. b. gambiense* infection (sensitivities for *T. b. rhodesiense* infection have been variable). However, due to their increased sophistication and cost, these assays are not widely used in non-endemic countries.[60,65] HAT *rhodesiense* screening is primarily based on a combination of clinical history, recognition of commonly associated inflammatory symptoms (e.g. fever), and other non-specific biological parameters (which are frequently highly abnormal in *T. b. rhodesiense* infection), such as severe anaemia and thrombocytopaenia.[60]

Parasitological assessment is usually done *via* microscopic examination of the blood or chancre aspirate, if present. As parasitaemia can often be low, as is typically the case for trypanosomiasis *gambiense*, blood concentration techniques, such as microhematocrit centrifugation or quantitative buffy coat extraction, are commonly performed prior to microscopic assessment.[60,65] This procedure is highly dependent on the expertise of the personnel, and may miss 20–30% of HAT patients.[58] Other more sensitive methods, including PCR to detect and amplify trypanosome DNA, have been used but these methods have not yet been standardised for diagnostic validation.[60]

Lastly, diagnostic staging is mandatory for all confirmed HAT cases as drug treatments differ significantly between the stages. This step involves a painful and technically challenging lumbar puncture, to aspirate cerebral spinal fluid (CSF), followed by its microscopic examination. Accord-

ing to the WHO, the patient has entered the late stage when a minimum of 5 parasites/µl and/or a white blood cell (WBC) count of at least 20 cells/µl are observed in the CSF.[67] Furthermore, an elevated Immunoglobulin M (IgM) concentration in CSF has also shown to be a robust marker of neurological involvement.[60] Controversies with the WHO guidelines also exist, with some investigators claiming an additional 'intermediate' stage (with 5–10 cells/µl WBC in CSF) in patients who could be treated with early stage drugs, yet treatment outcomes have been variable.[58,60]

As with Leishmaniasis, there is currently no vaccine available for HAT, and thus disease management relies on vector control, screening and treatment of identified cases. Concernedly, the limited number of available drugs are largely dated, scarce, highly toxic, have little known about their mechanisms of action, and in some instances, have been decreasing in efficacy due to emerging drug resistance.[64,68] There are currently only four licensed drugs available against HAT, all of which are currently donated to the WHO by drug manufacturers, listed in Table 1.1.

Table 1.1: Standard Anti-trypanosomal Drug Treatments for Ealy and Late Stage HAT

	Stage	**Route of Application**	**Dosing**	**Main Adverse Drug Reactions**
Trypanosoma brucei gambiense				
Pentamidine*	Early	Intramuscular	4 mg/kg bodyweight at 24 h intervals for 7 days	Hypoglycaemia, injection site pain, diarrhoea, nausea, vomiting
Eflornithine	Late	Intravenous Infusion of >30 min	100 mg/kg bodyweight at 6 h intervals for 14 days	Diarrhoea, nausea, vomiting, convulsions; anaemia, leucopenia, and thrombocytopenia
NECT: Eflornithine + Nifurtimox	Late	Eflornithine intravenous, Nifurtimox oral	Eflornithine at 400 mg/kg bodyweight IV in 2x 1 h infusions for 7 days + Nifurtimox at 15 mg/kg orally in 3 doses for 10 days[60]	Vomiting, gastrointestinal/abdominal pain, nausea, headache, seizures, malaise, musculoskeletal and connective tissue disorders, psychiatric disorders, anorexia[69]
Melarsoprol***	Late	Intravenous	2.2 mg/kg bodyweight at 24 h intervals for 10 days	Encephalopathic syndromes, skin reactions (pruritus, maculopapular eruptions), peripheral motoric (palsy) or sensorial (paraesthesia) neuropathies, thrombophlebitis
Trypanosoma brucei rhodesiense				
Suramin*	Early	Intravenous	Test dose of 4–5 mg/kg bodyweight at day 1, then five injections of 20 mg/kg bodyweight every 7 days (e.g. day 3, 10, 17, 24, 31); maximum dose per injection 1 g	Hypersensitivity reactions (acute, late); albuminuria, cylinduria, haematuria, peripheral neuropathy
Melarsoprol*	Late	Intravenous	Three series of 3.6 mg/kg bodyweight, the series spaced by intervals of 7 days; maximum dose per day 180 mg	Encephalopathic syndromes, skin reactions (pruritus, maculopapular eruptions), peripheral motoric (palsy) or sensorial (paraesthesia) neuropathies, thrombophlebitis

*Endemic countries: according to national legislature or guidelines.
**Only where eflornithine is not available or where melarsoprol is first-line treatment, according to national guidelines.
Abbreviations: IV, intravenous; NECT, Nifurtimox–eflornithine combination therapy. *Modified with permission from Brun, R. et al. (2010)[58] Copyright 2010 Elsevier*

Pentamididine isethionate (taken intramuscularly or intravenous for a week) is the first line drug

taken to treat HAT *gambiense*. It is effective against early stage but not late stage of disease, nor is it very potent against HAT caused by *T. b. rhodesiense*.[64] It is thought to interfere with parasitic nucleic acid metabolism and induce mitochondrial dysfunction, though the exact mechanisms are not known.[68] Suramin, a polysulphonated naphthalene derivative that possibly targets glycolytic enzymes, is another effective drug used for early stage HAT only, as it does not cross the blood-brain barrier.[68] Despite being effective at killing both *T. b. gambiense* and *T. b. rhodesiense*, it requires a long treatment period (typically 5 weeks) and causes severe allergic reactions in patients who happen to be co-infected with *Onchocerca* species (parasitic roundworms that are endemic in countries where HAT *gambiense* is present), and is thus the first-line drug for HAT *rhodesiense* only.[58]

Melarsopol, an arsenical, continues to be the most widely used drug for late stage HAT in resource-poor countries, where the newer (and safer) late stage drug, eflorinithine, is either not available or not affordable for HAT *gambiense*, and the is only drug for late stage HAT *rhodesiense*.[58] Amongst the multiple toxic side-effects — such as nephrotoxicity, hepatotoxitiy, neuropathy and various skin reactions — the main complication of melarsopol is encephalophatic syndrome.[60] This syndrome refers to a range of drug-induced neurological disorders (including coma, convulsions and psychotic disorders), which is present in approximately 5–10% of treated patients, of which a fatality rate of approximately 44% in HAT *gambiense* and 57 % in HAT *rhodesiense* is observed.[58] Additional drugs are thus given to try and manage encephalophatic syndrome (e.g. dexamethosone, prednisolone or diazepam), alongside careful disease monitoring.[60]

Eflornithine (either on its own or in combination with nifurtimox, a cheap oral drug originally used to treat American Trypanosomiasis) is recommended as a first-line drug against *T. b. gambiense*, due to its much improved safety profile over melarsonopol. Originally treatment with this drug was difficult (and expensive) to administer, requiring long treatment courses (56 days), yet more current formulation advances, and drug combinations, have helped shorten the time period to 2 weeks (on its own) or 7 days (with Nifurtimox, which is taken for 10 days).[58,60] Eflornithine trypanocidal properties are considered to be as a result of inhibiting polyamine and trypanothione biosynthesis, as well as inducing protein methylation.[68] Unfortunately studies have shown that *T. b. rhodesiense* parasites are naturally resistant to this drug, compared to *T. b. gambiense*[70], meaning that HAT *rhodesiense* patients have only the toxic Melarsopol to rely on. This is increasingly worrying as there have a growing number of reduced drug sensitivities reported, with mechanisms of drug resistance still unknown.[58]

1.3.3 The Life-cycle and Biology of Trypanosoma brucei ssp

Trypanosoma spp remain free-living throughout their life-cycle, inducing substantial and multiple structural changes, in order to survive and evade host defence mechanisms, summarised in Figure 1.6. Transmission begins with the introduction of metacyclic (non-proliferative) trypanomastigotes into the mammalian host bloodstream, following a blood meal with an infected tsetse fly (*Glossina spp*), the vector for *T. brucei*. These trypanomastigotes mature into long 'slender' forms that actively proliferate within the bloodstream and lymph vessels *via* asexual replication and maintain infection within the host, during early stage HAT.[71,72] In late stage HAT, these parasites invade other tissues, including the CNS, after crossing the blood-brain-barrier (Figure 1.6c). Importantly, trypanomastigotes are covered by millions of copies of variable surface glycoproteins (VSG's), that are attached to the cell *via* GPI-anchors and are repeatedly switched to change their antigenic surface, thus allowing continual evasion of host immune recognition whilst living openly, in the blood.[73]

Accumulation of the slender forms induces a density-dependent differentiation of some slender forms, to arrest their cell-cycle and undergo morphological transformation, into short 'stumpy' forms, in preparation for survival within the tsetse fly (Figure 1.6). The exact nature of driving signals for this transformation are not certain, though evidence suggests that it is due to a soluble parasite-derived factor (yet to be identified) that may be linked with the cyclic adenosine monophosphate (cAMP) metabolism.[71] Both slender and stumpy forms are taking up into the midgut of the tsetse fly in a subsequent bloodmeal.

The slender forms are destroyed in the fly gut, presumably due to the drop in pH, midgut proteases and/or restricted glucose availability (their primary source of energy *via* glycolysis).[71] In contrast, the stumpy forms survive (partially due to the fact that they reactivate their mitochondria and break down glucose *via* the TCA cycle instead of merely glycolysis, thus obtaining more energy) and differentiate into proliferative procyclic trypanomastigotes.[73] During differentiation, they lose their VSG coat (and have a thinner glycocalyx instead, composed of procyclins), and then undertake a long journey, migrating from the midgut, *via* the proventriculus (glandular part of the stomach), to the salivary glands (blue route in Figure 1.6b).

Whilst in the proventriculus, the procyclic trypanomastigotes transform into epimastigotes *via* asymmetrical division. Upon arrival to the salivary glands, the short epimastigotes attach to the epithelial lining of the glands and divide. Eventually these mature into non-dividing metacyclic trypanomastigotes, which re-acquire their VSG coat, and are released into the gland lumen, where they travel towards the mouth parts of the fly (red route in Figure 1.6b), ready to be transmitted

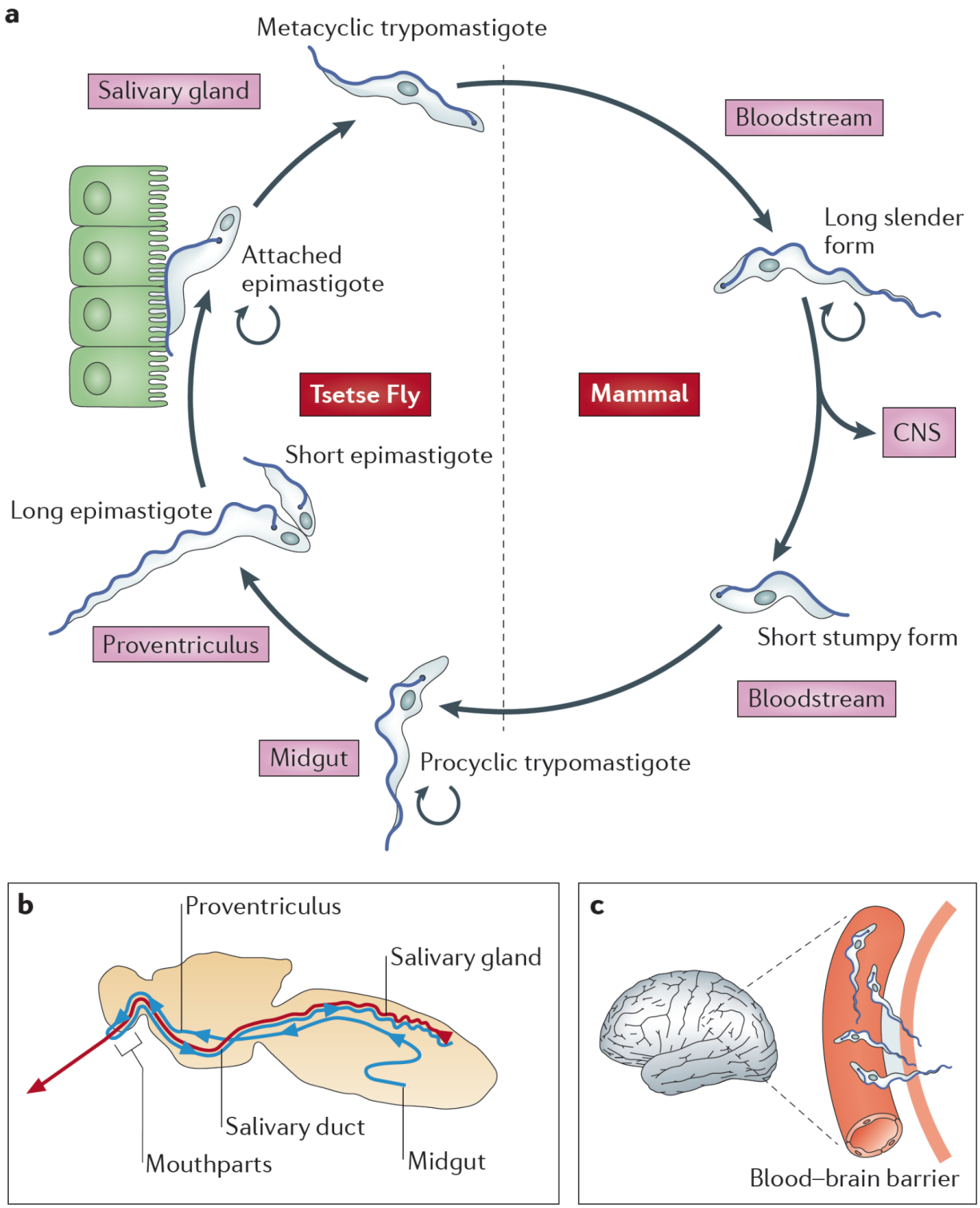


Figure 1.6: Life-cycle of Trypanosoma brucei Parasites

(**a**) Overall life-cycle of *Trypanosoma brucei spp.* Replicative forms of the parasite (*via* binary fission) are shown as circular arrows.(**b**) Parasite migration within tsetse fly, where blue route depicts the journey from the midgut (past the proventriculus, through the mouthparts and salivary ducts) to the salivary glands, and the red line shown the route from the salivary glands to the mouth parts. (**c**) Parasites crossing the blood-brain-barrier to penetrate the CNS.
Figure reproduced with permission from Langousis, G. and Hill, K. L. (2014)[72]
Copyright 2014 Nature Publishing Group.

to the next host.[71–73]

Trypanosoma brucei Metabolism

Bloodstream forms of *T. brucei* species rely on glycolysis for their energy requirements (taking part largely within their glycosomes). Glucose acts as their sole substrate and pyruvate as their major waste product.[32] In fact, evidence shows that up to 90% of their glycosome protein content consist of glycolytic enzymes alone.[6] This is perhaps not surprising, considering that the blood provides a continuous ample supply of glucose.[32] Mitochondrial activity in bloodstream *T. brucei* is largely suppressed, with the exception of an alternative oxidase system (similar to that used by certain plants) to maintain the redox balance.[74] In procyclic forms on the other hand, where mitochondrial function is reactivated, pyruvate is further broken down to acetate and succinate as major end-products, similar as *Leishmania spp.*[32,74] Interestingly, mitochondrial activity is also thought to regulate parasite surface expression of VSG.[71] Most metabolic analyses have largely been limited to bloodstream and procyclic forms of *T. brucei* (whilst relatively little is known of the other morphologies such as epimastigotes), as these two forms are easily cultured *in vitro* and have been comprehensively genetically characterised.[75,76]

1.3.4 Host Immune Factors in HAT

Major modifications of the immune system have been observed in HAT, leading to exacerbation of disease. Bloodstream trypanomastigotes undergo extensive antigenic variation by switching their VSG antigens, which causes the mammalian host to constantly generate new rounds of antibodies specific to the most recently displayed antigens, resulting in cycle waves of parasitaemia in the host, never being able to completely eliminate the parasite.[77] It is this ability to repeatedly switch its antigenic surface that has also severely hampered effective vaccine development. Several invariant surface proteins of *T. brucei spp* have been targeted for candidate vaccine development, yet so far have only generated poor protection *in vivo*.[59] Furthermore, VSG has been shown to induce auto-antibodies by molecular mimicry (causing tissue damage in the host), inhibit complement-mediated killing of the parasite, and induce the secretion of immune-modulating cytokines.[77]

Numerous studies have shown that *T. brucei* infection in humans is linked with a potent inflammatory response in the host, characterised by elevated pro-inflammatory cytokines TNFα, IFNγ and IL-6, as well as regulatory cytokine IL-10, in both CSF and plasma.[78–81] Similar plasma results are also observed in experimental *T. b. brucei* infection in mice, where increased in TNFα and IL-10 were measured in plasma[82], and augmented TNFα, IFNγ and IL-6, detected in murine

brain tissue.[83,84] Evidence suggests that the increase in pro-inflammatory cytokines is produced by activated astrocytes within the CNS. Resulting inflammation weakens the blood-brain barrier (promoting parasitic invasion) and damages brain cells, thus contributing to the neuropathology observed in HAT, particularly in late stage of disease.[83,84]

1.4 Metabolic Profiling of Parasitic Infections

Recent scientific advances in technology and mathematical tools have help steer experimental designs, from restrictive single read-out experiments, to more holistic multivariate approaches, that better reflect the multi-factorial (and often inter-dependent) processes within biological systems. Such approaches include the fields of genomics, transcriptomics, proteonomics, and more recently, the area of "metabonomics". Complementary to these other 'omics', metabonomics studies the systematic metabolic profile of the organism of interest, in response to pathophysiological stimuli, such diet, environment, drugs, etc., through simultaneous measurement of a large subset of metabolites *via* spectroscopic tools and multivariate statistical analyses.[85,86]

Untargeted metabolic profiling describes the qualitative and quantitative collection of "all" low molecular weight organic or inorganic chemicals (metabolites) present within a biological system (i.e. the 'metabolome'). Metabolites are chemicals involved in general metabolic reactions, required for the maintenance, growth and normal biological function.[87] This approach is in contrast with targeted metabolic analyses, which focus on a limited number of related metabolites investigated to test a hypothesis. Metabolic profiling does not have any specific assumptions before testing, and is a "hypothesis-generating" approach instead.[88] This approach is used to reveal a range of metabolites that may be differentially expressed between test groups and/or time-points.

The vast diversity in biochemical properties of the multitude of metabolites present in an analysed biological sample (from their molecular weight, to their polarity, hydrophobicity or volatility), make it currently infeasible for one metabolic profiling technology to reliably detect and measure "all" metabolites. Instead, different analytical techniques are used to measure a large subset of the metabolome, of which nuclear magnetic resonance (NMR) spectroscopy and mass-spectrometry (MS; typically combined with different separation techniques such as gas/liquid chromatography or capillary electrophoresis), are the most commonly used technologies.[88] Whilst some overlap can be found between the two platforms, they inherently measure a different set of molecules, and each have their benefits and shortcomings, discussed in more detail in the General Methods Chapter 2.

Given the many hurdles in current disease management of human parasitic trypanosomatid infections, such as CL and HAT, further research in their pathogenesis remains critical, in order to aid

the discovery of novel potential drug targets and/or diagnostic biomarkers for infection. Within the context of parasitology, metabolic profiling not only has the capability to reveal infection-associated markers, but also has the potential to provide key insights into complex dynamic interplay between metabolic processes of both the host and the parasite. Parasite-related perturbations in host metabolic pathways may highlight different mechanisms involved in either acute or chronic infection, host immune/inflammatory responses, and/or reveal other pathological processes that may have not been previously known. Furthermore, metabolic profiling also provides a comprehensive characterisation of 'normal' biology in uninfected hosts (as shall be presented in this thesis for example), that can give further insight into intrinsic differences in host backgrounds which may affect disease outcome.

Parasitic infections in general have been shown to induce striking alterations in host metabolism, reported in metabolic profiling studies from *in vitro* experiments[44,89,90] and a range of experimental *in vivo* models.[91–95] Infection-associated changes appeared to be wide-spread, affecting the metabolism of glucose, lipids, amino acids, and gut-microbial-associated molecules. In the case of *Leishmania* infection, such studies have been limited to *in vitro* studies only, comparing infection in differentially activated macrophages[44], and as yet have not been investigated *in vivo*. In contrast, metabolic profiling studies for *T. brucei* infection have already been conducted in mice[92], in addition to metabolic characterisation of the parasite cultures themselves.[96] Yet, humans paint a slightly more complex story as their metabolome can be strongly affected by a range of factors including age, gender, ethnicity, diet, weight, etc.[97], making it difficult to extract (often subtle) infection-related variation amongst the data.[98] So far, metabonomic investigations of human infections have largely been focused on bacterial infections (especially urinary tract infections), viral hepatitis B/C infections, or human immunodeficiency virus (HIV) infections.[99] To date, only two parasitic infections in humans have been metabolically phenotyped, both based in helminth infections only (*Onchocerca volvulus*[100] and *Schistosoma mansoni*[101,102]).

Metabolic profiling represented the core method used to investigate *Leishmania* and *T. brucei* infection in this PhD project. Due to its robust and non-destructive nature, ^{1}H NMR was chosen as the main method, to assure efficient initial biomarker characterisation and allow re-use of samples for alternative experimental analyses after NMR acquisition. MS was additionally used in the metabolic investigation of HAT, to provide more detailed information complementary to the NMR results.

1.5 Project Aims and Thesis Outline

The primary aim of this PhD project was to use an untargeted metabolic profiling approach to characterise parasitic trypanosomatid infection within the host. The secondary aim was to combine results from metabolic analyses with other host response measures (further discussed in General Methods Chapter), to explore potential metabolic interactions and, in doing so, obtain a holistic view of the complex pathological processes that occur during infection. In this thesis I will present work from both *L. major* and *T. b. rhodesiense* infections. Core aims for each study are summarised below, whereas specific objectives for each study are presented separately at the beginning of each results chapter. Further information on the common technologies and statistical analyses used in this project is given in the General Methods Chapter 2.

For *Leishmania* infection, investigations were performed using experimental *L. major* infection in two mouse strains, in order to model changes in both a self-healing disease setting and in a more aggressive form of leishmaniasis, as well as determine metabolic markers of infection. Such work included a comprehensive metabolic characterisation of the two uninfected mouse strains (i.e. control groups), to provide a reference baseline value for the infected animals, and to be able to differentiate between inherent strain differences and infection-associated changes between the two mouse models. In addition to metabolic profiling of a range of tissues and biofluids, cytokine measurements and faecal microbial diversity were measured, representing additional markers of host system responses to parasitic infection, which were each assessed individually and in correlation with each other, to investigate the possible interactions between them. Results from the control mice of this study are presented in Chapter 3, separately from those from the *L. major* infected mice, which are discussed in Chapter 4.

Infection-associated metabolic changes for murine *T. brucei* infection have already been characterised in-house[92], however these changes have not been validated in humans. The aim for this part of the PhD was to perform metabolic profiling of biofluids from HAT *rhodesiense* patients, compared with controls, in order to characterise infection associated changes in host metabolism, and determine any metabolic markers of disease. The secondary aims were to determine any metabolic differences between the different stages of HAT, to reveal markers with potential diagnostic staging capability, and investigate possible links between host metabolic and immune (cytokine) measures within patients. Results of the HAT study are given in Chapter 5.

Finally, the General Discussion Chapter 6 will present an overall summary of the major findings from each of the results chapters, their significance and potential impact on current research, as well as discuss general successes, challenges and limitations encountered during this PhD.

1.6 Additional Projects

In addition to my PhD project, I worked on the metabolic characterisation of *Plasmodium falciparum* gametocyte cultures, the sexual forms of malaria parasites that are responsible for disease transmission. The aim was to investigate metabolic changes that are associated with the differentiation from intra-eyrthrocytic asexual parasites into the sexual gametocytes, a process that remains poorly understood. Findings from this work have been published in Malaria Journal, entitled *"Changes in metabolic phenotypes of Plasmodium falciparum in vitro cultures during gametocyte development"*[103], a copy of which can be viewed in the Appendix Publication Section A.

Aside from laboratory-based work, I was also able to undertake field-work within Côte d'Ivoire during my PhD, where I assisted laboratory staff in Dabou hospital and Taabo hospital, by performing clinical assays, to help diagnose medical conditions (e.g. perform malaria thick blood smear tests). A mini-evaluation report on Ivorian clinical diagnostics that I had generated, using these two hospitals as case studies, is also available in the Publications Section A.

Chapter 2

General Materials and Methods

2.1 Methodology Overview 43
2.2 Sample Preparation for ^{1}H NMR Analysis 44
2.3 NMR Sample Acquisition and Data Processing 49
2.4 Clinical Sample Preparation for MS 51
2.5 Metabolic Analysis and Metabolite Identification 58
2.6 Immune and Gut Microbial Profiling 64
2.7 Correlation Analyses between Datasets 75

2.1 Methodology Overview

This chapter describes the key methods used in the presented PhD research. Additional protocol details specific to each study (such as experimental design) are outlined within each results chapter.

A non-targeted approach of metabolite screening was applied in this PhD for maximal capture of the different metabolic pathways that may be perturbed within the host, during parasitic infection. In both murine and clinical studies, this was achieved using primarily ^{1}H NMR spectroscopy. Clinical samples were additionally subjected UPLC-MS, for a more in-depth investigation of lipid metabolism. Sample preparation and spectral acquisition were performed in a randomised manner to minimise inherent systematic errors across samples that may arise from temperature effects, magnetic field drift within the NMR machine, or time-dependent reduction in detector sensitivity in the mass spectrometer.

Given the large amount of spectral information generated for each sample following NMR/MS measurement, multivariate statistical analyses were used to de-convolute and interpret the data.

Such models allowed the identification of peaks within NMR and MS spectra that were differentially expressed between groups, e.g. between control and infected samples. Statistical differences in metabolite levels between groups were then additionally compared using univariate non-parametric statistical tests, to assess the class discriminative capability of each particular metabolite.

Complementary to the metabolic profiling studies, additional measures of host responses were investigated to obtain a more holistic overview of the physiological changes as a result of infection. Such analyses were comprised of plasma cytokine measurements using Meso Scale Discovery plate technology for the murine experiments and Enzyme-Linked Immuno-sorbent assay for the clinical study. Furthermore, faecal microbial diversity assessment was performed for the murine studies using next-generation pyrosequencing technology. These data sets were individually analysed using univariate tests for significant differences, as was performed with metabolite integrals. Significance testing was followed by correlation analysis between the datasets, to co-analyse potential links between the different host responses that were generated towards the parasite.

2.2 Sample Preparation for ^{1}H NMR Analysis

2.2.1 Basic Principles of NMR Spectroscopy

Nuclear magnetic resonance is a physical property of certain types of atomic nuclei which, when placed within a magnetic field, are able to absorb and re-emit electromagnetic radiation. This particular phenomenon only occurs in isotopes with an odd number of protons and neutrons (i.e. they possess an intrinsic "non-zero spin").[104] The nucleus of hydrogen (i.e. a proton) represents the most commonly studied isotope in NMR spectroscopy, owing to its high abundance in biological systems and it sensitivity. Within an electromagnetic field, the nuclear spins of all hydrogens align so that their overall net magnetic vector aligns parallel to the axis of the magnetic field), either in the same direction as the field (low energy α state) or against the field (high energy β state).[1] In a non-perturbed system, the distribution of nuclei between the two energy states is at equilibrium (as defined by the Boltzmann equation).[105] The difference between them is proportional to the strength of the applied magnetic field.[104]

^{1}H NMR spectroscopy is based on the simultaneous excitation of ^{1}H nuclei within a sample placed inside a electromagnetic field of the spectrometer, which causes the αnuclei to "flip" to the β state, using a short radio-frequency pulse. The radio pulse is applied so that magnetic vector (B_1) of the nuclei tilts 90° away from the axis of the electromagnet's internal field (B_0, z-axis) into the

[1] In the absence of a magnetic field, both states have the same energy states.[104]

x–y plane, where the receiver coil is located. These nuclei then "relax" back into the α state, and in doing so, the change in magnetic spin state generates a current that is detected by the receiver coil of the NMR spectrometer. This signal, known as a free induction decay (FID) is typically 'Fourier' transformed, i.e. the time domain information is resolved into a frequency domain, plotted against NMR absorption intensity, resulting in a spectrum containing peaks corresponding to the different ^{1}H nuclei within the acquired sample.[105] Modern NMR machines use powerful superconducting electron-magnets that are cooled to low temperatures (close to 0 K/-273°C in a liquid helium tank, surrounded by another tank of liquid nitrogen), which results in zero electrical resistance and allows the generation of high strength magnetic fields, for increased sensitivity.

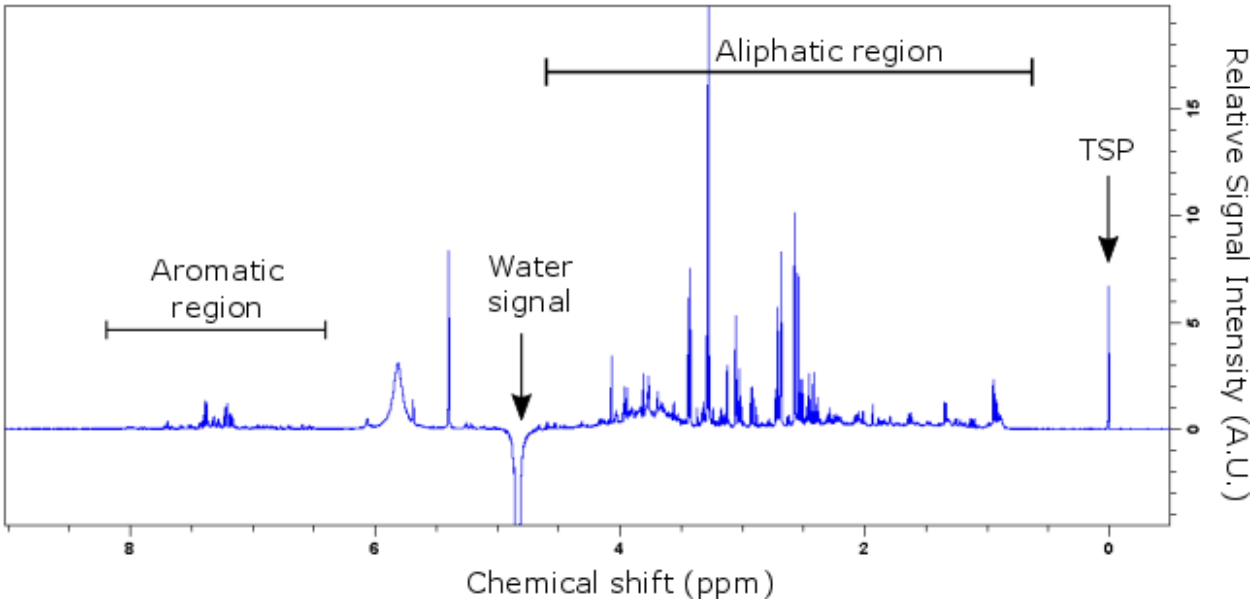


Figure 2.1: A Representative ^{1}H NMR Spectrum of Mouse Urine

A one dimensional ^{1}H NMR spectrum of a representative mouse urine sample, measured using a 600 MHz NMR Spectrometer. Chemical shift (δ, in ppm) is shown along on the x-axis and relative intensities of peaks (A.U.) on the y-axis. Regions which predominantly feature peaks from aliphatic (∼ 0.8–4.6 ppm) and aromatic molecules (∼ 6.4–8.2 ppm) are highlighted in the plot. The figure also shows the TSP peak, the reference chemical added to the buffers for calibration (δ 0.00). The spectrum additionally shows the water signal (inverted artefactually due to water irradiation as part of the spectral acquisition pulse sequence). This water peak, its flanking regions and the TSP peak are removed for final spectral analysis. Abbreviations: A.U., arbitrary units; ppm, parts per million; TSP, Trimethylsilylpropionic-2,2,3,3-d4 acid.

In this presented work, ^{1}H NMR spectroscopy was used to examine the metabolites present within complex biological fluids (e.g. plasma and urine) and tissue extracts. Differences in the physical and chemical environment of each hydrogen molecule (e.g. the size and flexibility of the molecule its located in, nature of chemical bonds, proximity of electronegative atoms, etc.), results in variances in detected currents. As a consequence, different peaks (and peak patterns) are produced, allowing the simultaneous detection of multiple distinct metabolites for each sample.

Furthermore, the areas under the peaks (i.e. integrals) are directly proportional to the number of protons that give rise to those signals, allowing for quantification of metabolite concentration within a sample.[104] An example of a typical mouse urine NMR spectrum is shown in Figure 2.1.

Due to the stable nature of metabolite chemical shifts (δ), NMR spectroscopy represents a highly robust and reproducible analytical method that is optimal for analysing multiple metabolites simultaneously within complex biological mixtures.[106,107] Sample preparation is relatively straight-forward and samples can be acquired, processed and analysed rapidly at a high throughput. Procedures for sample preparation and acquisition for this PhD were based on protocols described by Beckonert *et al.*[108], further outlined below. In contrast, acquisition with LC-MS requires optimisation prior to each sample run (e.g. chromatographic columns need to be conditioned before sample acquisition) and data processing, in particular metabolite identification, can be very time-consuming.[87] Furthermore, as NMR is a non-destructive method, samples can be subsequently stored and re-analysed at a later time-point, or even used for different analyses altogether (as was done in this study with mouse plasma samples). Nevertheless, it is important to note that in comparison with MS techniques, the sensitivity of NMR spectroscopy is generally much reduced, with only the most abundant metabolites in a sample being detected (typically 30–100 individual metabolites in a sample), whereas UPLC-MS can measure thousands of analyte ions.

2.2.2 NMR Buffer Reagents

The composition of NMR buffers used in these studies are outlined in Table 2.1. All chemicals and distilled water (H_2O) used for making solutions were of analytical grade, purchased from SIGMA-Aldrich. Trimethylsilylpropionic-2,2,3,3-d4 acid (TSP) sodium salt was used as the chemical shift reference at δ 0.00 for all samples with the exception of plasma, which was referenced to glucose instead, due to the high levels of plasma protein binding to TSP. All samples were acquired in solution containing D_2O (GOSS Scientific), which served as a field-frequency lock for the NMR spectrometer.

2.2.3 Preparation of Mouse Urine and Plasma Samples

Saline/buffer solutions were pre-chilled at 4°C before sample preparation. Approximately 30 µl of each plasma or urine samples was mixed 1:1 in Eppendorf tubes with either 0.9% saline solution or NMR phosphate buffer (see Table 2.1), respectively. Samples were centrifuged (Biofuge Pico, Heraeus, Germany) at 12,470 x *g* (13,000 rpm) for 5 minutes (mins) and transferred into 1.7 mm NMR capillary microtubes (Bruker, Germany) using a blunt-ended syringe. To determine any

Table 2.1: Buffer Preparations used for NMR Spectroscopy

Buffer	Phosphate buffer	0.9% Saline solution	Bruker urine buffer (10x)	Bruker plasma buffer
Composition	0.2 M Na_2HPO_4 0.043 M $NaHPO_4$ 0.58 mM TSP D_2O:H_2O = 7:3 (v/v) pH adjusted to 7.4	0.9% (w/v) NaCl D_2O:H_2O = 1:9 (v/v)	1.5 M KH_2PO_4 2 mM NaN_3 0.58 mM TSP 100% D_2O pH adjusted to 7.4	0.142 M $NaHPO_4$ 2 mM NaN_3 0.46 mM TSP D_2O:H_2O = 2:8 (v/v) pH adjusted to 7.4
Sample Use	Mouse urine Mouse tissue extracts Mouse faecal extracts Mouse lymph nodes (MAS-NMR)	Mouse plasma	Human CSF	Human plasma

Abbreviations: CSF, cerebral spinal fluid; MAS-NMR, magic-angle spinning NMR; w/v, weight of solute/total volume of solution; v/v, volume of added solution/total volume.

potential contamination products within the buffers, additional blank control tubes for plasma and urine samples were prepared by adding a 1:1 mixture of H_2O with either saline or phosphate buffer, respectively, to the microtubes. Samples were stored at 4°C until acquisition (no longer than 3 days), when they were placed into 5 mm outer diameter NMR microtube holders (Bruker, Germany) to be analysed.

2.2.4 Preparation of Mouse Tissue and Faecal Extracts

All faeces and tissues underwent solvent metabolite extraction for ^{1}H NMR spectroscopy except lymph nodes, which were analysed *via* Magic-Angle Spinning NMR (MAS-NMR) instead — see Subsection 2.2.5 for more details. Before extraction, solvents were pre-chilled on ice and 2 ml skirted bead-beater tubes (with twisting lids, VWR) were filled with $\sim$ 60 mg of 1 mm zirconia silica beads (BioSpec Products Inc, USA).

Solvent metabolite extraction was performed on samples using a ratio of HPLC-grade methanol 400 µl/285 µl H_2O/400 µl chloroform (all SIGMA-Aldrich) per 100 mg tissue, and volumes were adjusted for weights less than 100 mg. The same volumes for tissue extraction were used for faeces, yet for 50 mg instead. Tissues and faecal pellets were defrosted, weighed and cut into sections in order for each sample to have the same final weight, using tweezers and disposable surgical knife blades (which were rinsed in H_2O between samples). Samples were then added to bead beater tubes, pre-filled with zirconia beads and corresponding volumes of ice-cold methanol. Water and chloroform were subsequently added to the tubes, with the final tissue/faecal masses and volumes used for extraction shown below in Table 2.2.

Samples were run twice for 30 seconds at 6000 rpm in the Precellys 24 Tissue homogenizer

Table 2.2: Sample Weight and Solvent Volumes used for Metabolite Extraction

Tissue	**Mass of tissue section (g)***	**Volume of methanol and chloroform (µl)**	**Volume of H_2O (µl)**
Brain – cerebellum	30	120	86
Brain – cortex	45	180	128
Brain – hippocampus	100	400	285
Brain – stem	35	140	100
Heart	60	240	171
Faeces	50	400	285
Ileum	50	200	143
Kidney	60	240	171
Liver	100	400	285
Lung	75	300	214
Spleen	50	200	143
Thymus	30	120	86

*Shows typical tissue section mass that was used for analyses, subject to availability.

(Bertin Technologies, France) and left to incubate on ice for $\sim$ 20 mins, for optimal separation of the aqueous and organic layers. Tubes were subsequently centrifuged at 12,470 x *g* (13,000 rpm) for 10 mins and aqueous extracts (upper layer above the debris and organic layers) were carefully pipetted off and transferred into fresh Eppendorfs. Aqueous tissue and faecal extracts were then dried overnight at 45°C in a centrifugal evaporator (SpeedVac, Eppendorf, Germany), with the Eppendorf lids left open. In order to identify any contamination products that may be present within the buffers or may have arisen during sample preparation (e.g. from plastic tube degradation during homogenisation or in the SpeedVac), additional blank control tubes, containing solvents only without any tissues/faeces, were prepared alongside the samples, treated in the same way. All extracts were subsequently stored at -40°C until analysis. For acquisition, extracts were resuspended in 600 µl NMR phosphate buffer, ensuring maximal dissolution through vortexing. Samples were centrifuged at 12,470 x *g* for 5 mins and 550 µl of each sample was transferred into 5 mm NMR sample tubes (VWR, USA).

2.2.5 Mouse Sample Preparation for MAS-NMR

Due to their small tissue size and weight (1–5 mg), metabolite extraction could not be performed on mouse lymph samples — the final sample suspension in NMR buffer would have been too dilute to yield meaningful acquired spectra. Instead, these samples were analysed *via* MAS-NMR, as described by Beckonert *et al.* 2010.[109] Spinning intact tissues at a high speed (4,000–6,000 Hz) at the magic angle (54.7° relative to the static magnetic field) is known to dramatically increase spectral resolution and sensitivity whilst minimising line broadening effects that are observed, due to e.g. the heterogeneity of the sample, macromolecular and membrane binding and magnetic field inhomogeneity.[110]

Whole lymph nodes were weighed, rinsed in 50 µl D_2O and then directly placed into a 4 mm diameter MAS rotor, using tweezers. A few drops of D_2O was then added to avoid any air bubbles and tissue position was adjusted using the MAS positioning tool. A spherical insert was then added to the rotor (to help fix the position of the tissue) which was closed and tightened with the rotor grub screw and cap. Finally, the rotor was labelled around the bottom edge of the rotor for half a turn (aids the NMR spectrometer to count the rotation speed), ready for immediate acquisition.

2.2.6 Clinical Plasma and CSF Preparation for NMR

Human plasma and cerebral spinal fluid (CSF) samples were centrifuged for 10 mins at 12,470 x g to remove any initial debris. For plasma, 300 µl of sample was transferred into fresh Eppendorf tubes and mixed (vortexed) 1:1 with Bruker Plasma Buffer. For CSF, 280 µl per sample was mixed with 260 µl H_2O and 60 µl Bruker urine buffer in fresh Eppendorfs.[2] All samples were subsequently briefly centrifuged for ∼1 min at 14,000 x g (14,800 rpm) and 570 µl per sample was transferred to 5 mm NMR sample tubes in 96 rack (Bruker, Germany). Each sample tube was sealed with NMR tube caps and "POM" balls (Bruker, Germany), which are inserted into the tube caps to seal them and minimise potential contamination of vaporised solvents (e.g. methanol and/or ethanol that may be co-localised with the samples during fridge storage). Samples were kept at 4°C (for 1–3 days) until acquisition.

2.3 NMR Sample Acquisition and Data Processing

2.3.1 Mouse Sample Acquisition and Processing for NMR

^{1}H NMR spectra for mouse biofluids (plasma and urine) and tissue/faecal extracts were acquired on a Bruker Avance 600 NMR spectrometer with TXI probe head with automatic sample handling carousel (BACS60 autosampler, Bruker), operating at 600.13 MHz (14.1 T) ^{1}H frequency and at a temperature of 300 K, using XWIN-NMR software (Bruker BioSpin, Germany). Acquisition for all samples was performed using a standard one-dimensional (1D) solvent suppression pulse program, (noesypr1d, Bruker). The pulse sequence was as follows: recycle delay–90°–t-90°–tm-90°–ACQ, whereby relaxation delay (RD) is set at 2 seconds (s), 90° pulse length was set to between 11–16.5 µs (depending on sample matrix), t = short delay set to 2 s, tm = mixing time set to 100 ms and ACQ= data acquisition period at 2.73 s per scan. Water suppression was achieved through a secondary irradiation at the water resonance frequency during RD and tm. Number of dummy

[2] Initial protocols were for 400 µl CSF mixed with 140 µl water and 60 µl Bruker urine buffer but this was adjusted due to limited original sample volume.

scans (ds) was set to 8 for all samples, followed by 128 acquisition scans (ns) of ^{1}H FID's for tissue and faecal extracts or 256 scans for plasma and urine samples. Samples were acquired using a spectral width of 12 KHz over 65,536 (65 k) data points.

As plasma is generally rich in species with high molecular weights (including many proteins and lipids), an additional Carr–Purcell Meiboom–Gill (CPMG) spin echo pulse program was applied to all plasma samples (ns = 256), which reduces signals from large macromolecules such as lipids and proteins, to better examine metabolites with small molecular weights.[111] The pulse sequence for this was RD–90° pulse–(τ–180°)–τ_n–ACQ, with a fixed total spin echo time $2n\ \tau$ of 64 ms, number of loops n = 160 and τ = 200 µs. Scans were acquired over 65k data points (12 KHz spectral width).

For MAS-NMR of lymph nodes, samples in rotors were read on a Bruker Avance 600 NMR spectrometer, equipped with a triple-resonance MAS probe. The rotor spin rate was set to 5000 Hz at 283 K and samples were acquired using a standard 1D solvent suppression pulse program, followed by a CPMG program (total spin echo time $2n\tau$ of 120 ms). For both programs, ds = 8, ns = 128, spectral width 12 KHz and total FID was collected over 65 k data points.

Finally, all FID's were multiplied by an exponential function with a 0.3 Hz line broadening prior to Fourier transformation, used for spectral processing and analyses. Spectra were preprocessed using automatic phasing, baseline correction, and TSP reference peak calibration *via* nmrproc (courtesy of Dr Tim Ebbels and Dr Hector Keun), an in-house algorithm in MATLAB software (version R2012b, MathWorks, USA), and Topspin 3.1 software (Bruker BioSpin). TSP peaks, water peak and flanking regions (which have distorted baselines due to the dominant water signal) and urea regions (the latter for urine samples only) were removed prior to automatic spectral alignment[112] and probabilistic quotient normalization[113], using an in-house MATLAB script developed by Dr Kirill Veselkov.

2.3.2 Clinical Sample Acquisition and Processing

^{1}H NMR spectra for all clinical samples were acquired on a Bruker Avance 600 NMR spectrometer, fitted with a BB probe head (newer with higher sensitivity than the TXI probe) and refrigerated automatic sample handling carousel (SampleJet autosampler, Bruker), using TopSpin 3.1 software (Bruker BioSpin). Standard Bruker pulse programs for plasma and urine samples were used for acquisition of clinical plasma and CSF, respectively, whereby acquisition temperature was set to 300 K for plasma and 310 K for CSF. Samples were measured using a standard 1D pulse program with water suppression for general metabolite screening (ds = 4, ns = 64 for plasma and ns =

256 for CSF) applying a pulse program as described previously for mouse samples, with slight modifications due to the newer probe head, e.g. an increased size of spectrum (spectral width of 18 kHz over 132 k data points). Similar to mouse plasma acquisition, an additional CPMG pulse program (n = 128, τ = 300 µl, total spin echo time $2n\tau$ = 76.8 ms, ds = 4, ns = 64) was also measured for human plasma, to investigate small metabolite peaks. Furthermore, a standard two-dimensional (2D) short J-resolved pulse programme (ds = 16, ns = 2)[114] was performed, for additional information about multiplicity of spectral peaks, to aid metabolite assignment.

2.4 Clinical Sample Preparation for MS

2.4.1 Principle of UPLC-MS

Whereas NMR instruments can typically detect up to the nanomolar range, modern MS techniques can detect molecules with picomolar concentrations[115], giving thus a wider range of analysed molecules. It is important to note, however, that not all metabolites in NMR can be measured *via* UPLC-MS (and vice versa), highlighting the fact that both of these technologies provide complementary information. MS is based on the formation of (both negatively and positively) charged molecules (ions) and their separation and measurement of mass-to-charge ratio (m/z), in a gaseous state. Ion separation and detection is performed under high vacuum pressures, in order to minimise the number collisions between the ions and/or molecules, which can affect the final mass resolution, mass accuracy and sensitivity of mass-spectrometers.[88] Samples are typically subjected to additional separation step, prior to ion formation — commonly consisting of gas chromatography (GC), liquid chromatography (LC), hydrophilic interaction chromatography (HILIC) or capillary electrophoresis (CE) — to provide an additional dimension of sample separation which improves sensitivity and analyte ion identification.[88,116,117] HILIC and CE-MS are especially suitable for small and highly polar metabolites, whilst GC-MS is optimal for the examination of low-molecular weight species that are both volatile and thermally stable (the majority of endogenous compounds do not have such properties, e.g. complex lipids, though some simpler compounds can be still be analysed with GC-MS following hydrolysis and chemical derivatisation steps).[88,116]

In contrast, modernisation of liquid chromatography technology to HPLC, and more recently UPLC, allows the accurate detection of wide coverage of low and high-molecular weight compounds within complex biological mixtures, with high specificity.[116,118] Thus, HPLC/UPLC systems coupled with MS have become widely used in laboratory research and, more recently, in clinical settings.[87,117,119] Currently, a range of various MS and chromatography instrument combinations

are used, yet for the purpose of concision, I will focus on UPLC coupled to a Q-ToF MS machine, the technology used for the lipid profiling of plasma from the HAT study. An schematic overview of the main components involved during sample acquisition is shown in Figure 2.2. UPLC and MS conditions for lipid profiling were based on protocols as described by Isaac *et al.*[120] and Vorkas *et al.*[121]

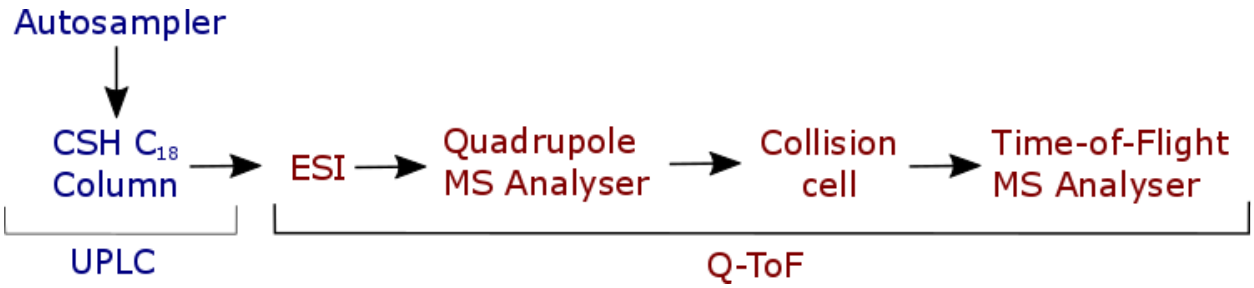


Figure 2.2: Main Instrument Components of UPLC-MS used for Lipid Profiling

Test and QC samples are automatically aspirated from MS vials in autosampler, mixed with the mobile phase and are run through the charged surface hybrid (CSH) C_{18} UPLC column. Samples are then ionised *via* electron spray ionisation (ESI), followed by ion selection in the quadrupole MS analyser. Analyte ions then pass through the collision cell, which contains inert gas molecules (nitrogen), whereby collision energy is set low for minimal fragmentation. In contrast, quality control samples were acquired both at low collision energy as well as using data dependent analysis, whereby increasing collision energies were applied to the collision cell, resulting in increasing number of generated fragment ions. All ions are finally acquired *via* the Time-of-flight MS analyser, which records m/z for each analyte ion hitting its detector.

An important limitation of metabolic profiling *via* LC-MS is that whilst general protocol guidelines can been generated[122,123], optimal parameters will vary depending both on the instruments (including the manufacturer and type of LC and MS instruments, on their combination, on the usage of the LC column and/or MS detector before a sample run, etc.) in addition to nature of analysed sample itself.[117] As a result, optimal operation conditions need to be determined on a case-by-case basis. This fact makes it highly difficult to generate generalised databases that combine both m/z and RT values and further slows down the metabolite identification, which represents one of the major bottlenecks within LC-MS data analysis.[124]

Ultra Performance Liquid Chromatography

Firstly, MS vials containing the samples are placed in the autosampler, where a controlled volume of each sample becomes aspirated *via* a syringe and suspended in a solution known as the 'mobile phase'. Mobile phases are comprised of volatile solvents (such as acetonitrile and water solution) which are suitable for subsequent coupling with MS *via* electron spray ionisation (ESI) as they easily evaporate, allowing the detection and analysis of just the endogenous substances within the

samples. They contain an acid and base (in this case formic acid [FA] and ammonium formate) which aid the ionisation process of the molecules within the sample. A combination of two mobile phase solvents are used in a gradient over the course of the run, with the first mobile phase solvent being more aqueous (used to flush out the polar molecules) whilst the second mobile phase is more organic (to elute the non-polar metabolites). Samples within the mobile phases are then ran through the UPLC column that is packed with porous particles (e.g. silica beads), known as the 'stationary phase'. In reverse phase UPLC (i.e. RP-UPLC, as is the case for lipid profiling), these particles are coated with non-polar molecules containing 18 carbons within their acyl chains (C_{18}). Separation of the molecules within the samples is based on their affinities with the mobile and stationary phases. For RP columns, polar substances pass through the column the quickest whilst non-polar molecules such as lipids will interact with the lipophilic beads and thus migrate slower through the column.[3] Furthermore, a charged surface hybrid (CSH^{TM}) UPLC column was used for the lipid profiling, which has been specifically developed by manufacturers for improved separation than conventional RP-UPLC, based on the lipid species' acyl chain length and number and position of saturated bonds.[120]

Electron Spray Ionisation

Following chromatographic separation, samples are transformed from a liquid state into a gaseous state and ionised, during a process known as electron spray ionisation (ESI). Unlike other ionisation techniques, such as electron ionisation (EI) — where samples are bombarded with high-energy electrons — ESI is known as a 'soft' ionisation technique, which allows the analysis of intact large macromolecules with relatively little fragmentation.[125] Molecules within the mobile phase are charged by the non-covalent addition or loss of chemical species e.g. H^+, NH_4^+ and Formate/$CHOO^-$ (aided by the acid/base within the mobile phase).[88] Samples are subjected to a high voltage and sprayed into small liquid droplets containing charged analyte ions in the mobile phase. Solvents within the mobile phase evaporate off ('desolvation') and the charged droplets becomes smaller and smaller, until finally the repulsive electrostatic forces between the charged molecules overcome the surface tension forces holding the droplet together, causing it to burst and release smaller aerosols and/or naked ionised molecules to be analysed by MS.[118,125] The applied voltage is interchangeable, and can thus be set to either positive mode (ESI+) or negative mode (ESI-), to analyse either positive or negative ions, respectively.[4] An important consideration in LC-MS analysis is ion suppression,

[3] In normal phase LC, the stationary phase is hydrophilic and thus non-polar molecules elute first instead of polar. This type of LC is used in HILIC, optimal for analysis of polar/charged molecules.

[4] In this case, all samples were first acquired in ESI+ and, following conditioning of the column, were then acquired in ESI-.

which affects ESI in particular. This phenomenon is thought to be caused by impaired droplet formation or evaporation due to presence of either non-volatile or poorly volatiles solutes, which result in reduced amounts of analyte ions in the gas phase that ultimately reach the detector.[126]

Mass Spectrometry Analysers: Quadrupole and Time-of-Flight

Analyte ions are then analysed and separated based on their m/z *via* two mass analysers, a quadrupole, used primarily as an ion filtering/selection system, followed by a time-of-flight (ToF) analyser, where analyte ions are finally detected. The quadrupole consists of four cylindrical, electrical rods that are parallel to each other. A direct current (DC) potential is applied to one of the opposing rod pairs whilst the other two are linked to an alternating radio-frequency (rf) potential, which switches the polarity of the rods and generates an oscillating magnetic field. The resulting effect is that certain ions with specific m/z ratios will be able to travel in a sinusoidal motion through the quadrupole (alternating between being attracted to or repulsed by the rods), whereas other ions will collide with the rods. The different potentials within the quadrupole can thus be ramped to select for the transmission of a wide range of ions.[125]

Ions are subsequently accelerated and analysed *via* the ToF analyser. For a given charge, species with different masses will separate according to their velocities. Unlike the quadrupole analyser, no filtering/selection of ions occurs as all species are detected. The ToF instrument used in this study also includes two reflectrons, which are electrostatic ion mirrors (consisting of a series of ring electrodes with very high voltage) that change the course of the ions within the instrument. Ions with higher kinetic energies (i.e. faster ions) penetrate deeper into the reflectron and are gradually repelled, whilst slower ions travel less deep, resulting in both fast and slow ions with identical m/z values reaching the detector simultaneously, causing improved resolution and more accurate measurement.[125]

Automatic Data Dependent Acquisition of MS/MS

Tandem mass spectrometry (MS/MS) involves the selection of sample ions by the first mass analyser (in this set-up, the quadrupole), followed by ion fragmentation and subsequent analysis by the second MS analyser (i.e. the ToF). The fragmentation is caused by the bombardment of sample ions with neutral gas molecules such as nitrogen gas, known as collision-induced dissociation.[127] The amount of collision energy applied will result in varying degrees of ion fragmentation, whereby each original sample ion ('parent') will generate characteristic fragments ('daughter ions') which aid its identification. A pooled quality control (QC) sample, consisting of an equal mix of all the

tested samples (thus representative of all samples) were subjected to both conventional MS as well as MS/MS analysis. The initial parent ions were selected and fragmented and recorded using an automatic instrument setting known as automatic data dependent acquisition (DDA).[127]

UPLC-MS Data Overview

The final output sample acquisition *via* UPLC-MS consists of two data dimensions. The data from the UPLC shows the intensities of substances that separated during the chromatographic process vs. the time it took to elute out of the column (i.e. retention time). Additionally there is the MS spectral data, which consists of the peak intensities vs. m/z for all the analysed ions. A representative total ion current chromatogram (TIC, showing the relative intensity of all analysed ions) and an example MS spectrum is provided in Figure 2.3.

2.4.2 Human Plasma Preparation for Lipid-MS

Complex samples with high protein content such as plasma and serum need to be treated to remove the protein prior to acquisition on the UPLC-MS, to minimise interference with separation and/or detection of metabolic analytes. A wide selection of common protein removal techniques of plasma for lipid LC-MS have been previously tested and compared in-house by Sarafian and colleagues (2014).[128] Their results demonstrated that protein precipitation by isopropanol (IPA) was the most efficient method that yielded the broadest coverage of different lipid species. Consequently, IPA precipitation was the method of choice used for the human plasma samples in the current study, following protocols as they described.[128] Firstly, 150 µl of HPLC-graded IPA (SIGMA-Aldrich), previously chilled to -20°C, was slowly added to 50 µl plasma in fresh Eppenorf tubes and vortexed for $\sim$ 20 s. Mixed IPA/plasma samples were left to incubate at room temperature for 10 minutes and then placed in the freezer to incubate at -20°C overnight. Two blank controls, containing 50 µl H_2O and 150 µl IPA, were prepared alongside samples. The next day, samples were defrosted and centrifuged (Eppendorf microfuge 5417-R, Germany) at 20,817 x *g* (14,000 rpm) at 4°C for 20 mins. 100 µl IPA/plasma mix from each sample was transferred from the Eppendorf into 100 µl polypropylene inserts with polyspring (Waters Ltd., UK), placed inside 1 ml glass MS sample vials. Vials were sealed with silicone lids (Waters Ltd., UK), ready for acquisition.

The QC was prepared by collecting a further 50 µl from each sample and combining into one fresh 2 ml Eppendorf tube. This was mixed through brief vortexing and centrifuged at 20,817 x *g* at 4°C for 5 mins. A volume of 1 ml of neat QC master sample was transferred into a 1 ml MS vial. The remaining QC sample was used to set up the following dilutions, directly in the polypropylene

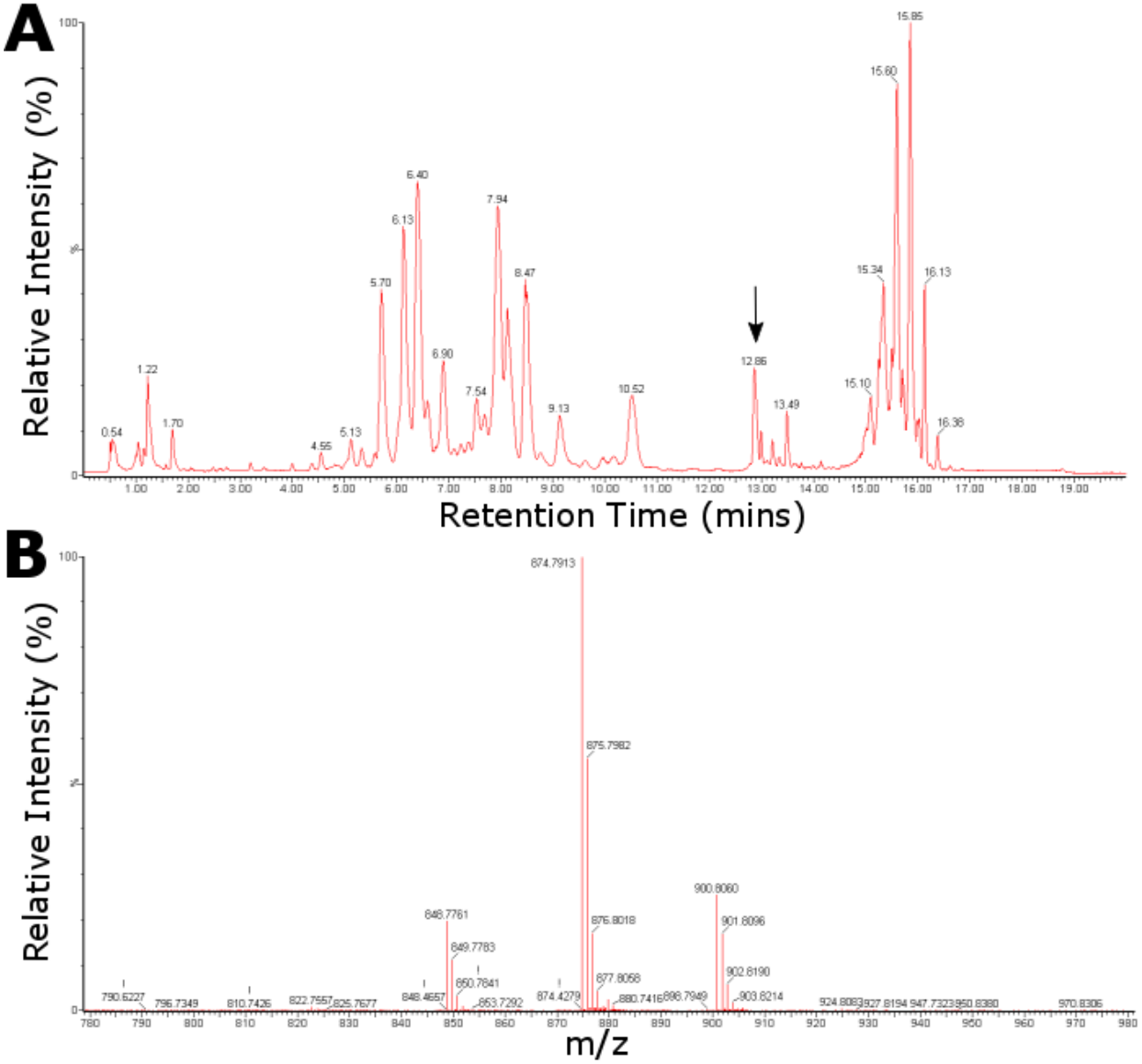


Figure 2.3: A Representative Chromatogram and MS Spectrum via UPLC-MS

(**A**) A total ion chromatogram (TIC) of a quality control sample acquired in ESI+. (**B**) The corresponding MS spectrum of the predominant analyte ions detected in the peak highlighted in the TIC with an arrow (retention time 12.86 mins).

inserts in MS sample vials:

- 1:1 (50 µl QC + 50 µl IPA)
- 1:2 (30 µl QC + 60 µl IPA
- 1:4 (20 µl QC + 80 µl IPA)
- 1:8 (10 µl QC + 90 µl IPA)

Finally, additional blank solvent control samples were prepared, transferring 100 µl of H_2O and 100 µl of IPA into separate MS vials.

2.4.3 Lipid-MS Acquisition and Pre-processing

Global lipid profiling of plasma samples was performed on an Acquity UPLC machine with an Acquity UPLC reverse-phase C_{18} column, coupled to a Quadrupole time-of-flight (Q-ToF) Premier^{TM} mass spectrometer (Waters Ltd., UK). Samples were placed in the autosampler, which was kept at 8°C throughout run. The QC was injected (minimum of 10 times for each ESI mode) before acquiring test samples, to optimise the injection volume and to condition the column. The QC was again injected every 8 samples, and after the test samples, for quality assurance of the acquisition. After the sample run, the different QC dilutions (neat, 1:2, 1:4, 1:8) were acquired, followed by the blanks, to aid identification of contaminant/artefactual peaks (that did not respond to dilution), as well as monitor dilution-related variability within our dataset.

Methods for separation conditions for reverse-phase UPLC-MS had been established by Isaac *et al.*[120] and Vorkas *et al.*[121] Gradient elution was performed using a C_{18} column (1.7 µm particles, 2.1 x 100 mm) with charged surface hybrid technology (CSH^{TM}; Waters Corporation, USA) which was kept at 55°C for the run. Mobile phase A consisted of a 60:40 (v/v) acetonitrile/H_2O mixture with 0.1% (v/v) formic acid (FA) and 10 mM ammonium formate. Mobile phase B was comprised of 90:10 (v/v) of isopropanol and acetonitrile, also containing 0.1% (v/v) FA and 10 mM ammonium formate. All solvents were purchased from SIGMA-Aldrich except HPLC-grade water which was purchased from Fischer, Germany. Flow rate was set to 0.4 ml/min and sample injection volume was 5 µl and 10 µl for ESI+ and ESI-, respectively.

The MS conditions were set to the following: capillary voltage at 3 kV for ESI+ and 2.5 kV for ESI-; sample cone voltage, 30 V (ESI+) and 25 V (ESI-); source temperature at 120°C; desolvation temperature at 400°C; desolvation gas flow at 800 L/hour and cone gas flow at 25 L/hour. A 0.2 ng/µl leucine encephalin (mass/charge ratio [m/z] 556.2771 in ESI+ and m/z 554.2615 in ESI-) solution at 20 µl/min was used as the lock mass reference to optimise instrument mass accuracy. Spectral data were collected over a scan range of 100–1200 m/z in centroid mode, with lock mass

scans collected every 30 seconds and averaged over 3 scans for mass correction. The total run time for each sample was 20 mins. For structural assignment of molecules (see Section 2.5.4), data-dependent acquisition (DDA) of the QC sample was additionally performed in each ionisation mode, with collision energies were automatically set to 5, 20, 30, 40 and 60 eV.

Chromatograms and spectra in both ionisation modes were displayed using MassLynx software (version 4.1; Waters Corp.), as shown in Figure 2.3. Data files were subsequently converted into NetCDF format for processing using Databridge (within MassLynx). XCMS package (in R software)[129] was used for the pre-processing of the data (including in-house algorithms for peak detection alignment, grouping, zero filling, and mean-fold change normalisation)[121], resulting in workable dataset based on median m/z ratios and median retention time (RT) integrals of all measured analyte ions, i.e. 'features' (courtesy of Dr. Panos Vorkas). The dataset was imported into Microsoft Excel (2013, Microsoft, USA), and the coefficient of variation (CV) was calculated for each feature integral (CV = [standard deviation/mean]). Features with a CV greater than 30% in the QC samples were removed, and the remaining data were imported into SIMCA software (v. 13.0, Umetrics, Sweden).

2.5 Metabolic Analysis and Metabolite Identification

2.5.1 Multivariate Analyses for NMR

Spectral data from both murine and clinical NMR experiments were imported into SIMCA software and scaled using standard unit variance (UV, where peak heights are divided by standard deviation) or pareto method (peak height divided by the square root of standard deviation)[130], depending on the species and biological matrix. Such pre-treatment of spectral data is required to help redistribute the weighting of unequally expressed metabolite levels before modelling (e.g. to take equal account of both very small peaks and large peaks within a spectrum), in order to optimise the extraction of biologically relevant discriminatory markers.

Spectral data were then modelled by Principal Components Analysis (PCA), using SIMCA software. PCA, first invented by Karl Pearson in 1901[131] — then further developed and named by Harold Hotelling[132] — represents an unsupervised model, i.e. where class assignments of samples are not considered in the modelling process. A PCA model reduces multi-dimensional data of each sample (in the context of NMR, the many peaks within a spectrum) and approximates and projects the data onto a new set of axes known as principal components (PC's)[133], as shown in Figure 2.4. The first principal component (PC1) is modelled in the direction of the greatest systematic variation

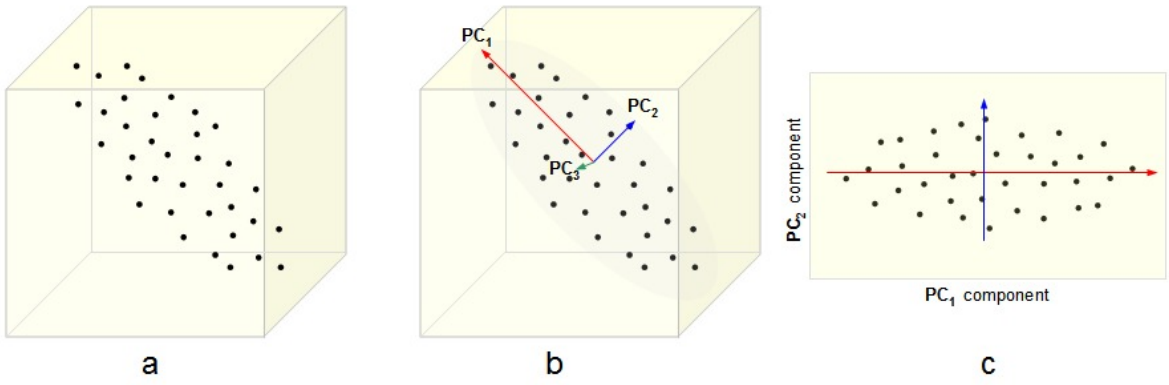


Figure 2.4: Geometric Representation of PCA

(a) Figure shows a schematic representation of sample data set in three-dimensional space. (b) First three orthogonal principal components (PC's) aligned and ordered in the direction of the maximal variation. (c) Data projection of the first two components only, i.e. PCA scores plot of PC1 vs. PC2.
Copyright 2003 held by Lydia E. Kavraki. Figure reproduced under Creative Commons Attribution License (by 1.0).[134]

within the data. The second component (PC2) is based on the second largest variation, orthogonal (in mathematical terms, thus linearly independent/uncorrelated) to PC1, creating a geometrical "plane", between PC1 and PC2 (Figure 2.4b). Further PC's can be added (creating more planes) to increase the descriptive ability of the model. The total explained variance of the model is denoted as R^2, also known as the parameter of the "goodness of fit". This can be shown as an individual value for each PC, or cumulative for all PC's used in the model. Samples are schematically represented as co-ordinates on PC planes, plotted on "scores" plots, whereby their relative positions to each other is dependent on their spectral similarity (i.e. clustered samples share many peaks whilst those segregated from each other show marked differences in their spectra). The variables responsible for their positioning in the scores plot are shown in the corresponding PCA "loadings" plot. PCA models were used to assess the variation that dominated within the samples and to identify sample outliers within the dataset (e.g. samples containing artefactual peaks, baseline distortions or too high signal:noise ratios).

Specific differences between sample groups (and/or time-points for a given biological matrix) were further examined using a variant of Partial Least Squares (also known as Project to Latent Structures, PLS) analyses in SIMCA. PLS was first introduced and developed by Herman and Svante Wold,[135,136] and represents a supervised model, where sample classification is incorporated into the model. Similar to PCA, multidimensional data of each sample in PLS is mathematically transformed and projected onto new planes on scores plots. Yet PLS is a multivariate regression

model based on correlating information between two data matrices, (e.g. the spectral data x and descriptive data y). The aim of PLS regression is to be able to predict y from the x-dataset, and to describe the relationship between them. Thus where PCA is modelled on maximal inherent variation of one dataset x, components in PLS are extracted for maximal covariance between data x and y. In the presented PLS analyses, the data y-matrix is categorical (consisting of a binary classification system e.g. infected 1 vs. uninfected 0). In this manner, PLS is used to find the maximal separation (or discrimination) between the classes in y, as described by information from the x matrix, known as PLS-Discriminatory Analysis (PLS-DA).[137]

A recent extension of PLS-DA model that has been widely used in metabonomics is Orthogonal PLS-DA (O-PLS-DA), which applies an orthogonal signal correction (OSC) filter onto the dataset.[138] Variation within the x data (spectra) is decomposed into two sections: firstly, into variation that is aligned with the variation of the descriptive y-data (i.e. classification). The remaining variation that is orthogonal to the y-data variation, represents inherent group variation that is uncorrelated to the difference between the classes. The differences between PLS-DA and O-PLS-DA can be visualised in Figure 2.5. Using O-PLS-DA, discriminatory loadings (i.e. spectral regions that contribute towards the class separation) are easier to extract as the systematic variation within the classes are minimised, facilitating the detection of more subtle spectral differences between groups.

An additional visualisation tool of the loadings has been developed by Cloarec *et al.* for NMR spectra in MATLAB software, in which the covariance of each loading is back-scaled onto a pseudo-NMR spectrum.[140] The resulting O-PLS-DA loadings plot show both the direction of the variation in relation to class as upwards and downwards peaks. The strength of the correlation is demonstrated by the colour of the peaks, based according to their correlation coefficient R^2 values, with red meaning high correlation and blue meaning none (see Figure 3.3 in the Chapter 3 as an example).

Validity of the discrimination of PLS-DA/O-PLS-DA models can be determined for the models' Q^2 (or Q^2Y) value, a parameter calculated from a 7-fold leave-one-out process, used to estimate the model's predictive capability for classifying samples correctly.[141] It is recalculated for each component that is added to the model, whereby the higher the value, the stronger the model is (1 being the absolute maximum). Components are added to models until the cumulative Q^2 value no longer increases, reaching the models maximal predictive capacity. A negative Q^2 value indicates that the model is not predictive of class discrimination. This parameter is mostly relevant to supervised discriminatory models such as PLS-DA/O-PLS-DA to assess the model robustness in classification, yet Q^2 is also useful for PCA models (i.e. Q^2X), to ascertain the maximal number

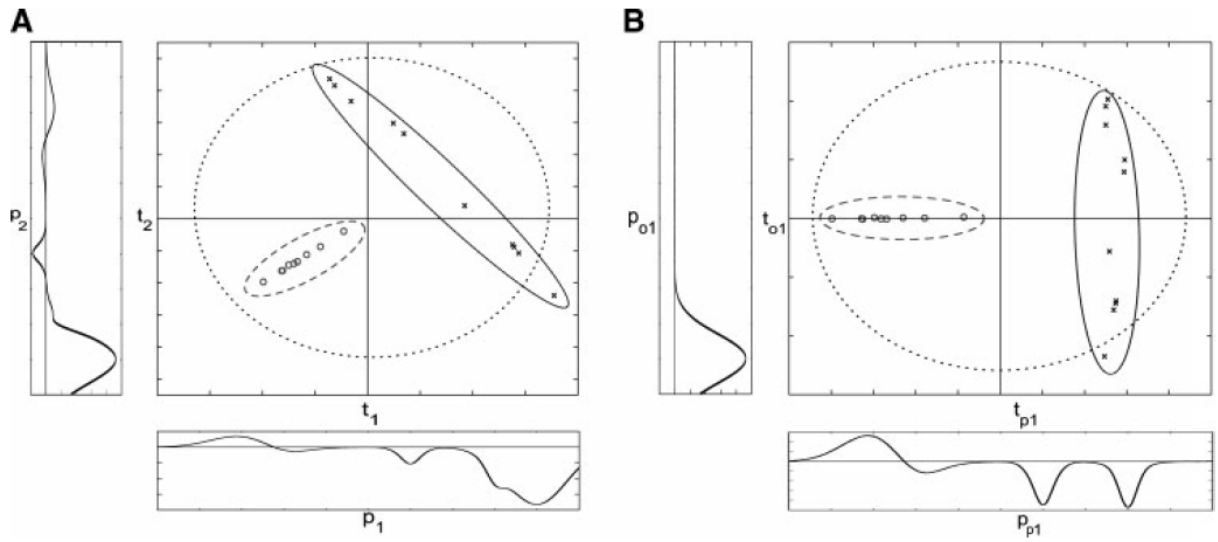


Figure 2.5: Comparison between PLS-DA and O-PLS-DA

Examples of scores and loadings plots showing simulated spectral data within a two-class system (class 1 in circles and class 2 as x's). (**A**) In PLS-DA model, discriminatory direction between the classes is a combination of both first (t_1) and second components (t_2), which is also confounded by the high within-class variation of class 2 in the corresponding loadings (p_1 and p_2). (**B**) In the O-PLS-DA models, where t_{p1} is the predictive (first) component and t_{o1} is the orthogonal (second) component, data is transformed so that the high within-class variation in class 2 is now perpendicular to class 1, thus facilitating discrimination between loadings that contribute to the class separation (p_{p1}), and inherent non-correlated variance (p_{o1}).
Copyright 2007 John Wiley & Sons, Ltd. Figure reproduced under Creative Commons Attribution License.[139]

of principal components that should be used for analyses, as described for supervised models.

Discriminatory metabolites from ^{1}H NMR data were designated as significant based on the following threshold criteria:

- metabolites were extracted from O-PLS-DA models with positive Q^2 values
- corresponding Pearson-product moment correlation coefficients of metabolite peaks from O-PLS-DA surpassed the minimum critical values for $p < 0.05$, for the given degree of freedom of that model[5], for selection of significantly differentially expressed metabolites
- metabolite peaks had a minimum Receiver Operating Characteristic (ROC) of 0.8, for extraction of metabolites with high sensitivity/specificity in differentiating between the two analysed classes

2.5.2 NMR Metabolite Assignment

Assignments of metabolites were performed using a combination of chemometric sources. Firstly, their 1D ^{1}H NMR resonances were matched with those within a 600 MHz standard library data-

[5]equal to the total number of analysed samples minus 2

base in Chenomx NMR suite 8.1 Profiler (Chenomx Inc., Canada) software, or if not present in Chenomx, using the online Human Metabolome Database (HMDB).[142–144] Additionally, corresponding chemical shift regions and multiplicities (peak shapes i.e. singlets, duplets, triplets, multiplets, etc.) were compared with in-house database library and published literature — the latter performed separately for human plasma[115,145] and CSF[146,147], and for mouse samples (urine[93,148–150], plasma[93,148–150], faeces[93,149–151], and tissues[152–155]).

Furthermore, statistical total correlation spectroscopy (STOCSY) was performed to aid metabolite assignment.[156] STOCSY analysis is based on the generation of a correlation matrix between all spectral datapoints across a set of 1D spectra. As all peak intensities arising from the same molecule will have identical ratios across the assessed spectra (providing the acquisition conditions were the same between samples), these peaks are shown as strongly correlated within the correlation matrix. This tool thus helps the identification of different spectral resonances for a given metabolite, as well as potentially revealing metabolites that are weaker or negatively correlated, arising from molecules that share a common metabolic pathways. Consequently, STOCSY proved particularly useful to aid identification of those metabolites that displayed more than one resonance signal within a spectrum. A list of ^{1}H NMR spectral assignments for murine biofluid and tissues can be found in Appendix Tables A.1 and A.2, respectively. Corresponding list of assignments for human plasma and CSF metabolites is shown in Appendix Table A.7.

2.5.3 MS Data Processing and Multivariate Analysis

Median integral values for lipid features were imported from Microsoft Excel into SIMCA software and pareto scaled. PCA (for each ESI mode) was initially performed on features from all test samples (i.e. samples from patients and controls), QC's, and blank controls. Features present only in the blank controls or in QC samples (and either absent or present in negligible levels across test samples) were deemed as contaminant peaks and selected and removed for final analysis, as described by Vorkas *et al* 2015.[121] Experimental reproducibility of both ESI+ and ESI- runs was then evaluated with separate PCA models, using data from only the test samples and QC's, whereby close grouping of the QC samples demonstrates analytical stability needed for metabolic profiling. Once this was confirmed, QC samples were excluded and final analyses was performed on the test samples only, using PCA, PLS-DA and O-PLS-DA. Model evaluation was performed as described previously for NMR data. The variables responsible for the separation observed in the O-PLS-DA scores plots were visualised in corresponding S-plots in which the integral of each metabolite feature is represented as one co-ordinate.[157] The covariance of variables (p[1]) is shown along the x-axis,

where variables furthest horizontally from the zero value display the highest contribution towards the model. The correlation coefficient of each feature with their predicted class is plotted along the y-axis (p(corr)[1]), representing the reliability of that feature (the most reliable being furthest vertically away from zero on the y-axis). Threshold criteria for selection of discriminatory markers for MS was as follows:

- features were extracted from O-PLS-DA models with positive Q^2 values
- using the O-PLS-DA S-plot for guidance, variables with a minimum correlation value p(corr)[1]>0.5 and covariance p[1]>0.05 were selected, to extract the most discriminatory and robust features
- features had a p-value < 0.05 when tested in Welch (unequal variance) T-test, followed by Benjamini–Hochberg multiple test correction (see Section 2.6.4), for final extraction of significant features
- features were identified as containing only one whole copy of molecular ions (see Section 2.5.4)

initial selection of markers were variables

2.5.4 UPLC-MS Metabolic Feature Assignment and Nomenclature

Identification of plasma metabolic features from the HAT study was based on matching the measured mass/charge (m/z) values against online MS database METLIN (`https://metlin.scripps.edu`) — ensuring mass accuracy was < 10 ppm — and literature[121,158,159], combined with comparison of chromatographic retention times with manufacturer's typical RT ranges for lipid subclasses.[120] This was complemented with MS/MS information obtained in the DDA experiments of QC samples, showing ion fragmentation patterns of the analysed features, aiding the identification of characteristic secondary fragments. For example, a fragment with 184.07 m/z in ESI+ typically represents a phosphorylcholine ion head group, present in phosphatidylcholines, lysophosphatidylcholines and sphingomyelin lipids species.[159] Only features identified as containing a complete single copy of the analysed molecular ion (M) were assigned as final discriminatory markers, e.g. $[M+H]^+$ and $[M+NH_4]^+$ in ESI+ and $[M+FA-H]^-$ in ESI-, whereas discriminatory fragments and adducts were omitted. General nomenclature for each lipid species was used, which consists of the abbreviation for the lipid subclass, followed by the sum of carbon atoms, a colon, and the number of unsaturated bonds present in the molecule, e.g. PC (34:2) indicates a phosphatidylcholine lipid, with a total of 34 carbons and 2 unsaturated bonds. In glycerolipids and glycerophospholipids, the presence of an ether bond between the fatty acid chain and the glycerol (as opposed to the more common ester bond) is indicated by the addition of either "O" (for O-alkyl-bond) or "P" (for O-alk-1-enyl-bond) before the total carbons:unsaturated bonds. Sphingolipids

are also annotated with the number of hydroxylgroups present in the sphingoid base (m for mono, d for di, t for tri), followed by the number of carbons:unsaturated bonds, e.g. SM (d38:1).[160]

2.6 Immune and Gut Microbial Profiling

2.6.1 Cytokine Measurements

Cytokines are small proteins released by a variety of cells, including immune cells such as T cells, B cells, mast cells and macrophages, as well as other cell types e.g. endothelial cells, fibroblasts and stromal cells. They are typically generated in response to an activation stimuli, e.g. growth factors, trauma, infection, etc.[161,162] They act as important signalling molecules, involved in a number of biological processes including cell differentiation, growth, activation, recruitment and cell death. Cytokines are critical in mounting an effective immune response towards the pathogenic agent by activating/deactivating different types of immune and inflammatory cells to fight off infection and resolve disease. Unlike other general clinical measures such as plasma immunoglobulin (Ig) levels, total white blood cell count or haematocrit, changes observed in local and/or systemic levels of certain cytokines can indicate the nature of inflammatory and/or immune responses present within the host, including signs of potentially pathological immune dysregulations.[162] As a result, cytokine measurements provide useful information that are used clinically to help assess disease outcomes, in addition to giving more insight into pathology in research. In the presented work, cytokines were measured in the plasma of both the murine models and in HAT patients, as measures of immune response towards parasitic infection.

Historically, cytokines have been detected based on their effect on a biological system, i.e. typically tested by measuring the induced (dose-depedent) proliferation, degranulation or chemotaxis of a range of responsive carcinoma cell lines following an incubation period, collectively termed as bioassays.[163,164] These assays have been shown to be highly sensitive and informative about the cytokine's functionality. However, many of them lack specificity as many cytokine actions overlap: KIT-225 for instance, a human chronic T-lymphocyte leukaemia cell line, used to test for IL-12 activity, will also proliferate in the presence of IL-2, IL-4, IL-7, IL-15, IL-21, and IL-23. Furthermore, different cytokines can produce both antagonistic or synergistic functional effects, causing difficulties in results interpretation, thus making bioassays unsuitable for the identification of specific cytokines within complex mixtures, such as plasma.[164,165]

With the advances in molecular biology, assays that make use of recombinant antibodies specifically targeting certain cytokines have become the norm, whereas the Enzyme-Linked Immune-

Sorbent Assay (ELISA), is by far the most commonly used and best validated method.[165,166] Where the target molecule (or antigen, Ag) of the antibody (Ab) is in solution — as in the case for measurement of cytokines in plasma — the "sandwich" or indirect format of ELISA is used. The sandwich ELISA concept is widely commercially available as ELISA kits and can be purchased for the majority of known cytokines (especially for human and mouse versions). ELISA kits are commonly based on plate-bound ("solid-phase") format. In this type of assay, the Ag in solution (i.e. the cytokine) is captured by a primary Ab, that has been immobilised on the plate surface. The capturing Ab is typically incubated with a blocking agent (e.g. a solution of bovine serum albumin, BSA) before or at the same time as the Ag is added, to minimise non-specific binding of the capturing antibody. Assay specificity is further increased through plate washes after each incubation step. A secondary Ab is then added, which binds to the Ag at a different site to the first capturing Ag. This secondary Ab is conjugated to an enzyme (e.g. streptavidin-horseradish peroxidase) that, in the presence of a suitable substrate (e.g. tetramethylbenzidine/TMB), will react in a dose-dependent manner to produce a visible signal. This is typically a colour or fluorescence that can be detected by a spectrometer when a certain wavelength is applied.[166] A titration of Ag standard with known concentration is added to the plate, including blank wells (where no Ag was added), to extrapolate Ag concentrations within the remaining tested samples and to determine the background signal, respectively. A schematic summarising the different layers of Ag/Ab in the sandwich ELISA is shown in Figure 2.6.

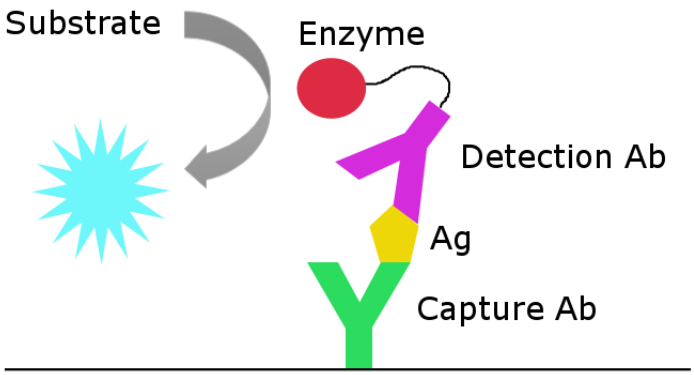


Figure 2.6: Solid-Phase "Sandwich" ELISA

Capture Ab is immobilised on the plate surface, which then binds the Ag in solution. A secondary detection Ab (which binds to a different site on the Ag from the first capturing antibody) is then added. The secondary Ab is conjugated to an enzyme, which in the presence of a suitable substrate, will react with it to produce a visible signal that can be measured by a spectrophotometer. Abbreviations: Ab, antibody; Ag, antigen

Mouse Cytokine Assessment via MSD

Evidence has demonstrated that a wide range of different cytokines were involved in the pathogenesis of leishmaniasis, in particular Th 1 and Th 2 linked cytokines.[31] Consequently, the concentration of selected Th 1, Th 2 and regulatory cytokines were tested in mouse plasma obtained from the murine *L. major* infection models. As sample volume was limited, cytokine measurements were assayed using a multiplex MesoScale Delivery (MSD, USA) assay kit.[167] MSD assays are similar to sandwich ELISA's, yet the kit uses special plates that contain separate spots coated with different capturing Abs within the same well and are thus able to measure multiple cytokines simultaneously. In addition, rather than conventional polystyrene plates typically used in ELISA's, the kit uses carbon-electrode plate surface for increased binding capacity. Finally, the secondary Abs are conjugated to SULFO-TAG^{TM}(instead of an enzyme), which emit light signals following electrochemical stimulation by the MSD plate reader, upon acquisition.[167] Manufacturer's claim this technology results not only in higher sensitivity but also superior dynamic range of analyte concentrations, when compared with conventional ELISA's.

Following NMR acquisition, plasma levels of IL-1β, IL-2, IL-4, IL-5, CXCL1 (Chemokine [C-X-C motif] ligand 1), IL-10, IL-12, and TNFα were tested in duplicate using the ultra-sensitive mouse Th 1/Th 2 Cytokine 9-plex assay kit, based on manufacturer's instructions with minor modifications, as outlined below.

Plasma samples were transferred from NMR tubes into 96-well plates and diluted 1:1 with 60 µl phosphate buffered saline (PBS, SIGMA-Aldrich). A volume of 25 µl per well of blocking buffer (MSD diluent 4) was added to all wells of the MSD assay plates, which were sealed with adhesive plate sealers and incubated for 30 mins at room temperature (RT), on a plate shaker (1000 rpm). During this time, 4-fold serial dilutions of calibrator (containing multiple cytokine standards) was set up from 10,000 pg/ml to 2.4 pg/ml) in blocking buffer, generating an 8-point titration curve (with the last well left blank with no calibrator). A volume of 25 µl of (diluted) plasma samples or prepared calibrator titrations were added to the wells, bringing the total well volume to 50 µl. Plates were sealed, returned to the plate shaker, and left to incubate at RT at 1000 rpm, for another 3 hours.[6] Plates were subsequently washed three times in PBS with 0.1% Tween-20 (SIGMA-Aldrich) solution and 25 µl per well of detection Ab mix (at 1X) was added to all wells. Plates were sealed and incubated for a further 2 hours (RT, 1000 rpm) and washed 3 times. Finally, 150 µl per well of Read Buffer T (at 2X) was added to the plates. MSD plates were then read on Sector Imager 2400 (MSD), using accompanying Discovery Workbench MSD software

[6] An extra hour was added to the incubation period of original MSD kit instructions, for increased detection of potentially dilute samples (as recommended by an MSD technical advisor).

to analyse and extrapolate cytokine concentrations.

Human Plasma and CSF Cytokine Assessment via ELISA

Previous research performed by our collaborator, Dr Jeremy Sternberg *et al.* (Aberdeen University) have shown that HAT has been linked with increases in pro-inflammatory cytokines TNFα, IFNγ, and IL-6, as well as the regulatory cytokines IL-10 and TGFβ, observed in both plasma[79] and CSF.[168] As a result, these particular cytokines were targeted and measured in CSF and plasma samples, performed by Sternberg's group. Concentrations had been tested using separate OptiEIA ELISA kits (BD Pharmingen, Cowley, UK) for each cytokine (except TNFα which was tested using BioLISA system, Bender MedSystems, Austria). Kits included cytokine standards, 96-well plates pre-coated with capture Abs, enzyme-conjugated secondary Abs, TMB substrate and wash diluent. All samples were tested in triplicate, following manufacturer's instructions. Out of the tested cytokines, only IFNγ, IL-6, IL-10 and TNFα were measured above minimum levels of detection in plasma, and only IFNγ, IL-6, and IL-10 were detected in CSF, which were used for final analyses.

2.6.2 Faecal Microbiota Diversity Analysis

An increasing amount of recent literature on gastrointestinal microbiota highlight the importance of their role in systemic physiological and pathological processes by regulating nutrient availability as well as influencing host inflammatory and immune responses.[169–171] They are essential for the extraction and generation of short-chain fatty acids and several vitamins (e.g K and B12), help maintain a defence barrier against pathogenic bacteria and aid towards intestinal maturation and development.[170] Dysregulation of the normal intestinal microbiome has been linked to a range of disorders, including increased susceptibility to pathogenic infections of the gut, gastric ulcers, diabetes and obesity.[171] Within the field of parasitic infections, several studies have shown an association between infection and an altered host intestinal microbiome. Changes in both composition of microbial species[172], as well as in the expression of microbiota-associated metabolites, have been observed in plasma and urine.[91–94]

Due to the ease of collection, microbial diversity analyses were performed on faecal samples, which were used as surrogates for gastrointestinal microbiota. No faecal material had been collected from HAT patients (or their corresponding controls) at the time of the study, therefore only the samples from the murine *L. major* infection (and associated control) studies were available for analyses. Faecal DNA extraction, followed by PCR amplification to measure microbial DNA quality for the samples, were both performed under the supervision of Dr Julian Marchesi (at Cardiff

University). The final analytical step, next-generating pyrosequencing of microbial DNA samples, was performed externally by Research and Testing Laboratory (Texas, USA).[173–175]

Faecal DNA Extraction Optimisation

DNA extraction was initially performed on practice samples (extra faecal material from control BALB/c and C57BL/6 mice), until methods were optimised for final extraction on main study samples. All DNA extractions were performed using the QIAamp® DNA Stool Mini Kit (QIAGEN, Germany), largely based on manufacturer's instructions.[176] The faecal samples that had been allocated for microbial analyses had been previously stored at -40°C in 10% formalin (a fixative) solution (in 5 ml tubes). Whilst such storage would have been optimal for microscopic microbial assessment, it was later determined as suboptimal for DNA extraction methods. It appeared that the formalin that had bound to the DNA was difficult to remove and introduced unwanted breaks in the DNA during the extraction and amplification processes, resulting in low DNA yields with high impurities. Several rounds of optimisation were required to try and remove the fixative and break down the faecal debris to isolate the DNA. The final protocols were partially based on methods described by Paglia and Visca (2004), amended from the original QIAamp instructions, specifically for formalin-fixed samples.[177] The modifications included adding wash steps to remove the formalin, increasing the time and temperature of the lysis and enzyme digestion incubation steps, and increasing amounts of digestive enzyme (proteinase K) for maximal DNA recovery. Additional changes involved the addition of a zirconia silica bead homogenisation step (similar as was done for metabolite extraction) to enhance bacterial cell lysis by the ASL buffer, and halving the elution volume in the final step, to further increase DNA sample concentrations.

Optimised DNA Extraction Protocol for Formalin-fixed Faeces using QIAamp Stool Mini Kit

Pipette tips, Eppendorf tubes, and PBS were sterilised (autoclaved) prior to extraction. All centrifugation steps were carried out at RT at 20,000 x g (14,000 rpm) in Eppendorf Microfuge 5417. Formalin-fixed faecal samples were left to defrost at RT. During this time, 2 ml skirted bead-beater tubes (with twisting lids, VWR) were filled with $\sim$ 0.5 g of 0.1 mm zirconia silica beads (Biospec). Once defrosted, $\sim$ 200 mg (3–8 pellets) of faeces (and surrounding formalin) were transferred into bead-beater tubes using a 1 ml pipette (with tip of the pipette tip cut off to facilitate suction of the pellet). Tubes were topped up to 1.5 ml with sterile PBS and centrifuged for 5 mins. Supernatant was subsequently pipetted off and discarded. Samples were then washed twice with 1.5 ml of sterile

PBS, and each sample was vortexed for $\sim$ 15 seconds and spun down for 5 mins. PBS supernatant was pipetted off and discarded, with care to not disturb the sediment of faeces and zirconia beads in the tube.

DNA extraction was performed using QIAamp DNA Stool Mini Kit (QIAGEN, Germany). A volume of 1.4 ml of ASL (lysis) buffer was added to each tube and samples were homogenised using a Precyllys® 24 Bead-beater (Bertin Technologies, France) for 2 x 45 seconds at 6000 rpm. Samples were then left to incubate at 90°C for 20 mins, vortexed for $\sim$ 15 seconds and then centrifuged at 20,000 x g (14,000 rpm) for 5 mins at RT. A volume of 1.2 ml of supernatant of each sample was pipetted and transferred into fresh 2 ml Eppendorf tubes, whilst the remaining debris was discarded. An InhibitEX® tablet (which quickly adsorbs compounds within the sample that can degrade DNA or interfere with subsequent extraction steps) was added to each Eppendorf and samples were immediately vortexed for 1 min. Samples were subsequently centrifuged for 3 mins and supernatants were transferred into fresh 1.5 ml Eppendorfs, whereas the tubes containing tablet debris were discarded. These were spun down again for 3 mins and 200 µl supernatant of each sample were transferred into new 1.5 ml Eppendorf tubes, each previously containing 3 µl of proteinase K. Another 200 µl of AL buffer was added to the tubes and samples were vortexed for $\sim$ 15 sec. Samples were then incubated at 70°C for 20 mins , to allow for enzymatic digestion of faecal proteins. After incubation, 200 µl of 99% pure Ethanol (VWR) was added to tubes and samples were briefly vortexed.

Sample lysate–ethanol mix were transferred into QIAamp spin columns (already placed in QIAamp 2 ml collection tubes), with care to avoid sample touching the rims of the columns. Columns were centrifuged for 1 min and then transferred into fresh collection tubes (first collection tubes containing the filtrate after centrifugation were discarded). A volume of 500 µl of AW1 buffer was added to all spin columns, which were then capped with lids and spun down again for 1 min. Collection tube with AW1 filtrate were discarded and columns were placed again in fresh tubes. AW2 buffer was added at 500 µl to each columns and samples were centrifuged for 3 mins. The filtrate was poured out of collection tubes and discarded, and columns were placed back in the same collection tubes. Columns were then re-spun for a further minute to ensure maximal removal of AW2 buffer from columns. Spin columns were subsequently transferred into 1.5 ml Eppendorf tubes and 50 µl of elution buffer AE was carefully added to the middle of the column (taking care not to touch the membrane with the pipette tip and avoiding column rim). Column caps were closed and samples left to incubate for 1 min at RT. Spin columns were then spun down for 1 min to elute the DNA into the Eppendorfs. Caps were opened and a further 50 µl of buffer AE was added to all

columns. Samples were incubated for another minute and then centrifuged again for 1 min. Spin columns were then discarded, and the final faecal DNA samples (100 µl eluates in the Eppendorf tubes) were subsequently stored at -80°C until shipment for next-generation sequencing.[7]

DNA Quantification via NanoDrop

Nucleic acids (RNA and DNA) absorb ultraviolet (UV) light (maximally at 260 nm) that can be detected and quantified by measuring their optical density during UV light exposure. Faecal DNA samples (either freshly eluted or briefly defrosted at RT) were quantified using NanoDrop technology, a UV spectrophotometer that combines fibre optic technology and surface tension properties.[178] Samples were read on the NanoDrop ND-100 (Thermoscientific, USA) at 260 nm with accompanying software, following manufacturer's instructions.[178] Firstly, 1.6 µl of nuclease-free water (VWR) was pipetted (using filter pipette tips) onto the spectrophotometer's optical surfaces and run twice on the instrument, once to rinse the machine and then to set the blank. This was followed by testing 1.6 µl of each DNA sample to measure their concentration and assess purity, wiping the machine sample holders with a clean tissue in between each measurement.

In addition to quantification of DNA concentration, NanoDrop also allows the assessment of sample purity as contaminants usually also absorb UV light, yet at different maximal wavelengths (e.g. protein absorbs maximally at 280 nm whilst other organic compounds such as phenol have their peak at $\sim$ 230 nm). The NanoDrop software automatically calculates 260:280 and/or 260:230 ratio as a measurement of sample purity. These values were not recorded for our analyses; however, their results were useful in aiding DNA extraction optimisation (preliminary rounds of DNA extractions yielded samples with high impurities, most likely due to the presence of formalin).

PCR for Confirmation of Microbial DNA

Microbial DNA presence was confirmed using polymerase chain reaction (PCR), an *in vitro* molecular biology technique used to amplify specific sections of DNA enzymatically — in this case, part of the prokaryotic 16S ribosomal gene, containing regions known to be highly conserved across bacteria. Typical PCR methods are based on a repeating number of cycles, comprised of a denaturing step, and annealing step, and an extension/elongation step, as shown below in Figure 2.7.

Firstly, double-stranded DNA becomes denatured into single strands at high temperatures (94–98°C), each being used as a template for enzymatic extension. This is followed by the annealing

[7] Samples were also briefly defrosted for NanoDrop quantification.

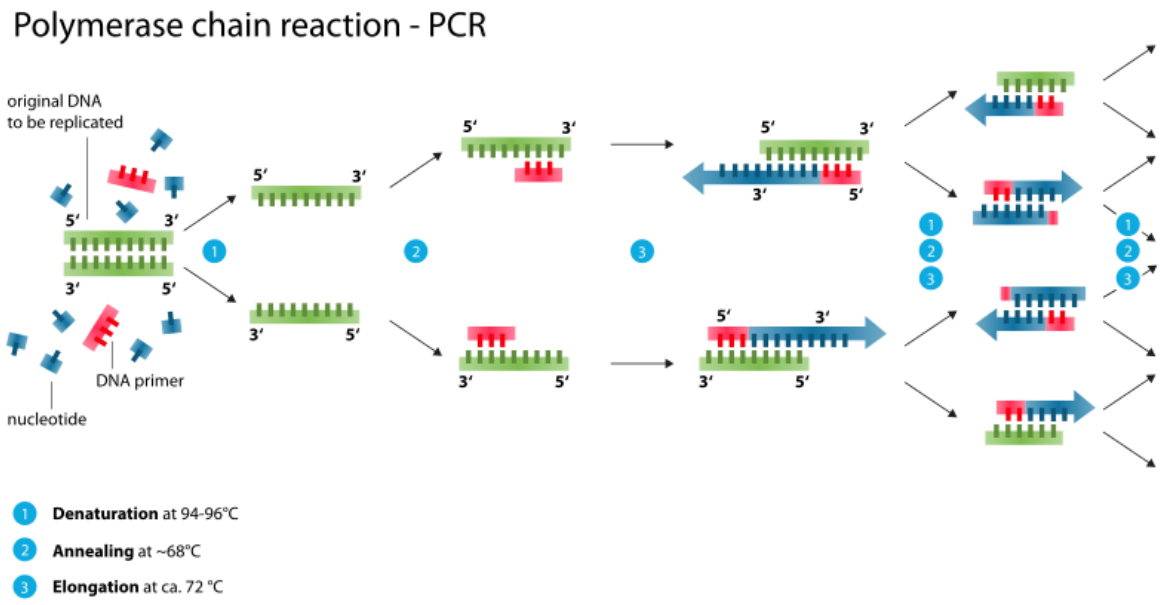


Figure 2.7: Schematic of Polymerase Chain Reaction Cycles

By Enzoklop (own work), via Wikimedia Commons, Wikepedia (2014).[179]
Figure reproduced under Creative Commons Attribution-Share Alike 3.0 Unported license.

of short oligonucleotide sequences, known as primers, to the template strands (at 32–68°C). These primers are specifically chosen to be complementary to the beginning (5' end, using "forward" primers) and terminating (3' end, using "reverse" primers) sequences of the targeted section of the DNA template. Lastly, heat-stable DNA polymerases (commonly *Taq* polymerase, isolated from the thermophilic bacterium *Thermus aquaticus*) bind to the primers and the template and incorporate nucelotides, to generate new DNA strands (extending on from the primers) that are complementary to the original two templates.[180] This set of steps is repeated over a number of cycles, resulting in rapid exponential amplification of targeted DNA sequences, until a maximal (plateau) concentration is reached, typically after ∼30 cycles.[180]

Similar to the DNA extraction procedures, multiple rounds of optimisation were required until bacterial DNA was detected in the mouse faecal DNA samples, whereas PCR protocols were initially based as described by Lewis *et al* (2013).[181] This was partly due to the DNA concentration of some of the samples being very low (< 5 ng/µl). Protocol modifications included changing the DNA sample volumes, the temperature and duration settings for the annealing and extension steps, and the final PCR cycle number. Preliminary rounds of PCR used the combination of primers 27f (5'-AGAGTTTGATCCTGGCTCAG-3') and 1492r (3'-GGTTACCTTGTTACGACTT-5'), amplifying ∼ 1500 kilo base pairs (kb) in regions V1–V9. As 454 next-generation pyrosequencing only requires a minimum of ∼ 500 kb (and results with the first primer set were sub-optimal) the reverse primer was subsequently changed to 534r (3'-ATTACCGCGGCTGCTGG-5'), targeting against V1–V3

regions of bacterial 16S rRNA genes.

All DNA samples were prepared for PCR on ice (typically 20 samples at a time), using pipette tips with filters for all pipetting steps, to avoid sample contamination/cross-over. PCR reaction for each sample was run in 20 µl, whereby a PCR master mix was initially prepared (enough was made for all reactions + 1 spare), containing all the reagents except the sample. The master mix was then pipetted into individual 0.2 ml polypropylene PCR tubes (generic), followed by the final addition of the sample. For every reaction, final reagent volumes were as follows (shown in order of PCR mix preparation):

1. 15.04 µl Nuclease free water (VWR)
2. 2.00 µl 10x reaction buffer (NewEngland Biolabs, NEB)
3. 0.20 µl of 100x Bovine Serum Albumen (NEB)
4. 0.16 µl of Deoxynucleotide mix (SIGMA-Aldrich)
5. 0.20 µl of fwd primer 27f (SIGMA-Aldrich)
6. 0.20 µl of reverse 534r (SIGMA-Aldrich)
7. 0.20 µl Taq polymerase (SIGMA Aldrich)
8. 2.00 µl of sample, containing $\sim$ 20–30 ng DNA [8]

Faecal microbial DNA extracted previously in-house (courtesy of James Evans, PhD student of Julian Marchesi) was used as a positive control DNA sample, whilst nuclease free water was tested as negative control. All samples were then added to a PCR machine (Thermocycler) for 35 cycles of PCR: [30 seconds at 95°C , 40 seconds at 55°C, 60 seconds at 72°C].

The integrity of the PCR products were assessed qualitatively *via* electrophoresis, as follows: firstly, 1.5 g of Agarose Powder (Severn Biotech Ltd, UK) was added in 75 ml of Tris/Borate/EDTA (TBE) buffer (Severn Biotech Ltd), in a 100 ml glass bottle and microwaved for 1–2 mins at 700 W. The resulting 2% Agarose gel solution was then allowed to cool down (until the bottle could be touched by hand). SafeView (NBS Biologicals Ltd., UK), a nucleic dye, was added into the cooled (yet still liquid) gel at 3.75 µl per sample, and mixed in by swirling around the bottle. Gel was slowly poured into a gel holder (15 x 6 cm) and 24-comb was added (to create loading wells for the DNA samples), making sure all air bubbles were removed. Gel was allowed to set at room temperature for 20–30 mins. Afterwards, the comb was taken out with care and the gel (in the gel holder) was added to electrophoresis bath, covered in TBE buffer. A volume of 10 µl of each sample (PCR product) was mixed with 3 µl gel loading dye (NEB) and ran for $\sim$30 mins at 100 Volts.

[8] Concentrated samples were diluted prior to addition to the PCR mix. For samples with concentrations < 9 ng/µl, 4 µl was added to the PCR mix instead (and the nuclease free water was reduced to 13.04 µl accordingly).

Gel was subsequently taken out of electrophoresis tank and visualised under UV light, whereby successfully amplified microbial DNA were shown as bands on the gel.

Next-generation Pyrosequencing of DNA Samples for Microbial Assessment

For maximal screening of diverse microbial populations, samples were analysed externally by next-generation pyrosequencing, using a 454 GS-FLX+ Genome Sequencer (Roche, Switzerland). Traditional culture-based methods of microbial identification are time-consuming, provide limited phenotypic information, and to date have only been able to culture approximately 1% of gastrointestinal bacteria.[174] Sequencing is an alternative and increasingly more wide-spread tool that is a much faster, can detect species without the need of culture, and can reveal more in-depth genomic information. Pyrosequencing is a sequencing-by-synthesis technique based on the detection of pyrophosphate release on nucleotide incorporation rather than chain termination with dideoxynucleotides as used in conventional "Sanger" sequencing.[182] Furthermore, the 454 GS-FLX+ Genome Sequencer represents a powerful modern molecular technology that is ultra high-throughput, able to cope with long reads with typically over 99.5% accuracy.[183]

Samples were shipped on dry ice to Research and Testing Laboratory (Texas, USA) for microbial screening, whereby data were processed and analysed according to their company guidelines.[184] Acquisition was based on initial PCR amplification of samples followed by sequencing of a section of the 16S ribosomal RNA genes, comparing generated data with their validated microbial sequencing database. Data processing steps included quality scanning, de-noising of data, and removal of poor reads and chimeric sequences (chimera), which are generated when an aborted sequence extension is misidentified as a primer and extended incorrectly in subsequent PCR cycles.[175,184] These steps were followed by sequence clustering, automatic allocation of each sequence to matching operational taxonomic units (OTU's, i.e. species classification/identification based on DNA sequence) and taxonomic assignment. Output from Research and Testing consisted of raw data files (in SFF and CSV format — the latter subsequently converted to Excel xls format and used for final analyses) which included all measured OTU counts and their relative percentage abundance, for each of the following taxonomic levels (in decreasing order): phylum, class, order, family, genus and species. The final preprocessed OTU sequence clusters assigned taxonomic levels as follows: >97% identity (<3% divergence) were resolved at the species level, 95–97% at the genus level 90–95% at the family level, 85–90% at the order level, 80–85% at the class, and 77–80% at phyla. Reads with less than <77% matching identify were termed unclassified OTU's.

2.6.3 Univariate Statistical Analyses of Cytokine and Microbial Data

Differences in cytokine concentrations as well as relative percentage abundance of faecal bacteria were both analysed using non-parametric statistical tests, which do not assume normal distribution nor equal group variances (required for parametric tests, e.g. T-tests and Analysis of Variance/ANOVA). In order to be able to distinguish between functionally distinctive bacterial groups, rather than merely genotypic differences when assessing changes at lower taxonomic levels, microbial abundances measured from murine *L. major* models were analysed at the class level (though trends were similar for other taxonomy levels). Infected and control groups were initially compared using Kruskal–Wallis ANOVA tests.[185] This was followed by testing pair-wise comparisons between infected and control groups *via* Mann-Whitney U-tests[186] (a non-parametric equivalent of Student T-test). Changes between time-points within groups were tested initially *via* Friedman Tests[187] (comparing all time-points at once, similar to a repeated measures ANOVA), followed by and Wilcoxon matched-pair tests[188] (for pair-wise comparisons, e.g. between the pre-infection time-point and various post-infection time-points).

2.6.4 Multiple Test Corrections

In order to account for the rising number false positives (also known as type I error) that arise when performing multiple statistical comparisons (e.g. the probability of seeing at least one significant result just due to chance, rejecting the null hypothesis), different forms of multiple test corrections were applied, depending on the dataset. For *L. major* mouse models, Bonferroni correction was used.[189] This method is known as a familywise error rate procedure and is a simple and commonly used form of multiple test correction. To apply this Bonferroni correction, the measured p-value was multiplied by the number of comparisons performed. Results were considered significant if the corrected p-value still remained below a critical value (α) — in this case, below 0.05. Thus the new critical p-value of significance is essentially equal to 0.05 divided by number of tests, which was used as the statistical criteria for cytokine and microbial univariate analyses used in Chapters 3, 4 and our recent publication of murine *L. major* infection model data (see Publications section A in the Appendix).

Yet, Bonferroni correction represents a very stringent method that in the process of minimising type I errors also concurrently augments the probability of not detecting true "positives" (type II errors).[190] As a result, this form of critical value adjustment can sometimes prove too conservative for showing biological significant differences between populations, especially if these are comprised of small sample sizes. An alternative and increasingly popular approach is to base the critical value

of significance not on reducing the probability of one false discovery but instead on controlling the number of expected false positives, termed the false discovery rate (FDR) by Y. Benjamini and Y. Hochberg (1995).[191] The Benjamini–Hochberg FDR method represents one of the most widely FDR method used in biological research, and was used for clinical HAT analyses. Lastly, another FDR method that was used was the Storey–Tibshirani FDR, a modification of the Benjamini–Hochberg FDR which tries to estimate what the proportion of truly null p-values is within the analysed population and scales the data accordingly.[192] Use of these two FDR methods is further outlined in the relevant results chapters.

2.7 Correlation Analyses between Datasets

Correlation analyses was performed between different data sets (e.g. between metabolic data and plasma cytokine concentrations and/or faecal microbial levels), in order to determine potential associations between them in the host, and assess these links over the course of infection. Pearson correlation (i.e. Pearson Product Moment Correlation) is the most common form of correlation analyses, which determines the strength of a linear relationship ("dependence") between two sets of data, comprised of continuous variables.[193] In other words, a change in one variable is associated with a proportional change in the other variable. The resulting Pearson correlation coefficient (r) is a value that ranges between -1 (negative correlation) and 1 (positive correlation) — a result of zero means that no (linear) correlation was observed in these analyses.

All correlation analyses were performed in MATLAB using either the standard toolbox scripts (e.g. "corr" function) which generated simple heatmaps and corresponding tables of correlation coefficients and p-values, or customised in-house scripts. A specialised MATLAB scripts for two-dimensional (2D) matrix correlation, based on STOCSY[156] and statistical heterospectroscopy (SHY)[194], was used to explore inherent covariance between signal intensities of multispectroscopic data in molecules measured by different techniques across cohorts of samples. The script was further adapted by Dr Judith Fonville, specifically for correlations between a matrix of numerical values (in this case, cytokine concentrations and/or microbial % abundance) and a matrix containing NMR spectral data.[195] The script embeds a peak picking step and a minimal intensity threshold criteria. The defined criterion was to surpass 100 x standard deviation of the first 500 spectral data points (which do not have any spectral peaks in them but consist of background baseline signals), to minimise artefacts and noise. The script also included a final 10,000-fold permutation step, whereby only correlative pairs with $p < 0.05$ were displayed, used for co-analysing metabolic data with cytokine data, and metabolic data and microbial data. Separate Pearson-based 2D matrix

correlations between time-matched cytokine concentrations and microbiota levels were analysed using standard MATLAB statistical toolbox, followed a 10,000-fold random permutation and p-values calculated similarly as for spectral correlation analyses (developed in-house by Dr Kirill Veselkov). 2D matrix correlation analyses was used to compare metabolic and cytokine (examples of these are shown in Results Chapter 4).

Correlation networks analyses were developed in-house by fellow PhD student Joram Posma, based on MetaboNetworks software.[196,197] In these analyses, pair-wise Pearson correlation was performed between variables (i.e. metabolite integrals from different biofluids, cytokine concentrations and faecal bacterial class abundance) across each sample, for each time-point, and each mouse model group — adjusted for multiple comparison using Storey–Tibshirani FDR.[192] In these analyses, correlation coefficients and p-values between variables are depicted in two formats. One, as 'wheels', where each variable is constrained in a fixed position and significant positive and negative correlations were shown as red or blue connecting lines (spokes), respectively. Secondly, correlation networks were also shown as 'maps', whereby variables that are highly correlated to each other are positioned near to each other in clusters. Whilst both formats are based on the identical correlations, the wheels allow direct comparisons of each variable between different groups/time-points, whilst the network maps facilitate the interpretation of major hubs and chains of correlation, and thus are more aligned with mapping with biological pathways.

Chapter 3

Results: Systemic Characterisation of Murine Strains

3.1 Background 77
3.2 Objectives 79
3.3 Experimental Methods 79
3.4 Results 84
3.5 Discussion 103

3.1 Background

Animal models, in particular mouse models, have played and continue to play a vital role in medical advances by providing useful tools to improve our understanding about health, disease mechanisms, and for the prediction and assessment of pharmacological interventions.[198] Despite many success stories, there are limitations of information animal models can provide to human studies, with only a small proportion of significant findings in animals finally translating into clinical research.[199]

Of all the murine inbred strains, the C57BL/6 mouse (also known as "Black 6") represents the most widely used strain in laboratory research[200] and was the first strain to have its genome sequenced.[201] It is used for a whole range of models, including drug safety and efficacy testing, diet-induced obesity, type II diabetes, and provides the genetic background for a multitude of transgenic and knock-out models.[200,202] Another commonly used mouse strain is BALB/c. These mice are well known for their use in the production of monoclonal antibodies and for their increased susceptibility to a range of infectious agents including *Listeria, Leishmania* and *Trypanosoma.*[203]

In this PhD project, BALB/c mice were used as a model for experimental cutaneous leishmaniasis, to characterise the metabolic phenotype attributed to *Leishmania* infection. Having identified metabolic changes relating to *L. major* infection, we further characterised and compared metabolic, immune and faecal microbial changes associated with *Leishmania* infection in the BALB/c mice and C57BL/6 mice, which have previously been established as non-healing and self-healing phenotype, respectively, further discussed in Chapter 4.

Identical experimental procedures, such as drug or diet interventions, gene manipulations, and environmental factors in the same animal species can produce vastly differing phenotypes, based primarily on the genetic background, highlighting strain and sub-strain differences as major contributors towards experimental outcome.[204] Characterisation of the reference baseline is fundamental for optimal data extraction and interpretation when comparing between different experimental groups and/or animal models. Thus, in order to be able to discriminate between the differential responses towards *L. major* infection, and the inherent genetically-linked phenotypic differences that occur between BALB/c and C57BL/6 strains, a comprehensive metabolic characterisation of each uninfected strain was performed. Such investigations were also required to identify other confounding factors that can introduce biological variation within the analysed data, e.g. any age/time-related changes or differences due to housing/caging conditions, diurnal rhythmic effects, etc.[204]

A characterisation and comparison of mouse physiology, plasma biochemistry, haematology, and glucose tolerance between common mouse strains, including C57BL/6 and BALB/c, have been well described by Champy *et al.* (2008)[205], who tested both male and female animals at 3, 6 and 12 months of age. The authors found that most strain differences were independent of gender, whereby significant physiological findings included differences in food intake, body weight, fat content, bone mineral density, with BALB/c being significantly higher than C57BL/6 mice in most cases. Further plasma analyses revealed significantly higher levels of cholesterol, triglycerides, and iron in BALB/c mice, whilst C57BL/6 had higher platelet counts, similar to major animal suppliers' haematological and biochemistry data of these two strains.[206–209] Furthermore, C57BL/6 mice were shown to be one of the least tolerant of the tested strains in oral glucose tolerance tests, sustaining higher levels of glucose in their blood after dosing for longer periods of time than the other strains.

Further efforts have been made to assess metabolic differences related to rodent strain and confounding factors such as gender, age and batch, using NMR and/or MS-based techniques.[204,210–212] Others have expanded to incorporate both metabolic characterisation and other datasets, such as microbiome, genomics and/or proteomic data, though these have been limited to the analysis of

one strain at a time.[213,214] For laboratory mice, a standard set of publicly available data now encompasses physiological, biochemical and haematological, phenotypic and functional genetics on a whole range of sub-strains, as well as transgenic and knock-out (KO) mouse models, provided in part by suppliers (Mouse Genome informatics: `http://www.informatics.jax.org`) or research institutions on a global scale (Europhenome Mouse Phenotyping Resource: `http://www.europhenome.org`). However, a comprehensive multi-level metabolic comparison between different mouse strains, has not been conducted so far.

In this PhD project, metabolic profiling of plasma, urine, faeces and multiple tissue sections, taken from C57BL/6 and BALB/c mice, were performed. Furthermore, plasma cytokines and faecal microbial composition were measured, to provide additional indices of general immune status and gut health for each strain. The overall study design is shown in Figure 3.1. Results from the uninfected animals are presented in this chapter — see Chapter 4 for the analyses on *L. major* infected mice.

3.2 Objectives

The aim of the murine characterisation was to provide a comprehensive overview of the biological background of C57BL/6 and BALB/c mice and identify the inherent phenotypic differences that separate these two strains. The following specific objectives applied:

- To characterise systemic and local metabolic profiles of two mouse strains by establishing metabolic profiles of their biofluids and tissues *via* ^{1}H NMR spectroscopy
- To assess the immune status of both strains based on their plasma cytokine measurement *via* multiplex MSD assays
- To determine the faecal microbial diversity of each strain
- To investigate the metabolic, cytokine and microbial differences between strains through multivariate and/or univariate statistical analyses

3.3 Experimental Methods

3.3.1 Animal Study Design and Sample Collection

All animal handling and sample collection were performed by Dr Jasmina Saric at Imperial College London, adhering to local and national animal handling guidelines, under project license PPL70/6997 — license held by study collaborator, Dr Karin Seifert, London School of Hygiene

& Tropical Medicine (LSHTM). The protocols were approved by the animal ethics committees at LSHTM, Imperial College London and the UK Home Office.

Samples were collected from two studies, as summarised in Figure 3.1. The first study represented the experimental cutaneous leishmaniasis study, in which *L. major* infection was compared in two mouse strains. Overall, this study was composed of four groups in total: infected and control BALB/c mice, and infected and control C57BL/6 mice. Samples from this study shall be referred to as 'batch 1' samples, which were used for metabolic, cytokine and faecal microbial analyses. The second study was performed as a replicate of the first study to increase control group sample sizes (from $n = 5$ to $n = 10$ for each strain) for metabolic analyses only. Experimental conditions were as described for the first study, though using a different set of animals (batch 2 samples) .

In both studies, 7–8 week old female C57BL/6 mice and BALB/c were purchased from Charles River Ltd, UK. Mice were housed in macrolon cages (5 animals of the same strain per cage) under environmentally-controlled conditions (temperature: $\sim$ 25°C, relative humidity: 60–70%; light–dark cycle: 12–12 hours). Mice had free access to local tap water and standard rodent diet (Rat and Mouse No. 1 Maintenance diet, Special Diets Service, UK). Following one week of acclimatisation, animals were randomly allocated into two control groups (labelled BC and CC for BALB/c and C57BL/6 mice, respectively) and infection groups (BT and CT). On 'day 0' of the experiment, mice within the infection groups received 100 µl *L. major* parasites, suspended in Schneider's Insect Medium (Sigma-Aldrich) without added Faecal Calf Serum (FCS), into the shaved rump of their tail, using a disposable hypodermic needle (25G X 5/8, Terumo®, Japan). For control groups, mice were injected with 100 µl of blank Schneider's medium instead.

Biofluids (blood, urine, and faecal samples) were taken from each mouse 2 days prior to injection (day -2) and post-injection at days 2, 6, and at the terminal day of the study, which was designated to be day 12 for control animals[1]. Samples collected on day -2 from mice in the infected group shall from this point onwards be referred to as 'pre-infection' samples.

Urine collection was performed at similar times each day (between 8–10 am), to avoid diurnal variations in urinary metabolites. Individual mice were temporarily placed into individual plastic cages (that had been previously been rinsed thoroughly with distilled water) until they released a few drops of urine (typically $>$ 80µl) — urine was subsequently transferred into Eppendorf tubes, snap frozen in liquid nitrogen and stored at -40°C. This was followed by faecal pellets collection in the same cages.[2] A portion of the collected pellets were fixed in 10% formalin solution for

[1] Terminal day for infected animals was defined as the day the lesion had developed in the tail before ulceration occurred (7–13 days post-infection, depending on each animal).

[2] Urine and faecal collection took place in the same cage unless there was cross-contamination between urine and faeces and the samples were rendered useless. In those cases, animals were transferred into clean cage.

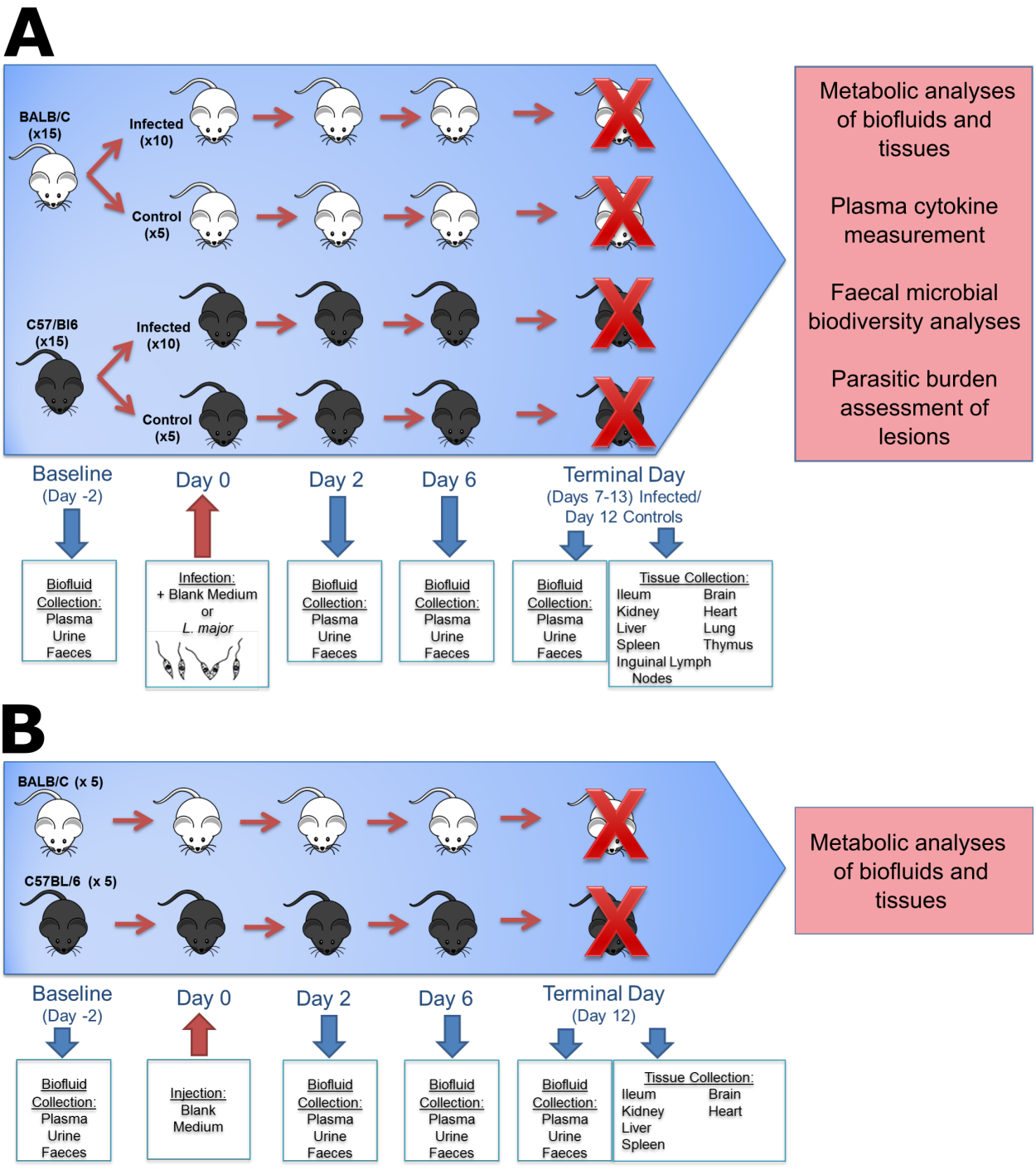


Figure 3.1: Overview of Experimental Murine Studies

Schematic Time-line of the murine experiments and sample collection. (**A**) Main study , comparing experimental *L. major* infection within two strains, namely in BALB/c mice (non-healing disease model) and C57BL/6 mice (self-healing model). Study was comprised of 4 groups in total, i.e. infected (Inf) vs. uninfected controls (Ctrl) for each of the two strains. Biofluids were taken throughout study whilst tissue samples were collected at the end. (**B**) An additional control study was performed using the same experimental set up as for control mice in the first study.

microbiological assessment in Eppendorf tubes, and then stored at -20°C. The remaining pellets were snap frozen in liquid nitrogen, transferred into Eppendorfs and stored at - 40°C, for metabolic analyses.

Approximately 50 µl of blood was collected from each mouse (after urine and faecal collection) following a tail snip directly into a (sodium-heparin coated) haematocrit tube, sealed with a soft wax plug. Tubes were centrifuged for 4 mins at 10,500 x *g* to separate plasma from the erythrocytes. The packed cell volume (PCV, also known as the haematocrit) was determined by measuring the height of packed erythrocytes in the tube relative to the total tube length (expressed as a percentage), as a measure for anaemia (assessing the pathophysiological stage of a host animal). Plasma was then transferred into small Eppendorf tubes, snap frozen in liquid nitrogen and subsequently stored at -40°C.

Animals were also weighed on days -2, 2, 6 for both studies, and additionally on day 12 for study 2, as an additional assessment of their general well-being (particularly important for the animals in the infected group).

On the terminal day of each study, mice were sacrificed by spinal dislocation. Animals were dissected for tissue collection, using surgical blades and tweezers, which were rinsed in distilled water in between each animal. Tissues that were collected in both studies included liver, spleen, ileum, heart and four brain sections — frontal cortex, cerebellum, brain stem, and hippocampus. Thymus and inguinal lymph nodes were additionally taken for animals from the first study. Tissues (or tissue sections) were immediately snap frozen and stored at -40°C for metabolic analysis.

3.3.2 Data Acquisition and Analyses

Sample preparation, acquisition for metabolic profiling, cytokine measurements and faecal microbial measurements were performed as described previously in Methods Chapter 2. The data analysis work-flow was also as described in the Methods Chapter, which included multivariate analyses (PCA and O-PLS-DA) for ^{1}H NMR data, and univariate tests (Kruskal–Wallis tests, Mann–Whitney U-tests, Friedman tests and Wilcoxon matched pair tests) with multiple test correction for cytokine and faecal microbial data (based on bacterial class). Microbial operational taxonomic unit (OTU) data were additionally re-processed by Dr Julian Marchesi (Cardiff University & Imperial College London) using MOTHUR software (`http://www.mothur.org/`), setting the minimum threshold of 300 reads per OTU for dataset inclusion, to confirm validity of data output from Research and Testing Laboratory, LLC. Microbial OTU counts (as opposed to % abundance) data were further analysed *via* PCA using STAMP software (`http://kiwi.cs.dal.ca/Software/STAMP`)[215],

which allowed simultaneous data input and results comparison between multiple taxonomic levels. Differences between strains were analysed *via* White's t-test, a non-parametric test incorporated within STAMP, optimally designed for metagenomic data that is based on count data[216], followed by multiple test correction.

Metabolic differences between the two strains were assessed initially only for batch 1 samples *via* PCA and O-PLS-DA, followed by reassessment with the pooled data of both batches. Differences between the batches were evaluated *via* PCA and O-PLS-DA. In order to account for strong batch variation which was subsequently observed in both biofluids and tissues (further outlined in results section of this chapter), data were adjusted by subtracting the median spectrum for each corresponding batch (combined across both strains) from each spectra. Thus, batch 1 median spectra were subtracted from all spectra from samples in batch 1, and batch 2 median spectra were subtracted from all spectra from batch 2 samples. This method was performed separately for each time-point for the biofluid data. The batch-adjusted data was then used for final analysis, to assess variation due to strain differences, as well as time-related changes, using PCA and O-PLS-DA models. Metabolic strain differences in biofluids were determined using models based on spectral data of all assessed time-points collectively, in order to determine discriminatory metabolic markers that were robust throughout the course of the study.

In addition to identifying the individual discriminatory metabolites/cytokines/faecal bacteria that differentiate between the two strains *via* aforementioned statistical methods, multiblock hierarchical PCA (H-PCA) was performed to determine which assessed biological compartments contributed the most towards strain discrimination, as described by Li *et al.* (2009).[155] PCA models were performed separately using metabolic data for each biofluid and tissue, cytokine concentrations and faecal microbial (family) percentage abundance in SIMCA, each representing different 'blocks' of data. Metabolic data were comprised of calculated integral values of metabolite peaks that were determined as significantly discriminatory between the two strains, based on O-PLS-DA analyses (and for the biofluids, combining all four time-points). The first two scores (t_1 and t_2) generated from each PCA model were then used as loading variables to construct subsequent PCA models (i.e. H-PCA). An advantage of this technique is that it reduces the dimensions of the original larger data matrices into fewer variables, which can facilitate the interpretation of correlations between the scores and loadings[217] — in this case, the visualisation of differences and/or similarities in the metabolic responses of multiple biofluids and tissues alongside cytokine and microbial responses.[3]

[3]Note that H-PCA analysis could only be performed on batch 1 data as only these animals had corresponding metabolic, cytokine and microbial datasets.

3.4 Results

3.4.1 Experimental Physiological Measurements

Animal weights and blood PCV from control groups are recorded in Table 3.1 for the first murine study and in Table 3.2 for the second study. Average PCV values were higher in BALB/c mice than C57BL/6 mice in both studies at all recorded time-points, yet this was only found to be significant in the first study on day -2 (D-2), day 6 post injection (D6) and day 12 post injection (D12). No significant differences in either weight or PCV were observed between the time-points for each strain, in either study. Additionally, no differences in weight or PCV were found when comparing time-matched results between the two batches (Table 3.2).

Table 3.1: Weight and PCV Measures from Control Animals in Mouse Study 1

		Group Median (range)				P-value	
	Time-point	Control BALB/c	Pre-infection BALB/c	Control C57BL/6	Pre-infection C57BL/6	Difference between strains	Difference over time
Weight (g)	Day -2	18.0 (17.1–19.0)	18.2 (17.2–19.4)	16.9 (16.5–19.2)	18.7 (17.3–19.5)	NS	NS
	Day 2	18.2 (17.9–19.7)	–	17.6 (15.5–18.8)	–	NS	NS
	Day 6	19.4 (18.2–20.4)	–	17.5 (15.7–18.9)	–	NS	NS
	Day 12	–	–	–	–	NT	NT
PCV (%)	Day -2	–	62.5 (56.9–66.1)	–	54.0 (47.5–58.7)	0.032 (B>C)	NS
	Day 2	65.1 (61.9–68.8)	–	53.1 (50.8–54.7)	–	NS	NS
	Day 6	56.5 (51.6–59.7)	–	52.4 (47.6–55.4)	–	0.032 (B>C)	NS
	Day 12	58.7 (56.5–60.9)	–	52.4 (44.4–54.8)	–	0.001 (B>C)	NS

Abbreviations: –, no data (results for animals post-infection shown in Chapter 4); B>C, levels in BALB/c mice higher than in C57BL/6; NS, not statistically significant; NT, not tested; PCV, packed cell volume

3.4.2 Batch Adjustment of Metabolic Datasets

Characterisation of biofluids and tissues *via* ^{1}H NMR spectroscopy demonstrated differences in the metabolic phenotypes between the two mice strains in both studies. However, compilation of the results from both studies revealed even stronger data variation due to the different batches of animals used. Separation of sample spectra due to batch differences were observed in the first two components of PCA models in all biofluids and the majority of tissues, as well as in the first two components of PLS-DA models across, as summarised in Table 3.3.

An example of the dominant batch variation observed in the plasma samples is shown in Figure

Table 3.2: Weight and PCV Measures from Mouse Study 2

		Group Median (range)		P-value		
	Time-point	Control BALB/c	Control C57BL/6	Difference between strains	Difference over time	Difference between batches
Weight (g)	Day -2	18.4 (16.6–7)	18.5 (16.4–19.9)	NS	NS	NS
	Day 2	19.4 (17.9–20.7)	18.9 (17.5–20.4)	NS	NS	NS
	Day 6	19.8 (17.6–20.6)	19.6 (17.5–20.6)	NS	NS	NS
	Day 12	19.4 (17.3–20.6)	19.5 (18.3–21.7)	NS	NS	NS
PCV (%)	Day -2	65.6 (56.7–72.6)	56.3 (47.7–60.9)	NS	NS	NS
	Day 2	63.1 (52.9–67.7)	51.5 (50.0–53.1)	NS	NS	NS
	Day 6	63.5 (58.1–68.8)	52.4 (50.8–53.1)	NS	NS	NS
	Day 12	61.9 (51.6–65.1)	53.2 (48.4–54.0)	NS	NS	NS

Key: NS, not statistically significant; PCV, packed cell volume

Table 3.3: Batch Differences Observed Before and After Batch-Adjustment of Data

		Separation Observed in Pooled Data		**Separation Observed in Batch-Adjusted Data**	
		via PCA	via PLS-DA	via PCA	via PLS-DA
Biofluids*	Urine	Yes (1 vs. 2)	Yes (1 vs. 2)	No	No (-)
	Plasma	Yes (1 vs. 2)	Yes (1 vs. 2)	No	No (-)
	Faeces	Yes (1 vs. 2)	Yes (1 vs. 2)	No	No (-)
Tissues	Liver	Yes (1 vs. 2)	Yes (1 vs. 2)	No	No (-)
	Kidney	Yes (1 vs. 2)	Yes (1 vs. 2)	No	No (-)
	Spleen	Yes (1 vs. 2)	Yes (1 vs. 2)	No	No (-)
	Ileum	No (-)	Yes (1 vs. 2)	No	No (-)
	Heart	No (-)	Yes (1 vs. 2)	No	No (-)
	Brain hippocampus	Yes (1 vs. 2)	Yes (1 vs. 2)	No	No (-)
	Brain cortex	No (-)	Yes (1 vs. 2)	No	No (-)
	Brain cerebellum	No	Yes (1 vs. 2)	No	No (-)
	Brain stem	Yes (1 vs. 2)	Yes (1 vs. 2)	No	No (-)

Key: *, Collective analyses for all time-points performed for each biofluid; (-), negative value for model predictive parameter Q^2 already in the first component; 1 vs. 2, batch separation observed in the first two components in PCA/PLS-DA models.

3.2, where separation between the two batches is explained largely by the first two components when using PCA. In contrast, separation between the two strains only becomes apparent in the 5th component, indicating the subtlety in these metabolic differences in plasma. To minimise the effect of batch variation and facilitate identification of metabolites that significantly differ between strains, a spectral subtraction method was used to adjust the data. For each biological compartment (and for each time-point for biofluids), the median spectra for each batch (across all animals) were calculated and subtracted from all samples of that batch. The batch-adjusted spectra were then re-analysed *via* PCA which showed that the previous spatial separation between the batches has been removed and was no longer apparent in the first two components. Separation between the two stains was now observed when modelling the third vs. fourth component (Figure 3.2). Further tests using PLS-DA models to compare batch variation yielded negative Q^2 predictive parameters, confirming that any previous significant batch-differences have now been removed in the adjusted data set (Table 3.3). Compiled batch-adjusted spectral data were thus used for subsequent analyses.

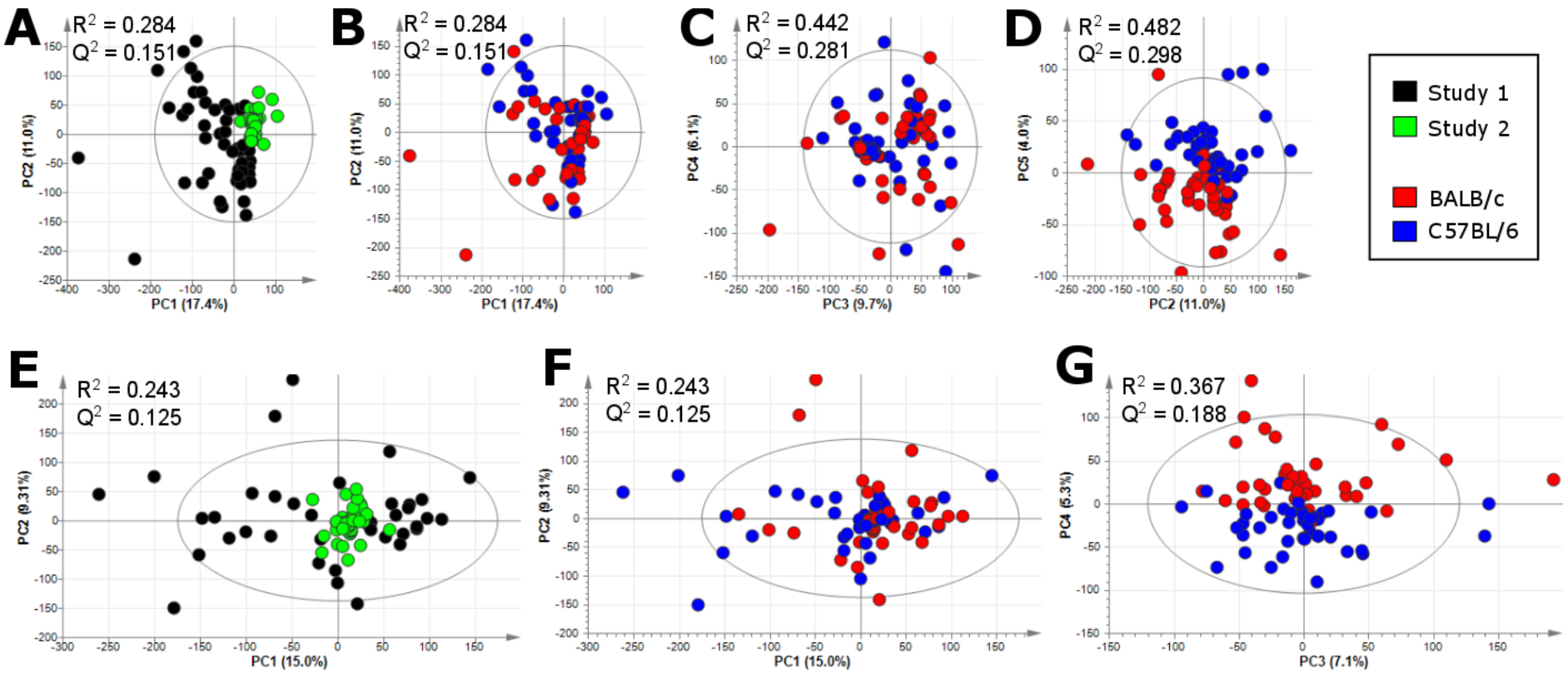


Figure 3.2: Batch and Strain Variation in Plasma Metabolic Profiles of Plasma across the Two Murine Studies

PCA scores plots based on ^{1}H NMR spectral data of control mouse plasma combined from the two studies. For batch unadjusted data (top row), PC1 vs. PC2 are shown where samples have been coloured according to the study batch, in (**A**) or according to strain, in (**B**). Further added components are shown in (**C**), namely PC3 vs. PC4, and in (**D**), PC2 vs. PC5, where strain separation becomes apparent along PC5. Similarly for batch-adjusted data (bottom row), (**E**) shows PC1 vs. PC2 highlighting samples from different batches, and the different strains in (**F**). Strain differences are visible in (**G**), PC3 vs. PC4. Cumulative R^2 and Q^2 values each model are shown on each plot, whereas individual component contribution of R^2 are shown on the axes, as percentages. Key: PC, principal component; R^2, model fit parameter for variation in spectral data; Q^2, parameter for model predictive capability.

3.4.3 Metabolic Characterisation of Biofluids

^{1}H NMR analysis of the three biofluids confirmed that the highest number of distinct spectral peaks were found in urine, yet in the presented study, the total number of metabolites that could be confidently assigned was equal for urine and plasma (31 in each case). Faecal spectra provided the least amount of detected metabolites (17 assigned), partly due to superimposition of broad resonances from larger proteins and lipids. Lactate, acetate and succinate represented the only metabolites that were identified in all three biofluids. Plasma contained several amino acids which included alanine, glutamine, lysine, branched-chain amino acids (BCAA: valine, leucine and isoleucine) and other molecules linked to energy metabolism (e.g. creatine), and lipid metabolism (3-hydroxybutyrate, cholesterol, choline and its derivatives, and several lipid moieties). Urine, on the other hand, contained a large number of gut-microbiota associated molecules including indoxyl-sulphate, phenylacetylglycine, trimethylamine, and trimethylamine-*N*-oxide, as well as a range of degradation end-products such as allantoin and ureidopropanoate. The highest resonances in faeces originated from lipid moieties (particularly from methylene [CH_2] and methyl [CH_3] groups), as well as the short-chain fatty acids (SCFA's), namely acetate, butyrate, and propionate. In addition to the SCFA's, other microbial-linked molecules detected in the faeces included dimethylamine, 5-aminovalerate and 3-hydroxyphenylpropioniate. A complete list of assigned (and unassigned) ^{1}H NMR peaks can be found in Appendix Table A.1.

Strain Differences in Metabolic Phenotypes of Biofluids

Analysis demonstrated spectral differences between BALB/c mice and C57BL/6 mice, across the two batches. PCA revealed partial separation between the two strains — as shown previously in Figure 3.2 for plasma — which was followed by O-PLS-DA analyses for extraction of the differential metabolites. An example of an O-PLS-DA of plasma, showing the strain-discrminatory metabolite peaks, is shown in Figure 3.3.

The complete list of the discriminatory metabolite markers measured in urine, plasma and faeces has been summarised in Table 3.4, showing results from both pooled (batch-adjusted) data from both studies, as well as just from the first batch. Results based on both batches showed that C57BL/6 mice had higher levels of bile acids and citrate in their plasma and urine than BALB/c mice. They also had increased concentrations of glucose, glutamine, pyruvate and certain amino acids (e.g. lysine, leucine and valine) in their plasma when compared with BALB/c, as well as increased levels of propionate in their faeces, and higher urinary 2-oxodipate concentrations. In contrast, BALB/c mice displayed higher concentrations of choline and glycerol in their plasma,

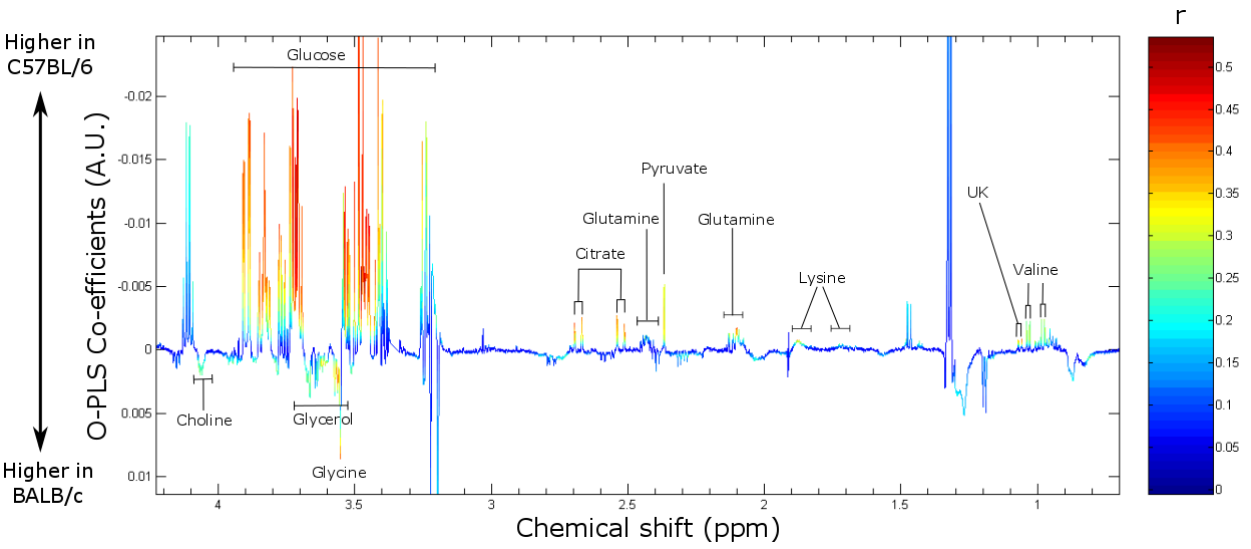


Figure 3.3: O-PLS-DA Plot of Strain Differences in Plasma

O-PLS-DA coefficient plot showing strain-discriminatory peaks within the aliphatic region of mouse plasma (0.7–4.2 ppm). Peaks pointing upwards positively correlate with C57BL/6 mice whilst those pointing downwards positively correlate with BALB/c mice. The colour of the peaks represent the strength of this correlation, with red peaks meaning high correlation and blue meaning no correlation observed. Cumulative values for R^2 and Q^2 were 0.805 and 0.695, respectively. Key: A.U., arbitrary units; ppm, parts per million; r, correlation coefficient; UK, unknown (doublet).

higher lipid fraction CH_3 in faeces, and increased levels of allantoin and creatinine in the urine. The discriminatory metabolites also included a large range of gut-microbial associated molecules in the urine e.g. concentrations of indoxylsulphate and 3-phenylpropionate were higher in C57BL/6 and 2-oxoisocaproate levels were higher in BALB/c. In general, results based on the first study alone were similar to that obtained from both batches, though significantly higher levels of lipid moieties in BALB/c plasma were also determined (Table 3.4).

Investigation of Intra-Strain Variation in Biofluids

In addition to evaluating strain-dependent variation, differences between time-points, as well as between pre-infection animals and control animals, were investigated for each strain, *via* PCA and O-PLS-DA. No significant differences were observed between control and pre-infection samples (predictive parameter Q^2 values for O-PLS-DA models were negative already in the first component for all cases). Furthermore, no separations between time-points were observed in any of the three biofluids *via* PCA (summarised in Table 3.5). O-PLS-DA showed significant differences between time-points in urinary metabolic profiles only (albeit with low Q^2 for both strains), whereas Q^2 values for faeces and plasma were all negative (Table 3.5).

In both BALB/c and C57BL/6 mice, urine from D12 was significantly different from the other time-points for a group of mice (containing roughly equal numbers from both batches), as can be

Table 3.4: Strain Discriminatory Metabolites in Biofluids

	Strain Differences across both Batches		**Strain Differences based on Batch 1 Only**	
	Positively correlated with BALB/c	Positively correlated with C57BL/6	Positively correlated with BALB/c	Positively correlated with C57BL/6
Urine	3-MOV 2-oxoisocaproate 2-oxovalerate* Allantoin Creatinine Glycine Guanidoacetate UK 1.13(d) UK 3.16(s) UK 3.17(s) UK 7.68(s)	2-oxoadipate 3-HV* 3-PP* Bile acids Citrate Glycolate* Indoxylsulphate NAG *o*-cresol* Succinate UK 1.15(d)	2-oxovalerate* 2-oxoisocaproate 2-oxoisovalerate Allantoin Glycine Guanidoacetate PAG UK 3.16(s) UK 3.17(s) UK 7.68(s)	2-oxoadipate 3-HIV Bile acids Indoxylsulphate NAG Trimethylamine *o*-cresol* UK 1.15(d)
Plasma	Choline Glycerol Glycine	Bile acids Citrate Glucose Glutamine Leucine Lysine Pyruvate Valine UK 1.07(d)	Lipid CH_3 Choline Glycerol Lipid CH=CH-CH_2-CH=CH Lipid CH=CH-CH_2-CH_2 Lipid CH_2	Citrate Glucose Glutamine Lysine Pyruvate UK 1.07(d)
Faeces	4-HPA* Lipid CH_3 NAG Phenylacetate*	3-HPP* Propionate UK 7.87(s)	4-HPA* Lipid CH_3 NAG Phenylacetate*	Propionate UK 7.87(s)

Key:, *, tentative assignments; (d), doublet; (s), singlet; 3-HIV, 3-hydroxyisovalerate; 3-HV, 3-hydroxyvalerate; 3-MOV; 3-methyl-2-oxovalerate; 3-PP, 3-phenylpropionate; 3-HPP; 3-hydroxyphenylpropionate; 4-HPA, 4-hydroxyphenylacetate; NAG, *N*-acetylglycoproteins; PAG, phenylacetylglycine; UK, unknown metabolites (accompanied with chemical shift and peak multiplicity).

Table 3.5: Evaluation of Time-related Differences Metabolic Profiles of Biofluids

	Differences between Time-points					
	***via* PCA**			***via* O-PLS-DA**		
Biofluid & Strain	Q^2 Positive?	Separation observed?	No. of Components	Q^2 Positive?	Separation observed?	No. of Components
Urine BALB/c	Yes	No	1–4	Yes	Yes	1
Urine C57BL/6	Yes	No	1–7	Yes	Yes	1
Plasma BALB/c	Yes	No	1–3	No	–	1
Plasma C57BL/6	Yes	No	1	No	–	1
Faeces BALB/c	Yes	No	1–2	No	–	1
Faeces C57BL/6	Yes	No	1–3	No	–	1

Key: –, not applicable (model invalid due to negative Q^2)

observed in the O-PLS-DA score plot in Figure 3.4. Results from the corresponding loadings plot demonstrated that in BALB/c mice, D12 samples had higher levels of 2-oxoadipate, fumarate and lactate compared with earlier time-points. In contrast, D12 urine samples in C57BL/6 mice displayed increases in 2-oxoisocaproate, 2-oxovalerate, 2-oxoglutarate, allantoin, phenylacetylglycine, indoxylsulphate and an unknown singlet at 7.68 ppm.

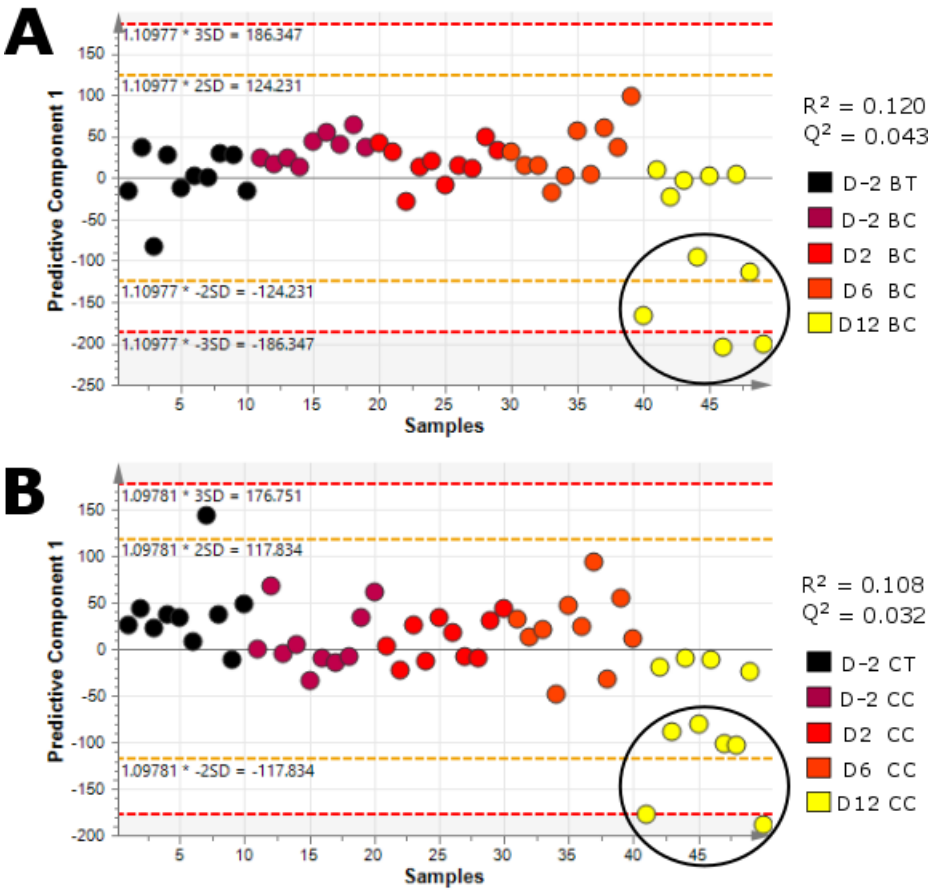


Figure 3.4: Urinary Metabolic Profiles from Day 12 Different from Other Time-points

(**A**) O-PLS-DA scores plot for urine ^{1}H NMR spectra from BALB/c mice, using one predictive component. (**B**) O-PLS-DA scores plot for urine ^{1}H NMR spectra from C57BL/6 mice, using one predictive component. In both strains, a group of animals within D12 sample set (circled in black) separate vertically along the first component from the rest. Key: BC, control BALB/c (BC); BT, pre-infection BALB/c; CC, control C57BL/6; CT, pre-infection C57BL/6; D; day (of study).

3.4.4 Metabolic Characterisation of Tissues

A total of 62 metabolites were identified in the aqueous extracts of brain sections, heart, ileum, kidney, liver, and spleen samples from batches, as well as thymus, lung and lymph nodes, measured in batch 1 only, listed in Appendix Table A.2. The BCAA's, lactate, alanine, acetate, choline, phosphocholine, and glycerophosphocholine were clearly identified in all tissues. Other commonly detected metabolites (present in at least 6 tissues) included 3-hydroxybutyrate, betaine, ethano-

lamine, glutamate, glutamine, glutathione, glycine, inosine, nicontamide, succinate, and adenosine monophosphate (AMP[4]). Energy stores glycogen and mannose represented liver-specific metabolites, whilst 4-aminobutyrate (also known as GABA, a neurotransmitter) was only found in the brain.

PCA models using the first two components, based on the the tissues common to both batches, were constructed for each strain separately, as demonstrated in Figure 3.5. Despite the high degree of variation across the data, ileum was found to be clearly differentiated from the other tissues in BALB/c mice. In fact, the plots show that both ileum and liver represented two of the three tissues furthest from the origin for both strains (highlighted in the figure), thus contributed the most towards the variation observed the PCA model. The third model apex represented the hippocampal brain region for BALB/c mice, and brain cerebellum for C57BL/6 mice. Extracts from the four brain sections and kidney were localised closely together in BALB/c mice, indicating spectral similarity, not observed in C57BL/6 mice.

Strain Differences in Metabolic Phenotypes of Tissues

Metabolic analysis of both strains revealed significant strain-discriminatory metabolites in the ileum, liver, spleen, brain hippocampal region and stem, as shown in Table 3.6. The hippocampus yielded the largest list of discriminatory metabolites of the tissues, which included increased levels of *myo*- and *scyllo*-inositol, creatine, glutamate, lipid fractions and *N*-acetylaspartate (NAA) in BALB/c when compared with C57BL/6 mice. The spleen represented the next most discriminatory tissue, which displayed augmented concentrations of choline and several lipid fractions and lower levels of adenosine mono-/di/tri-phosphate, glutamate, lactate, phosphocholine and glycerophosphocholine in BALB/c mice than the C57BL/6 strain. In contrast, C57BL/6 mice had higher concentrations of acetate and trimethylamine (TMA) in the ileum, and higher levels of 2-aminoadipate and BCAA's valine and isoleucine, in the liver.

No significant differences were observed in the heart, kidneys, brain cortex and cerebellum when analysing both batches, though when analysing batch 1 data on its own, significant differences in the kidney metabolic phenotypes were found. Tissue differences appeared more pronounced when analysing batch 1 data alone, generally resulting in a larger list of discriminatory markers, shown in Table 3.6. Results included higher phosphocholine concentrations in the kidneys, spleen and brain (hippocampus and stem) of C57BL/6 mice than BALB/c, which in turn had increased succinate in their ileum and liver extracts.

[4]Tentatively assigned

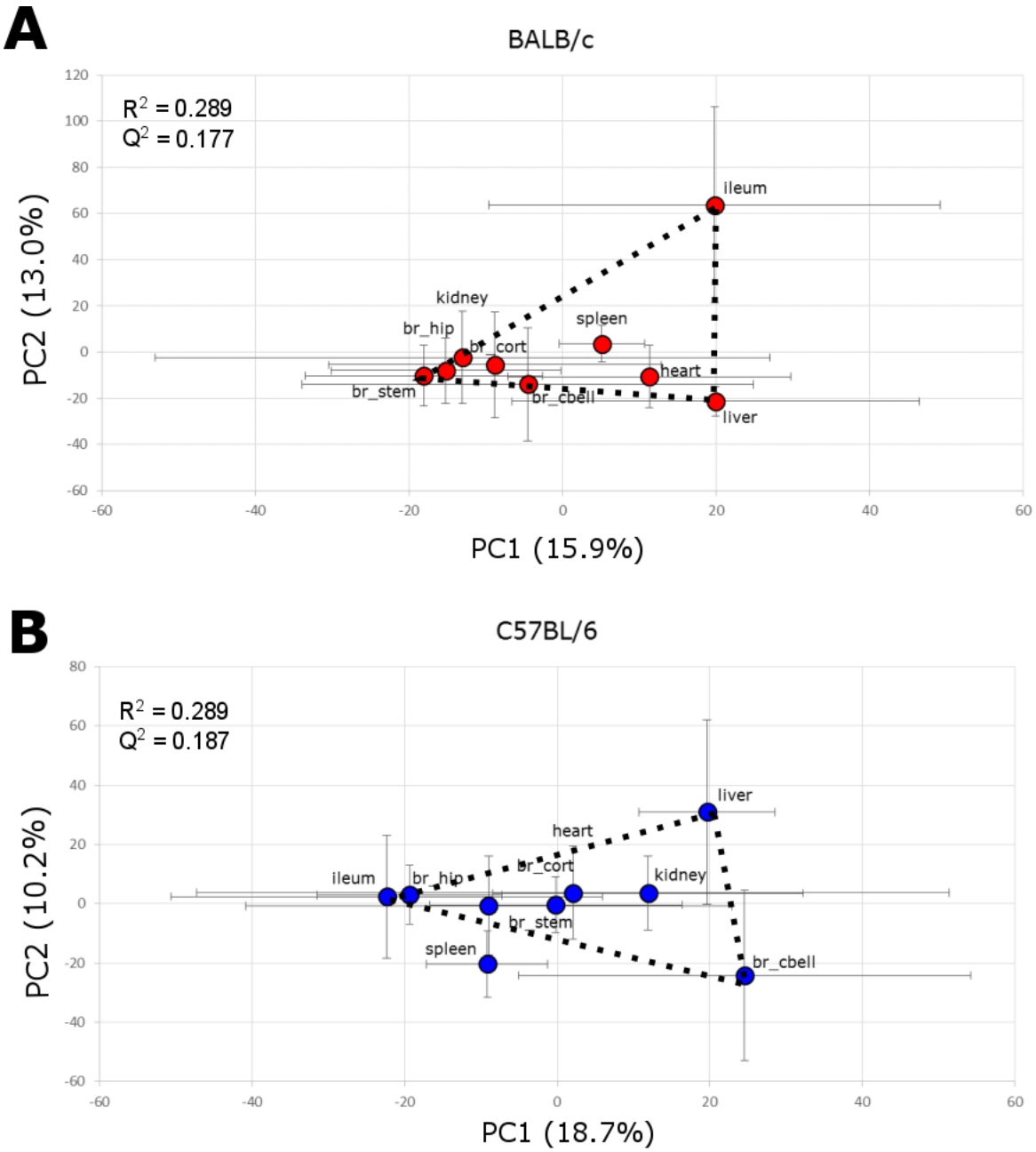


Figure 3.5: Metabolic Variation across Tissue Extracts

PCA scores plots for tissues analysed *via* ^{1}H NMR taken from BALB/c mice (**A**) and from C57BL/6 mice (**B**). Each circle represents the mean average PC score for each tissue section, with standard error of the mean (SEM) shown as error bars. The 3 tissues furthest away from the origin have been connected with a hashed line, representing the tissues that most contributed towards the model. Cumulative R^2 and Q^2 values each model are shown on each plot, whereas individual component contribution of R^2 are shown on the axes, as percentages. Key: br, brain; cbell, cerebellum; cort, cortex; hip, hippocampus.

Table 3.6: Strain Discriminatory Metabolites in Tissues

	Strain Differences seen in both Batches		**Strain Differences based on Batch 1 Only**	
	Positively correlated with BALB/c	Positively correlated with C57BL/6	Positively correlated with BALB/c	Positively correlated with C57BL/6
Brain, cerebellum	NS	NS	NS	NS
Brain, cortex	NS	NS	NS	NS
Brain, hippocampus	3-hydroxyvalerate* Acetate Adenosine* Choline Creatine Ethanolamine Glutamate Glycerol Lipid CH_3 *myo*-Inositol NAA NAD^+ *scyllo*-Inositol UK 7.08(s)	Glycine Phophocholine	GABA Alanine AMP Creatine Glutamate Glycerol Lipid CH_2 Lipid CH_3 *myo*-Inositol NAA NAD^+ *scyllo*-Inositol	ADP* Glycine Lactate Phophocholine
Brain stem	UK 7.08(s)	Adenosine* ATP* Lactate UK 1.45(d)	–	Aspartate ATP* Fumarate Glutamate Lactate NAA Phophocholine
Heart	NS	NS	NS	NS
Ileum	AMP*	Acetate Fumarate Hypoxanthine TMA	AMP* Glutathione Succinate	Fumarate TMA
Kidney	NS	NS	–	Glutamate Glutamine Glycine Isoleucine Phosphocholine Succinate Tyramine Uridine Valine
Liver	Glutamine Inosine 4.29(m)	2-aminoadipate AMP* Isoleucine Tyrosine Valine UK 8.07(s)	Dimethylamine Glutamine Inosine Succinate Uridine	2-aminoadipate AMP* ATP* Tyramine Tyrosine Valine UK 8.07(s) UK 8.74(s)
Lungs	N/A	N/A	Choline Creatine Ethanolamine Glutamate Hypoxanthine Inosine Sarcosine*	Betaine Glutamine Glutathione GPC PEA Phosphocholine
Lymph nodes	NS	NS	NS	NS

Continued on next page

Table 3.6: (continued)

	Strain Differences seen in both Batches		**Strain Differences based on Batch 1 Only**	
	Positively correlated with BALB/c	Positively correlated with C57BL/6	Positively correlated with BALB/c	Positively correlated with C57BL/6
Spleen	Choline Lipid CH_2 Lipid CH_3 Lipid CH_2CH_2CO	ADP AMP ATP* Glutamate GPC Lactate PEA Phoshocholine Taurine Valine	Lipid CH_3 Lipid CH_2 Choline Inosine	ADP AMP ATP Aspartate Betaine Glutamate Glutamine Glutathione GPC Lactate PEA Phoshocholine Succinate Taurine Valine
Thymus	N/A	N/A	AMP Glutamine Lactate	ADP Aspartate Betaine NAA

Key as described for Table 3.4, with following additions: –, no data; ADP, adenosine diphosphate; AMP, adenosine monophosphate; ATP, adenosine triphosphate; GABA, 4-aminobutyrate; GPC, glycerophosphocholine; N/A, not applicable (lung and thymus only collected in batch 1); NAA, *N*-acetylaspartate; NAD^+, nicotinamide adenine dinucleotide (oxidised ion); PEA, *O*-phospho-ethanolamine TMA, trimethylamine.

3.4.5 Plasma Cytokine Profiles

Plasma concentrations of eight cytokines (IL-1β, IL-2, IL-4, IL-5, CXCL1, IL-10, IL-12, and TNFα) were tested using the ultra-sensitive mouse Th 1/Th 2 Cytokine multiplex MSD assay kit (see Methods Section 2.6.1). Only IL-1β, IL-10, IL-12 and TNFα displayed detectable levels, shown in Figure 3.6. No significant differences in cytokine concentrations were observed between the time-points of the control groups, or between the pre-infection group and the control group of the same strain, as expected. When cytokine levels were compared between time-point matched strains, IL-1β was found to be significantly higher in BALB/c mice than C57BL/6 mice for D-2 in the pre-infection groups. As no significant time-related changes were determined, differences between the two strains were re-analysed, pooling all time-points, as well as the pre-infection and control animals (i.e. [BT and BC] vs. [CT and CC]). Such analyses further revealed significantly elevated concentrations of both IL-1β ($p = 0.017$) and CXCL1 ($p = 0.001$) in the BALB/c strain when compared with C57BL/6 (Figure 3.6).

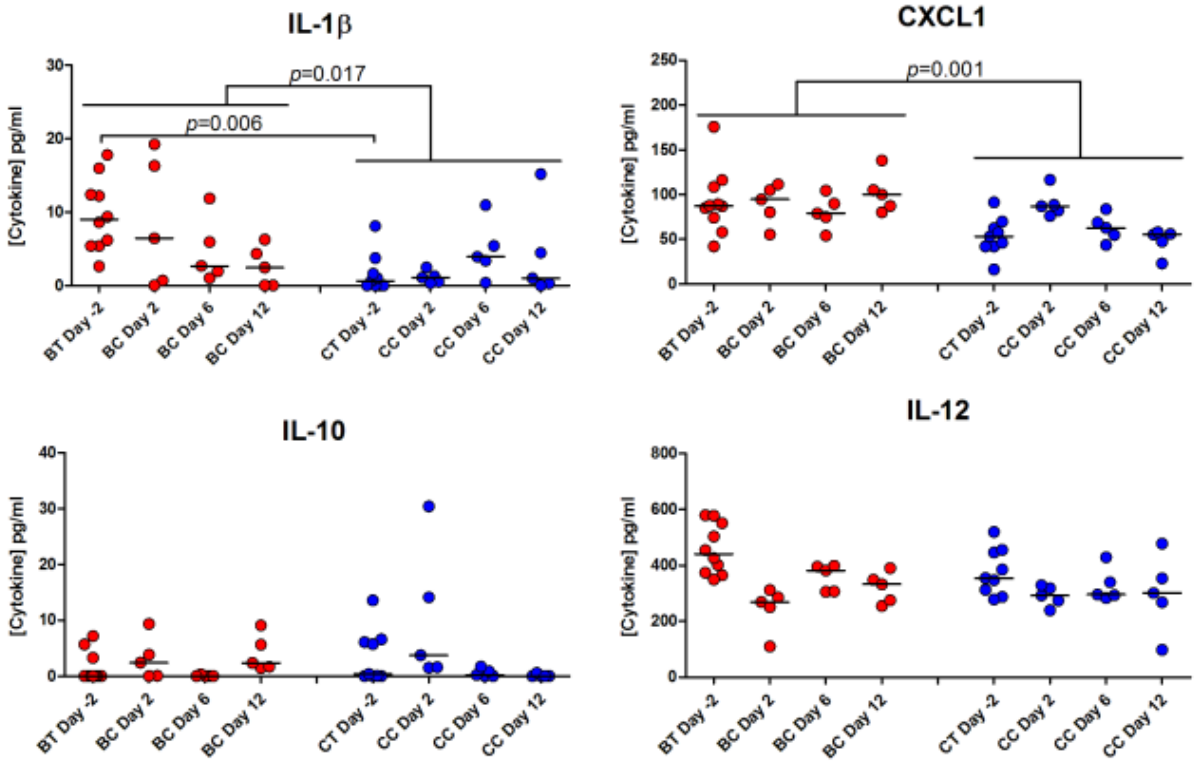


Figure 3.6: Elevated Pro-inflammatory Cytokines in Plasma of BALB/c Mice

Each circle represents one sample and horizontal lines show median of each group. Changes between time-points for each strain, were tested for *via* repeated measures Friedman tests and Wilcoxon matched -pair tests (followed by Bonferroni correction). Significant differences between strains were tested separately between matched time-point data, as well as collectively for all time-points (pooling pre-infection mice with the control animals for each strain), using Kruskal–Wallis tests and Mann–Whitney tests (and Bonferroni correction). Key: BC, BALB/c controls; BT, BALB/c pre-infection group; CC, C57BL/6 controls; CT; C57BL/6 pre-infection group. (No available data for BC/CC Day -2).

3.4.6 Faecal Microbial Results

Overall relative percentage abundances of faecal microbiota were examined at multiple taxonomic levels, including phylum, class, order, family, and genus and species (shown in decreasing order). Slight discrepancies existed between microbial datasets obtained from Research & Testing Laboratory and from Dr Marchesi's MOTHUR analyses, which was generated subsequently.[5] Data from Research & Testing at the microbial class level were selected for analyses of the *L. major* infection study, in order to determine high order microbial changes in response to infection and discriminate between functionally distinctive groups of bacteria, as opposed to purely genotypic differences observed at the lower taxonomic levels. Thus, for consistency purposes, the top eight classified bacterial classes with highest relative abundance were analysed for both control mice and infected

[5]Notably, all unclassified OTU's had been excluded from Research & Testing datasets and no *Bacteroidia* were determined across the control animals with the MOTHUR data.

mice. These top eight classes were comprised of *Clostridia, Gammaproteobacteria, Erysipelotrichia, Bacteroidia, Alphaproteobacteria, Betaproteobacteria, Bacilli, and Actinobacteria*, which in total represented 99.1% of all (classified) bacterial classes.

Preliminary analyses revealed 13 samples that displayed unusually high levels of *Gammaproteobacteria*, coupled within lower than average abundance of the *Clostridia* class (see Appendix Figure A.1). These samples were BC08_D12, BC09_D12, BC10_D12, BT07_D-2, BT08_D-2, BT09_D-2, BT10_D-2, CT01_D-2, CT02_D-2, CT03_D-2, CT04_D-2, CT05_D-2, and CT06_D-2. As these samples were substantially different from the remaining samples, they were all excluded from further analyses. Subsequent analyses using non-parametric tests showed that there were no statistical differences between time-points per strain, or between pre-infection groups and the control animals. Out of the eight classes, *Erysipelotrichia* represented the only one that was significantly different between the two strains, as shown in Figure 3.7. This class of bacteria was significantly higher in C57BL/6 mice than BALB/c when combining pre-infection with control groups from all time-points ($p = 0.0008$), yet when comparisons were broken down for each time-point, the p-values were no longer significant. Additional trends indicated apparent higher levels of *Alphaproteobacteria* in BALB/c mice when compared with C57BL/6 mice.

The MOTHUR dataset was generated in such a format to facilitate direct comparison of the microbial OTU data at different taxonomic levels and was thus used for subsequent analyses. PCA analyses were performed on phylum, class, order, family, genus data sets, incorporating both classified and unclassified OTU's, which demonstrated no visible separations between strains within the first three components in the phylum, class and order data in either scenario. Partial separations between strains were observed at the family level, which were most pronounced when the dataset also included unclassified OTU's, shown in Figure 3.8. No obvious separations were observed between the pre-infection groups and the control animals. Similar results were also determined at the genus level, though the family level was selected for final analyses since the majority of the data at the genus level consisted of unclassified (and thus unknown) OTU's.[6] A list of detected microbial families and genera detected across both strains is summarised in Appendix Table A.3.

Discriminatory bacteria were compared between strains using White's non-parametric tests (collectively for all time-points), followed by multiple test correction. Considering the high number of comparisons at the family taxonomy level (including unclassified OTU data), Bonferroni correction was considered too stringent for this data set. Analyses with this type of multiple test correction only revealed four significant discriminatory markers with corrected p-values < 0.05

[6] 73.1% unclassifieds were present at genus level, compared with 42.0% at family level. (Proportion of unclassified OTU's at the higher taxonomic levels, order, class and phylum, were 24.1%, 22.2%, and 14.2%, respectively).

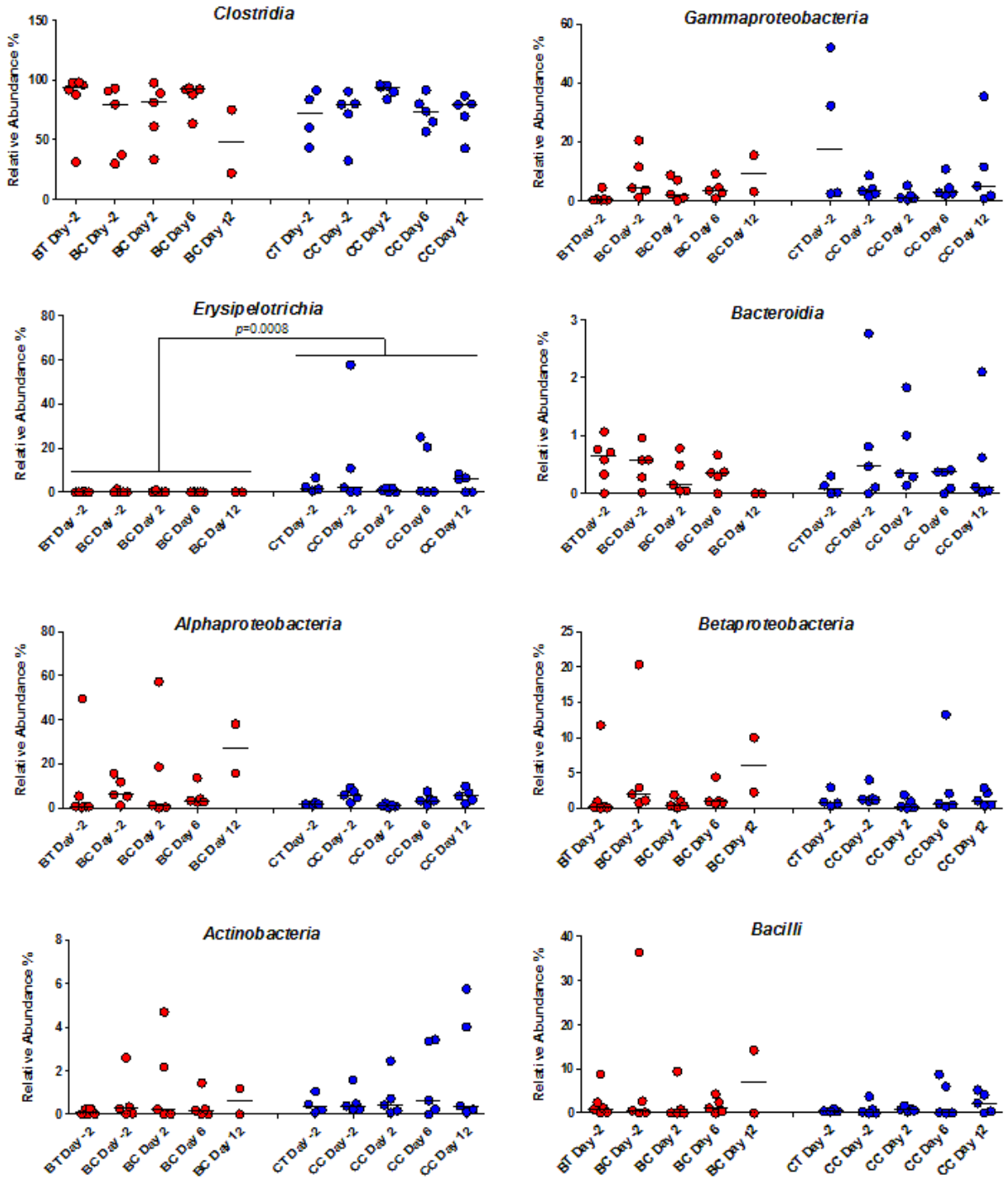


Figure 3.7: Erysipelotrichia Class Found Significantly Higher in C57BL/6 than in BALB/c Mice

Relative percentage abundances of eight faecal microbial classes measured in the two mouse strains (excluding outliers). Each circle represents one sample, and horizontal lines show median of each group. Differences between strains, and between the pre-infection animals and controls, were tested using Mann–Whitney U-tests for each bacterial class, followed by Bonferroni multiple test correction. Abundances between time-points per class were compared using Wilcoxon matched-paired test, also with Bonferroni correction. Key as described for Figure A.1.

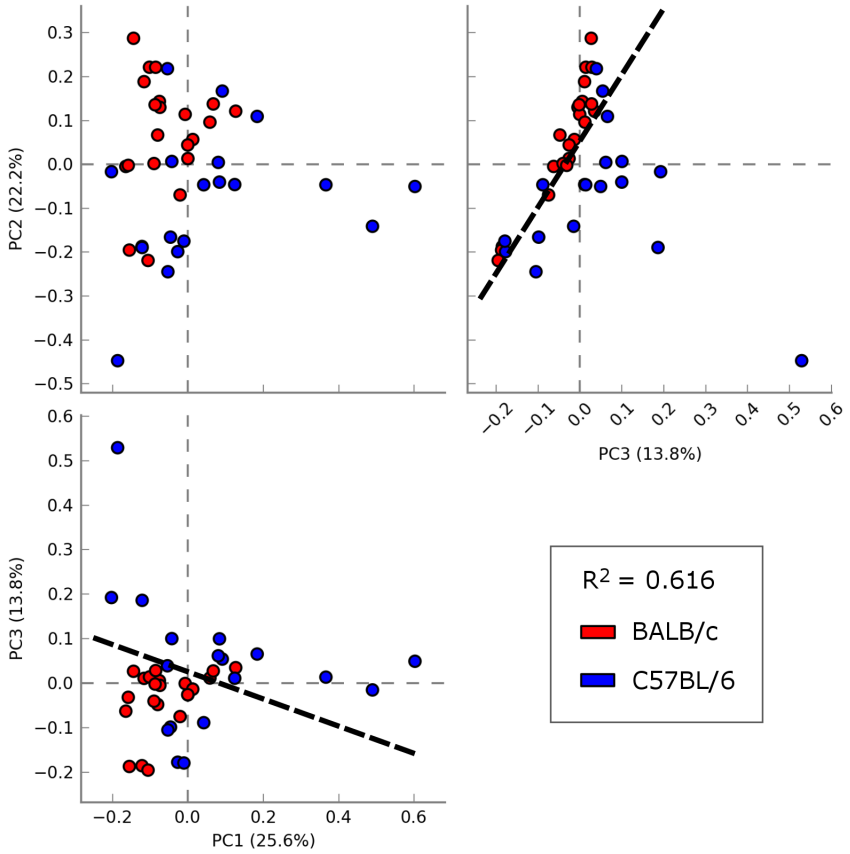


Figure 3.8: Variations in Faecal Bacterial Family Composition Show Partial Strain Separation

PCA scores plot showing the first three components, based on microbial family OTU counts measured in the faeces of BALB/c and C57BL/6 mice. Each circle represents one sample. Cumulative R^2 value shown for 3PC's shown on figure, whereas individual component contributions are shown as percentages on axes. Key: PC, principal component; R^2, proportion of variance explained by model.

with average differences $< 0.2\%$ (shown in Appendix Figure A.2). Instead, Benjamini–Hochberg was selected as an appropriate multiple test correction, which demonstrated 21 significant strain-discriminatory bacterial makers (15 of which represented unclassified OTU's), summarised in Figure 3.9. Results demonstrated that the relative levels of the *Lachnospiraceae* family (order *Clostridiales*, class *Clostridia*, phylum *Firmicutes*) displayed the largest difference between strains, which were higher in BALB/c mice. Other unclassified *Firmicutes* were also significantly higher in BALB/c mice. The second largest differential microbial family were members of the *Erysipelotrichaceae* (order *Erysipelotrichales*, class *Erysipelotrichia*, phylum *Firmicutes*), of which *Turicibacter* represented the most dominant genus (see Appendix Table A.3). This family was more abundant in the C57BL/6 strain, in line with previous analyses on bacterial classes. Bacteria belonging to the *Peptostreptococcaceae* family (particularly in *Clostridium* cluster XI, order *Clostridiales*, class *Clostridia*, phylum *Firmicutes*) were also significantly higher in C57BL/6.

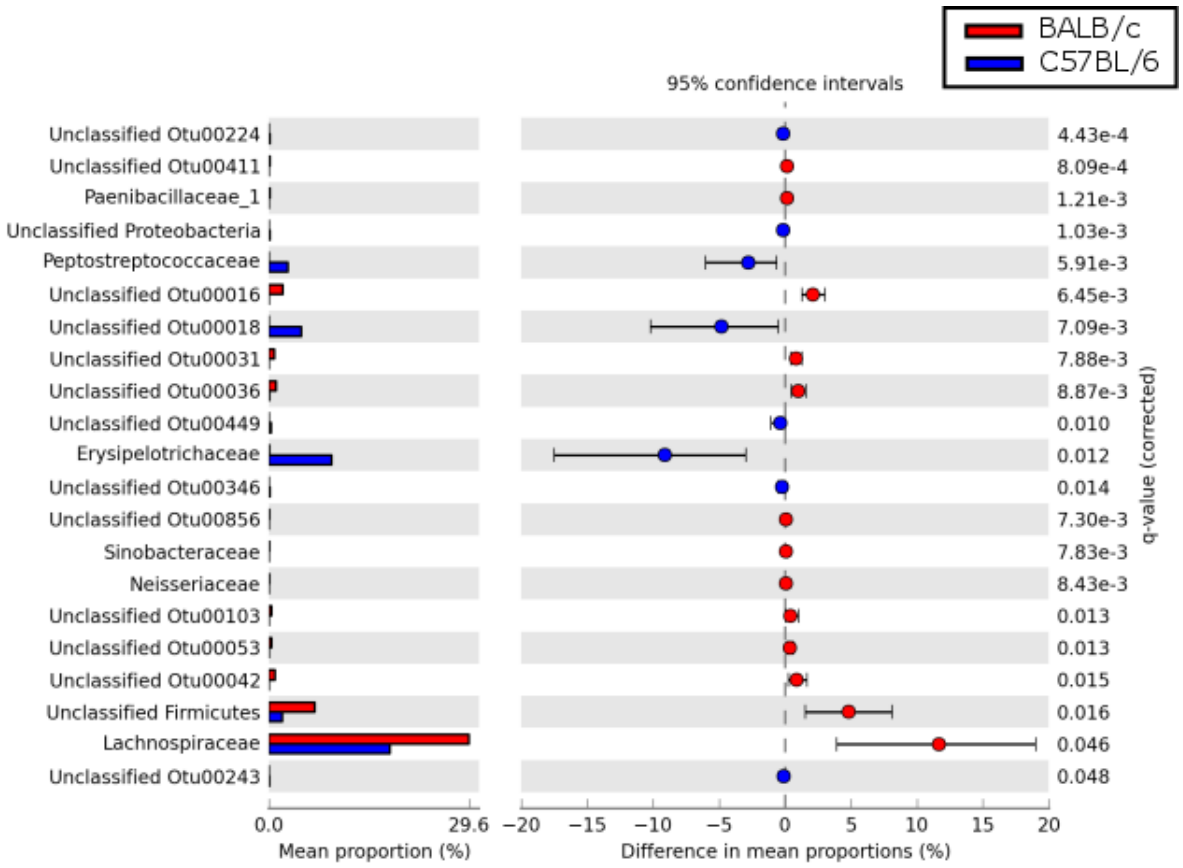


Figure 3.9: Differentially Expressed Faecal Microbial Families in the Two Mouse Strains

Extended error plot showing average percentage relative abundance (proportion) of faecal microbial OTU's on the left plot, and average difference between the group means and their confidence intervals on the right. Results were based on OTU count data resolved at the family taxonomic level. Significant differences between BALB/c mouse strain and C57BL/6 strains were tested using White's non-parametric test, with Benjamini–Hochberg false discovery rate of $q < 0.05$.

3.4.7 Hierarchical PCA of Metabolic, Cytokine and Microbial Profiles

Separate PCA models were generated for metabolites of urine, plasma and faeces (collectively for all time-points), as well as cytokine concentrations and relative abundances of faecal microbial families. Scores from the first two components of each of these "lower level" PCA models were then loaded as variables into a newly constructed "upper level" PCA model, known as H-PCA. H-PCA provides a broad overview of the relationships between the different lower level variables, used in this case to determine the relative contribution of each of the biological compartments to the overall variation in phenotypes observed across the two mouse strains. As shown in Figure 3.10, the first component obtained from PCA on urine (urine_PC1) contributed the most towards the strain separation observed in the H-PCA model. Figure 3.10 also shows the lower level PCA scores and loadings plots for urine, highlighting the urinary metabolites with the most discriminatory power. These metabolites included an unassigned singlet peak at 3.16 ppm and glycine, which

were higher in BALB/c, and 2-oxoadipate, which was higher in C57BL/6 mice. Following these labelled metabolites, the next most differential metabolites included 2-oxoisocaproate and creatinine (higher in BALB/c) and *O*-cresol (higher in C57BL/6).

The first components of faeces provided the next largest impact towards strain separation in the H-PCA models; *N*-acetylglycine and phenylacetate (higher in BALB/c) and an unknown singlet peak at 7.87 ppm (higher in C57BL/6) were most the most differentially expressed metabolites in the faeces. The first component of several tissues' PCA models represented the next most differential phenotypes, as well as the second component of faecal microbiota PCA. The bacteria that contributed the most towards PC2 were *Rhodobacteracea*e and a several unclassified OTU's (especially Otu00016 and Otu00031), higher in BALB/c, as well as *Erysipelotrichaceae* and *Peptostreptococcaceae*, which were both higher in the C57BL/6 strain. The most differential tissues included lungs, liver, thymus, and ileum (in order of decreasing contribution). Betaine, phosphocholine and glutamine were the most discriminatory metabolites in the lungs, all of which were higher in the C57BL/6 mice. Betaine was also one of the most discriminatory metabolites within the thymus, in addition to *N*-acetylaspartate, also higher in C57BL/6. Uridine, glutamine and inosine were the most strain-selective metabolites in the liver, all of which were higher in BALB/c mice. Lastly, succinate, glutathione and AMP contributed the most towards strain separation in the ileum PCA, all of which were higher in BALB/c mice.

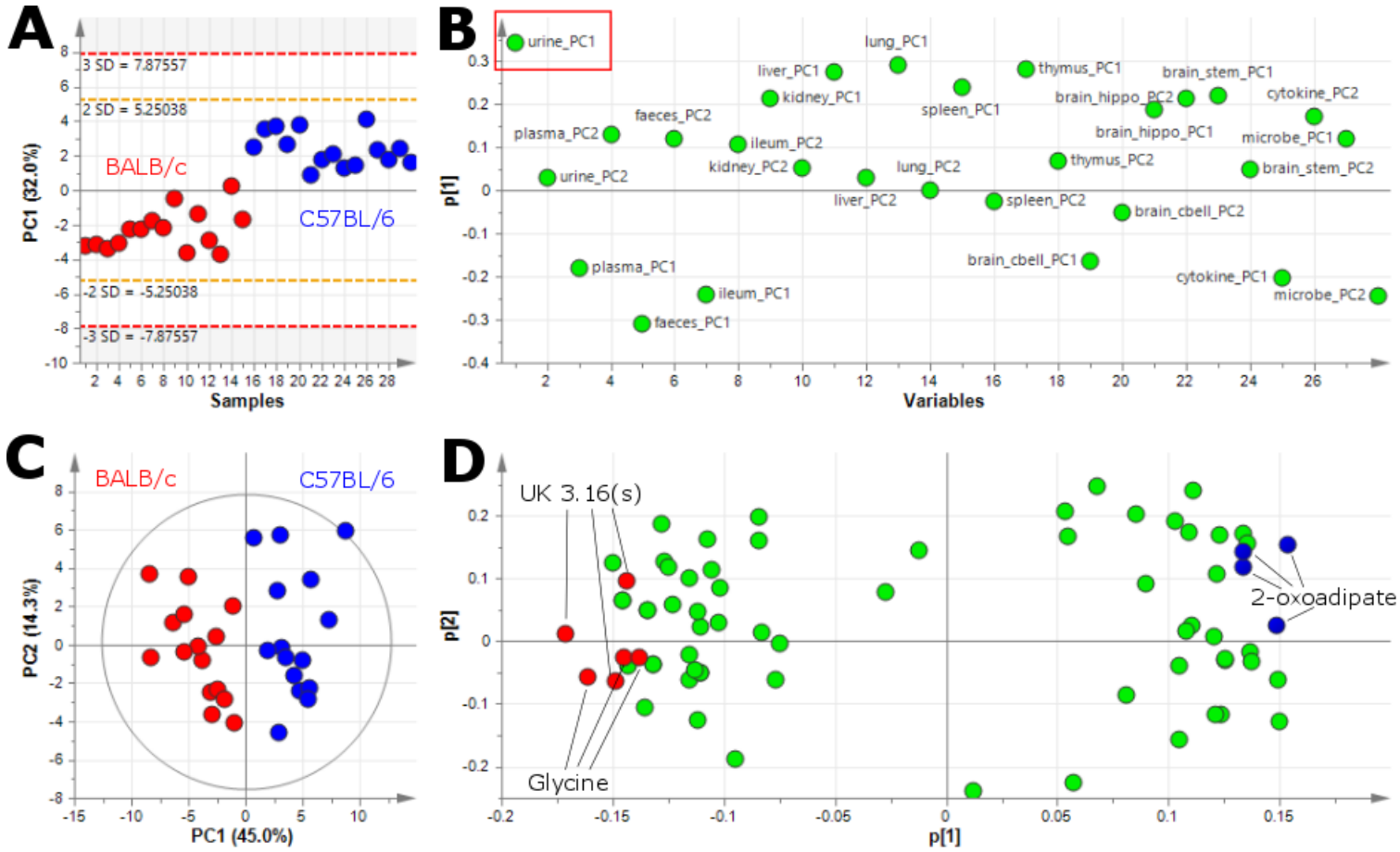


Figure 3.10: Hierarchical PCA Demonstrates Urinary Metabolites as Most Discriminatory Markers between Two Mouse Strains

(**A**) H-PCA scores plot (using one PC) showing vertical strain separation between samples (red/blue circles). (**B**) The corresponding loadings plot of the H-PCA model, showing the relative contribution (p[1]) of each of the variables (green circles), revealing PC1 from urine PCA model (shown in **C** and **D**) as the most discriminatory component. (**C**) The original PCA scores plot for urinary metabolites (combining all time-points), with horizontal strain separation along PC1. (**D**) The loadings plot corresponding to the urine PCA plot, showing metabolite contributions to PC1 (p[1]) and PC2 (p[2]). Selected metabolites with the highest p[1] have been labelled. Model R^2 values (in %) for each component shown on axes of scores plots.

3.5 Discussion

3.5.1 Inherent Differences in Haematocrit between Strains

The groups of BALB/c and C57BL/6 mice characterised in the presented studies had been designated as control animals, as part of a larger experimental leishmaniasis study, which lasted nearly two weeks in total. Whilst the dietary intake of the mice was not officially monitored, no significant changes in weight were detected during the course of the study, minimising finding potential metabolic discrepancies due to e.g. weight gain/increased dietary intake. These findings were to be expected as the studies only lasted 12 days and the animals were not subjected to any serious intervention. Furthermore, no significant differences in weight were observed between the strains, which was also as anticipated as animals had been weight-matched at the start.

In contrast, significant differences between in haematocrit were apparent between strains, already from the first time-point. BALB/c mice consistently had higher PCV's than C57BL/6 mice ($p = 0.011$ for D-2, D6, and D12), similar to suppliers' information.[208,209] Yet no changes in the PCV were observed between time-points for each strain, signifying that the strain variation is most likely due to stable intrinsic differences.

^{1}H NMR spectroscopy was subsequently applied, where a large range of metabolites across biofluids and multiple tissues were identified, providing a comprehensive multi-compartment metabolic characterisation of BALB/c and C57BL/6 mice. Further plasma cytokine measurements and faecal microbial screening provided additional insights into the their overall health, as indicators of immune status and intestinal dysbiosis, respectively.

3.5.2 Metabolic Characterisation Reveals Strain-Differences on a Systemic and Local Level

Metabolic analyses of biofluids revealed that the majority of the metabolites that were detected were specific to one or two biofluids, highlighting the complementary information gained by analysing all three. Plasma gave an important insight into systemic carbohydrate, lipid and (in particular) amino acid metabolism. Urine and faeces gave additional information on gastrointestinal microbial activity, with detection of a number of microbiota-associated metabolites in both biofluids, though the number of identified metabolites was greatly increased in urine. Co-analyses of variation across metabolic biofluid and tissue profiles, cytokine concentrations and faecal microbial abundances *via* H-PCA further demonstrated that urine displayed the largest strain differences. As metabolites in urine are to be excreted and thus not directly under strict homoeostatic regulation as other body

components[218], their concentrations can vary to a much higher degree, allowing the detection of much increased/decreased concentrations of metabolites e.g. in disease relative to healthy controls. Together with ease of collection, such properties have promoted the use of urine as a major source of several clinical pathological biomarkers.[219]

The screening of multiple tissue samples provided an additional layer of metabolic information on a local scale, for a more complete overview of mouse physiology (and in the case for the *L. major* infected animals, pathological changes linked with the parasite). Assessment of metabolic variation across tissue extracts revealed that liver and ileum extracts displayed differential phenotypes from the other tissues, observed in both BALB/c and C57BL/6 strains. Furthermore, they represented major strain-differentiating compartments in H-PCA. This finding is not unexpected, considering that both of these tissues are key organs involved in metabolism of nutrients. Lungs and thymus were the two most strain-discriminatory of all assessed tissues. Metabolic dissimilarities between the strains in thymus, a major immune organ responsible for T cell development[162], may reflect the varying capabilities of these two strains in mounting adaptive immune responses, observed as their differing susceptibilities to various infectious agents in experimental models. However, the reason for the differential phenotypes of lungs between BALB/c and C57BL/6 remains unclear, especially as both mice were exposed to similar venting/housing conditions.

Comparative multivariate analyses between the two strains revealed metabolic disparities in all three biofluids and the majority of the assessed tissues. C57BL/6 mice generally displayed higher levels of metabolites linked with the tricarboxylic acid (TCA) cycle[220] than BALB/c mice, including citrate (in urine and plasma), fumarate (in the ileum) and succinate (in kidneys and spleen). They also contained significantly higher glucose plasma concentrations, in agreement with published results[205] and supplier information.[206–209] In contrast, augmented levels of several lipid moieties were determined in plasma, faeces, splenic extracts and brain hippocampal regions of BALB/c mice. This was also the case for glycerol, another lipid-associate metabolite that is a component of triglycerides, which was higher in plasma and hippocampus. BALB/c mice have been shown to have overall higher fat content than C57BL/6[205], as well as have higher concentrations of triglycerides and cholesterols.[206–209] Consequently it could be hypothesised that BALB/c mice rely more on lipid metabolism for energy, due to their increased bioavailability, whilst the TCA cycle is more important in C57BL/6 mice.

Other metabolic discrepancies between the two strains included increased concentrations of choline in BALB/c plasma, spleen, hippocampus, and lungs, whilst phosphocholine was higher in these aforementioned tissues in C57BL/6 mice. Choline is a major component of biological

membranes, complexed in its phosphorylated form with fatty and glycerols as phospholipids, and is responsible for stabilising membrane integrity and membrane signalling functions.[221] Consequently it would appear that C57BL/6 mice have higher phospholipid content within their cell membranes than BALB/c mice, which may result in varying degrees of cell communication within/between tissues such as the brain and spleen, compared to BALB/c mice. Another key function of choline is its role as a neurotransmitter across the central nervous system, in the form of acetylcholine[221], suggesting that cholinergic transmission may be increased in BALB/c mice.

Strain-related differences in BCAA and purine metabolism were also highlighted across tissues profiles. C57BL/6 mice displayed higher concentrations of BCAA's, in particular valine, in the plasma, liver, and spleen. On the other hand, BALB/c mice had higher urinary levels degradative metabolites of BCAA's, such as 2-oxoisovalerate (break-down product of valine), 2-oxoisocaproate (leucine metabolite) and 3-methyl-2-oxovalerate (product of isoleucine catabolism)[222], signifying a higher metabolic turn-over. Furthermore, multiple tissues displayed contrasting concentrations of nucleosides/nucleotides and their precursors: adenosine and/or its phosphates (AMP, ADP, ATP) were found to be significantly different (in most cases, higher in C57BL/6 mice), whilst inosine was found in higher concentrations in BALB/c lungs and liver. Levels of hypoxathine were found to be higher in lungs of BALB/c, yet was relatively increased in the ileum of C57BL/6 mice.

Lastly, variations in a number gut microbial related metabolites were detected in the excreted biofluids, affecting phenyl derivatives[171] in particular: urinary 3-phenylpropionate and faecal 3-hydroxyphenylpropionate were relatively elevated in the C57BL/6 mice, whereas concentrations of faecal phenylacetate and 4-hydroxyphenylacetate were higher in the BALB/c strain. Concentrations of urinary indoxylsulphate (derived from microbial break down of dietary tryptophan) and faecal propionate (a SCFA) were also augmented in C57BL/6. Thus, despite identical diet, these metabolic changes implied divergences in gastrointestinal microbial populations, results that were further examined with microbial pyrosequencing analyses.

3.5.3 Strain-related Variation in Faecal Microbial Families

Previous studies have demonstrated variability in gut microbial composition due to both environmental (e.g. variation due to diet, batch and housing/cages) and genetic factors.[223–225] Faecal microbial analysis *via* pyrosequencing have shown that both the class and family level demonstrated increased proportion of *Erysipelotrichia*, strict anaerobes that contained primarily bacteria of the *Turicibacter* genus), present in C57BL/6 mice. Limited and diverse literature exists on the physiological and/or pathological functions of these bacteria: increased *Erysipelotrichia* have been linked

with choline-induced fatty liver in humans[226] and tumorigenesis in mice[227], whilst decreases have been associated with diabetes[228] and mammalian acute haemorrhagic diarrhoea.[229]

Whilst *Firmicutes* were generally increased in BALB/c mice, results at the family levels showed that the *Peptostreptococcaceae* family within the *Firmicutes* were in fact higher in the C57BL/6 mice. *Lachnospiraceae* (also *Firmicutes* phylum) were in contrast significantly higher in the BALB/c mice. These results show that investigations at phylum or class show a global overview of microbial populations, analyses at the lower taxonomic levels can provide more subtle variations between groups, and thus a more accurate reflection of the diversity within the gastrointestinal microbial communities.

3.5.4 Cytokine Measurements Demonstrates Intrinsic Pro-inflammatory Profile in BALB/c mice

Contrasting immune responses of mouse strains towards a range of experimental pathological infections have been well documented. The majority of such studies have been based on *ex vivo* experiments, often measuring proliferation, secreted cytokine or activation markers from stimulated/infected splenic cells[230–234], with only a limited number of studies measuring circulating cytokine levels.[235] This is partially due to difficulty in measuring the typically low basal concentrations of many of these cytokines, especially as the majority of them act locally.[236] This characteristic, is in addition to the dynamic and transient nature of cytokine responses, where different cytokines vary in concentrations in response to a stimuli, and have differing incubation periods.[166]

These aforementioned complications may explain why only four out of the nine tested cytokines were able to be detected in the plasma. Nevertheless, results from these four cytokines revealed that BALB/c mice displayed increased levels of IL-1β, a well-studied pro-inflammatory cytokine with pleiotrophic roles, e.g inducing the expression of the a multitude inflammatory-linked genes including cytokines and their receptors, acute-phase proteins, adhesion molecules, growth factors and tissue-modelling enzymes.[237,238] Furthermore, concentrations of CXCL1 (also known as GRO1 oncogene, GROα, KC, or neutrophil-activating protein 3), another pro-inflammatory cytokine known for its neutrophil chemoattractant abilities[239], were also higher in BALB/c mice. These findings are the first to my knowledge that have shown strain-associated variations in circulatory cytokine concentrations in the absence of experimental stimuli. These results imply that unlike C57BL/6 mice, BALB/c mice have continual on-going pro-inflammatory processes that are inherent to their strain, which in turn may contribute to their varied susceptibility to a range of experimental

infections.[203]

3.5.5 Study Limitations

One of the major limitations of this study was the relatively low numbers of control animals used for each strain in the first main study. Whilst additional animals were recruited in a separate batch for metabolic analyses, cytokine and faecal microbial assessment were not performed due to limited time and resources. This was particularly unfortunate for the microbiome data, where sample numbers were even further reduced due to the exclusion of outliers. The reason for presence of these unconventional high–*Gammaproteobacteria*–low–*Clostridia* samples is unclear, possibly owing to sample cross-contamination during DNA extraction.

Furthermore, acquisition of samples obtained from the second study introduced additional variation due to batch differences which had to be accounted for (using median batch spectral subtraction), prior to final analyses. The issue of low sample numbers in metabolic, cytokine and microbial data were circumvented for the biofluids through pooling of all time-points and of the pre-infection animals with the control groups, as no significant metabolic differences had been determined between them. Nonetheless, a repeat of this experiment with a larger number of animals would be required to validate the presented findings.

Another important consideration is that as these animals represented controls animals for experimental *L. major* infection, these mice received subcutaneous injections of medium into their tails on Day 0 of the study. Thus it may be possible that biological effects associated with the pain or stress of the needle, or responses towards the cell culture medium, may have occurred in these mice. However, according to the presented results, it appears that these effects were minimal as no significant differences were observed between the time-point before (D-2) and those after the injection (D2/D6/D12), in any of the assessed measurements.

3.5.6 Conclusions

Characterisation of the genetic background of control animals provides fundamental information, required to establish a reference point for comparison with groups subjected to experimental intervention. The presented research represented a proof-of-concept study for the tandem characterisation of metabolic, immune and microbial profiles of two mouse strains. For the purpose of this study, the investigation of biofluid and tissue metabolic profiles, circulating cytokine concentrations and faecal microbiome were successfully performed on C57BL/6 and BALB/c mice. Strain-associated dissimilarities were identified, which included increased turn-over of BCAA's, more dominant lipid

metabolism and a pro-inflammatory cytokine profile in BALB/c mice, whilst C57BL/6 had increased abundance of faecal *Erysipelotrichaceae*. Results obtained from this work have facilitated the extraction of differences in these three assessed biological measures (of metabolism, cytokine and microbiota) in response to experimental *L. major* infection between these two strains, that were independent of the inherent genetic strain-related variation (presented in the next Chapter).

Chapter 4

Results: Systemic Characterisation of Contrasting Murine Models of Cutaneous Leishmaniasis

4.1 Background 109
4.2 Objectives 111
4.3 Experimental Methods 111
4.4 Results 116
4.5 Discussion 139

4.1 Background

Parasitic infections such as *Leishmania* induce a cascade of physiological responses, including changes in metabolism and activation of the immune system. Infection with *L. major* protozoa causes cutaneous leishmaniasis (CL) in humans, a neglected tropical disease. There are severe limitations to the current management of CL, including suboptimal diagnosis and prognosis, and variable efficacy of drugs. Gaining a better understanding of the pathogenesis is fundamental in order to pursue new opportunities for more efficient diagnosis and treatment.

Studies have shown that infection with *L. major* for experimental CL models results in a slowly developing, self-healing form of disease in most laboratory mouse strains, including C57BL/6 mice, yet susceptible BALB/c mice develop a non-healing lesion, accompanied by internal dissemination of the parasite that eventually proves fatal.[31] Evidence indicates that these differential disease

outcomes depend largely on the extent of the immune-phenotype of these mice, with a dominant Th 1 response being pivotal in effective parasite elimination in C57BL/6 mice[37–39], in contrast to a dominant Th 2 response and disease progression in BALB/c mice[240–242] (refer to Introduction Section 1.2.4 for more detail). Furthermore, results from a murine *L. major* infection model has demonstrated that mice lacking natural gastrointestinal microbiota have increased pathological symptoms, despite generating appropriate an immune Th response to the parasite[54], adding to the complexity of leishmaniasis pathology.

Evidence from experimental infection models has shown that metabolic profiling provides a comprehensive approach for the detection of parasite-specific signatures and their effects on global host metabolism. Such studies have been typically based on ^{1}H NMR spectroscopy of host's biofluids. This approach has expanded successively to include tissues, and more recently, to combine metrics from both the metabolic and immune-systems measures, that were integrated in order to enable a more systemic characterisation view on of a host response to a pathogen and to characterise the immune-metabolic interface during infection.[243] Additionally, the importance of gut microbial involvement in parasitic infection has been confirmed by the discovery of significant changes of multiple microbial-associated metabolite concentrations in plasma and/or urine of numerous parasite rodent models.[91–94,243]

To date, metabolic profiling studies for *Leishmania* species have been limited to *in vitro* characterisation of either parasite promastigote cultures[244–247] or, more recently, of infected macrophages[44], the target host cells of *Leishmania*. This study presents, for the first time, metabolic characterisation of *Leishmania* infection *in vivo* and will firstly discuss the systemic metabolic changes associated with the early phase of *L. major* infection in BALB/c mice, the most severe and commonly studied model of experimental CL. Furthermore, in order to gain a better understanding of the multiple complexities of the differential host responses that have been observed in CL, *L. major* infection was additionally examined in C57BL/6 mice, a self-healing model of CL. The study focused on the investigation of the host response in the early phase of infection, which was terminated shortly before the bursting and inflammation/ulceration of the lesion for both mouse models, in an attempt to identify infection-specific responses as opposed to generic inflammatory responses.

In the presented research, differences in parasite-associated changes in host metabolism between the two mouse models were investigated by comparing the metabolic phenotypes of urine, plasma, faeces, and multiple tissues of the infected animals. Circulating cytokines concentrations were measured as read-outs of current host immune status, to reflect the allegedly opposing immune responses of these two mouse models. Furthermore, following the discovery of significant differences

in gut-microbiota associated metabolites in parasite infection models, microbial diversity analyses were also performed on faecal samples and co-analysed with the cytokine and metabolic data. Parasitic burden was assessed in selected tissues at the end of the study to confirm established infection in both models and to investigate potential parasitic dissemination in the host. Lastly, to determine differences between the two models for each biological measure, correlation network analysis was performed as a novel approach to integrate and co-analyse interactions between the different responses.

Results from this study presented in this chapter have since been published and a copy of the journal article can be viewed in the Appendix Publications Section A, entitled *"Metabolic, Immune, and Gut Microbial Signals Mount a Systems Response to Leishmania major Infection"*.[197]

4.2 Objectives

The primary aim of this study was to gain a better understanding of the interactions that lead to the complex pathology of CL, by using a combined set of response measures within a non-healing and self-healing model. This investigation included the first demonstration of an untargeted metabolic characterisation of *Leishmania* infection *in vivo*. The secondary aim was to identify host biological parameters that were associated with successful disease resolution, by taking the self-healing model as reference system, which may give rise to future alternative treatment strategies. These aims were addressed by applying the following specific objectives:

- To assess systemic metabolic changes in plasma, urine and faeces associated with *L. major* infection in non-healing BALB/c mice
- To characterise and compare *L. major* infection-related host responses of metabolic, cytokine and microbial systems that separate the non-healing mice from self-healing mice
- To investigate potential interactions between host metabolism, immune system and gut microbiota in response to *L. major* infection through correlation analyses, for each mouse model
- To identify biological parameters of the host that are linked with either disease resolution in self-healing mice or disease propagation in the non-healing strain

4.3 Experimental Methods

4.3.1 Parasite Culture Methods

Parasite cultures were prepared by Dr Jasmina Saric, at Imperial College London, in collaboration with Dr Karin Seifert, LSHTM. A stock from LSHTM of *L. major* promastigotes (strain

MHOM/SA/85/JISH118 at passage 1, suspended in a glycerol solution) was cultivated in 5 ml Schneider's insect medium (SIGMA-Aldrich, Germany), supplemented with 10% foetal bovine serum (FBS, SIGMA-Aldrich), in a small tissue culture flask. The culture was left to grow at log phase at 26°C and counted daily. On day 2, an aliquot of 2.32 x 10^7 parasites/ml was expanded 1:10 with 10 ml of Schneider's insect medium (with 10% FBS) in a new culture flask. The parasites were further incubated at 26°C and counted daily until parasite numbers remained stationary, by day 6, which represents the phase that contains the most infectious forms of promastigotes (metacyclics).[248] The final parasite count on day 6 was 3.14 x 10^8 promastigotes in 10 ml (passage 2). An aliquot containing 2 ml parasite suspension was taken and centrifuged at 560 x *g* (2,400 rpm) for 5 mins. The supernatant was discarded and the remaining pellet was resuspended in Schneider's insect medium (excluding FBS) to a final concentration of 2 x 10^7, that was used for *in vivo* infection. The final volume for infection inocula was 100 µl for each mouse, containing 2 x 10^6 stationary-phase *L. major* promastigotes.

4.3.2 Murine Leishmaniasis Study Overview

Experimental set-up of *L. major* infection models and sample collection are described in the previous Chapter, Section 3.3, which includes a summary diagram (Figure 3.1A). In brief, the study was comprised of four groups, namely *L. major* infected BALB/c mice (BT, $n = 10$), *L. major* infected C57BL/6 mice (CT, $n = 10$), control BALB/c (BC, $n = 5$), and control C57BL/6 (CC, $n = 5$). The day of lesion onset and lesion size in infected animals was recorded. Biofluids (plasma, urine and faeces) were collected on D-2, D2, D6 and the terminal day (Dterm), which for the infection groups was defined as the day the lesion had developed, shortly before ulceration occurred (7–13 days post-infection, depending on each infected animal).[1] On Dterm, animals were euthanised and multiple tissues were collected for metabolic analyses. Additionally, lesion and control skin samples were collected from infected and control mice on Dterm, respectively, and fixed in 10% formalin for subsequent analysis of parasitic burden. Metabolic characterisation of biofluids and tissues were performed using ^{1}H NMR spectroscopy, details of which have been outlined in the General Methods Chapter, Section 2.2. Plasma cytokines were assessed using Th 1/Th 2 multiplex Meso Scale Discovery technology and faecal microbial diversity was investigated using next-generation pyrosequencing, further described in Methods Chapter, Section 2.6.

[1] For control animals, day 12 post-infection was designated as the terminal day.

Parasitic Burden Assessment

Investigation of parasitic burden was performed with the assistance of Dr Emilie Giraud, under the supervision of Dr Matthew Rogers at the LSHTM. Steps included DNA extraction from relevant tissues, followed by real-time PCR, for DNA amplification and quantification.

Firstly, DNA was extracted from lesion, liver and spleen samples using DNeasy® Blood & Tissue Kit (QIAGEN, Germany), following manufacturer's instructions (spin-column protocol)[249] with modifications for formalin fixed samples, identified following several rounds of optimisations (similarly to that described for the DNA extraction faecal microbial analyses, General Methods, Section 2.6.2). Furthermore, lesion samples were rigid and required double the reagents for many of the steps compared to other tissues, based on recommendations from collaborators. Protocols for extraction were as follows: ∼25 mg of lesion and liver tissue, and ∼15 mg of spleen were placed into 2 ml microtubes (supplied with kit) and further cut into small pieces using tweezers and scissors. Samples were washed and vortexed twice for ∼10 seconds in 1.2 ml of PBS (spinning down at 9500 x *g* [10,000 rpm] for 5 mins each each vortex) to remove formalin. Supernatant was discarded and 360 µl (for lesions) or 180 µl (for liver and spleen) of ATL buffer was added, followed by 40 µl proteinase K. Samples were vortexed twice, each for 20 seconds at a time, and then left to incubate overnight at 56°C for enzymatic digestion. Samples were vortexed twice again. A volume of 800 µl (for lesion) or 400 µl (for liver/spleen) of 1:1 premixed ATL buffer:ethanol solution was added to each sample. Samples were vortexed for 20 seconds and transferred into spin-columns, placed in fresh collection tubes. Samples were centrifuged at 9500 x *g* (lesion samples were additionally spun down a second time) and flow-through was discarded. A volume of 500 µl AW1 buffer was added to all columns and samples were spun at 9500 x *g* for 1 min. Flow-though was discarded and samples were washed again with 500 µl AW2 buffer and centrifuged at 17,640 x *g* (13,400 rpm) for 3 mins. Flow-through was discarded and samples were centrifuged again at 17,640 x *g* for 3 mins, for maximal removal of wash buffers. Flow-through and collection tubes were subsequently discarded and columns were placed in fresh microtubes. DNA was eluted in 2 x 25 µl AE buffer, allowing the DNA to incubate with the buffer for ∼2-3 mins in the spin-column before spinning down at 9500 x *g*. DNA concentration and purity were assessed *via* NanoDrop® (Thermo Scientific) as previously described in General Methods, Section 2.6.2.

Real-time PCR was performed with triplicate samples using SYBR-Green kit (Applied Biosystems).[250,251] Multiple rounds of PCR were required for identification of optimal parameter settings, sample and primer volumes, sample dilutions, and selection of primers. The final primers (Guaranteed Oligo^{TM}, Sigma-Aldrich) that were selected targeted the mouse arginase gene (for detection of

genomic DNA) and Leishmania-specific RV1/RV2[252] to detect parasites (tested separately for each sample). Samples were tested in triplicate, with each PCR reaction performed in 10 µl on MicroAmp® Optical 384-Well Reaction PCR Plates (Applied Biosystems). Each PCR mix consisted of 3.5 µl H_2O, 0.25 µl forward primer, 0.25 µl reverse primer, 5 µl SYBR-Green reagent and 1 µl DNA sample. The final PCR program for sample acquisition included 45 cycles of [10 seconds denaturation at 95°C , 25 seconds annealing at 55°C, and 60 seconds extension at 72°C]. Fluorescence of SYBR Green was measured at end of the elongation step, visualised in amplification plots by the Applied Biosystems SDS2.4 software which automatically generated Cycle threshold values (Ct) for each well.[250] This PCR program was followed by an additional melting curve program with continuous fluorescence measurement from 70°C to 95°C (ramping rate was set to 0.1°C/second). Data from this step was visualised on temperature dissociation curves, which allowed the identification of non-specific binding.[250][2] Ct values were imported into Qbase Plus 2.4 software (Biogazelle, Belgium), an open source program for qPCR data management and analysis.[253] Parasitic load for each sample was assigned as average fluorescent intensity of parasitic DNA, calculated as a multiple of the lowest positive reading (lesion sample CC7), and normalised to arginase levels. Calculations were performed using Qbase Plus 2.4 software (Biogazelle, Belgium), which includes the Pfaffl correction to account for inequalities in primer efficiencies.[254] Final output of relative fluorescent intensity for parasitic DNA was plotted in GraphPad PRISM and levels were compared between strains using Mann–Whitney U-Tests.[186]

4.3.3 Data Acquisition and Analyses

Samples were prepared, acquired and analysed for metabolic characterisation, cytokine measurements and/or faecal microbial diversity analyses, as described previously in Methods Chapter 2. Spectral data were analysed separately for each biofluid/tissue, at each time-point, using PCA and O-PLS-DA models. Metabolic differences between time-point-matched *L. major* infected BALB/c and control BALB/c were investigated, to determine metabolic changes induced by the parasite. In order to identify disparities in infection-related responses between two mouse models, as opposed to strain-specific variation (as has been highlighted previously in Chapter 3), spectra were adjusted to their respective control strain, by subtracting the group median spectra level of the controls from infected spectra. The new strain-adjusted metabolic data were further analysed using PCA and O-PLS-DA, to determine differences in infection-associated metabolic responses between the two CL models, and used for all presented correlation analyses.

[2] Final selection of primers was based on the maximal detection of parasitic DNA (with lowest Ct values on the amplification plots), combined with the least about of non-specific binding.

Changes in infection-related cytokine concentrations, and in relative percentage abundance of faecal bacterial classes, were assessed in BALB/c mice, by comparing control levels with those of infected, using Kruskal–Wallis tests and Mann–Whitney U-test, followed by Bonferroni multiple testing correction. Similar to metabolic data, cytokine and microbial data were strain-adjusted through subtraction of control group median levels from infected samples and differences in infection-related cytokine and microbial responses between the two models were tested with further Kruskal–Wallis and Mann–Whitney U-tests (with multiple test correction). Baseline plasma samples from the control groups were lacking thus no control-strain adjustment could be performed for plasma metabolic and cytokine data at this time-point, required for spectral O-PLS-DA and cytokine univariate analyses, respectively.

Differences in parasitic burden between the two infected strains, as measured by relative fluorescence of parasitic DNA, were tested separately for lesion, liver and spleen data, using Mann–Whitney U-tests. As results from the lesion represented the most complete and robust dataset, lesion data were used for subsequent correlation analyses.

Two-dimensional (2D) matrix Pearson-based correlation analyses were used to determine associations between two (strain-adjusted) data sets, namely between ^{1}H NMR spectral data and cytokine concentrations, between spectral data and class percentage abundance of faecal microbiota, and cytokine concentrations and faecal microbial class percentage, which were performed separately for each time-point and for each CL model.

Subsequently, 12 metabolites were chosen each from urine, plasma, and faeces, based on their ranked discriminatory role between the two models (from O-PLS-DA) and number of correlations with host cytokines and/or microbiome (from 2D matrix analyses), and their integrals were calculated. Metabolite integrals, time-matched cytokine concentrations and faecal microbial class % abundance (all strain-adjusted) were co-analysed using PCA and PLS-DA, to assess the relative contributions of the different datasets (comprised of 48 variables in total) in driving the separation between the two infection models. Parasitic burden data were correlated with each of the 48 variables from the Dterm time-point (most closely matching the time the lesion/skin samples were collected) using Pearson based analyses for each strain, in order to investigate direct associations between *L. major* infection and host measures.

Furthermore, Pearson-based correlation network analyses were performed on the 48 variables, for each time-point in both strains individually, to determine potential interactions in the host between variables, in response to infection. Correlation networks were generated in-house by Joram Posma, using a MATLAB code based on MetaboNetworks software[196], which incorporated Storey–

Tibshirani False Discovery Rate[192] for multiple test correction. A generous FDR cut-off of a q-value (i.e. FDR corrected p-value) < 0.3 (i.e. 30% expected false positives) was applied for both parasite burden–host measure correlation analyses and correlation network analyses, for maximal exploration of host interactions with potential biological relevance.

4.4 Results

4.4.1 CL Model Physiological and Pathological Measures

A summary of animal weight, PCV, lesion onset and lesion size can be found in Appendix Table A.4. No significant differences in weight were observed between the infected strains, or between time-points per strain, similar to previous findings in control mice. PCV was significantly higher in infected BALB/c than infected C57BL/6 ($p = 0.001$ for all time-points), also consistent with findings in the controls. Furthermore, the infected BALB/c mice had significantly increased PCV when compared with their controls on D6 ($p = 0.044$). Whilst average day of lesion onset were similar between the two strains, the size of the lesions in non-healing BALB/c mice generally appeared larger than in infected self-healing C57BL/6 mice, though this finding was not statistically significant.

Trends Showed Higher Parasitic Burden in Non-healing Model

Parasitic-specific DNA was detected in all three analysed tissues, yet was found to be most prevalent in the lesion itself, as shown in Figure 4.1. Spleen had the least amount of samples that tested positive for parasitic DNA, particularly in BALB/c mice. Results showed apparent higher levels of parasites in the lesions and livers of BALB/c mice, when compared with C57BL/6 mice, yet these were not found to be significant. A clear difference was observed between background fluorescence in the skin of control mice, and lesion samples of infected mice; 3 control liver samples from C57BL/6 mice displayed unusually high background DNA levels. Consequently, parasitic burden results from lesion samples were selected for subsequent analyses.

4.4.2 Effect of L. major Infection on Host Metabolism in Non-healing CL Model

Systemic changes in host metabolism associated with *L. major* infection were investigated *via* PCA and O-PLS-DA models in the plasma, urine and faeces of BALB/c mice, which display the most severe form of disease. Representative figures of these models are shown in Figure 4.2. Separation between infected mice and control animals were not apparent in the first component of PCA

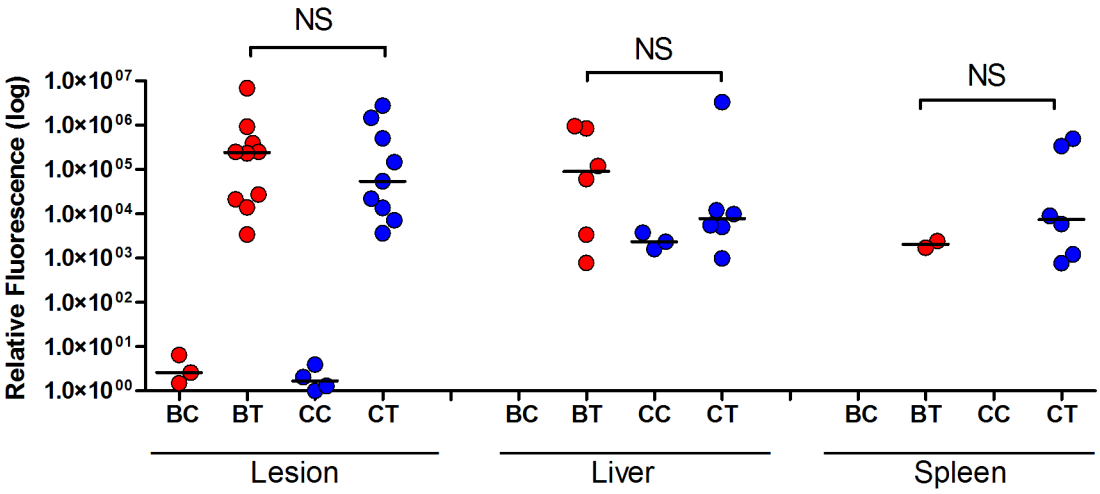


Figure 4.1: Parasitic Levels Measured in the Lesion, Liver Spleen of CL Models

Plots show relative fluorescence of *L. major* DNA, as measured *via* SYBR-Green PCR, in the lesion, liver and spleen samples. Samples from BALB/c mice and C57BL/6 mice are shown as red and blue circles, respectively, with group median levels shown as horizontal bars. Levels in control groups (CC/BC) represent background levels (data normalised to minimum value, which was determined as lesion sample CC7). Samples where no parasite DNA were detected were omitted from the graph. No significant differences (NS) were determined between BT and CT values. Key: BC, control BALB/c (BC); BT, infected BALB/c; CC, control C57BL/6; CT, infected C57BL/6.

models, which had low Q^2X values (e.g. 0.061 for plasma), indicating a low predictive capability of the model component, due to high variability within the dataset. Yet differences between infection status of mice were determined along the first predictive component of O-PLS-DA models (showing controls in the top half of each plot and the majority of infected on the lower half).

Significant differences in metabolite concentrations were observed in plasma (at all three post-infection time-points) and urine (on D2 and D6), yet no changes were observed in faecal extracts. A list of discriminatory metabolites is summarised in Table 4.1. Plasma levels of 3-hydroxybutyrate were increased on D2 and D6 following infection, whilst plasma glucose and urinary hippurate decreased for both of those time-points. Other metabolites that were only altered at one of the time-points in plasma included reductions in several lipid moieties on D2, and decreases in citrate and glycerophosphocholine on D6. The vast majority of the changes in urine were observed on D6, which included increases in succinate and indoxylsulphate, and decreases in creatine, taurine, trimethylamine-*N*-oxide. Few significant changes were determined on Dterm for plasma (increased glutamine to glutamate turnover) and none were observed in urinary profiles.

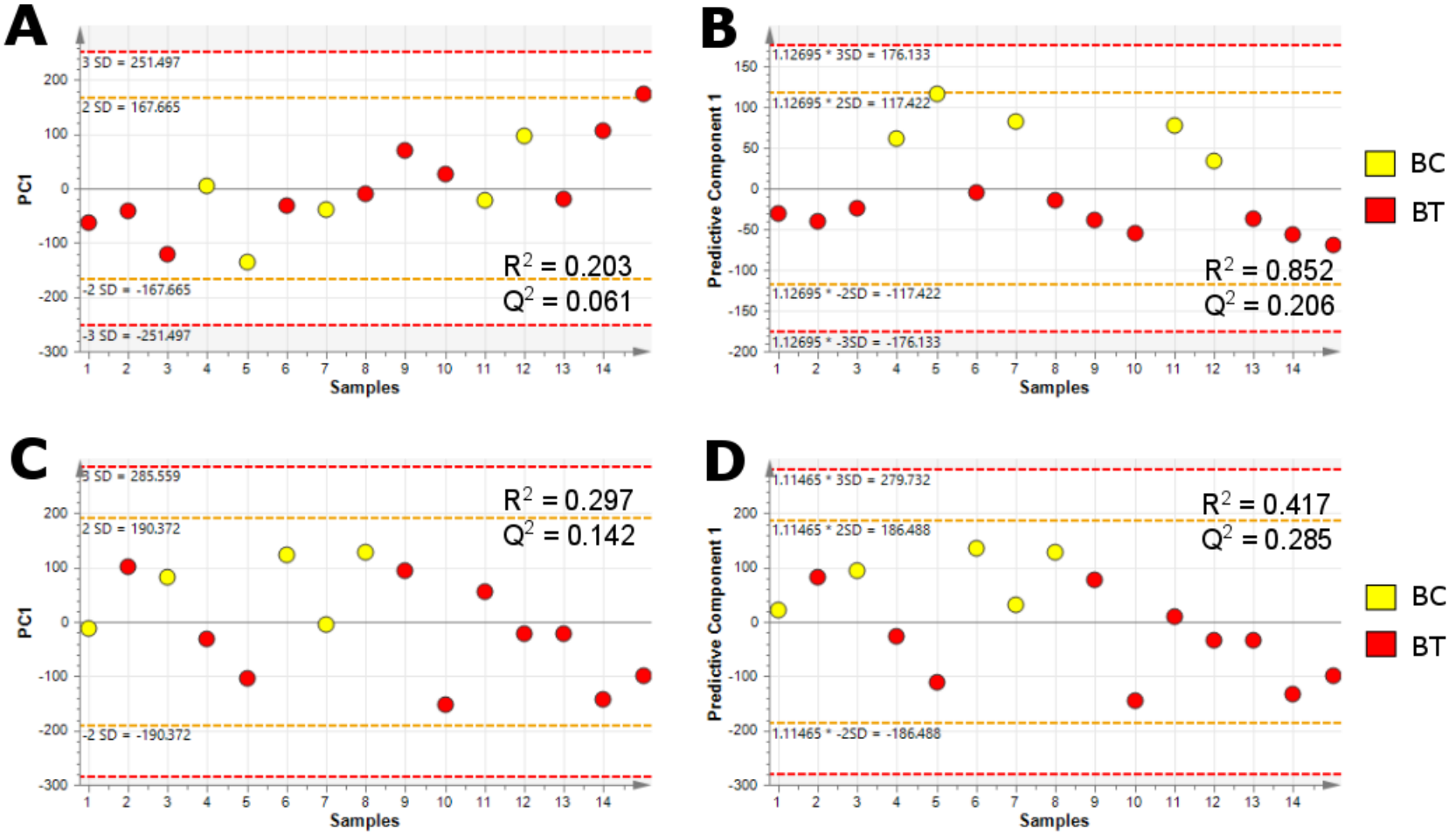


Figure 4.2: Metabolic Variation Linked with *L. major* Infection in Plasma and Urine of BALB/c Mice

PCA and O-PLS-DA scores plots of plasma (top row) and urine (bottom row), comparing infected BALB/c mice (BT) with controls (BC) on D6. Models are based using the first component. (**A**) PCA model of plasma. (**B**) O-PLS-DA model of plasma. (**C**) PCA model of urine. (**D**) O-PLS-DA model of urine. Horizontal hashed lines represent two and three standard deviations (SD) from the mean.

Table 4.1: Metabolic Changes Associated with *L. major* Infection in the Non-Healing Model

Biofluid	Metabolites	Time-point D2	D6	Dterm
Plasma	3-Hydroxybutyrate[a]	+	+	
	Acetate[b]	+		
	Alanine		–	
	Citrate		–	
	Glucose[a]	–	–	
	Glutamine			–
	Glutamate			+
	Glycerophosphocholine		–	
	Lactate[b]	–	+	
	Lipid CH_2	–		
	Lipid CH_2CH_2CO	–		
	Lipid CH_2CO	–		
Urine	2-oxoisocaproate	+	–	
	cis-Aconitate		+	
	Creatine		–	
	Hippurate	–	–	
	Indoxylsulphate		+	
	Lactate		+	
	Phenylacetylglycine	+		
	Succinate		+	
	Taurine		–	
	TMAO		–	

Table shows plasma and urinary metabolites that significantly differed between *L. major* infected BALB/c mice and their time-point matched controls, as tested by O-PLS-DA. Results are based on ^{1}H NMR spectra from standard 1D pulse program, with additional discriminatory metabolites from CPMG pulse programs for plasma. Key: +, increased in infection; –, decreased in infection; D2, day 2 post-infection; D6, day 6, post-infection; Dterm, terminal day of study; TMAO, trimethylamine-*N*-oxide. [a] metabolite discriminatory in both standard 1D and CPMG programmes; [b] metabolite only discriminatory in CPMG data

4.4.3 Strain Adjustment of Data to Determine Infection-related Responses

Strong biological variation linked to differences in murine background strain were observed across metabolic, cytokine and faecal microbial datasets, as has been discussed in the previous Chapter. PCA models of all four groups (BT, BC, CT and CC) for the vast majority of different biofluid and tissue metabolic phenotypes revealed that strain-related differences were more dominant than infection-related changes within each strain (a representative model of plasma is shown in Appendix Figure A.3).

This was followed by further comparisons of metabolic phenotypes between either two control strains, or between the two infected strains *via* O-PLS-DA. The original strategy was to exclude any discriminatory metabolites that were common to both comparisons, from the results from the infected animals, thus leaving only infection-related metabolites. However, this task proved difficult to execute as several metabolites were correlated in the controls in an opposite direction to results observed in the infected strains, confounding final analyses. Additionally, any metabolites that

only differentiated between the two control animal strains (but not between infected strains) would not have been incorporated for final interpretation. Lastly, some of the metabolites that were discriminatory between the strains in controls also differentiated between the infected strains (e.g. plasma glucose and lipids), yet to a much greater extent in the latter, and thus this information would also have been discarded, if following the initial strategy.

Consequently, in order to account for strain-related variation and better extract infection-associated differences between the two mouse models, data were adjusted for strain by subtracting median values of the control groups from their respective infected groups, for each strain (similar to the batch adjustment that has been discussed in the previous chapter, Section 3.4.2). This adjustment was performed for metabolic, cytokine and faecal microbial datasets, conducted separately for each time-point.

4.4.4 Comparison of Individual Infection-Associated Host Responses between Non-healing and Self-healing CL Models

Differences in Infection-related Metabolic Phenotypes between Two Models

Disparities in infection-associated metabolic phenotypes between BALB/c and C57BL/6 mice were observed across biofluids and tissues, summarised in Tables 4.2 and 4.3, respectively. Metabolites that were positively correlated with the non-healing BALB/c strain included plasma glucose and 3-hydroxybutyrate (3HB), urinary phenylacetylglycine (PAG) and ureidopropanoate, and trimethylamine in the faeces and urine. Additionally, level of alanine (in plasma, urine and multiple tissues) and succinate (in urine and faeces) were also found to be higher in infected BALB/c mice than C57BL/6. In contrast, markers that were positively correlated with C57BL/6 included several fatty acid moieties from lipoproteins in plasma and tissues (primarily in ileum extracts), glycogen (in the liver), glycerophosphocholine in urine and plasma, urinary hippurate, and plasma *scyllo*-inositol. Please refer to Appendix Tables A.5 and A.6 for a breakdown per time-point for biofluids, and for each individual tissue, respectively.

Differences in Infection-related Cytokine Responses between Two Models

As had been outlined in the previous chapter for results from the control mice, only IL-1β, IL-10, IL-12 and CXCL1 were measured at detectable levels in plasma. Strain-adjusted concentrations of these cytokines for both infected strains are shown in Figure 4.3. Following strain-adjustment, infected BALB/c mice had higher concentrations of pro-inflammatory cytokines CXCL1 and IL-1β at all time-points, reaching significance for IL-1β at D6. BALB/c mice initially had higher levels

Table 4.2: Discriminatory Infection-Associated Metabolites in Biofluids

	Discriminatory Metabolites in Biofluids		
Biofluid	Positively correlated with BALB/c	Positively correlated with C57BL/6	Strain association varied between time-points
Plasma	3HB Alanine Citrate Creatine Glucose Glutamine Pyruvate Valine	Glycerophosphocholine Glycine Lipid CH_2CH_2CO Lipid $CH_2C{=}C$ Lipid CH_2CO *scyllo*-Inositol*	Acetate Lactate Lipid CH_3 Lipid CH_2
Urine	2-Oxoglutarate Creatinine Dimethylamine Fumarate Lipid CH_2 Phenylacetylglycine Succinate Trimethylamine Ureidopropanoate UK 1.88(dd)	2-Oxoisocaproate 3-Hydroxyisovalerate* Acetate Creatine Dimethylglycine Hippurate Glycerophosphocholine Taurine	2-Oxoadipate* 2-Oxovalerate* 3-Methyl-2-Oxovalerate* cis-Aconitate Citrate Indoxylsulphate Lactate
Faeces	Alanine Glycine Lactate Lipid CH3 Succinate Trimethylamine TMAO Tyramine UK 7.87(s)	Glutamate	Acetate Bile Acid Butyrate Cholate* Glycocholate* Tyrosine

Metabolites that were significantly discriminatory between the two CL models, for at least one time-point per biofluids, from infected mice (strain-adjusted), as measured *via* O-PLS-DA. Metabolites were classified according to the infected strain that they were predominantly positively correlated with out of the assessed time-points, whereas those that displayed alternating correlation patterns classified separately. See Appendix Table A.5 for a full breakdown of results for each biofluid, per time-point. Key: *, tentative assignments; (dd), doublet of doublets; (s), singlet; 3HB, 3-hydroxybutyrate; TMAO, trimethylamine-*N*-oxide; UK, unknown metabolites.
Modified with permission from Lamour et al.(2014)[197] *Copyright 2014 American Chemical Society.*

of IL-12 than C57BL/6 yet these appeared to reduce by the terminal time-point. In contrast, IL-12 was significantly increased in the healing C57BL/6 strain by the end of the study. Whilst no significant disparities in anti-inflammatory cytokine IL-10 were observed between the two models, D2 concentrations in BALB/c tended to be higher than that of C57BL/6 mice.

Differences in Infection-related Faecal Microbial Composition between Two Models

Infection-associated changes in host gastrointestinal microbial composition were explored by analysing the strain-adjusted relative percentage abundance of eight bacterial classes measured in faecal samples, which representing on average 99.1% of all identified classes across the dataset.[3] *Clostridia* and *Gammaproteobacteria* were the most dominant classes of faecal bacteria in this

[3]Raw abundance data from Research and Testing Laboratory used for analyses.

Table 4.3: Discriminatory Infection-Associated Metabolites in Tissues

Discriminatory Metabolites across Tissues		
Positively correlated with BALB/c	Positively correlated with C57BL/6	Strain association varied between tissues
Adenosine*	Glycogen	3-Hydroxybutyrate
ADP	Lipid CH_2CH_2CO	4-Aminobutyrate
Alanine	Lipid $CH_2C{=}C$	Acetate
ATP*	Lipid CH_2CO	AMP
Betaine	Lipid $C{=}CCH_2C{=}C$	Aspartate
Choline	NAD^+*	Creatine
Creatinine	Valine	Glutamine
Dimethylamine		Glutathione
Ethanolamine		Glycerophosphocholine
Fumarate		Glycogen
Glutamate		Lipid CH_2
Glycerol		Lipid CH_3
Glycine		*N*-aceytlaspartate
Histamine		Phosphocholine
Hypoxanthine		Phosphoethanolamine
Inosine		Succinate
Lactate		UK 7.68(s)
myo-Inositol		
Nicotinamide		
Taurine		
UDP-glucose		
UK 3.48(d)		

Metabolites that were significantly discriminatory between the two CL models, for at least one tissue, from infected mice (strain-adjusted), as measured *via* O-PLS-DA. Metabolites were classified according to the infected strain that they were predominantly positively correlated with out of the assessed tissues, whereas those that displayed alternating correlation patterns classified separately. See Appendix Table A.6 for a full breakdown of results for each tissue. Key as for Table 4.2, with following additions: ADP, adenosine diphosphate; AMP, adenosine monophosphate; ATP, adenosine triphosphate; NAD^+, nicotinamide adenine dinucleotide (oxidised ion); UDP, uridine diphospho.

study, accounting for over 85% relative abundance of the total data. Bacterial levels of *Clostridia* class (comprised largely of species belonging to *Clostridiaceae* and *Lachnospiraceae* families) were consistently higher in the non-healing strain than self-healing mice, which reached significance on D2 and Dterm. In both strains, levels of *Gammaproteobacteria* (predominantly members of the *Enterobacteriaceae* family) appeared lower on Dterm than the pre-infection baseline, though still significantly higher in C57BL/6 than BALB/c, as shown in Figure 4.4.

The third most abundant bacterial class detected in this study was *Erysipelotrichia* (mainly *Turicibacter* genus), which was largely absent in the faeces of BALB/c yet detected in considerable proportions in that of C57BL/6 mice (up to 60% abundance in some animals). Levels of *Erysipelotrichia* on D6 and Dterm appeared higher than D-2 and D2 in C57BL/6 mice, and were significantly higher than that of BALB/c mice on D6. Relative abundance of *Bacilli* were significantly decreased by Dterm in the self-healing model and were also significantly lower than BALB/c at this time-point (Figure 4.4).

The remaining classes were less prevalent, detected at relatively low levels (typically less than

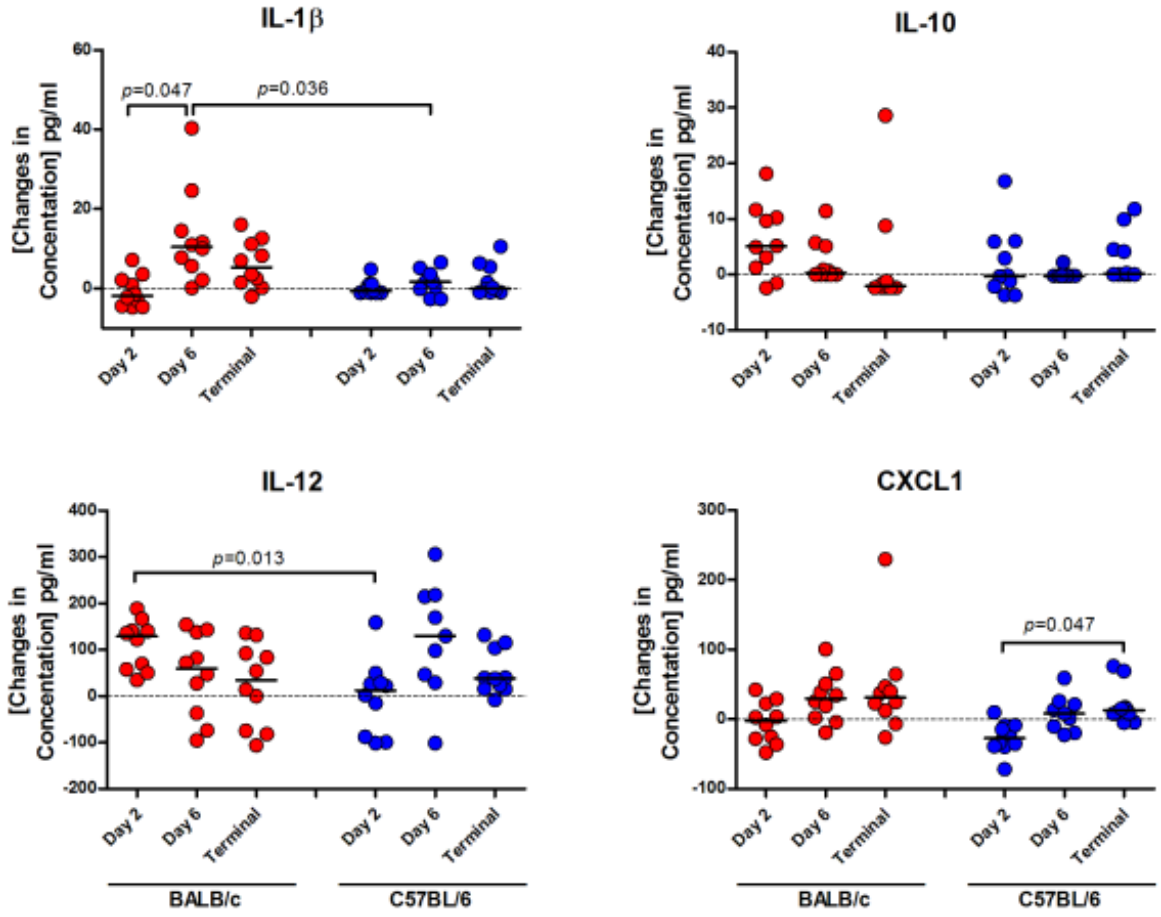


Figure 4.3: Elevated Pro-inflammatory Cytokines in Plasma of BALB/c Mice

Plot shows strain-adjusted, plasma cytokine concentrations with the group median shown as horizontal line and each sample represented as one circle. Significant differences between models and/or time-points have been highlighted. *Modified with permission from Lamour et al.(2014)*[197] *Copyright 2014 American Chemical Society.*

5%), and are shown in Figure 4.5. Results showed that levels of *Alphaproteobacteria* and *Bacteroidia* were significantly increased relative to baseline, in C57BL/6 mice on D2 and Dterm, respectively. No significant changes were observed within the non-healing BALB/c model.

In addition to model and time-associated changes, variation due to cage housing conditions were observed, particularly evident with data from *Gammaproteobacteria* and *Clostridia*, as highlighted in Appendix Figure A.4. For these two classes, the baseline levels differed greatly between the cages of infected animals (e.g. cage 1 with animals 1–5 vs. cage 2 housing animals 6–10), yet partially converged by the final time-point, a feature that was noted in both models. Cage-associated data variation was not observed in either the metabolic or cytokine data.

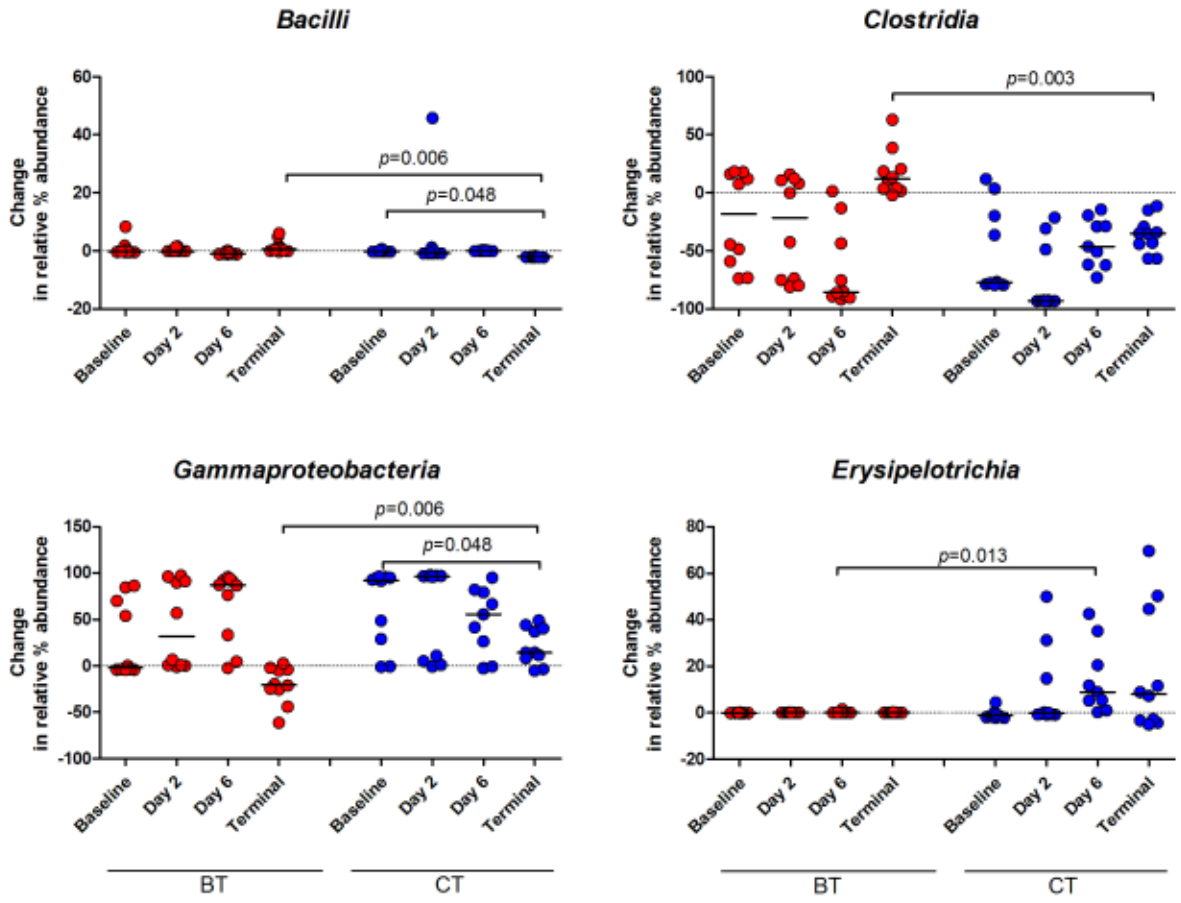


Figure 4.4: Differences in Dominant Faecal Bacterial Composition between CL Models

Plot shows infection-related changes in relative abundance of microbial classes detected in the faeces of *L. major* infected BALB/c (BT) and C57BL/6 (CT) mice, with group medians highlighted as horizontal lines. Significant differences between models and/or time-points have been highlighted. *Modified with permission from Lamour et al.(2014)*[197] *Copyright 2014 American Chemical Society.*

4.4.5 Interactions between Two Sets of Infection-Related Host Responses via 2D Matrix Correlation Analyses

Correlations between Cytokine and Metabolic Spectra

2D matrix correlation analyses between cytokine concentrations and metabolic spectra demonstrated multiple interactions between host immunity and metabolic markers, in response to *L. major* infection, observed in both CL models. A representative example of co-analyses of cytokines and D2 urine spectra for each infected strain is shown in Figure 4.6, which shows overall disparate correlation patterns between the two models. Lists of all significant correlative immune-metabolic pairs for biofluid and tissue data are summarised in Tables 4.4 and 4.5, respectively. A few correlates were common to both BALB/c and C57BL/6, such as a positive association between IL-12 and faecal propionate, and a negative correlation between IL-10 and urinary hippurate. IL-1β displayed

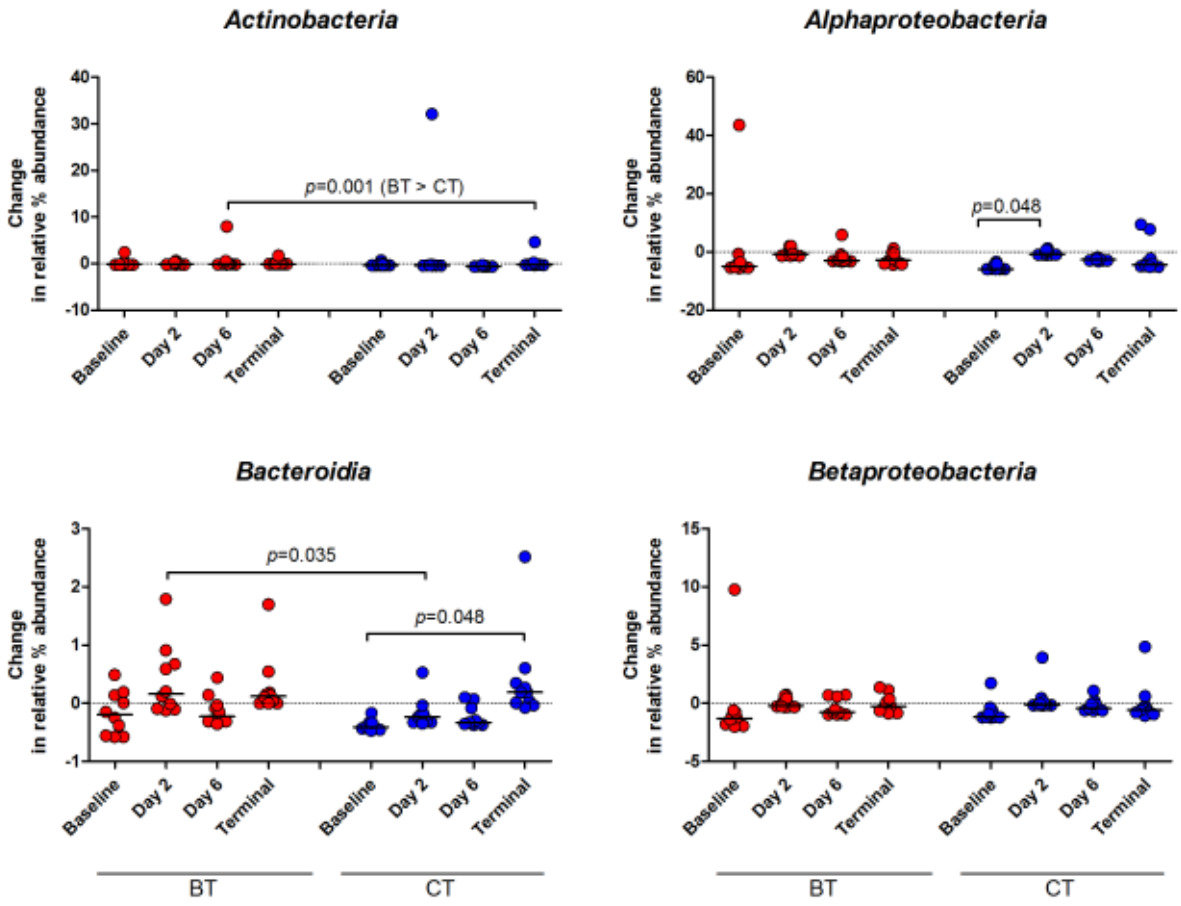


Figure 4.5: Changes in Relative Abundance of Less Prevalent Faecal Microbiota Observed in Self-Healing Model

positive correlation patterns with a set of tissue metabolites that were observed in both models, namely alanine, choline, lactate, *myo*-inositol, and taurine. Others correlative pairs were model-specific, such as urinary ureidopropanoate, which was negatively correlated with IL-12, IL-1β and CXCL1 in non-healing mice yet produced no correlations in the self-healing strain.

On numerous occasions, metabolites that were correlated with cytokines IL-1β, CXCL1 and IL-12 displayed opposing correlative associations with IL-10, if at all, indicating a partial counter-regulatory effect of IL-10 towards the other cytokines on the host metabolism. Such disparate correlative patterns included lactate, PAG and 2-oxoisocaproate (2-OIC) in urine (Table 4.4), and betaine, *myo*-inositol, lactate, dimethylamine, glycine, phosphocholine and taurine in tissues (Table 4.5). On the contrary, CXCL1 and IL-1β shared several metabolic correlative associations across both models, especially evident in tissues of infected BALB/c mice.

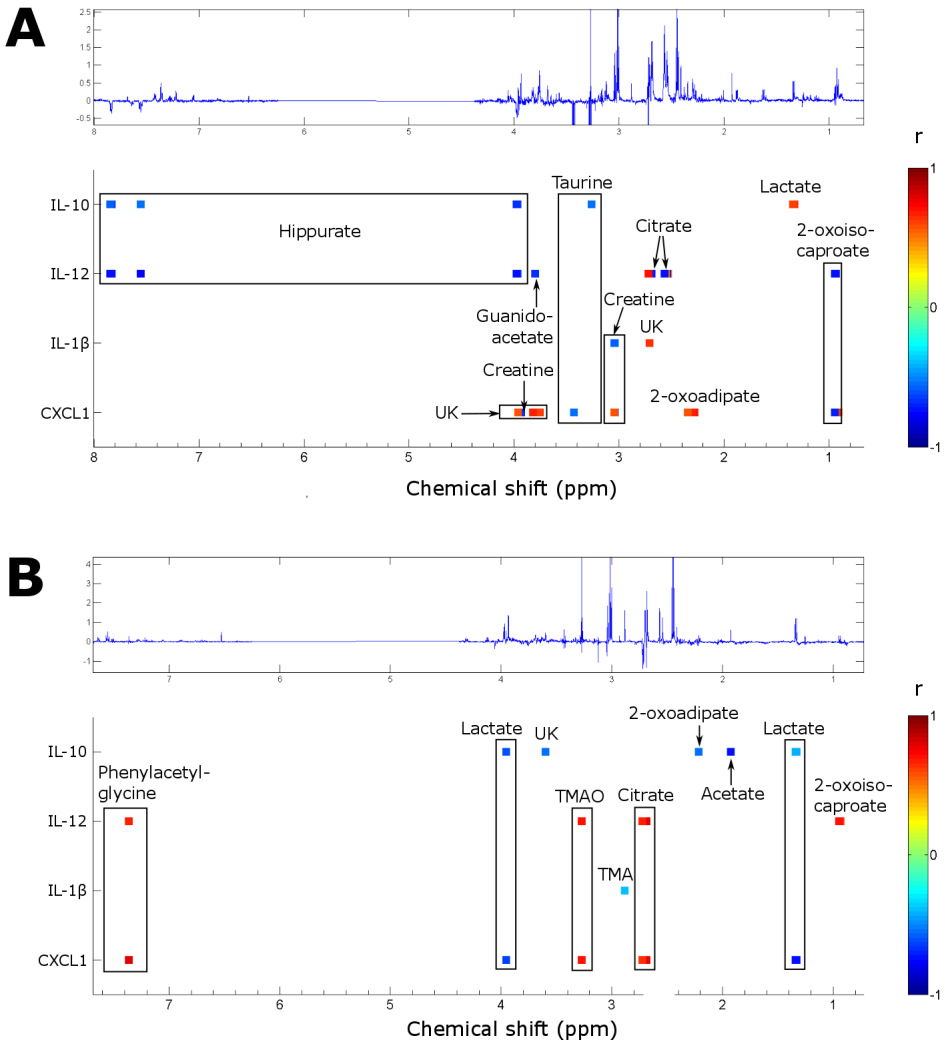


Figure 4.6: Variation in Immune-Metabolic Correlation Patterns Associated with Infection in two CL Models

Plot shows 2D matrix correlation plots between plasma cytokines concentrations and ^{1}H NMR urine spectra from D2 for *L. major* infected BALB/c mice (**A**) and C57BL/6 mice (**B**). Inset along the top of each plot shows average urine spectrum. Statistically significant correlations ($p < 0.05$) between cytokine concentrations and metabolite peak intensities are shown as coloured squares: positive correlations in red/orange and negative correlations in blue. Key: r, Pearson correlation coefficient; TMA, trimethylamine; TMAO, trimethylamine-*N*-oxide; UK, unknown metabolite.

Table 4.4: Immune–Metabolic Correlations in Plasma, Urine and Faeces of CL Models

Biofluid	Metabolic Correlate	Infected BALB/c IL-10	IL-12	IL-1β	CXCL1	Infected C57BL/6 IL-10	IL-12	IL-1β	CXCL1
Plasma	3HB								+
	Choline		+						
	Glucose	+	+		–	+/–		+	+/–
	Glycine					–			
	GPC		–						
	Lactate	+/–	–	–				+	
	Lipid CH_2		–			+			+
	Lipid CH_3		–			+			+
	scyllo-Inositol*			–					
Urine	2-Oxoadipate					–		–	
	2-Oxoglutarate							–	
	2-Oxoisocaporate	+	–	–	–		+		
	Acetate				–	–			
	cis-Aconitate			–					
	Citrate		–				+		+
	Creatine				–	+/–			+
	Creatinine			–					
	Fumarate				–				
	Guanidoacetate		–		+				
	Hippurate	–	–			–			
	Lactate	+	–	–		–		–	
	PAG		+		+/–		+	+	+
	Succinate				–			–	
	Taurine	–	–		+/–	+/–	–	+	+
	Trimethylamine							–	–
	TMAO					+	+		+/–
	Ureidopropanoate		–	–	–				
Faeces	Butyrate		–		–	+			
	Glutamate						–		
	Lactate						+		+/–
	Lipid CH_3						+		
	Propionate		+				+		+
	Succinate							–	+
	Taurine						+		

Significant correlations between metabolites' spectral peak intensities and plasma cytokine concentrations, using 2D matrix correlation analyses ($p<0.05$). Unassigned metabolic-correlates have been omitted. Key: +, positive correlation for one time-point; –, negative correlation for one time-point; +/–, direction of correlation alternated between different time-points; *, tentative assignment; 3HB, 3-hydroxybutyrate; GPC, glycerophoshocholine; PAG, phenylacetylglycine; TMAO; trimethylamine-*N*-oxide. *Reprinted with permission from Lamour, et al. (2014)*[197] *Copyright 2014 American Chemical Society.*

Table 4.5: Immune–Metabolic Correlations in Tissues of CL Models

Metabolic Correlate	Infected BALB/c				Infected C57BL/6			
	IL-10	IL-12	IL-1β	CXCL1	IL-10	IL-12	IL-1β	CXCL1
4-Aminobutyrate		–						
ADP			+					
Alanine	+/–		+	+/–	–	+/–	+	
AMP							+	–
Aspartate	–					+	+	–
Betaine	+	–			–	+	+	
Choline	+/–	+	+	+/–	–	+/–	+	
Creatine	+/–		+	+	–	+		+
Dimethylamine	+				–	+	+	
Glucose			+	+				
Glutamate	+			–	+/–	+	+	+/–
Glutamine	–				–			
Glutathione	+			–		+	+	
GPC	+/–	–			–	+	+/–	–
Glycine	+/–	–		–	–	+	+	
Glycogen			+					
Inosine	+							
Lactate	+	–	+	+	–	+/–	+	+
Lipid/Sugar CH=CH	+							
Lipid C=CCH_2C=C	+							
Lipid CH_2C=C	+				–			
Lipid CH_2	+			+	+	–		
Lipid CH_2CH_2CO	+							
Lipid CH_2CO	+							
Lipid CH_3	+					–	–	
myo-Inositol	+		+	+	–	+	+	
N-Acetylaspartate			+	+	+/–		–	+
Phosphocholine	+/–	+			–	+	+	
Phosphoethanolamine	+/–					+	+	–
Succinate		+/–			+		+	
Taurine	+/–	–	+	+/–	–	+	+	

Significant correlations between metabolites' spectral peak intensities and plasma cytokine concentrations, using 2D matrix correlation analyses ($p<0.05$). Results have been summarised from immune-metabolic correlations performed separately for each of the CL-model-discriminatory tissues, namely liver, spleen, ileum, thymus, inguinal lymph nodes, heart and four brain sections (frontal cortex, cerebellum, brain stem, and hippocampus). Unassigned metabolic-correlates have been omitted. Key: +, positive correlation for one tissue; –, negative correlation for one tissue; +/–, direction of correlation alternated between different tissues; *, tentative assignment; ADP, adenosine diphosphate; AMP, adenosine monophosphate; GPC, glycerophoshocholine.

Correlations between Faecal Microbial Abundance and Metabolic Spectra

Following on from the immune-metabolic analyses, similar 2D matrix correlations were performed between spectral NMR data and percentage abundance of faecal microbial classes, to investigate microbial-metabolic interactions that associated with infection for each CL model. Microbial-metabolic correlates have been summarised in Table 4.6 for plasma and faeces, and Table 4.7 for urine.

Analyses show that *Clostridia* and *Gammaproteobacteria* classes displayed opposing correlative patterns with metabolites across both infected strains, i.e. the majority of metabolites that were positively linked with one, were negatively linked with the other, and vice versa. Examples of these metabolic correlates that were common to both models included choline, glycerophosphocholine

Table 4.6: Microbial–Metabolic Correlations in Plasma and Faeces of CL Models

	Metabolic Correlate	Clos	Gamma	Erysi	Alpha	Beta	Bacil	Actino	Bacter
BALB/c Plasma	3HB			+	+			+	
	Choline	+	–	+	+			+	+
	Glucose	–		+	– –	–	+	–	+
	GPC	–	+	+	–	–			–
	Lactate				+/–		–	–	
	Lipid CH_2		+	+	+/–	–		+	
	Lipid CH_3			+	+/–	–		+	
	scyllo-Inositol*				–	+/–		+	–
BALB/c Faeces	Acetate			–					
	Butyrate	–	+	–		+	+		–
	Lactate						–		
	Propionate		+	–			–		
	Succinate						–		
C57BL/6 Plasma	3HB	–	+	–	–				
	Choline	–	+						
	Glucose			–	–			– –	+ +
	GPC	–	+						
	Lactate	–						–	
	Lipid CH_2	–	+	–					
	Lipid CH_3	–		–					
C57BL/6 Faeces	Acetate					+	+	–	
	Butyrate							–	
	Dimethylamine			–					
	Glutamate			–			–		
	Lactate	–	+	–	–				+
	Propionate	–	+					–	
	Succinate	–	+	–	–				

Significant correlations between plasma and faecal metabolites' spectral peak intensities and relative percentage abundance of faecal microbial classes, using 2D matrix correlation analyses ($p<0.05$). Unassigned metabolic-correlates have been omitted. Key as described for Table 4.4, with the following additions: + +, positive correlations for at least two time-points; – –; negative correlation for at least two time-points; Actino, *Actinobacteria*; Alpha, *Alphaproteobacteria*; Bacil, *Bacilli*; Bacter, *Bacteroidia*; Beta, *Betaproteobacteria*; Clos, *Clostridia*; Erysi, *Erysipelotrichia*; Gamma, *Gammaproteobacteria*. *Reprinted with permission from Lamour, et al. (2014)*[197] *Copyright 2014 American Chemical Society.*

(GPC), and lipid methylene (CH_2) moiety in plasma, faecal propionate and urinary creatinine, with many more correlates that were model-specific. Additionally, *Erysipelotrichia* exhibited similar correlations with many of the host urinary and faecal metabolites as *Clostridia* in both models, as well as with plasma metabolites in C57BL/6 mice.

Plasma glucose and urinary creatinine displayed the highest number of correlations with microbial classes collectively across both strains, highlighting the link between the host microbiome and energy metabolism. Unlike glucose, clear disparities between the models were observed in the correlation patterns with creatinine, which was found to be positively correlated with the majority of microbial classes in the non-healing model, yet was negatively associated with most classes in the self-healing model.

Table 4.7: Microbial–Metabolic Correlations in Urine of CL Models

	Metabolic Correlate	Clos	Gamma	Erysi	Alpha	Beta	Bacil	Actino	Bacter
BALB/c Urine	2-Oxoadipate	–	+			–			
	2-Oxoglutarate			–	–	+			
	2-OIC	+	–		+	+	–	+	+ +
	Acetate	–	+	–		+			
	cis-Aconitate	+	–			+			–
	Citrate	– –	+ +		–	– –			
	Creatine			–					
	Creatinine	+	– –	+		+		+	+ +
	Dimethylamine			+/–				+	
	Fumarate	–	+						
	Glycine	+	–			+			
	Guanidoacetate			+	–			+/–	+
	Hippurate	+	–		+		–		–
	Lactate	–	+		–	+			
	PAG	+ +	–		+	–			+
	Succinate			+/–					
	Taurine	–	+		–		+	–	+
	Trimethylamine	+	–		+	+			
	TMAO	+	–		+	+			
	Ureidopropanoate								+
C57BL/6 Urine	2-Oxoglutarate							+	
	2-OIC							+/–	+ +
	Acetate				+				
	cis-Aconitate						–	–	–
	Creatine		–					+	
	Creatinine	–	+	–	–		–	– –	–
	Dimethylamine					–	–		
	GPC							–	
	Lactate	+/–			+				–
	PAG	–			–		–		
	Taurine				–	–	– –	+/–	–
	Trimethylamine	+		+					–
	TMAO	+/–	+ +	+/–	+				
	Ureidoproanoate								+

Significant correlations between urinary metabolites' spectral peak intensities and relative percentage abundance of faecal microbial classes, using 2D matrix correlation analyses ($p<0.05$). Unassigned metabolic-correlates have been omitted. Key as described Key as described for Table 4.6 with the following addition: 2-OIC, 2-oxoisocaproate. *Reprinted with permission from Lamour, et al. (2014)*[197] *Copyright 2014 American Chemical Society.*

Correlations between Cytokine Concentrations Faecal Microbial Abundance

Correlation analyses between infection-related faecal microbial class abundances and plasma cytokine concentrations for each model have been summarised in Table 4.8. Results revealed that IL-1β exhibited the biggest number of associations with the host microbiome in the self-healing model, being positively correlated with all microbial classes for at least one time-point. In contrast,the same cytokine was only correlated with *Bacilli* and *Gammaproteobacteria* in the non-healing BALB/c strain. Additionally, IL-12 shared multiple microbial correlations in both CL models yet did not correlate with *Gammaproteobacteria* in either strain. IL-12 and IL-10 displayed the highest number of microbial correlates out of the assessed cytokines in the non-healing strain, although unlike observations with immune-metabolic correlations, IL-10 did not display correlation patterns opposing

the other pro-inflammatory cytokines with the host microbiota. One of the strongest immune-microbial correlates was a positive association between IL-10 and *Actinobacteria* in both mouse models, which had a correlation coefficient (r) > 0.6 for two out of the 3 post-infection time-points, in each strain. *Bacteroidia* was also correlated with IL-10 in both models, implying that, in addition *Actinobacteria*, these commensal bacteria may influence IL-10-dependent anti-inflammatory effects in the host. CXCL1 displayed the least amount of correlations with bacterial classes in these analyses, indicating only minimal interaction with gut microbiome.

Table 4.8: Immune–Microbial Correlations in CL Models

	Infected BALB/c				**Infected C57BL/6**			
Microbial Class	IL-10	IL-12	IL-1β	CXCL1	IL-10	IL-12	IL-1β	CXCL1
Clostridia	+	++				+	++	
Gammaproteobacteria			+/–	+	+		+	+
Erysipelotrichia		+		++	+	+	+	
Alphaproteobacteria	++	+				+	++	
Betaproteobacteria	+/–					+	++	
Bacilli		+	+		+	+	+	
Actinobacteria	++	+			++		+	
Bacteroidia	++	+			+	+	+	

Significant correlations (p<0.05) between plasma cytokines concentrations and relative percentage abundance of faecal microbial classes using 2D matrix correlation analyses. Key as described previously. *Reprinted with permission from Lamour, et al. (2014)*[197] *Copyright 2014 American Chemical Society.*

4.4.6 Discriminatory Co-analyses of Metabolic, Cytokine and Microbial Datasets that Differentiate CL Models

To evaluate the relative contribution of metabolic, immunological and microbial markers towards the separation between the two mouse strains' responses to parasitic infection, metabolite integrals, cytokine concentrations and bacterial class percentage abundances were co-analysed using O-PLS-DA, shown in Figure 4.7. Metabolites from plasma, urine and faeces (which represent the biological matrices of most clinical relevance) were selected based on criteria outlined previously in Section 4.3.3.

Analyses demonstrated that urinary and plasma metabolites represented the major drivers of separation between the between the models, especially towards the beginning of infection, e.g. urinary hippurate and plasma glycine, which were higher in the self-healing C57BL/6 mice, and plasma glucose, which was higher in the non-healing BALB/c strain. Hippurate was found to be consistently higher C57BL/6 throughout the study, along with urinary creatine, faecal glutamate and levels of bacterial classes *Betaproteobacteria* and *Erysipelotrichia*. In contrast, markers that were consistently higher in non-healing mice, post-infection, included plasma 3HB, alanine, and

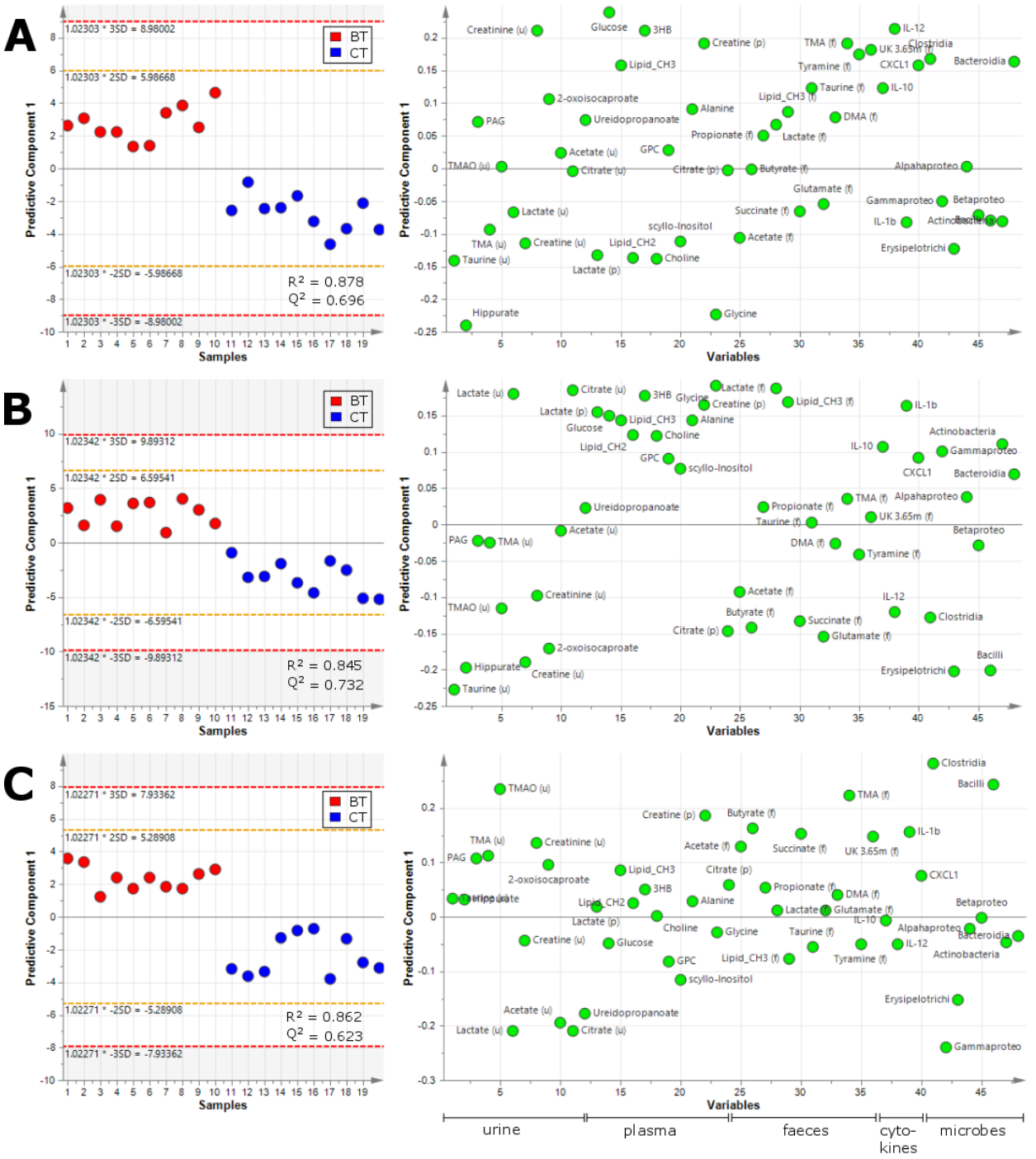


Figure 4.7: Differences in Metabolic and Microbial Responses to L. major Infection Dominate Separation between Self-Healing and Non-healing Models

Corresponding O-PLS-DA scores plots (left) and loadings (right) for time points D2 (**A**), D6 (**B**) and Dterm (**C**), based on the first predictive component. Scores plots show vertical CL model separation between infected BALB/c (BT, in red) and C57BL/6 (CT in blue), where each circle represents a sample. Loading plots show discriminatory variables (green circles) that contribute towards the separation in the model. Variables consist of 48 markers in the following order: 12 urinary (u), 12 plasma (p), and 12 faecal metabolites (f), 4 cytokines, and 8 eight microbial classes. Origin of metabolites present in more than biofluid are additionally labelled with (u), (p), or (f). Abbreviations as described previously.

lipid methyl moieties, faecal trimethylamine and propionate, and CXCL1 cytokine.

With study progression, the faecal microbiota became the most discriminatory variables between the two infection mouse models, with the *Gammaproteobacteria* class being strongly linked with the self-healing mice and *Clostridia* and *Bacilli* classes linked with the non-healing model. Multiple markers varied their association over time, which included most cytokines. Despite this alternating model affiliation, IL-1β and IL-12 were always opposite to each other, and were the two strongest cytokine contributors in differentiating between the two strains.

4.4.7 Parasitic Burden Associated with a Set of Metabolic, Immune, and Microbial Markers

The 48 variables presented in previous analyses were correlated with the relative fluorescence intensities of lesion-derived *L. major* DNA to determine potential links between parasitic burden and host metabolism, immune status, or faecal microbiota, as shown in Figure 4.8.

Following multiple test correction, none of the parasite-host-response correlations displayed a false discovery rate (FDR) q-value less than 0.05. Yet in the interest of maximal exploration of potentially biologically relevant associates, the threshold was increased to $q < 0.3$, which showed positive associations between faecal succinate and IL-12 in the self-healing strain. In the non-healing strain, plasma *scyllo*-inositol, faecal acetate, and *Betaproteobacteria* were positively associated with parasite counts, whereas urinary 2-oxoisocaproate was negatively associated.

4.4.8 Correlation Network Analyses for Interactive Systems Host Response towards L. major Infection

The final assessment of the presented data was to perform correlation network analyses between the aforementioned 48 variables for each time-point in each CL model, to investigate potential links between host systemic metabolism, immune status and gut microbiome, throughout the course of infection. Correlation networks were evaluated using two formats: firstly using in the 'wheel' format, where positions of each variables are fixed, to facilitate direct comparisons of connections between strains and/or time-points, as shown in Figure 4.9. Significant associations ($q < 0.03$) between variables are shown as connecting red and blue lines, for positive and negative correlations respectively, and nodes for each variable are colour-coded according to their biological matrix.

Contrasting results were obtained when assessing the degree of 'inter-connectivity' between the variables, which differed greatly between two models. Even prior to infection, little correlations were observed in non-healing BALB/c mice (56 correlations) which were all limited to connections within

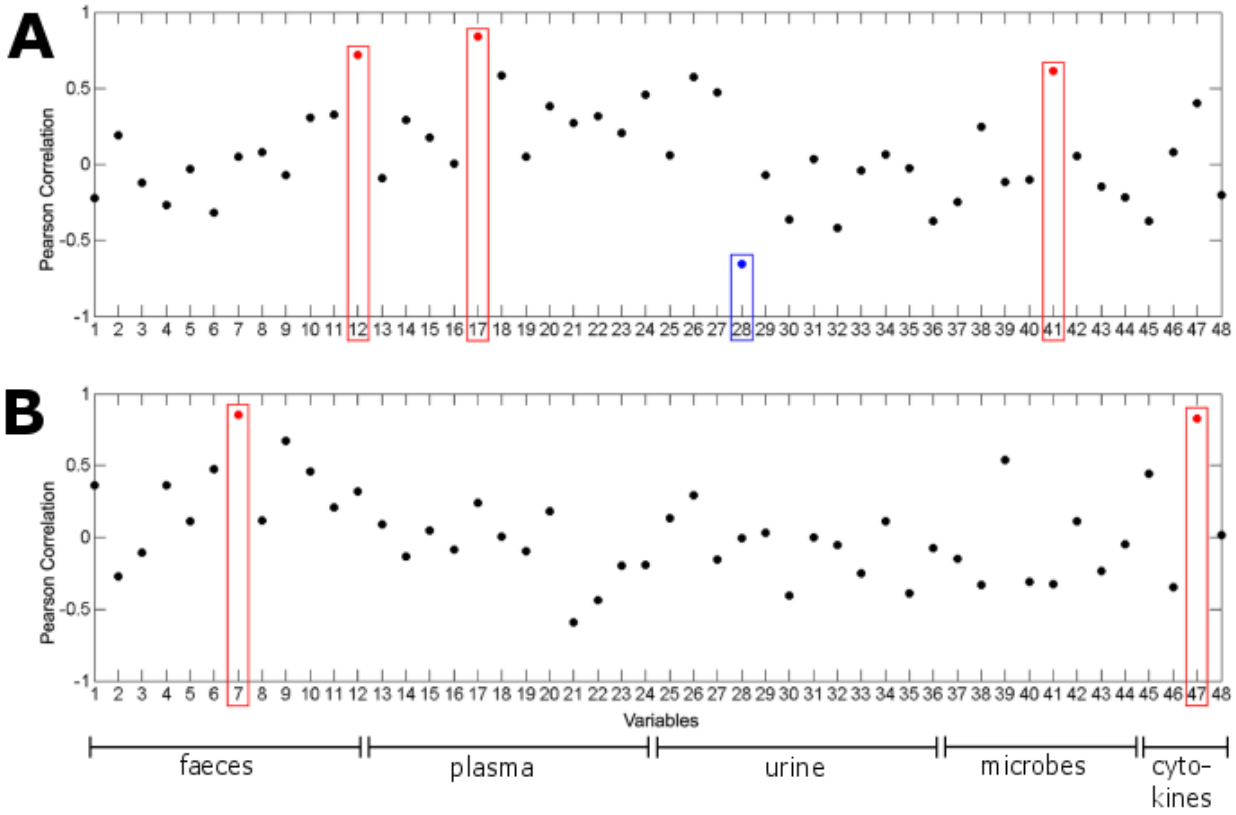


Figure 4.8: Correlation Analyses between Parasitic Burden and Host Response Markers in Non-healing and Self-healing CL models

Plot shows Pearson correlation between infection-related metabolic, immune and faecal microbial variables (Dterm) and parasitic burden in lesions from *L. major* infected BALB/c (**A**) and C57BL/6 mice (**B**). Each variable is represented by a small circle, where those with a FDR corrected q<0.3 have been highlighted as red (positively correlated with parasites) or blue (negatively correlated). Key: faecal metabolite: 1, unassigned multiplet peak at 3.65 ppm; 2, tyramine; 3, trimethylamine; 4, dimethylamine; 5, glutamate; 6, taurine; 7, succinate; 8, lipid CH_3; 9, lactate; 10, propionate; 11, butyrate; 12, acetate. Plasma metabolites: 13, citrate; 14, glycine; 15, creatine; 16, alanine; 17, *scyllo*-inositol; 18, glycerophosphocholine; 19, choline; 20, 3-hydroxybutyrate; 21, lipid CH_2; 22, lipid CH_3; 23, glucose; 24, lactate. Urine metabolites; 25, ureidopropanoate; 26, citrate; 27, acetate; 28, 2-oxoisocaproate; 29, creatinine; 30, creatine; 31, lactate; 32, trimethylamine-N-oxide; 33, TMA; 34, phenylacetylglycine; 35, hippurate; 36, taurine. Bacterial classes: 37, *Clostridia*; 38, *Gammaproteobacteria*; 39, *Erysipelotrichia*; 40, *Alphaproteobacteria*; 41, *Betaproteobacteria*; 42, *Bacilli*; 43, *Actinobacteria*; 44, *Bacteroidia*. Cytokines: 45, CXCL1; 46, IL-1β; 47, IL-12; 48, IL-10. *Reprinted with permission from Lamour, et al. (2014)*[197] *Copyright 2014 American Chemical Society.*

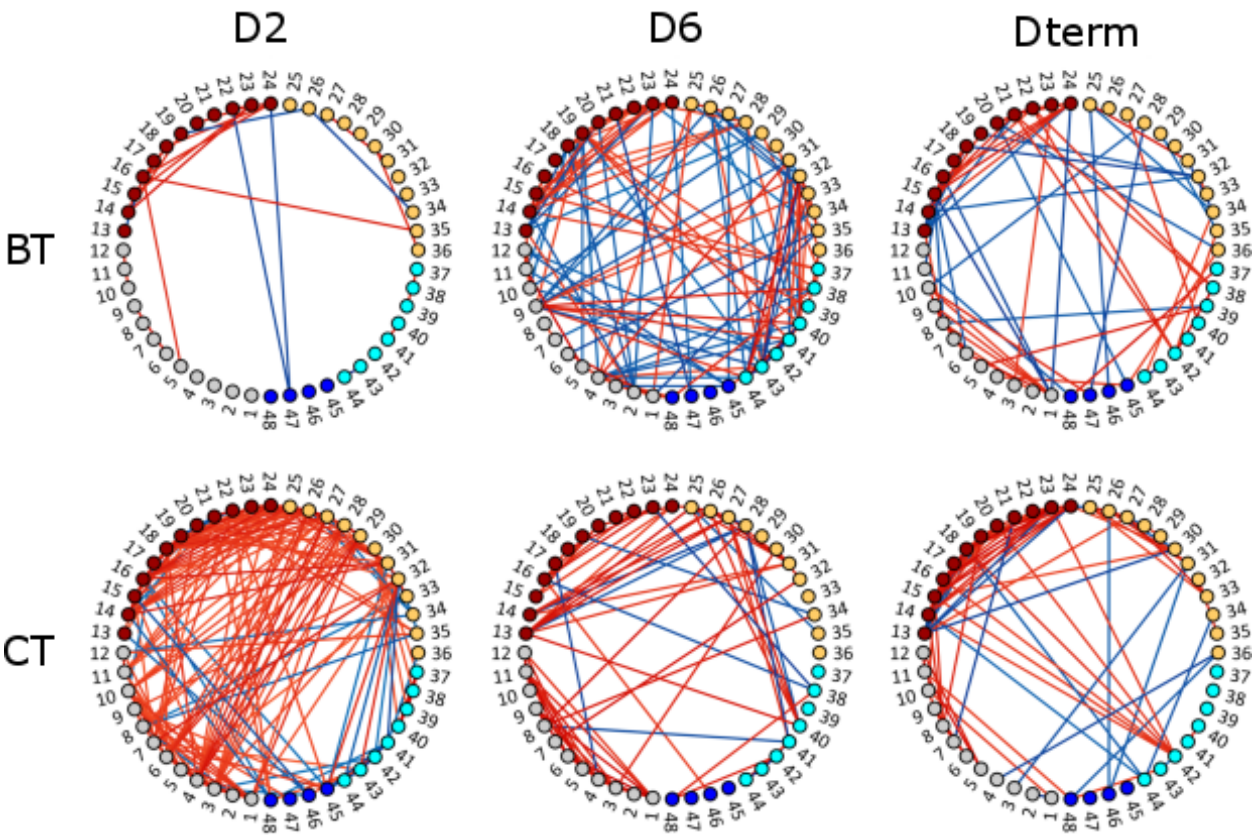


Figure 4.9: Correlation Network Wheels of Host Responses to Infection Show Differences in Inter-connectivity between CL Models and Time-points

Correlation network analyses shown in wheel format, performed separately for each infected strain, for each time-point, based on pair-wise Pearson correlation between variables. Significant positive and negative correlations between variables are shown as red lines and blue lines, respectively ($q < 0.3$). Each variable has a fixed position, with nodes along the circumference that are colour-coded according to biological matrix in the following order: faecal metabolites in grey (1–12); plasma metabolites in red (13–24); urinary metabolites in yellow (25–36); microbial classes in cyan (37–43), cytokines in dark blue (45–48). Complete list of variables as in Figure 4.8.
Modified with permission from Lamour et al.(2014)[197] *Copyright 2014 American Chemical Society.*

each assessed biological matrix (i.e. intra-matrix) when compared to the C57BL/6 (114 correlations — see Appendix Figure A.5). On D2, infected BALB/c mice continued to display a low number of correlations between the majority of response markers (only 48 correlations observed), where the majority (79%) of these were comprised of intra-matrix associations, e.g. plasma markers that correlated with other plasma markers. Inter-matrix correlations included IL-12 being negatively associated with plasma lipids and lactate (Figure 4.9). On the other hand, a remarkable increase in connections was observed on D2 in the infected self-healing mice, which not only represented the highest amount of correlations (346) in this study, where every assessed variable was associated with at least one other variable, but also showed the highest proportion of inter-matrix correlations (69% of 346). The vast majority of inter-matrix associations were between metabolic markers from the three biofluids, of which urinary metabolites displayed the most correlations with other

compartments.

By D6, more correlations were observed between variables in the non-healing strain (244 in total) than at earlier time-points. Similar to D2 in the C57BL/6 mice, infected BALB/c displayed a majority of correlations between biological compartments (62% of 244): collectively, faecal metabolites, followed by urinary metabolites, displayed the most inter-matrix associations, whilst plasma metabolites were largely correlated amongst themselves. Additionally, an overall higher proportion of negative correlations were observed with the non-healing model than self-healing mice, which was particularly apparent on D6. From D6 in the self-healing model, correlations became increasingly more confined to their individual biological matrices (intra-matrix correlations represented 60% of 144 for D6 and 69% of 170 for Dterm).

Correlation analyses were also depicted as network 'maps', shown in Figure 4.10, where positions of variables are grouped according to correlation clusters. This format allows the visualisation of major hubs and chains of correlation within the networks, and is thus more aligned with mapping to biological pathways. Results highlighted that plasma and urinary metabolites represented the core frameworks within the networks throughout the course of infection in both CL models, and exhibited the highest number of inter- and intra-matrix correlations. Whilst the majority of plasma metabolites were largely correlated with each other in this study, plasma choline was separated from the rest of the plasma network in the majority of the time/train-specific correlation maps. Instead, it was associated with faecal metabolites and/or urinary metabolites including TMAO, a gut-microbial associated metabolic product derived from choline. Unlike plasma, urinary metabolites displayed a higher degree of inter-matrix associations, especially with faecal metabolites and bacterial classes.

In self-healing mice, acetate, citrate, and alanine represented major network hubs, displaying the some of the highest amount of correlations with other variables across all three post-infection time-points. In contrast, network hubs in the non-healing strain were less stable between time-points and alternated throughout the course of the study. Plasma lactate represented the only hub that displayed a high number of correlations consistently throughout the study. Other hubs that were present in BALB/c for two time-points (D6 and Dterm) included urinary trimethylamine-*N*-oxide (TMAO) and plasma alanine.

Apart from D2 in the non-healing strain, IL-10 was consistently connected with multiple host markers within the core network of both models, highlighting its systemic regulatory function in the host. CXCL1 was directly associated with a high number of bacterial classes and metabolic markers from different biological matrices in both CL models. Furthermore, despite the higher

concentrations of IL-1β levels observed in infected BALB/c, IL-1β displayed a more interactive function the C57BL/6 mice, where overall a higher number of direct correlations were observed in this model. IL-12 was directly correlated with a multitude of variables in the non-healing strain yet produced negligible correlations in the self-healing mice C57BL/6.

Examination of microbial markers demonstrated that *Clostridia* and *Gammaproteobacteria* were negatively correlated with each other across multiple time-points, in both models, in-line with their opposing correlation patterns observed in earlier findings. On the other hand, *Alphaproteobacteria* and *Betaproteobacteria* were either positively correlated with each other or displayed similar correlative patterns across several time-points. *Erysipelotrichia* represented one of the most interconnected bacterial classes, particularly apparent on Dterm in the self-healing strain, where it was positively linked with several plasma metabolites. *Bacilli* and *Bacteroidia* classes were typically disconnected from the major network, correlating largely with themselves, if at all.

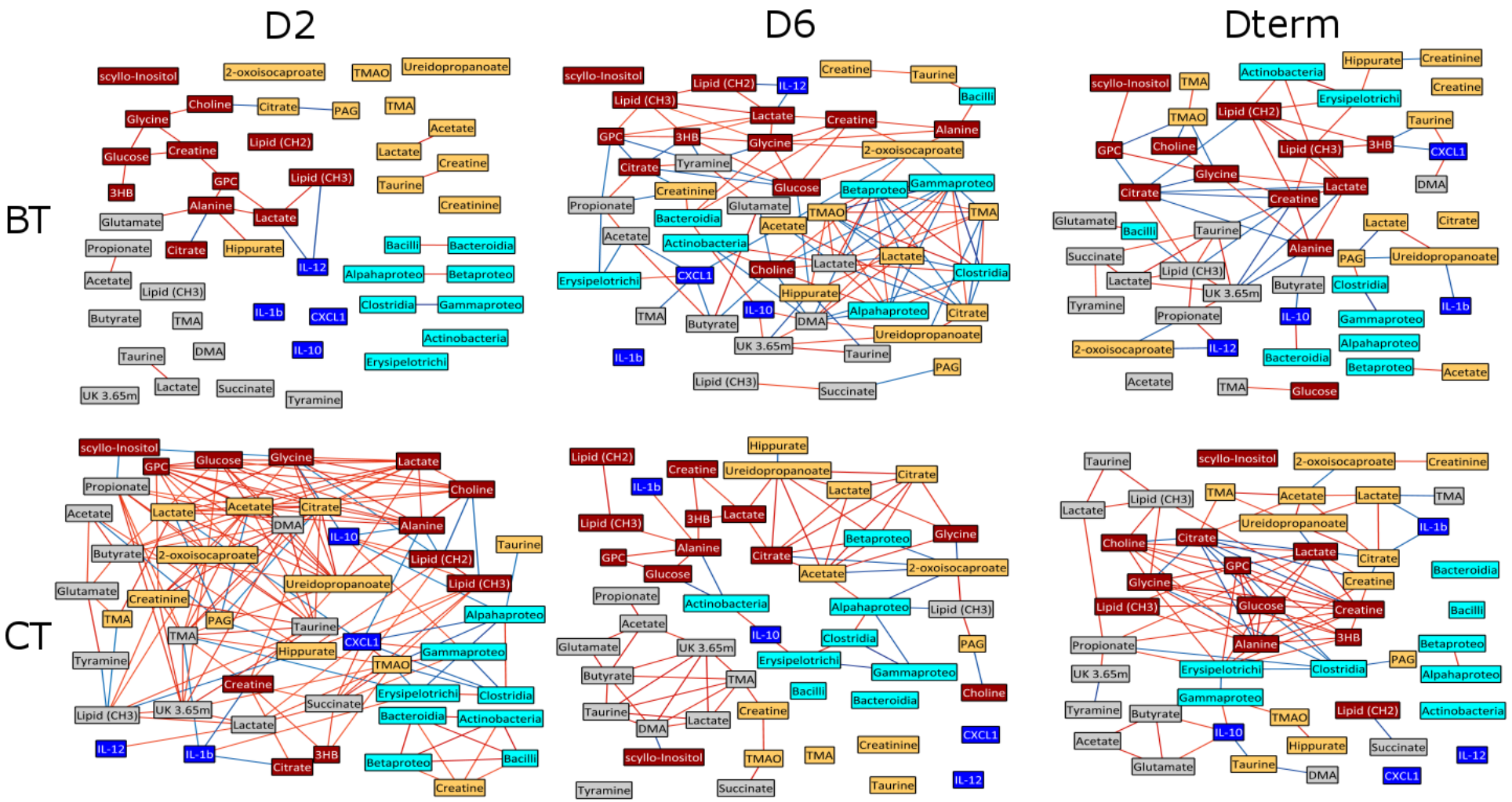


Figure 4.10: Correlation Network Maps of Host Responses Highlight Hubs in CL Models

Correlation network analyses shown in map format, performed separately for each infected strain, for each time-point, based on pair-wise Pearson correlation between variables. Significant positive and negative correlations between variables are shown as red lines and blue lines, respectively ($q < 0.3$). Markers are grouped according to correlation clusters and are in boxes, colour-coded according to their biological matrix: urine in yellow, plasma in red, cytokines in dark blue, and bacterial classes in cyan. Key: 3HB, 3-hydroxybutyrate; DMA, dimethylamine; GPC, glycerophosphocholine; PAG, phenylacetylglycine; TMA, trimethylamine; TMAO, trimethylamine-N-oxide; UK 3.65(m), unassigned multiplet at 3.65 ppm.
Modified with permission from Lamour et al.(2014)[197] *Copyright 2014 American Chemical Society.*

4.5 Discussion

The presented work represents the first demonstration of using a metabolic profiling approach to characterise to *L. major* infection *in vivo*, using ^{1}H NMR spectroscopy, to determine the early changes in host metabolism linked with parasitic infection. Setting the immunological context of this study, the first assessed post-infection time-point (D2) represents a response that is primarily driven by innate immune system, whereas the two later time-points (D6 and Dterm) are generally considered to be driven by the host's adaptive immunity.[162] This study also uniquely presents a systemic characterisation and integrative network analyses of global metabolic phenotypes, plasma cytokine profiles and faecal microbial composition in both non-healing "susceptible" BALB/c and self-healing "resistant" C57BL/6 mouse models of experimental CL.

4.5.1 Pathological Measures Indicate Contrasting Phenotype between Two CL Models already in Early Infection

Experimental *L. major* infection in C57BL/6 mice leads to lesion development at the site of administration that usually resolves from 8 weeks post infection, whereas the same infection induces uncontrolled lesion growth and visceral parasitic dissemination in the susceptible BALB/c mice.[240,241] Although our study focused only on the early stages of infection (up to 13 days post-infection depending on animal), trends in increased lesion size and higher parasitic burden in lesion and liver tissue were already apparent in the non-healing BALB/c model. Additionally, infected BALB/c mice also displayed significantly higher haematocrit levels than infected C57BL/6 mice, suggesting the presence of inflammation (also confirmed by cytokine results), as was also observed in the uninfected strains (see previous chapter for results of control mice).

4.5.2 L. major Infection Associated with Changes in Glucose and Lipid Turnover and in Gut-Microbial Metabolism

As *L. major* infection in BALB/c mice represents the most extreme disease model of CL (and thus in theory show the biggest parasitic-related effects), infection-associated changes in host metabolism were investigated in this strain. Whilst no changes in faecal metabolic phenotypes were apparent, infection was linked with changes in glucose and lipid metabolism in plasma profiles, where glucose was reduced and several lipid moieties were depleted. Combined with the significant increases observed in 3-hydroxybutyrate (3HB), a ketone body largely generated from hepatic fatty acid catabolism[255], the results indicate an increased lipid turnover during infection. Interestingly, *in vitro* studies have shown that intracellular *Leishmania* amastigotes rely heavily on the breakdown

of fatty acids rather than carbohydrates for energy and growth, whilst also increasing glycolysis of the host macrophage[34,256], implying that the depletion in plasma lipids may be directly due to parasitic uptake. Decreases in plasma lipid concentrations have also been observed in mice following infection with *Trypanosoma brucei (T. b.) brucei*, a species related to *Leishmania*, also belonging to the *Trypanosomatida* order.[32] It has already been established that BALB/c mice inherently display high levels of circulating lipids (discussed in Chapter 3). Results thus suggest that due to their high bioavailability, the break down of fatty acid represents a major alternative for ATP generation, presumably to accommodate the high energy requirements needed to elicit an active immune response towards the parasite.

Infection-related differences in the urine included changes in several gut-microbial-associated metabolites, such as a significant decrease in hippurate and increased phenylacetylglycine (PAG). Reductions in urinary hippurate have been observed in several parasitic murine infection models (including infections with *Echinostoma (E.) caproni*[93], *Schistosoma (S.) mansoni*[150], and *T. b. brucei*[92,257]). Similarly, increased urinary PAG has been observed following experimental murine infection with *Plasmodium berghei*[258], *E. caproni*[93] and *S. mansoni*[150], indicating that these two features may be general markers of parasitic infection. Metabolic changes that appear unique to *L. major* infection are increases in urinary indoxylsulphate, a dietary protein metabolite thought to be uraemic toxin[259] and succinate, a known major end-product of *Leishmania* metabolism.[34,260]

4.5.3 Infection-related Differences between CL Models Observed in Every Assessed Response System

Comparisons between metabolic, cytokine and faecal microbial phenotypes of non-healing and self-healing mouse models were performed using data from infected animals, that had been adjusted for strain-related differences observed in the control groups. This strategy allowed for the identification of differences in infection-related metabolite concentrations that were superimposed on top of genetic variation between BALB/c and C57BL/6 (as has been discussed in the previous chapter). Results revealed significant differences in infection-associated responses between the two CL models on every assessed system, whether metabolic, cytokine or microbial. Further discriminatory analyses showed that metabolites particularly from the plasma and urine contributed the most in differentiating between the two infected strains.

Infection-associated Metabolic Changes Pronounced in Non-Healing Model

Many of the metabolites that were discriminatory between the two infected CL models were also shown to differentiate between control and infected animals, e.g. levels of plasma 3HB, urinary PAG, and faecal succinate, all of which were increased due to infection, were also higher in the non-healing BALB/c strain than the self-healing model. Concentrations of urinary hippurate, which was reduced in mice infected with *Leishmania* (in this study) and with several other parasitic infections, were also found to be lower in the non-healing model. Additionally, several lipids were lower in the plasma and tissues of in infected BALB/c mice compared with infected C57BL/6 mice. These had been shown to be decreased as result of infection in our study, despite having established that uninfected BALB/c normally have higher circulating lipids than control C57BL/6 mice (see Chapter 3). Plasma glucose on the other hand, which was depleted due to infection, was found to be positively correlated with the non-healing strain. This is contrasting to the observation that self-healing C57BL/6 mice are known to have generally higher plasma glucose levels.[205–209] These findings suggests that glucose consumption in infected C57BL/6 mice outstrips that of infected BALB/c mice, which as data imply, rely heavily of lipid break-down for energy needs instead. Following infection, self-healing strains also displayed higher hepatic levels of glycogen, which may reflect the increased pressure on the liver to generate more glucose reserves, to meet this high consumption rate.

Furthermore, major *Leishmania* waste products alanine and succinate[34,260] were found in higher concentrations across multiple compartments (in faeces and tissues for alanine and in urine for succinate) in the non-healing phenotype. Succinate levels were also increased due to infection in the non-healing model and were found to be directly correlated with parasitic burden. Earlier *in vitro* work showed significantly higher concentrations of alanine and succinate in cell media from *L. major* infected macrophages, when compared to uninfected cells, highlighting their potential role as biomarkers of *Leishmania* infection.[44] Thus it is plausible that these metabolites may in fact be directly derived from the parasite.

Pro-inflammatory Phenotype Associated with Non-healing State

As previously alluded to, in reference to the by haematocrit results, infected BALB/c mice displayed a stronger pro-inflammatory cytokine profile than the self-healing model. They displayed significantly higher plasma concentrations of IL-1β, an important inflammatory mediator produced by macrophages and involved in a variety of cellular activities, with trends also showing higher levels of neutrophil chemoattractant, CXCL1. Thus instead of finding anticipated discrepancies

in Th 1 vs. Th 2 cytokine profiles between the two models, results confirmed the importance of the innate immunity in driving the opposing response to the parasite within the early stages of parasitic infection, as has become increasingly recognised as in the context of *Leishmania*.[261] Interestingly, Voronov and colleagues recently showed that infection with *L. major* in BALB/c mice, engineered to lack physiological inhibitor of IL-1 (knock-out of IL-1 receptor antagonist [IL-1Ra] gene), displayed more serious form of disease, linked with increased levels of other inflammatory cytokines and suppressed T cell responses.[262] Together, evidence strongly suggests that an excessive pro-inflammatory phenotype already exists in the early stages of infection and not only hinders the generation of an effective immune response but actively promotes disease propagation.

In contrast to published evidence[31], defining the role of Th 1 cytokine IL-12 in the current study proved difficult, not only owing to the variable concentrations measured over the course of infection in the self-healing strain, but due to its observed positive correlation with parasitic burden (in the non-healing model). Earlier research has shown that shortly after *Leishmania* infection, IL-12 is increased more in BALB/c mice than C57BL/6 mice yet thereafter, rapidly declines in BALB/c mice only.[233] A similar finding was observed in the presented work, where IL-12 was significantly higher in the non-healing strain than C57BL/6 mice two days post-infection, and appeared to decrease over time in infected BALB/c mice. However, it is still unclear whether this apparent decline in IL-12 is linked to disease propagation.

Infection Associated with Dynamic Microbial Changes in Self-healing Mice

Infection was associated with significant differences in faecal bacterial composition between the two CL models, notably in the expression levels of the three most dominant classes observed in this study: *Clostridia*, which was higher in the non-healing mice and *Gammaproteobacteria* and *Erysipelotrichia*, which were higher in self-healing model. Analyses showed that the relative abundance of *Clostridia* and *Gammaproteobacteria* were inversely correlated, a result that was also reported following weight-loss due to bariatric surgery in rats by Li *et al.*[263] Following infection, significant changes in faecal microbial class abundances were only observed in the self-healing mice, largely between the pre-infection time-point and the terminal day of the study. These changes included reductions in levels of *Bacilli* and *Gammaproteobacteria*, as well as an increase in *Bacteroidia* abundance.

4.5.4 Links between Metabolic, Immune and Microbial Markers Highlight a Systematic Response to L. major Infection

To determine potential links between metabolic, cytokine and faecal microbial measures in each CL model, two statistical approaches were performed. Firstly 2D matrix correlation analysis was used to correlate between two out of the three measures at a time. This was followed by correlation network analysis, using a more targeted approach for metabolic data, as only integrals of selected metabolites were used. This method uniquely allowed the direct statistical analyses between all three host response datasets simultaneously. In both cases, significant associations between metabolic, cytokine and faecal microbial markers were observed, indicating the presence of complex interactions between these systems in order to generate an overall integrative response towards *Leishmania* infection.

Two-dimensional matrix analyses highlighted the contrasting correlation patterns of *Clostridia* and *Gammaproteobacteria* with metabolic markers: these two dominant bacterial classes were also often correlated with the same metabolites yet with opposing directionality, a phenomenon that was observed across both infection models in this study. Furthermore, levels of *Erysipelotrichia* displayed similar metabolic correlates with *Clostridia*, suggesting that these two classes share many metabolic pathways.

A strong association between IL-10 concentration and *Actinobacteria* levels was observed in both CL models *via* 2D matrix correlation analyses. The *Actinobacteria* class includes microbial species within the *Bifidobacterium* genus, commonly used in probiotics for general gastrointestinal well-being[264], and have been shown to increase IL-10 production both *in vitro*[265] and *in vivo*.[266] Analyses also showed that several metabolites that were positively correlated with IL-10 were negatively linked associated with the other inflammatory or Th 1 cytokines and vice versa (evident particularly in the tissues), confirming its systemic regulatory role in the host.[267]

Yet the most remarkable differences in this study was the varying degree of inter- and intra-matrix connections between the two mouse models and sampling time-points, determined using correlation network analyses. The self-healing model displayed the maximal number of correlation between metabolites, cytokines and microbiota, as well as the highest proportion of inter-matrix correlations, two days post-infection. In this strain, acetate, citrate and alanine represented the primary network hubs, indicating that these metabolites may prove significant in orchestrating a systemic response towards *Leishmania.* Acetate is central to both carbohydrates and fatty acid metabolism, and is largely used endogenously for formation of acetyl coenzyme A. It is also the main short-chain fatty acid generated by the gut microbiota that is readily reabsorbed back into

systemic circulation, and hence, is additionally used as a marker of colonic fermentation of dietary carbohydrate.[268] Citrate, an intermediate of the TCA cycle, and alanine, an amino acid with gluconeogenic functions (*via* the Cahill metabolic pathway)[269], both represent markers associated with host energy metabolism. Thus results from the presented work emphasise the central role of carbohydrate metabolism during infection in the self-healing C57BL/6 strain, presumably to meet the high energy demand required to mount a successful system response towards the parasite.

The non-healing BALB/c strain displayed the highest number of correlations between variables only 6 days post-infection, with few associations seen at earlier time-points. Additionally, the observed correlations appeared to be less ordered (connected at random), most evident on D6, as opposed to the functionally stable network hubs presented in C57BL/6 mice. Plasma lactate represented the only variable that consistently displayed a high number of correlations across all sampling time-points in infected BALB/c mice (and was also well connected in the core networks in self-healing strain), largely connected with other plasma metabolites, particularly lipids.

Studies from Beil *et al.* (1992) have shown that following the first 12 days post-infection with *L. major*, recruitment of immune cells to the site of infection was limited to polymorphonuclear cells (mainly neutrophils) in BALB/c mice for the majority of sampling time-points.[270] On the other hand, they found that infected C57BL/6 mice recruited a diversified repertoire of leukocytes (including macrophages), already from the first day post-infection, which resulted in a higher rate of parasite degradation. Consequently, it could be hypothesised that the inter-connected interactions between the multiple host compartments observed in our study are vital for generating a rapid, diverse and effective innate response, as observed in the self-healing model of CL. In contrast, the delay observed in non-healing strain results in an inefficient response to the infection, and poorer communication between the immune, metabolic and bacterial systems, thus allowing parasitic survival.

4.5.5 Study Limitations

Unlike metabolic or cytokine data, marked differences depending on animal caging was observed in the faecal microbial data. The use of group average microbial abundance for final data analyses helped reduce the disparities to identify overall trends within the data, yet this did not alter the high degree of variation across each infected group for each time-point. This feature may have impaired the detection of more subtle differences between and within groups, particularly in the most dominant bacterial classes, where the cage-related variation was most pronounced. The housing conditions used for this study (i.e. 5 animals per cage, all belonging the same group and

strain), which covered such a multi-disciplinary scenario, was based on common house caging and available resource allocation. However, a different caging allocation may be more suitable for future studies similar to this that are emphasised on gut microbial results, e.g. where each cage holds one animal from each infection group and each strain (in this case, four animals per cage), or where mice are housed separately in individual cages, to minimise this confounding variation.

Another limitation specific to this study was the absence of baseline (D-2) metabolic and cytokine data for plasma control samples. The lack of this data meant that no control-strain adjustment could be performed for this time-point, resulting in omission of that time-point for statistical comparisons between time-points for both metabolic profile and cytokine data, and thus loss in information. An alternative would have been to simply use the pre-infection time-point as the control group, and compare all post-infection time-points to the former. Yet as revealed in the previous chapter, subtle metabolic changes were also observed in uninfected mice over time, confirming use of time-matched control animals instead as more appropriate reference points for data analyses.

Furthermore, the q-value cut off of 0.3 for correlation network could be considered overly generous to draw conclusive results from a statistical point of view. Yet it should be recognised that this method represented a proof-of-principle approach of co-analysing multiple diverse host response markers, using *in vivo L. major* infection as a model. Additional experiments would be needed to confirm the findings observed in this study.

4.5.6 Conclusions and Future Work

The systemic approach of concurrent metabolic, immune and microbial characterisation of two experimental cutaneous leishmaniasis models adds a novel aspect to the assessment of host defence, that is applicable to other disease models. Clear differences between the self-healing and non-healing models included disparities in energy and lipid metabolism, and gut microbial composition, many of which were also discriminatory markers for infection status. Markers of inflammation were also increased in the non-healing mice, already in early stages of infection. Furthermore, this work uniquely demonstrates the accumulated presence of parasitic-linked metabolites *in vivo* in the infected state, which were increased in the non-healing model, and shown to correlate with parasitic burden. Future clinical assessments of alanine and succinate levels in leishmaniasis patients would be required to investigate whether these metabolites would prove useful as candidate diagnostic biomarkers of *Leishmania* infection.

Analyses imply that already during the innate immune response, an efficiently organised network of metabolic, immune, and gut microbial markers are associated with successful defence to infection.

A comprehensive and structured correlation network between multiple biological responses was present in the self-healing model following infection that was not observed in the non-healing mice, which may explain the superiority of this strain in controlling *Leishmania*. Finally, a set of metabolic candidates were determined as major network hubs linked to a successful immune response to *L. major*. Further research is needed to validate the indicated potential of these metabolites: an example of this would be the employment of *in vitro* supplementation studies (as has already been initiated in-house)[271–273], testing both directly on promastigote cultures for direct leishmanicidal effect, and on (infected) host cells, for immune-modulatory effects.

Chapter 5

Results: Metabolic and Immune Characterisation of Human African Trypanomiasis (*rhodesiense*)

5.1 Background 147
5.2 Objectives 148
5.3 Experimental Methods 149
5.4 Results: Plasma Analyses 154
5.5 Results: CSF Analyses 168
5.6 Discussion 174

5.1 Background

Diagnosis for Human African Trypanosomiasis (HAT), also known as African sleeping sickness, remains a major challenge for disease diagnosis and treatment, as patients typically display a variety of non-specific inflammatory clinical symptoms that are shared with other endemic diseases, e.g. enteric fever and malaria.[65,66] Infection in humans is caused by *Trypanosoma brucei (T. b.) rhodesiense*, which causes a short acute febrile illness within south/west Africa, and *T. b. gambiense*, which results in a more chronic form, and is prevalent in east and central Africa. Infection with the related species, *T. b. brucei*, is limited to domestic and wild animals, and cause disease in cattle, goats, and sheep, across Africa but is generally not infectious to humans.

The disease is diagnostically categorised into two stages: in early (also known as haemo-lymphatic) stage, parasitic infection is limited to the bloodstream and peripheral organs, where patients usually present with inflammatory symptoms. Without intervention, the disease progresses to the late (meningoencephalitic) stage, where the parasites have crossed the blood brain barrier, and the disease is accompanied by neurological clinical manifestations in patients and, ultimately, death, if left untreated. WHO defines this late stage as the presence of at least 5 parasites or 20 leukocytes in 1 µl of CSF. Correct assignment of staging is critical in HAT as appropriate drug regiments (and notably, their accompanying toxic side-effects) differ between the stages.[58] Diagnostic staging still relies on lumbar puncture for CSF collection, followed by microscopic examination, yet this highly invasive procedure is technically difficult and often inaccurate. Furthermore, unlike for *T. b. gambiense* infection, no general screening assays — such as the Card Agglutination Test for Trypanosomiasis (CATT) or new generation lateral-flow rapid diagnostic tests — currently exist for *T. b. rhodesiense* infection, highlighting the urgent requirement of novel and more effective diagnostic tools for trypanosomiasis *rhodesiense*.[274]

Metabolic profiling has already been performed on a murine experimental model of HAT, infected with *T. b. brucei*, which has revealed a range of metabolites that were significantly changed, already from one day post-infection.[92] These included increased plasma concentrations of lactate, creatine and acetylglycoproteins, and decreases in branched-chain amino acids valine, leucine and isoleucine. Urinary changes highlighted perturbations within the gut microbiota, as concentrations of several microbial associated metabolites were altered following infection, e.g. hippurate levels decreased whilst trimethylamine and 4-hydroxylphenylacetate increased. Whether similar changes have been observed in humans, however, has not yet been investigated.

In search for novel diagnostic candidates for human infection, a comprehensive metabolic profiling approach was used for investigation of plasma and cerebral spinal fluid (CSF) samples from trypanosomiasis *rhodesiense* patients and matched uninfected controls (from Uganda), using a combination of NMR- and MS-based analytical platforms. Results from the plasma analyses have since been published, with a copy of the research article attached in the Appendix Publications Section A, entitled *"Discovery of Infection Associated Metabolic Markers in Human African Trypanosomiasis"*.[275]

5.2 Objectives

The primary aim of this study was to use an untargeted metabolic profiling approach to determine potential biomarker candidates that are associated with *T. b. rhodesiense* infection in humans. The

secondary aims were to determine metabolite markers with diagnostic staging capability, as well investigate of possible interactions between host metabolism, immunity and clinical pathology in HAT patients. These aims were addressed using the following specific objectives:

- To identify differences in metabolic phenotypes of plasma and CSF samples between HAT *rhodesiense* patients and controls
- To assess metabolic differences in plasma and CSF between early and late stage HAT patients
- To explore potential links between host plasma and CSF metabolites and HAT patients' clinical symptoms
- To investigate correlations between host metabolic markers and cytokine concentrations for plasma and CSF samples of HAT patients

5.3 Experimental Methods

5.3.1 Ethics

Sample collection was performed using protocols approved by ethics committees in Uganda (UNCST for Lwala and Serere samples and Ministry of Health for Tororo Samples) and UK (North of Scotland Research Ethics Committee), conforming to the principles of the Declaration of Helsinki. Consent forms and informations sheets were written in English and further translated into local languages. Consent was recorded as a signature or a thumb-print from the subject, as approved by the UNCST, following verbal explanation. Consent was also obtained from legal guardians for subjects aged under 18.

5.3.2 Subject Information and Sample Collection

For the main study, 70 HAT patients and 21 controls with similar ethnolinguistic background were recruited from two neighbouring districts in Eastern Uganda, Lwala Hospital (Kaberamaido District) and Serere Health Centre (Serere District), between November 2008 and March 2010, organised by collaborator Dr Jeremy Sternberg (Aberdeen University). Details of study sites, recruitment protocols, clinical examinations and sample collections have all been described by MacLean *et al.*(2010).[276] Confirmation of HAT diagnosis was performed through detection of parasites by microscopic examination of blood films/buffy coats, following micro-haematocrit centrifugation. Clinical history was obtained for each subject and a clinical examination was performed upon admission by a physician or experienced nurse. Patients which also had inter-current infections of

malaria, filariasis or schistosomiasis were excluded from the study. Diagnostic staging of trypanosome infection was assigned in accordance with WHO criteria: late stage was differentiated from the early stage by the detection of parasites in the lumbar CSF and/or a white blood cell (WBC) count > 5 µl.[67]

Plasma and CSF samples were collected from all patients prior to treatment, as part of standard diagnostic procedure.[79] Blood samples were collected in EDTA Vacutainer collection tubes (Greiner, United Kingdom) which were centrifuged at 3000 x g for 10 mins for plasma isolation. CSF samples were collected following a lumbar puncture, conducted by a qualified physician or experienced health centre nurse. Aliquots of 1–2 ml plasma and CSF for each sample were immediately frozen and stored in liquid nitrogen until shipment to the UK (*via* air-freight for up to 24 hours on dry ice) and subsequently stored at -80°C until analysis. For the presented work, 51 CSF and 45 plasma samples from HAT patients were available for metabolic investigation. Samples from healthy individuals that were confirmed as non-infected with either *Trypanosoma sp.* or *Plasmodium sp.* following parasitological assessment, were used as controls. A summary of demographic and clinical characterisations of these subjects is shown in Table 5.1. No significant differences in age distribution or gender ratio between patient and control groups, were found.

Additional CSF samples that had been collected with consent, using similar protocols from an earlier study between 1998–1999 from 9 healthy individuals from neighbouring Tororo District (LIRI Hospital) and stored at -80°C, were also used for metabolic profiling, as controls.[1]

5.3.3 Data Acquisition and Analyses

Metabolic Analyses

Details for sample preparation, analyses and metabolic assignments are described in the General Methods Chapter 2. Plasma and CSF samples were each analysed *via* ^{1}H NMR , with additional UPLC-MS acquisition for plasma. Plasma and CSF cytokine data had been provided by collaborator, Dr Sternberg. Plasma (n=67) and CSF samples (n=60) were prepared and analysed *via* ^{1}H NMR spectroscopy on a 600 MHz Spectrometer (Bruker), using a standard 1D pulse program (ns=64 for plasma and 256 for CSF). Plasma samples were additionally acquired using a CPMG pulse sequence for suppression of broad resonances from lipids/proteins (typically in high abundance in blood/plasma) and improved examination of metabolites with small molecular weights. Separate aliquots of a subset of these plasma samples (n=30) were used for lipid profiling, performed in both positive and negative ESI mode on an Acquity UPLC reverse-phase C_{18} column,

[1] Unfortunately there is no demographic information available for these samples apart from their ages being above 18 years

Table 5.1: Patient and Control Cohort Demographic and Clinical Information

	Plasma Samples (n=67)		**CSF Samples** (n=60)	
	Controls (n=21)	**Patients** (n=46)	**Controls** (n=9)	**Patients** (n=51)
Age, median	31	26	UK	23
Age (IQR) in years	(22–50)	(17–43)	18+	(17–41)
Gender, number of males:females	10:11	19:27	UK	19:26*
Diagnostic Stage, numbers of early:late	–	10:31*	–	13:31*
Inflammatory Symptoms (used for plasma analyses)				
Fever, numbers present:absent	–	24:13*	–	29:14*
Enlarged lymph nodes, number present:absent	–	18:19*	–	21:22*
Hepatomegaly, numbers present:absent	–	22:15*	–	21:22*
Splenomegaly, numbers present:absent	–	22:15*	–	32:11*
Neurological Symptoms (used for CSF analyses)				
Gait, numbers impaired:normal	–	15:25*	–	17:26*
Tremors, numbers present:absent	–	15:25*	–	14:29*
Somnolence, numbers present:absent	–	13:27*	–	14:29*

The table shows patients and sample numbers used in metabolic profiling analysis. An additional 24 patient plasma samples were only used for cytokine measurements (total sample n=69, of which n=55 had measurements for all assessed cytokines).
*Staging and/or clinical information missing from certain patients.
Key: –, not applicable; IQR, interquartile range; UK, unknown (demographic information missing from control CSF samples, though all were taken from adults aged over 18 years)
Modified with permission from Lamour, et al. (2015)[275] Copyright 2015 Lamour, et al.

coupled to a Q-ToF Premier^{TM}(Waters) mass spectrometer.

Metabolic data were initially analysed using PCA, to examine variation in metabolic phenotypes and identify strong outliers. Such analyses resulted in the removal of one NMR patient sample which displayed an acquisition artefact, thus reducing the final number of patient samples to n=45 for NMR data. Initial analysis was followed by PLS-DA and/or O-PLS-DA, to investigate systematic metabolic differences between patients and controls. PLS-DA models were further validated using permutation testing, whereby sample classifiers were randomised 999 times and models re-calculated to establish the likelihood of achieving the same result (i.e. R^2 and Q^2 model parameters) by chance.

For NMR data, all assigned metabolites in plasma and CSF that displayed distinct spectral peaks were used for integral calculation and subsequently compared between patients and controls using Welch (unequal variance) T-test, to extract significantly discriminatory markers (p<0.05 following Benjamini–Hochberg FDR correction). For UPLC-MS data, the first round of selection of HAT discriminatory features was based on O-PLS-DA S-plots, performed separately for each ionisation mode, where each feature surpassed the minimum covariance threshold of 0.05 and minimum correlation coefficient of 0.5 for correlation. This was followed by statistical comparisons between the features' integrals from patient and controls, also using Welch T-test and Benjamini–Hochberg FDR correction.

In order to assess the combined accuracy for HAT classification of discriminatory metabolites for biomarker potential, their integrals were re-imported as loading variables for additional O-PLS-DA models (performed separately for NMR and MS data). Two sets of O-PLS-DA models were calculated for each dataset: firstly, using integrals of all significant discriminatory metabolites, and secondly using only integrals of the top 5 discriminating metabolites. The latter was defined as those which displayed the greatest percentage difference between patients and controls, which represented markers with most practical diagnostic potential in the clinic. The total numbers of correctly and falsely assigned samples (to either HAT or controls) were counted, based on separation between the two groups along the model's first predictive component. These numbers were used to calculate combined sensitivity (true positives/[true positives + false negatives]) and specificity (true negatives/[true negatives + false positives]) — expressed as percentages — of assessed discriminatory variables in patient/control classification. 95% binomial proportion confidence intervals (CI) were subsequently calculated for each measured sensitivity/specificity value, by multiplying the standard error by 1.96 (approximation of the 97.5 percentile point of the normal distribution), and subtracting or adding this value to the mean (upper limit capped at 100%).

Further potential differences between categorical patient groups in plasma and CSF metabolite integrals, i.e. between early and late stage HAT, presence and absence of inflammatory symptoms (for plasma)[2] or neurological symptoms (for CSF), gender, etc., were examined *via* PCA, O-PLS-DA, and Welch T-tests (with Benjamini–Hochberg correction). Associations between metabolite integrals and numerical data (e.g. white blood cell count and subject age) were assessed using Pearson correlation, with those that displayed $p<0.05$ after Benjamini–Hochberg FDR correction, being assigned as statistically significant.

Cytokine Measurements and Correlation Analyses

In addition to metabolic analyses, cytokine concentrations had been measured in plasma and CSF samples *via* ELISA, as detailed in Methods Chapter Section 2.6.1. Plasma cytokine measurements of IL-10, IL-6, IFNγ, and TNFα were acquired for $n=81$ samples: 12 controls and 69 HAT patients, though only 55 of patient samples contained complete results for all four cytokines. IL-10, IL-6 and IFNγ concentrations were also measured in the CSF of $n=48$ of patients. Statistically significant differences between cytokine concentrations from patients and controls (for plasma samples only), and between late and early stage patients (for both plasma and CSF) were assessed using the Mann–Whitney U-test, followed by Benjamini–Hochberg FDR correction ($p<0.05$).

[2] Discriminatory analyses between diagnostic stage and presence/absence of fever were not performed for plasma UPLC-MS data, due to low numbers ($n=2$) in one of the analysed groups.

To further explore possible links between markers observed across the different platforms, correlation analyses were performed between two datasets (i.e. 2D matrix correlation) for each biofluid separately, namely between NMR integrals and cytokine concentrations (for plasma and CSF), NMR and UPLC-MS integrals, and UPLC-MS integrals and cytokine concentrations (for plasma only). Correlation analyses could only be performed in the subsets of patients for which results were complete across the two analysed datasets; corresponding sample numbers have been summarised in Figure 5.1. Plasma metabolite integrals were additionally correlated with CSF metabolite integrals, to investigate metabolic associations between the two biofluids (n=22 HAT patients). These analyses were based on the ^{1}H NMR results as the data were acquired using the identical spectrometer at similar times (i.e. continuous acquisition of plasma samples, immediately followed by CSF samples). Furthermore, metabolite integrals in ^{1}H NMR spectra are truly quantitative in nature, thus allowing for direct comparisons and extrapolations of metabolite concentrations in CSF and plasma. All measured correlations that exhibited Benjamini–Hochberg FDR corrected p<0.05 were considered as statistically significant.

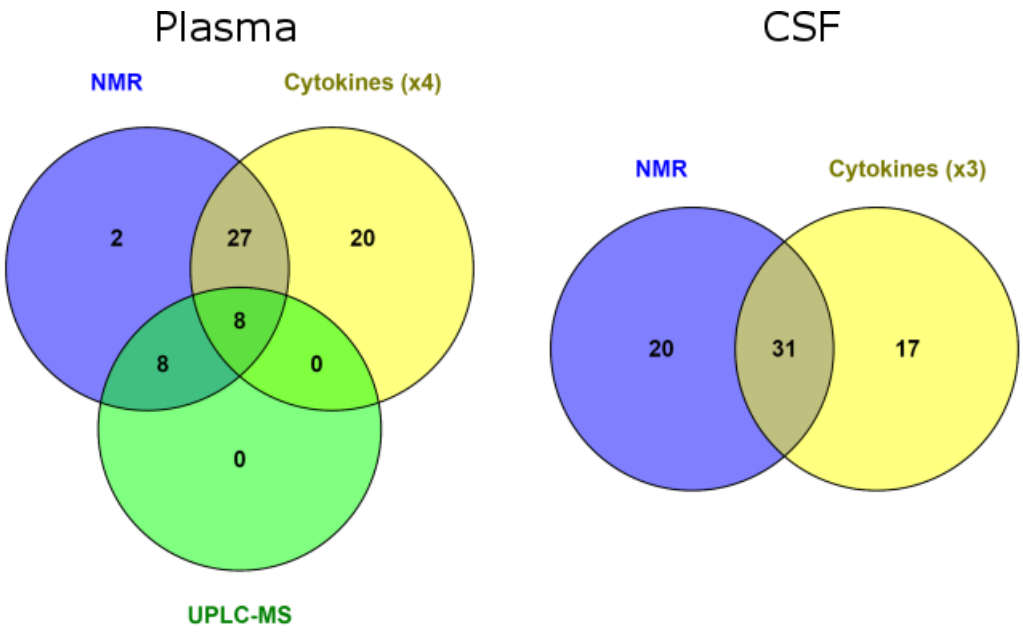


Figure 5.1: Overview of HAT Patient Sample Numbers Analysed Across the Three Platforms

Venn diagrams showing number of plasma and CSF samples from HAT patients that were co-analysed *via* ^{1}H NMR , cytokines, and (for plasma only), UPLC-MS. For plasma, final of number of analysed samples n=45 for NMR, n=16 for UPLC-MS and n=55 for cytokines (complete for all 4 cytokines); for CSF, n=51 for NMR and n=48 for cytokines (complete for all 3). HAT samples with artefactual errors (outliers) or incomplete data were excluded from correlation analyses between datasets. *Figure generated using Venny 2.0*[277]

5.4 Results: Plasma Analyses

5.4.1 Metabolic Profiling via ^{1}H NMR

Each plasma sample was acquired using both a standard 1D pulse programme and CPMG pulse sequence: a representative spectrum for each of these methods is shown in Figure 5.2. The figure confirmed that the multiple broad and/or indiscriminate lipid/protein resonances observed using the standard program — which for the vast majority of samples were expressed as a higher baselines driven by macromolecular components — have been largely suppressed in the CPMG spectrum, facilitating the identification (and calculation of integrals) of lower-molecular metabolites. Consequently, CPMG spectra were solely used for subsequent data analyses and interpretation. In total, 36 endogenous plasma metabolites were confidently assigned, as listed in Appendix Table A.7. In addition, peaks from two exogenous compounds were identified across numerous samples; firstly ethanol, a contaminant that was detected in blank buffer controls but which may have additionally arisen due to its use as a skin disinfectant prior to lumbar puncture and blood collection. The second set of peaks were attributed to a blood collection anticoagulant agent, namely ethylenediaminetetraacetic acid (EDTA)[278], which were all subsequently removed from the spectra for final analyses (see Appendix Table A.7 for spectral resonances of both compounds), to focus on inherent metabolic variation. Furthermore, one HAT patient sample was excluded from final analyses due to the presence of artefactual spectral peaks (unusual dips in the baseline), bringing the final total number of HAT patients to 45 (vs. 21 controls) for NMR analyses.

Metabolic Differences Linked with HAT

Investigations using PCA models showed that the inherent differences between HAT patients infected with *T. b. rhodesiense*, and uninfected controls dominated the overall variation across plasma samples, as shown in Figure 5.3A. Partial segregation between these two populations was observed already along the first component in PCA, implying that HAT infection generated a strong host metabolic response. Differences between patients and controls were further assessed using O-PLS-DA, which demonstrated complete separation between HAT patients and controls, also along the first (predictive) component, shown in Figure 5.3B. This model was shown to be highly robust, as indicated by the high model predictive capability parameter (Q^2Y) of 0.708. Additional validation through permutation testing showed that the likelihood of achieving the same discriminatory result by chance was zero (p=0 for both R^2Y and Q^2Y), as presented in Figure 5.4.

Examination of the corresponding loadings plots revealed that multiple lipid peaks accounted

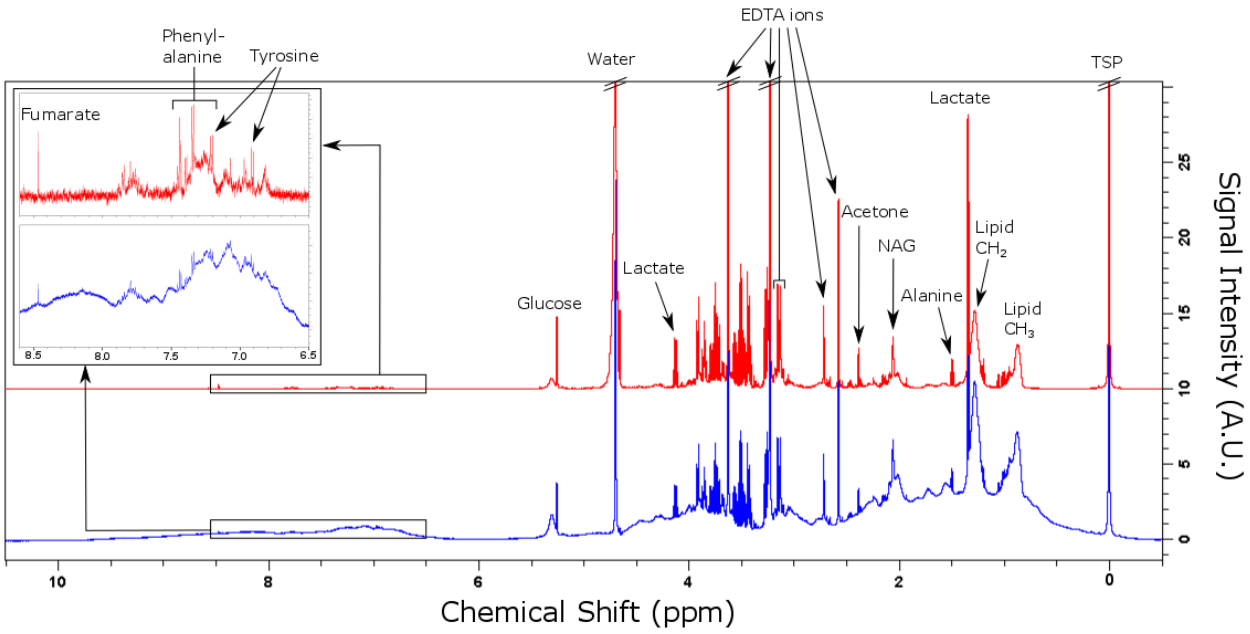


Figure 5.2: Representative ^{1}H NMR Spectra of Human Plasma

Plots show CPMG spectrum (in red) above the corresponding standard 1D spectrum (in blue), at the same magnification, of a representative plasma sample (from a HAT patient). Peaks within aromatic region have been magnified, shown in boxed inset. Key: A.U., arbitrary units; EDTA, ethylenediaminetetraacetic acid; NAG, *N*-acetylglycoprotein; TSP, trimethylsilylpropionic-2,2,3,3-d4 acid (reference standard).

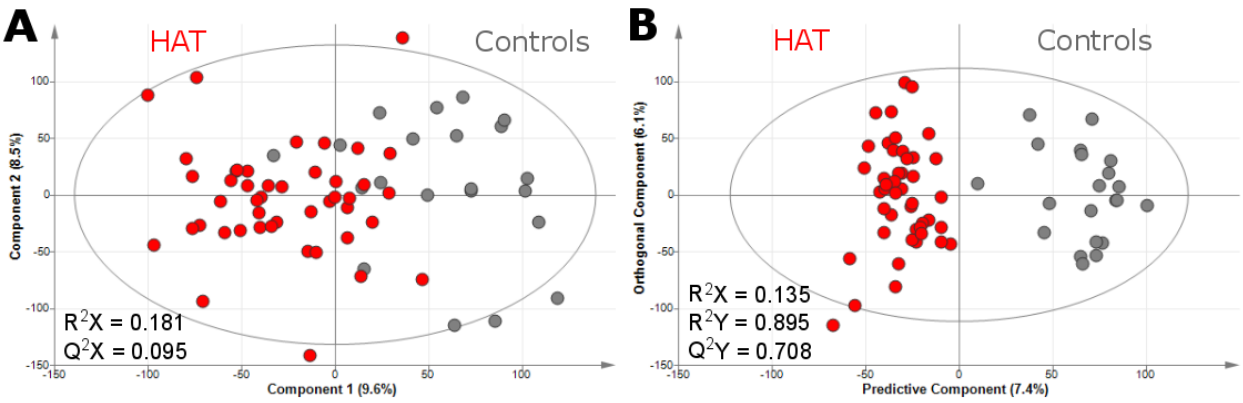


Figure 5.3: Strong Metabolic Response Associated with HAT

PCA model (**A**) and O-PLS-DA model (**B**) score plots of plasma ^{1}H NMR spectra across HAT patients and controls. Each circle represents a spectrum from one sample, whereby patients are presented in red (n=45) and controls in dark grey (n=21). Cumulative values for model parameters are shown on each plot, whereas individual component contribution of R^2X are shown on the axes, as percentages. Key: R^2X, model fit parameter for variation in spectral data; R^2Y, model fit parameter for variation in classifier data (for O-PLS-DA); Q^2, model predictive parameter for spectral data in PCA (Q^2X) and for classifier data in O-PLS-DA (Q^2Y).
Modified with permission from Lamour, et al. (2015)[275] Copyright 2015 Lamour, et al.

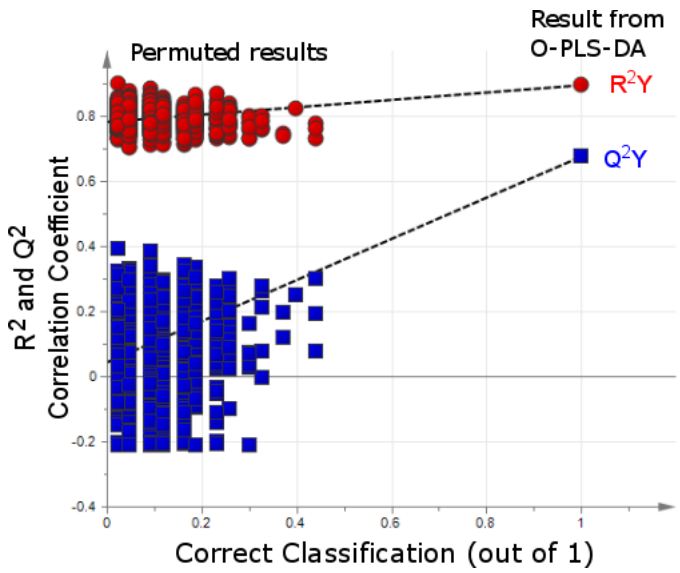


Figure 5.4: Permutation Analyses of PLS-DA Model Comparing HAT Patients with Controls

Descriptive classifying y-matrix data of plasma samples (i.e. identification as HAT patient or control) were randomly permuted 999 times and newly generated R^2Y and Q^2Y values were plotted and compared to values from the original PLS-DA model. P-values were subsequently calculated as the number of times the permuted values exceeded the original model value.

for the majority of the differentiation between the two groups. Thirty-one of the 36 identified endogenous metabolites displayed distinct peaks (without being superimposed by other overlapping resonances, e.g. from lipids), that were used for integral calculation as a measure of relative metabolite concentration, listed in Table 5.2. Integrals were subsequently compared using Welch-T-test to determine differences between patients and controls, revealing 16 significant discriminatory metabolites. Such metabolites included several plasma amino acids, e.g. phenylalanine (which displayed the largest percentage difference), found in higher concentrations in HAT patients than controls, and alanine and histidine, which were lower in patients. Marked perturbations in lipid metabolism were observed in HAT patients, which displayed reduced concentrations of the majority of lipid moieties, including those which contained unsaturated bonds (CH=CH and $C{=}CCH_2C{=}C$), lysyl group in albumin, and cholesterol. On the other hand, levels of lipid ketene (CH_2CO) and glyceryl groups were higher HAT than in controls. Other significant differences included higher concentrations of creatinine, formate, *N*-acetyl glycoproteins, and *myo*-inositol in patients, whereas citrate concentrations were higher in controls.

In order to determine the collective capability of these metabolites, in correctly discriminating between patients and controls, additional O-PLS-DA models were generated. The first model was based on integrals of all 16 discriminatory metabolites as loading variables, and a second O-PLS-DA model was generated using integrals of only the top 5 metabolites[3]: phenylalanine, formate and

[3]The latter were defined as those which displayed the biggest measured percentage difference between HAT and

Table 5.2: Plasma Metabolite Integrals Measured via ^{1}H NMR Spectroscopy

Metabolites	Integrated Peak (multiplicity)	Average Integrals in Controls (±SEM) in A.U.	Average Integrals in HAT (±SEM) in A.U.	HAT - Control	
				% Difference	P-value
3-Hydroxybutyrate	2.32(m)	29.66 (± 0.89)	32.09 (± 1.31)	8.20%	NS
Acetate	1.94(s)	6.60 (± 0.38)	6.71 (± 0.46)	1.70%	NS
Acetoacetate	2.25(s)	8.93 (± 0.72)	10.85 (± 0.85)	21.50%	NS
Alanine	1.49(d)	57.65 (± 2.18)	46.84 (± 1.97)	-18.70%	1.60E-03
Citrate	2.68(d)	24.97 (± 0.64)	22.87 (± 0.43)	-8.40%	2.01E-02
Creatine	3.94(s)	8.17 (± 0.38)	10.24 (± 0.85)	25.40%	NS
Creatinine	4.07(s)	7.26 (± 0.35)	8.61 (± 0.30)	18.70%	1.30E-02
Formate	8.47(s)	1.00 (± 0.13)	1.40 (± 0.08)	41.00%	2.13E-02
Glucose	5.25(s)	76.87 (± 4.92)	79.49 (± 3.71)	3.40%	NS
Glutamine	2.14(m)	67.08 (± 2.09)	67.64 (± 1.58)	0.80%	NS
GPC	4.32(m)	19.97 (± 1.06)	18.27 (± 0.33)	-8.50%	NS
Glycine	3.58(s)	21.79 (± 1.34)	19.20 (± 0.68)	-11.90%	NS
Histidine	7.08(s)	3.12 (± 0.16)	2.58 (± 0.10)	-17.40%	1.79E-02
Lactate	1.34(d)	531.15 (± 97.44)	503.81 (± 55.89)	-5.10%	NS
Lipid C=CCH_2C=C	2.78(m)	53.59 (± 2.71)	36.07 (± 1.08)	-32.70%	1.35E-05
Lipid CH_3	0.88(m)	718.89 (± 33.00)	495.75 (± 13.50)	-31.00%	8.46E-06
Lipid CH_2-C=C	2.02(m)	161.76 (± 7.13)	147.48 (± 3.07)	-8.80%	NS
Lipid CH_2CH_2CO	1.51(m)	14.21 (± 0.39)	10.63 (± 0.20)	-25.20%	5.28E-08
Lipid CH_2CH_2-C=C	1.58(m)	59.26 (± 7.19)	74.80 (± 2.74)	26.20%	NS
Lipid CH_2CO	2.23(m)	36.29 (± 4.35)	54.84 (± 1.89)	51.10%	1.60E-03
Lipid CH_2	1.28(m)	992.66 (± 79.79)	981.73 (± 33.16)	-1.10%	NS
Lipid Cholesterol	0.68(m)	6.50 (± 0.21)	4.52 (± 0.15)	-30.40%	5.28E-08
Lipid Glyceryl CH	5.20(m)	4.53 (± 0.39)	6.28 (± 0.22)	38.80%	1.48E-03
Lipid ε-CH_2 lysyl	2.93(m)	23.85 (± 0.81)	18.82 (± 0.49)	-21.10%	2.53E-05
Lipid CH=CH	5.30(m)	154.75 (± 8.53)	101.50 (± 3.89)	-34.40%	2.10E-05
myo-Inositol	4.087(t)	6.98 (± 0.32)	8.21 (± 0.29)	17.60%	1.58E-02
N-acetylglycoproteins	2.06(m), 2.09(m)	233.03 (± 4.68)	270.13 (± 4.67)	15.90%	7.02E-06
Phenylalanine	7.45(t)	4.92 (± 0.25)	7.44 (± 0.46)	51.20%	4.18E-05
Pyruvate	2.38(s)	17.64 (± 2.05)	26.99 (± 4.27)	53.00%	NS
Tyrosine	6.92(d)	6.19 (± 0.38)	6.87 (± 0.31)	11.00%	NS
Valine	1.05(d)	21.90 (± 1.37)	22.10 (± 0.79)	0.90%	NS

Key: (d), doublet; (m), multiplet; (s), singlet; (t), triplet; GPC, glycerophosphocholine; NS, not stastistically significant; SEM, standard error of the mean.
Modified with permission from Lamour, et al. (2015)[275] Copyright 2015 Lamour, et al.

lipid moieties CH_2CO, CH=CH, and glyceryl. Results are shown in Table 5.3, which demonstrate that both models were robust (with Q^2Y values greater than 0.5) and that the collection of these markers had the power to correctly assign HAT patients from controls with up 95.24% sensitivity and 88.9% sensitivity.

Metabolic Differences Linked with Disease Stage and Clinical Symptoms

In addition to investigating HAT associated differences in host metabolism, possible disparities between plasma profiles of early stage and late stage patients were also analysed. No significant differences between the two patient groups were observed, using either PCA or O-PLS-DA (which was not predictive with a negative Q^2Y value), as shown in Appendix Figure A.6. Further comparisons of metabolite integrals between the two patient groups *via* Welch T-test initially

controls, representing metabolites of most practical diagnostic value.

Table 5.3: Accuracy of the Collection of Plasma NMR Discriminatory Markers for HAT

	Using All 16 Markers (R^2Y=0.736, Q^2Y=0.657)		**Using Top 5 Markers** (R^2Y=0.641, Q^2Y=0.582)	
	Actual HAT Patients	Actual Controls	Actual HAT Patients	Actual Controls
Predicted HAT Patients	40 (TP)	1 (FP)	38 (TP)	1 (FP)
Predicted Controls	5 (FN)	20 (TN)	7 (FN)	20 (TN)
	Sensitivity = TP/[TP+FN] = 88.89%	Specificity = TN/[TN+FP] = 95.24%	Sensitivity = TP/[TP+FN] = 84.44%	Specificity = TN/[TN+FP] = 95.24%
Standard error	4.68%	4.65%	5.40%	9.35%
Upper 95% CI	98.07%	100%*	95.03%	100%*
Lower 95% CI	79.71%	86.13%	73.85%	86.13%

Confusion matrix showing calculations for sensitivity and specificity, based on sample HAT/control assignment from O-PLS-DA models, using NMR peak integrals as loading variables. * Calculated upper CI was capped to absolute maximal value of 100%. Key: CI, confidence interval; FN, false negatives; FP, false positives; TN, true negatives; TP, true positives.
Modified with permission from Lamour, et al. (2015)[275] Copyright 2015 Lamour, et al.

showed higher levels of creatine (+49.7% relative to early stage), formate (+43.9%) and lactate (+50.7%) in late stage patients, yet none of these differences retained their statistical significance after Benjamini–Hochberg FDR correction.

Similar analyses were also performed with patient plasma metabolite integral data, to determine potential differences between patients that presented with inflammatory clinical manifestations — namely fever and swelling/enlargement of lymph nodes, liver, and spleen (which were observed in approximately half of all patients — see Table 5.1 for patient numbers) — and those that did not. No significant differences were observed between the tested patient groups *via* PCA or O-PLS-DA (which all displayed negative Q^2Y values). Additional Welch T-tests demonstrated no significant differences in metabolite integrals after FDR correction. Furthermore, of the metabolites that displayed $p<0.05$ prior to FDR correction, none included HAT discriminatory metabolites, e.g. enlarged lymph nodes were linked with increased concentrations of 3-hydroxybutyrate (3HB; +22.0%, relative to non-symptomatic patients), acetoacetate (+54.6%), tyrosine (+22.1%) and valine (+15.1%).

Metabolic Variation Associated with Confounding Factors

Other variation across all plasma NMR data were also investigated, in order to determine possible effects due to confounding factors such as gender, age and sample collection site. No differences were determined between age groups or gender, as measured by PCA or O-PLS-DA models (the latter were not predictive, with negative Q^2Y values). Subtle differences between the two hospital sites

was determined, though these did not include HAT discriminatory metabolites: samples collected in Serere hospital had significantly increased lactate, glutamine, and branched-chain amino acids whilst those taken from Lwala displayed higher glutamate concentrations (O-PLS-DA model R^2Y was 0.898 and Q^2Y was 0.609).

5.4.2 Lipid Profiling via UPLC-MS

Following the discovery of pronounced differences in lipid metabolites between HAT patients and controls, lipid plasma profiles were further examined *via* reverse-phase UPLC-MS. In total, 8771 distinct metabolic features were detected in ESI+ mode and 2543 features in ESI-. Following exclusion of peaks with high variance (CV>30% across the QC's) and contaminants (detected in either the blank controls or QC samples only), this number was reduced to a final of 6952 features in ESI+ and 1714 features in ESI-. Assessment of overall variation across test samples (i.e. patient and control samples) and QC's *via* PCA showed close clustering of QC samples, compared with the rest of the samples, as shown in Figure 5.5, thus confirming the robustness of the experimental runs in both modes.

Lipid Changes Linked with HAT

As seen with the NMR data, lipid profiling PCA results also showed a dominant separation between HAT patients and controls already in the first two components, observed in both ionisation modes, demonstrated in Figure 5.5. Differences in profiles between the two groups were also further enhanced in corresponding O-PLS-DA models and additionally validated using permutation tests (permutation results for ESI+, p=0.004 for R^2Y and p=0 for Q^2Y; for ESI-, p=0 for both R^2Y and Q^2Y). To facilitate visualisation of the variables that contributed the most towards O-PLS-DA class discrimination (i.e. HAT patients vs. controls), the lipid features were plotted as S-plots, as demonstrated in Figure 5.6, where cut-off values of 0.5 for correlation and 0.05 for covariance were chosen to select the variables with the most robust discriminatory power. These amounted to 53 features in ESI+ and 44 features in ESI- in total. The differences between HAT patients and controls in their integrals were tested *via* Welch T-Test and Benjamini–Hochberg correction, which showed that all were significant (p<0.05). Selected discriminatory features that have been highlighted in the S-plots, reveal that these features included lysophosphatidylcholines (LysoPC's) 16:0 and 18:0, which were significantly lower in HAT patients than controls, and phosphatidylcholines (PC's) 32:0 and 34:1, which were higher in HAT.

Following the exclusion of adducts, fragments, and molecules containing non-dominant isotopes,

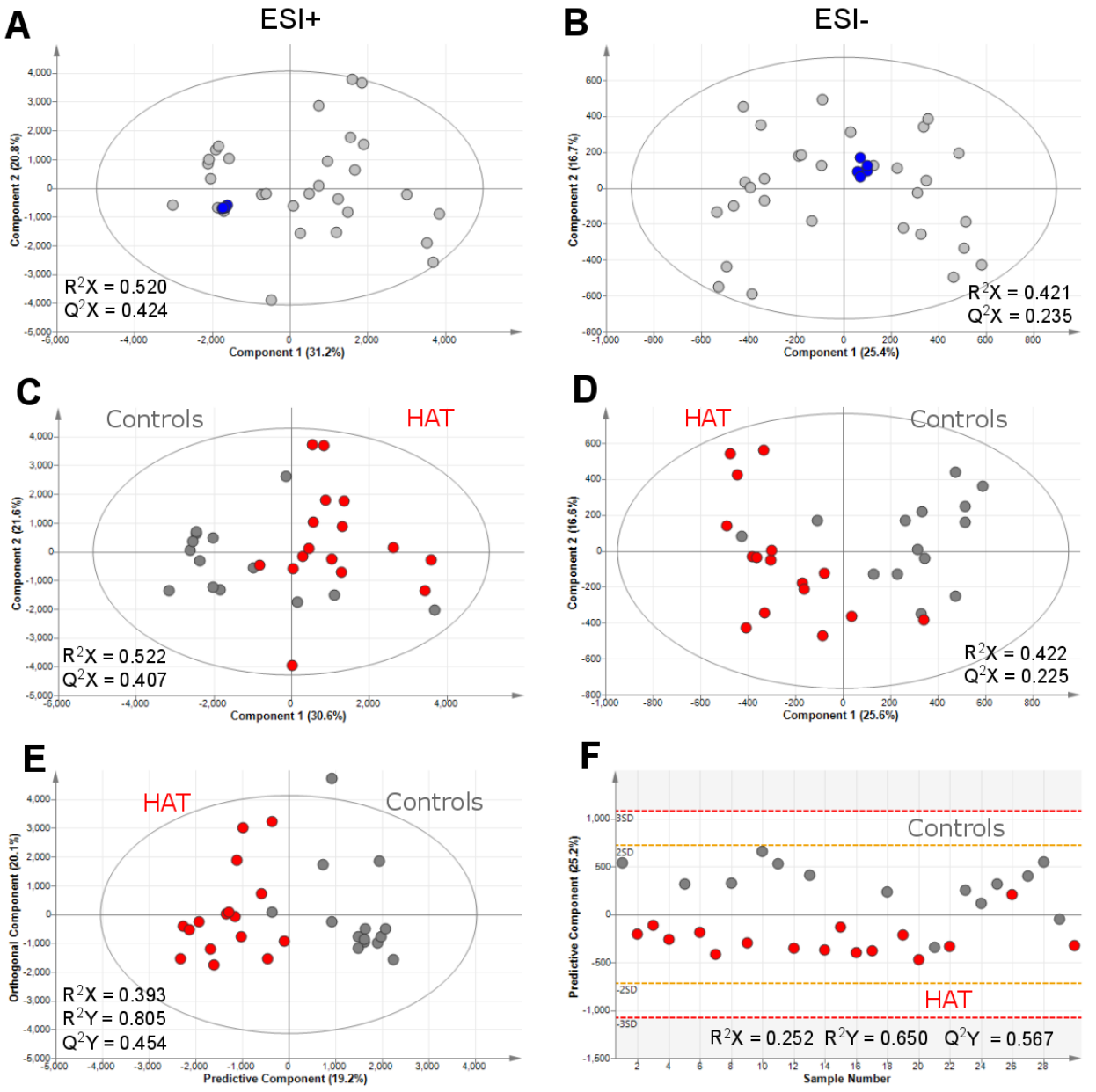


Figure 5.5: Marked Disparities in Plasma Lipid Profiles between HAT Patients and Controls

Figures show scores plots plasma samples base on lipid UPLC-MS profiles, each constructed using up to two components. PCA models show all plasma test samples (in light grey, n=30) and QC's (in blue, n=6) in positive ionisation (ESI+) mode (**A**) and negative (ESI-) mode (**B**). PCA models of plasma test samples only, comparing HAT patients (in red, n=16) vs. controls (in dark grey, n=14), in ESI+ (**C**) and ESI- (**D**). O-PLS-DA comparing HAT samples vs. controls (colours and group sizes identical to PCA's) for ESI+ (**E**), using both predictive and orthogonal components, and for ESI- (**F**), which only used the first predictive component (orthogonal component reduced Q^2 and was thus omitted). Key as described for Figure 5.3, with the following addition: SD, standard deviation.
Modified with permission from Lamour, et al. (2015)[275] Copyright 2015 Lamour, et al.

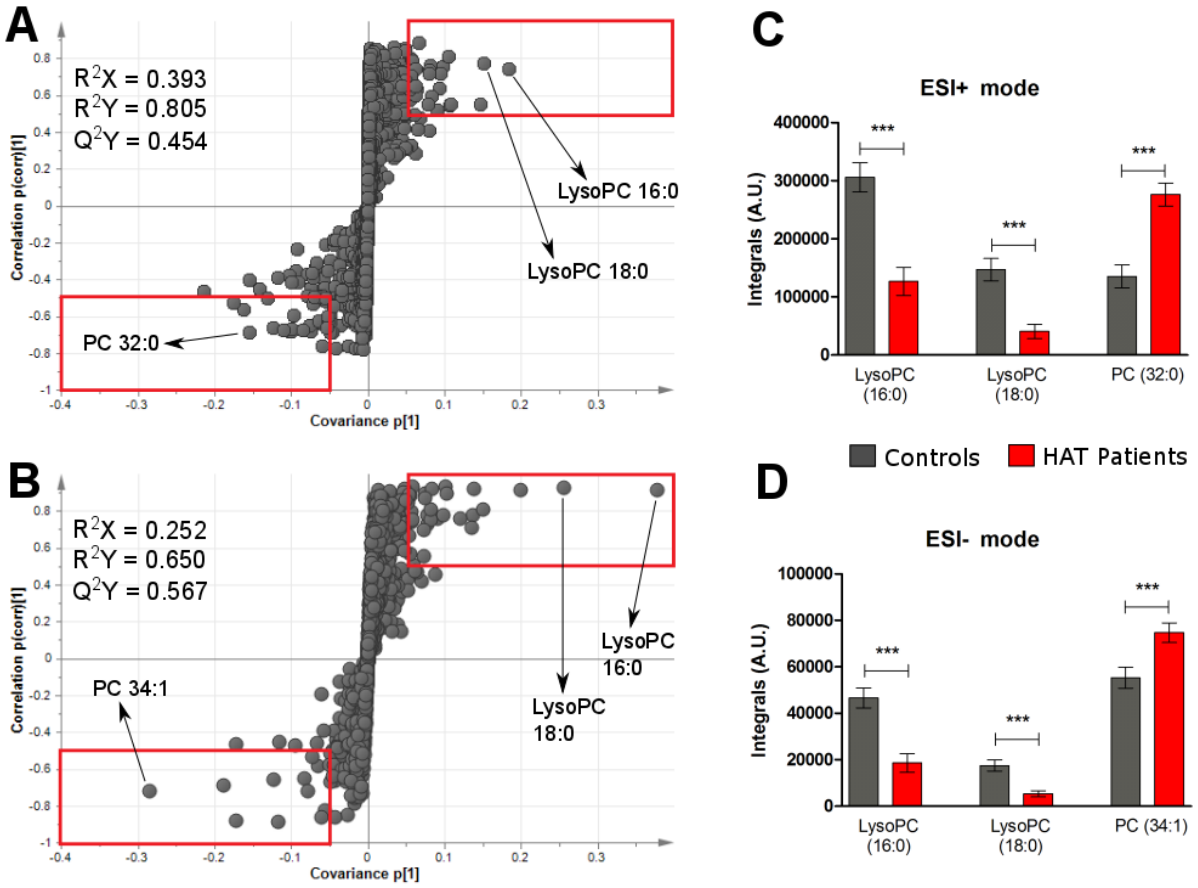


Figure 5.6: Selection of HAT Discriminatory Lipid Markers

S-plots for O-PLS-DA models plasma UPLC-MS features measured in ESI+ mode (**A**) and ESI- mode (**B**) — see Figure 5.5 for corresponding scores plot. Each circle represents one lipid feature with a unique combination of m/z and retention time. Discriminatory features that surpassed correlation and covariance threshold criteria are highlighted in red boxes, with the integrals of those labelled in the S-plots additionally shown as bar-charts (**C**) for ESI+, and (**D**) for ESI-. Charts showing mean averages (± standard error of the mean as error bars), with HAT patients in red (n=16) and controls in dark grey (n=14). Significant differences are labelled with asterisks, where *** p<0.001. Key: A.U., arbitrary units; LysoPC, lysophosphatidylcholine; PC, phosphatidylcholine.
Modified with permission from Lamour, et al. (2015)[275] Copyright 2015 Lamour, et al.

a final of 37 unique lipid species were identified as statistically significant discriminatory makers, 12 of which were determined as significantly different in both ESI modes, listed in Table 5.4. Such features included six LysoPC's (16:0, 18:0, 18:1, 18:2, 20:3 and 20:4), four PC's (32:0, O-34:1, 38:3 and 38:5) and two sphingomyelin (SM) species (d40:1 and d41:1). The majority of lipid classes analysed were glycerophospholipids (namely LysoPC's and PC's), followed by sphingolipids (mainly SM's) and triglycerides (TG's). Free fatty acids were detected using our MS methods yet none of these significantly differed between patients and controls. A complete list of all discriminatory lipid UPLC-MS assignments is shown in Appendix Table A.8, which contains information on lipid (sub-)classes, retention times, m/z ratios, chemical formulae, observed ions, and mass accuracies from METLIN database (which were all within 6 ppm).

Results showed that all six analysed lysoPC's were significantly lower in patients than controls, as well as eight out of the nine analysed sphingolipids, including seven sphingomyelins and nervonic ceramide (Cer d42:2). Only levels of SM d34:1, the smallest of the analysed SM species, were higher in HAT patients than controls. All analysed TG species (6 in total) were elevated in HAT, in contrast to cholesterol ester (CE) 18:2, which was lower in HAT than controls. This result is similar to the presented NMR findings, where HAT patients had higher lipid glyceryl moieties and lower cholesterol levels.

The observed increases and decreases in the lipid species' integrals between patients and controls were generally similar across the lipid subclasses, with the exception of the phosphatidylcholines. Instead, UPLC-MS results indicate that PC's containing a high number of saturated bonds (e.g. PC's 38:5, 38:6, and 40:5) were significantly lower in patients, whilst those with two or fewer unsaturated bonds (e.g. PC's 32:0, 32:1, and 34:1) were higher. This finding is also mirrored the NMR results, where lipid moieties with unsaturated bonds (CH=CH and $C{=}CCH_2C{=}C$) were significantly lower in patients than controls (Table 5.2).

As with NMR data, additional O-PLS-DA models were generated using the integrals all 37 discriminatory markers from both ionisation modes, followed by further models using only the top 5 features with the greatest percentage difference (shown in Table 5.4). As the majority of the largest differences were observed in the positive mode, the top 5 features from this mode only were chosen, which were as follows: PC(32:0), TG(51:2), SM(d34:1), PC(O-34:1/P-34:0), and TG (53:2). The accuracy in HAT vs. control discrimination of these markers were high, giving identical sensitivity and specificity values (93.8% and 85.7% respectively) for models using either all 37 discriminatory markers or top 5, as shown in the confusion matrix in Table 5.5. For exploratory purposes, the top 5 UPLC-MS markers were also combined with the top 5 NMR markers, to determine whether the

Table 5.4: Disease Discriminatory Lipid Species Measured in Plasma via UPLC-MS

	Lipid Species	ESI+		ESI-	
		P-value	% Difference	P-value	% Difference
Higher in HAT	Cer (d42:2)	–	–	2.51E-03	54.30%
	PC (32:0)	2.49E-05	104.20%	2.58E-03	111.70%
	PC (32:1)	–	–	1.29E-04	62.50%
	PC (34:1)	–	–	8.21E-03	35.00%
	PC (34:2)	–	–	1.29E-04	40.20%
	PC(O-34:1)/ PC(P-34:0)	3.93E-04	86.20%	2.00E-03	84.30%
	PE (36:2)	–	–	2.51E-03	54.40%
	PS (39:3)	–	–	4.42E-03	37.50%
	SM (d34:1)	1.75E-02	96.10%	–	–
	TG (51:2)	6.52E-05	98.00%	–	–
	TG (52:3)	4.40E-03	45.40%	–	–
	TG (53:2)	6.73E-04	86.10%	–	–
	TG (53:3)	4.02E-04	77.50%	–	–
	TG (54:2)	3.74E-04	70.30%	–	–
	TG (54:5)	2.20E-03	68.20%	–	–
Lower in HAT	CE (18:2)	1.12E-04	-53.10%	–	–
	LysoPC (16:0)	1.93E-05	-58.60%	5.42E-04	-60.10%
	LysoPC (18:0)	1.11E-04	-72.60%	4.42E-03	-69.50%
	LysoPC (18:1)	1.98E-05	-66.50%	1.00E-04	-62.50%
	LysoPC (18:2)	3.28E-04	-80.50%	3.18E-04	-78.20%
	LysoPC (20:3)	5.97E-05	-75.00%	7.00E-04	-77.90%
	LysoPC (20:4)	4.36E-06	-75.60%	2.62E-04	-73.10%
	PC (40:5)	1.13E-02	-39.20%	–	–
	PC (36:3)	6.56E-03	-33.90%	–	–
	PC (38:3)	1.93E-03	-43.00%	2.51E-03	-30.70%
	PC (38:5)	1.71E-02	-40.90%	1.29E-04	-28.80%
	PC (38:6)	1.63E-02	-33.30%	–	–
	PC (40:4)	3.25E-03	-47.90%	–	–
	PC (40:6)	2.94E-02	-32.70%	–	–
	PC(O-34:3)/ PC(P-34:2)	8.43E-04	-52.80%	–	–
	SM (d38:1)	–	–	3.46E-03	-37.60%
	SM (d39:1)	–	–	4.42E-03	-52.10%
	SM (d40:1)	1.01E-03	-43.90%	3.46E-03	-48.60%
	SM (d40:2)	–	–	2.58E-03	-60.20%
	SM (d41:1)	1.06E-03	-58.60%	2.00E-03	-48.20%
	SM (d41:2)	–	–	4.42E-03	-55.30%
	SM (d42:1)	–	–	3.92E-04	-52.50%

List of lipid species that significantly discriminated between HAT patients and controls. Species that were not detected in a particular ESI mode have been omitted. See Appendix Table A.8 for further information of molecular species. Key: –, no discriminatory marker detected in that ionisation mode; Cer, ceramide; CE, cholesterol ester; LysoPC, lysophosphatidylcholine; LysoPE, lysophosphatidylethanolamine; NS, not statistically significant; PC, phosphatidylcholine; PE, phosphatidylehtanolamine; PS, phosphatidylserine; RT, retention time; SM, sphingomyelin; TG, triglyeride. *Modified with permission from Lamour, et al. (2015)[275] Copyright 2015 Lamour, et al.*

collective accuracy would increase, but sensitivity and specificity values were the same as using the MS markers on their own.

Table 5.5: Accuracy of the Collection of Plasma UPLC-MS Discriminatory Markers for HAT

	Using All 37 Markers (R^2Y=0.780, Q^2Y=0.555)		**Using Top 5 Markers** (R^2Y=0.558, Q^2Y=0.505)	
	Actual HAT Patients	Actual Controls	Actual HAT Patients	Actual Controls
Predicted HAT Patients	15	2	15	2
Predicted Controls	1	12	1	12
	Sensitivity	Specificity	Sensitivity	Specificity
	93.75%	85.71%	93.75%	85.71%
Standard error	6.05%	9.35%	6.05%	9.35%
Upper 95% CI	100%*	100%*	100%*	100%*
Lower 95% CI	81.89%	67.38%	81.89%	67.38%

Confusion matrix showing calculations for sensitivity and specificity, based on sample HAT/control assignment from O-PLS-DA models, using integrals of discriminatory lipid UPLC-MS features as loading variables. Key and calculations as described for Table 5.3.
Modified with permission from Lamour, et al. (2015)[275] Copyright 2015 Lamour, et al.

Variation in Lipid Profiles Linked with Clinical Symptoms and Confounding Factors

Potential differences between discriminatory lipid features associated with the presence/absence of clinical inflammatory symptoms were investigated as was done for NMR data. Significant differences were observed between patients with or without swollen lymph nodes, with those with swollen glands displaying significantly higher levels of SM (d34:1) and PS (39:3) — both of which were also shown to be higher in HAT than controls — as well as lower levels of SM (d39:1 — which was also previously shown as lower in patients than controls. No significant differences were determined between patients who presented with or without splenomegaly, and hepatomegaly. Differences in diagnostic stage and fever status were not tested due to insufficient numbers of early stage patients and non-feverish patients, respectively (n=2 in both cases). Furthermore, no significant associations were determined between discriminatory lipid features and patient age, gender, or sample collection site.

Correlations between NMR and Lipid UPLC-MS Data

Metabolite integrals obtained from ^{1}H NMR analyses in HAT patients were correlated with corresponding UPLC-MS discriminatory markers, *via* 2D Pearson-based matrix correlation, to investigate associations between the two complementary metabolic profiling platforms. Whilst several correlations were observed, none of these correlations remained statistically significant following Benjamini–Hochberg FDR correction. Correlations with p<0.05 prior to FDR correction have

been presented in Figure 5.7, which highlight the (albeit statistically weak) positive associations between lipids detected *via* NMR and lipid MS species, e.g. between lipid cholesterol and cholesterol ester (18:2) and between lipid glyceryl and triglycerides species. Lipid species that were detected in both ionisation modes displayed similar correlations, e.g. lysoPC's (16:0 and 20:4) and PC (38:5), as expected. Similar correlation patterns were observed within lipid classes (e.g. sphingolipids ceramide and sphingomyelins) and amongst polyunsaturated PC's. Correlative pairs with a corrected $0.05>p<0.1$ have been highlighted in Appendix Figure A.7, to show the strongest trends that may still be of biological interest. These results highlighted that PC 38:5 (in ESI-) displayed the highest number of associations with lipids measured *via* NMR.

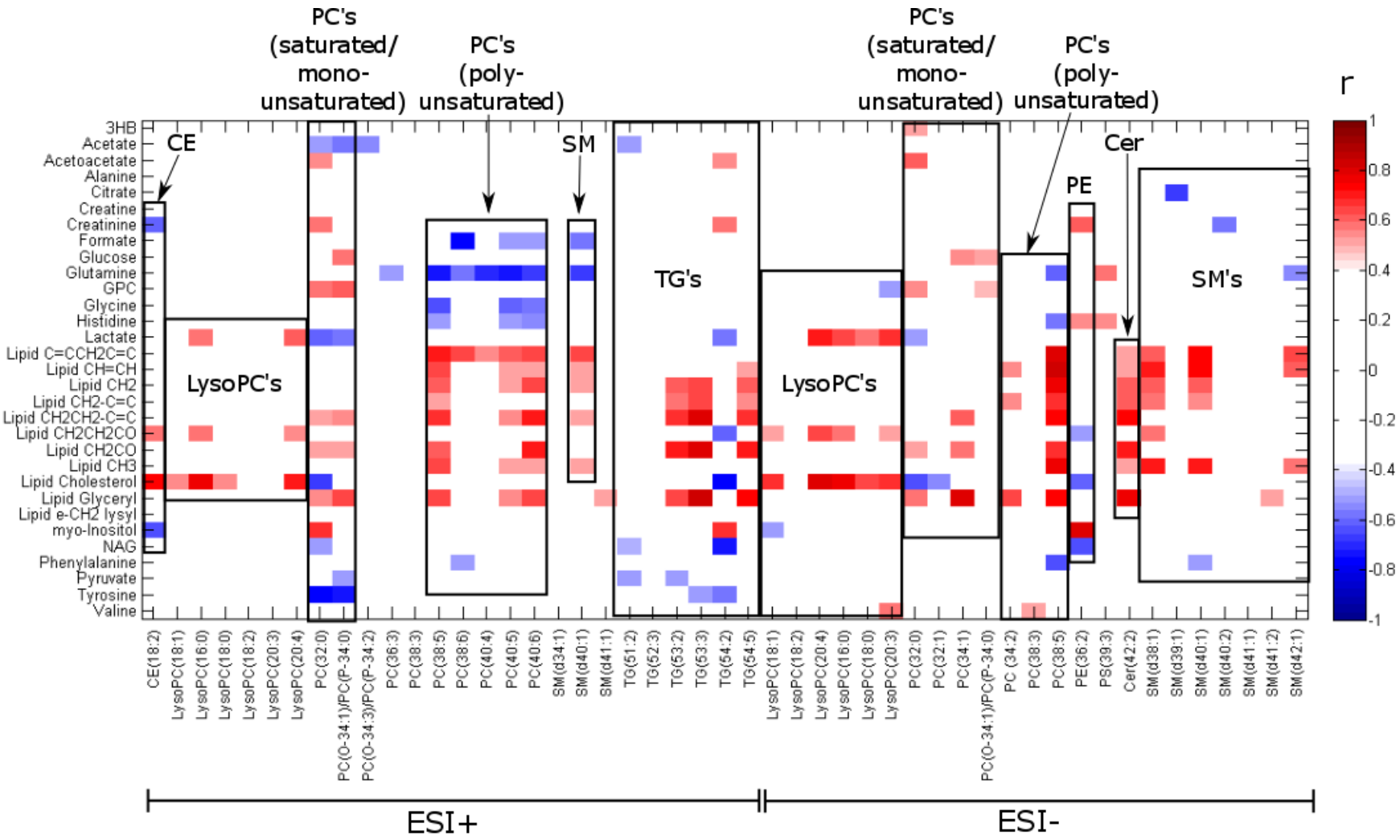


Figure 5.7: Weak Correlations Observed between Plasma ^{1}H NMR and UPLC-MS Markers

2D Pearson-based correlations performed between integrals of ^{1}H NMR metabolite peaks and lipid UPLC-MS features, measured in plasma of HAT patients (n=16). Positive associations are shown in red and negative ones in blue. Presented correlations all displayed p<0.05 prior to multiple test correction, yet were no longer significant following Benjamini–Hochberg FDR (see Appendix Figure A.7 for the strongest of these links, showing correlations with corrected p<0.1). Abbreviations as described previously.

5.4.3 Plasma Cytokine Analyses

Plasma cytokine concentrations were measured for IL-10, IFNγ, TNFα and IL-6 in HAT patients and controls, which revealed IL-10 as the most dominant cytokine, shown in Figure 5.8A.

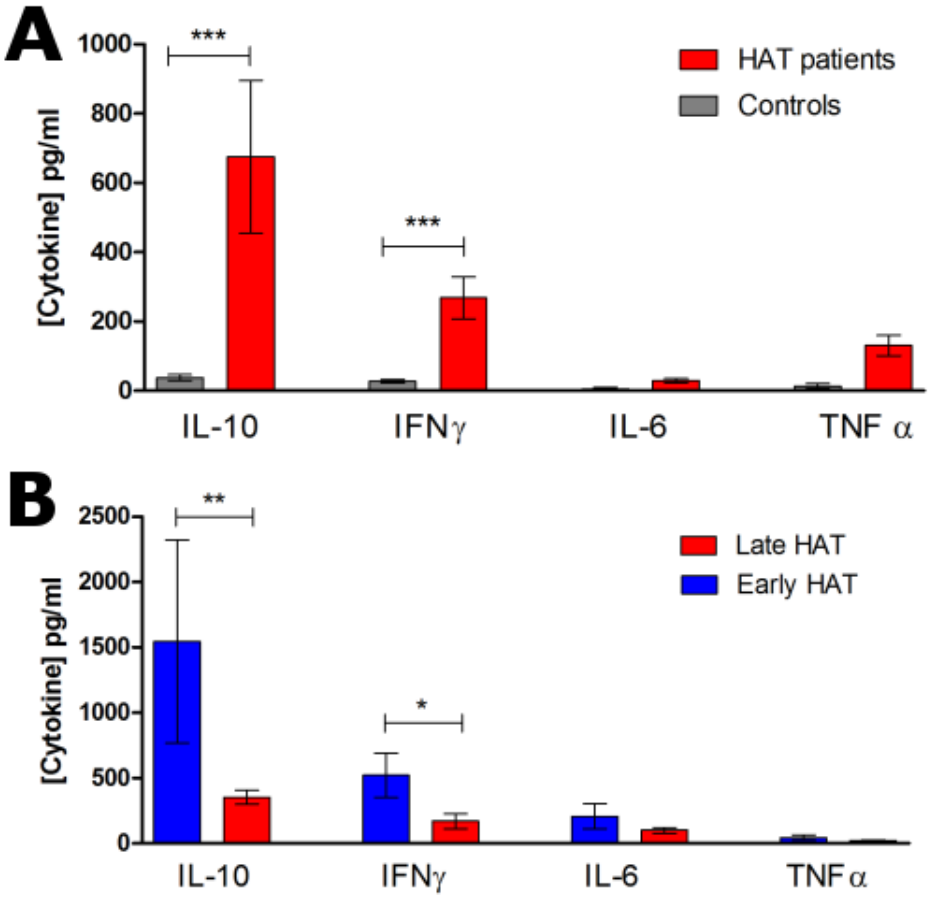


Figure 5.8: Elevated Concentrations of Selected Cytokines in Plasma HAT Patients

Bar charts showing mean average plasma cytokine concentrations with SEM as error bars. (**A**) Comparisons between HAT patients (red bars) and controls (grey). Numbers of samples analysed varied for each cytokine (for HAT patients: IL-10 n=59, IFNγ n=59, TNFα n=56, IL-6 n=66; for controls: IL-10 n=12, IFNγ n=11, TNFα n=5, IL-6 n=4). (**B**) Comparisons between patients with early stage HAT (in blue, n=16 for all except TNFα n=14) vs. late stage HAT (in red, IL-10 n=42, IFNγ n=47, IL-6 n=44 and TNFα n=40). Patient samples with missing diagnostic staging information (n=6) have been omitted. Significant differences between HAT patients and controls (tested for IL-10 and IFNγ only) and between early and late stage patients (for all four cytokines) were analysed *via* Mann-Whitney U-tests with Benjamini–Hochberg FDR correction, where * p<0.05; ** p<0.01; *** p<0.001.

Statistical differences between patients and controls were only tested for IL-10 and IFNγ (as the other two cytokines had insufficient numbers of controls for valid comparisons), which were both significantly higher in samples from HAT patients (p=0.0004 for each cytokine). Concentrations of IL-10 (+337%) and IFNγ (+206%) were also found to be significantly higher in patients with early stage disease (p=0.004 and p=0.011, respectively) than those with late stage HAT (Figure 5.8B).

Immune-Metabolic Correlations in Plasma

In order to assess potential immune-metabolic links in HAT patients, plasma cytokine concentrations were correlated with plasma ^{1}H NMR metabolite integrals (that were listed in Table 5.2) using a Pearson-based 2D matrix analyses as described previously. Results showed that out of the four cytokines, only IL-6 displayed significant correlations with NMR metabolites. IL-6 was positively associated with acetate, acetoacetate, and pyruvate, as well as with HAT-discriminatory metabolites *myo*-inositol and histidine. IL-6 also displayed the highest number of correlations with lipid UPLC-MS features (including negative associations with lysoPC's 16:0, 18:0 and 18:1) — see Appendix Figure A.8 — yet none of the correlations observed between any of the cytokines and MS markers were found to retain their statistical significance after multiple test correction.

5.5 Results: CSF Analyses

5.5.1 Metabolic Profiling via ^{1}H NMR

Metabolic profiling of CSF using standard 1D ^{1}H NMR pulse revealed 27 individually assigned endogenous metabolites, the majority of which were also observed in the plasma, as listed in Appendix Table A.7. Metabolites that were uniquely detected in CSF included 2-hydroxybutyrate (2HB), 2-hydroxyisovalerate, 3-hydroxyisovalerate, acetone and gluconate. Peaks from drug metabolite paracetamol (also known as acetaminophen) were also identified in a few of the patients' CSF samples, as well as contaminant ethanol resonances (the latter were also detected in the plasma samples and blank buffer controls). Both paracetamol and ethanol resonances (listed in Appendix Table A.7) were subsequently excluded from final analyses.

Metabolic Differences Linked with HAT and Disease Diagnostic Stage

PCA was used to examine metabolic differences between CSF samples from HAT patients and controls, yet unlike plasma results, variation linked to disease status was unable to separate the two groups, as demonstrated in Figure 5.9A. Additional O-PLS-DA was also unable to robustly differentiate between HAT patients and controls (displayed negative predictive Q^2Y value). Further examination of variation within patients also showed no significant differences in profiles between late and early stage of disease, using either PCA (Figure 5.9B) or O-PLS-DA (also negative Q^2Y).

Despite not determining any dominant disease-linked metabolic response in CSF *via* multivariate analyses, integrals were calculated for all assigned metabolites and compared using Welch T-test, and Bonferroni–Hochberg FDR correction, to investigate whether any metabolite was differentially

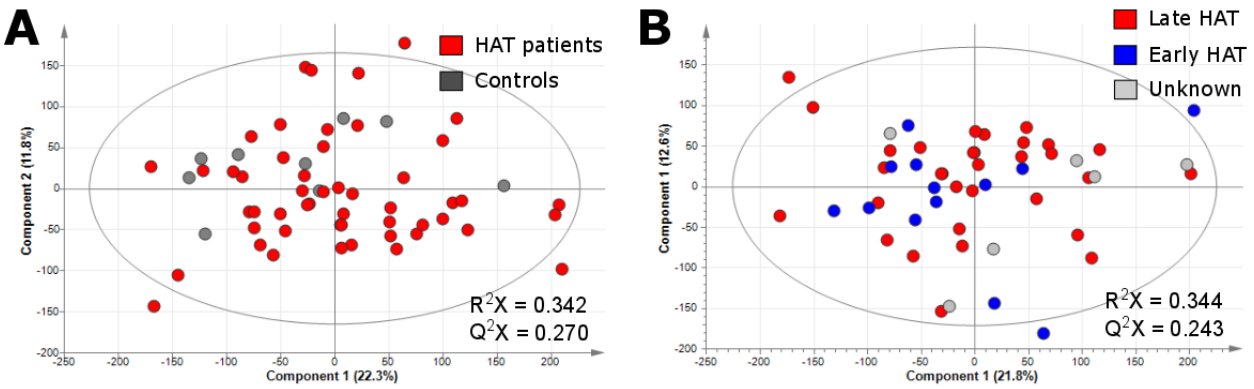


Figure 5.9: No Dominant Metabolic Variation in CSF ^{1}H NMR Phenotypes Associated with HAT Disease or Diagnostic Stage

PCA scores plots based on first two components where each circle represents one sample. (**A**) Comparison between HAT patients (in red, n=51) vs. controls (in dark grey, n=9). (**B**) Comparison between late stage patients (in red, n=31) vs. early stage patients (in blue, n=13). Staging information was missing for 7 patients (in grey). Both PCA models were tested up to 3 components yet no differences between either groups were observed. Additional O-PLS-DA models for both analyses were not predictive in either case.

expressed between either patients and controls, and between late and early stage patients. Metabolites that were significantly different between the first two groups have been highlighted in Figure 5.10.

Results showed that CSF concentrations of glucose (-29.0%), glutamine (-23.0%) and histidine (-14.7%) and *scyllo*-inositol (-47.8%) were significantly lower in HAT patients, whilst 2HB levels were higher (+35.2%); all p=0.040 after FDR correction. The collective discriminatory accuracy of these 5 metabolites was assessed in additional O-PLS-DA models (as described previously for plasma markers), which demonstrated 100% specificity, yet poor sensitivity (60.8%, 43.78–74.18% CI), for correct HAT classification (see Appendix Table A.9 for confusion matrix). A complete list of results for all integrals has been summarised in Table 5.6.

Comparisons between early and late stage patients also showed an apparent increases in 3HB (+12.8%) and lactate levels (+36.5%) in late-stage patients, relative to early stage — similar trends were also observed between patients and controls — yet both these differences lost significance after FDR correction (Table 5.6). As diagnostic staging criteria are based on both the presence of parasites and/or white blood cell (WBC) counts in CSF, metabolite integrals of HAT patients were further tested for differences between parasitaemia status (*via* O-PLS-DA) and were correlated with WBC counts (using Pearson-based correlation). Mirroring earlier findings, no significant differences or associations were determined in CSF for either case following multiple test correction, though trends reiterated apparent increases in 3HB and lactate, as well as decreased mannose, in samples that tested positive for parasites.

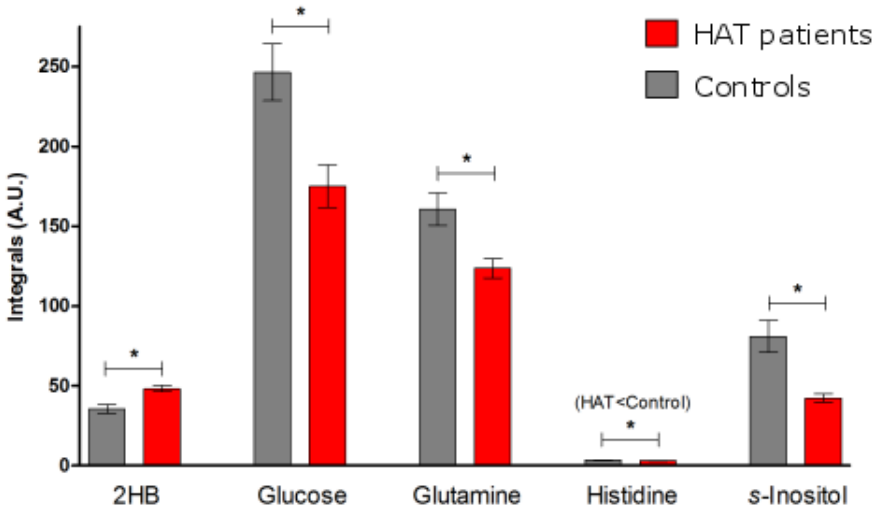


Figure 5.10: Specific Set of CSF Metabolites Differentially Expressed between HAT Patients and Controls

Bar-charts showing mean averages (± standard error of the mean as error bars), with HAT patients in red (n=51) and controls in dark grey (n=9). Significant differences (as tested via Welch T-test following Bonferroni–Hochberg correction) are labelled with an asterisk, where * p<0.05. Key: A.U., arbitrary units; 2HB, 2-hydroxybutyrate; *s*-Inositol, *scyllo-inositol*.

Additionally, in order to assess whether metabolites detected in plasma were predictive of those measured in patient CSF, Pearson correlation analysis was performed between metabolites integrals of the two biofluids. Followed by FDR correction, only one metabolite was found to be significantly linked between the two biofluids, namely phenylalanine (p=0.004), which was previously determined discriminatory for HAT in plasma.

Table 5.6: CSF Metabolites Measured via ^{1}H NMR Spectroscopy

Metabolites	Integrated Peak (multiplicity)	Average Integrals in Controls (± SEM) in A.U.	Average Integrals in HAT (± SEM) in A.U.			HAT - Control		Late - Early HAT	
			Early stage	Late stage	All	% Diff	P-value	% Diff	P-value
2-Hydroxybutyrate	0.90(t)	35.64 (± 2.91)	50.27 (± 3.83)	46.84 (± 2.33)	48.19 (± 1.84)	35.22%	0.040	-6.82%	NS
2-Hydroxyisovalerate	0.84(t)	17.58 (± 1.87)	21.83 (± 1.77)	20.37 (± 1.60)	21.05 (± 1.19)	19.73%	NS	-6.71%	NS
3-Hydroxybutyrate	1.23(d)	18.45 (± 0.94)	18.58 (± 0.69)	20.96 (± 0.66)	21.14 (± 0.75)	14.59%	NS	12.84%	NS
3-Hydroxyisovalerate	1.27(s)	10.39 (± 1.08)	15.48 (± 3.05)	10.24 (± 0.61)	11.78 (± 0.93)	13.33%	NS	-33.86%	NS
Acetate	1.93(s)	34.48 (± 2.91)	31.34 (± 3.79)	43.07 (± 4.69)	42.69 (± 4.52)	23.80%	NS	37.41%	NS
Acetone	2.24(s)	15.46 (± 2.80)	12.62 (± 0.92)	10.81 (± 0.65)	11.43 (± 0.52)	-26.06%	NS	-14.39%	NS
Alanine	1.48(d)	41.52 (± 3.02)	42.22 (± 2.27)	45.79 (± 1.65)	45.69 (± 1.30)	10.05%	NS	8.47%	NS
Citrate	2.67(d)	95.42 (± 7.77)	80.40 (± 5.87)	86.16 (± 2.88)	86.06 (± 2.56)	-9.80%	NS	7.16%	NS
Creatine	3.04(s)	25.14 (± 1.81)	22.70 (± 1.75)	22.81 (± 0.77)	23.40 (± 0.72)	-6.89%	NS	0.47%	NS
Creatinine	3.05(s)	38.83 (± 2.93)	34.92 (± 3.10)	32.57 (± 2.46)	32.89 (± 1.81)	-15.30%	NS	-6.72%	NS
Dimethylamine	2.72(s)	2.77 (± 0.18)	3.32 (± 0.43)	2.89 (± 0.25)	2.99 (± 0.19)	8.08%	NS	-13.07%	NS
Formate	8.46(s)	6.04 (± 0.52)	6.46 (± 0.50)	7.00 (± 0.34)	7.49 (± 0.71)	24.00%	NS	8.36%	NS
Gluconate	4.04(t)	6.04 (± 0.52)	6.46 (± 0.50)	7.00 (± 0.34)	7.49 (± 0.71)	24.00%	NS	8.36%	NS
Glucose	5.24(d)	246.55 (± 17.72)	203.38 (± 22.95)	161.50 (± 17.00)	175.05 (± 13.27)	-29.00%	0.040	-20.59%	NS
Glutamine	2.46(m)	160.67 (± 10.20)	136.85 (± 13.22)	120.30 (± 7.20)	123.66 (± 6.04)	-23.03%	0.040	-12.10%	NS
Histidine	7.09(s)	3.48 (± 0.14)	2.98 (± 0.12)	2.92 (± 0.08)	2.97 (± 0.06)	-14.67%	0.040	-1.85%	NS
Lactate	1.34(d)	1,049.19 (± 83.65)	986.14 (± 85.83)	1,346.08 (± 83.48)	1,293.51 (± 64.82)	23.29%	NS	36.50%	NS
Lysine	3.03(t)	12.30 (± 1.12)	11.79 (± 0.96)	11.57 (± 0.37)	11.75 (± 0.36)	-4.46%	NS	-1.89%	NS
Mannose	5.19(d)	7.08 (± 0.94)	8.13 (± 0.92)	6.21 (± 0.54)	6.62 (± 0.44)	-6.49%	NS	-23.66%	NS
myo-Inositol	4.08(t)	35.07 (± 3.22)	27.86 (± 2.58)	27.98 (± 1.96)	28.12 (± 1.48)	-19.81%	NS	0.41%	NS
scyllo-Inositol	3.37(s)	80.99 (± 9.99)	45.73 (± 7.62)	40.90 (± 2.58)	42.24 (± 2.53)	-47.84%	0.040	-10.56%	NS
N-Acetylglycoprotein	2.06(m)	19.13 (± 1.16)	20.99 (± 1.22)	19.19 (± 0.96)	19.69 (± 0.71)	2.93%	NS	-8.60%	NS
Phenylalanine	7.33(d)	8.36 (± 0.71)	8.67 (± 0.44)	8.32 (± 0.42)	8.44 (± 0.30)	0.98%	NS	-4.01%	NS
Pyruvate	2.34(s)	34.80 (± 4.37)	34.22 (± 4.07)	29.15 (± 3.17)	32.82 (± 2.63)	-5.67%	NS	-14.80%	NS
Tyrosine	7.20(d)	6.63 (± 0.46)	6.26 (± 0.29)	6.37 (± 0.20)	6.41 (± 0.16)	-3.44%	NS	1.76%	NS
Urea	5.80(m)	147.32 (± 11.76)	177.67 (± 21.02)	157.54 (± 24.36)	164.27 (± 17.41)	11.51%	NS	-11.33%	NS
Valine	1.05(d)	13.75 (± 0.94)	14.32 (± 0.97)	13.79 (± 0.50)	14.25 (± 0.46)	3.61%	NS	-3.70%	NS

Key: (d), doublet; (m), multiplet; (s), singlet; (t), triplet; % Diff, percentage difference (relative to either controls or early stage patients); A.U., arbitrary units; NS, not statistically significant; SEM, standard error of the mean.

Metabolic Differences Linked with Neurological Symptoms

Additional PCA and O-PLS-DA models were generated using metabolite integrals to investigate potential links between the CSF markers and the presence of neurological symptoms in patients. Discriminatory potential of metabolites were tested in relation to patient gait (impaired/unable to walk vs. normal), presence or absence of tremors, and drowsiness. Results for patient drowsiness have been highlighted in Figure 5.11, which demonstrated a partial segregation between the two groups *via* PCA and O-PLS-DA. Examination of the corresponding O-PLS-DA loadings revealed that lactate, 3HB and alanine were all positively linked with drowsiness, though out of these, only alanine was significantly higher, as tested using Welch T-test and Benjamini–Hochberg FDR correction (+24.2%, p=0.001). In contrast, mannose concentration was significantly lower in patients with drowsiness compared with non-drowsy patients (-37.9%, p=0.003), alongside levels of acetone (-24.9%, p=0.009) and urea (-44.3%, p=0.017).

As observed with drowsiness, differences between patients with impaired gait (either unable to walk or required assistance) vs. those who walked normally, were also observed using PCA, O-PLS-DA and Welch T-test (with FDR correction) — see Appendix Figure A.9. Patients with impaired gait had significantly higher concentrations of alanine (+24.4%, p=0.004) and 3HB (+17.9%, p=0.045), an lower levels of mannose (-41.8%, p=0.001) in their CSF, than those who could walk normally. No significant differences were observed between patients that presented with tremors and those that did not.

Metabolic Variation Associated with Confounding Factors

Following the investigation of potential associations between pathological states and CSF metabolites, the possible effect of other confounding factors on CSF metabolism were also assessed: namely gender, age, and study site. No differences in metabolic profiles were determined between females and males, or between the two sites (Lwala Hospital and Serere Health Centre), as assessed by PCA an O-PLS-DA (the latter generated negative Q^2Y values in both cases). Furthermore, no metabolite integrals were found to be significantly associated with the patient age, through trends indicated that levels of dimethylamine were positively correlated with age.

5.5.2 CSF Cytokine Analyses

CSF samples from patients were additionally assessed for cytokine measurements. Similar to plasma results, IL-10 displayed the most elevated concentrations out of the tested cytokines. Yet in contrast to plasma, all tested cytokines displayed significantly higher concentrations in late stage samples

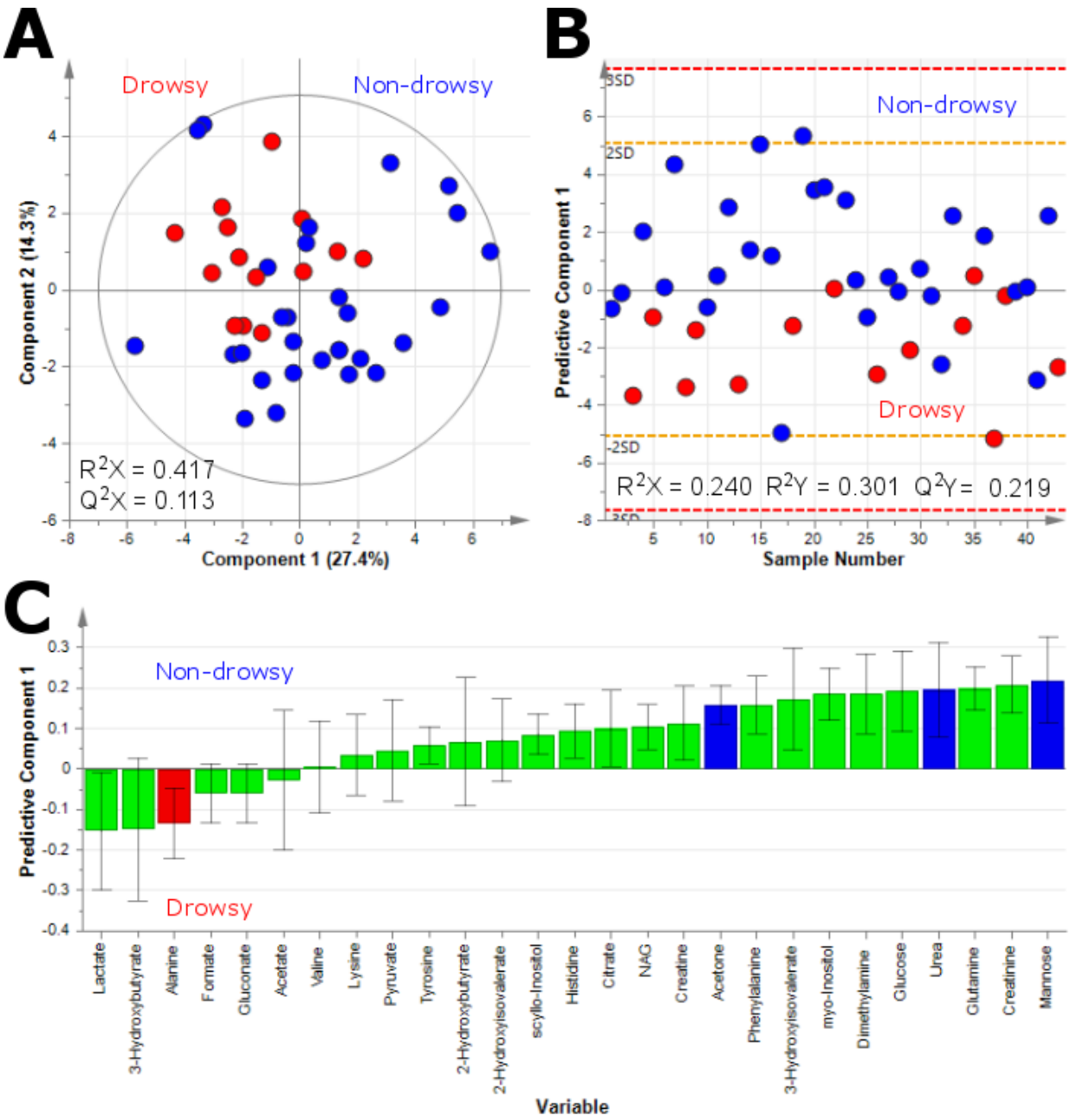


Figure 5.11: Metabolic Changes in CSF Associated with Drowsiness

(**A**) PCA scores plot, comparing HAT patients with drowsiness (in red, n=14) vs. non-drowsy patients (in blue, n=29). (**B**) O-PLS-DA scores plot and corresponding loadings plot (**C**), showing differences between the two groups. Metabolites that also tested as significantly different *via* Welch T-test and FDR correction (p<0.05) are shown in red (higher in drowsy patients) and blue (higher in non-drowsy).

than early stage samples (IL-10 +163% p=0.016, IFNγ +625% p=0.003 and IL-6 +480% p=0.016), as shown in Figure 5.12.

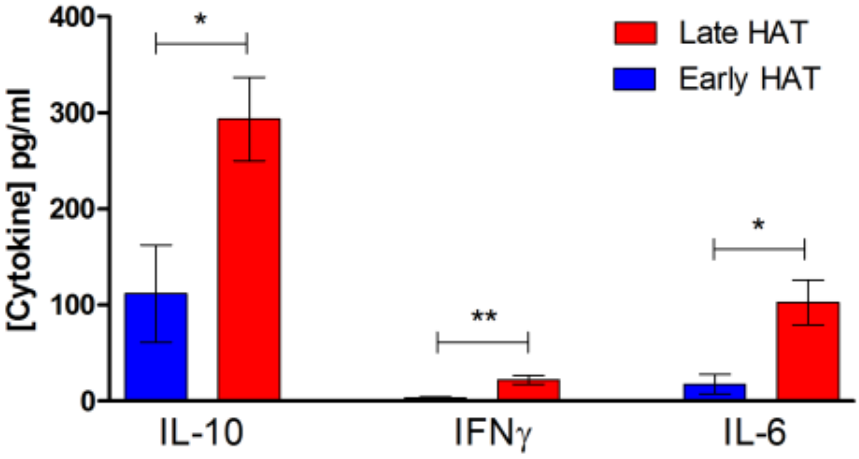


Figure 5.12: Elevated Concentrations of Selected Cytokines in Plasma HAT Patients

Bar charts comparing CSF cytokine concentrations between patients with early stage HAT (in blue, n=13) vs. late stage HAT (in red, n=33). Samples with missing diagnostic staging information (n=2) have been omitted. Key as described for Figure 5.8.

Immune-Metabolic Correlations in CSF

Several strong correlations in CSF patient samples were observed between ^{1}H NMR metabolite integrals and cytokine concentrations, summarised in Figure 5.13. Results show that all three cytokines were positively correlated with 3HB and lactate, and were negatively associated with glucose. IL-10 and IL-6 were also both negatively correlated with glutamine, histidine and mannose.

5.6 Discussion

Striking effects on host metabolism have been reported in a range of experimental parasitic infections in rodents.[91–94,197,243] In humans, however, only two studies to date have reported diagnostic metabolic profiles for parasitic infections, namely for infection with *Onchocerca volvulus*[100] (causative agent for river blindness) and *Schistosoma mansoni*[102][4], an intestinal helminth. Given the diversity of genetic and environmental exposures of humans, the fact that polyparasitism is the norm rather than the exception, and that the parasitic burden encountered in clinical practice is generally much lower than experimentally induced infection, it is unsurprising that there have been few clinically useful diagnostics developed to date. The presented study represents the first to successfully apply a comprehensive metabolic profiling approach to examine human *Trypanosoma*

[4]Though this study has not been replicated by other *Schistosoma* studies in humans.[101]

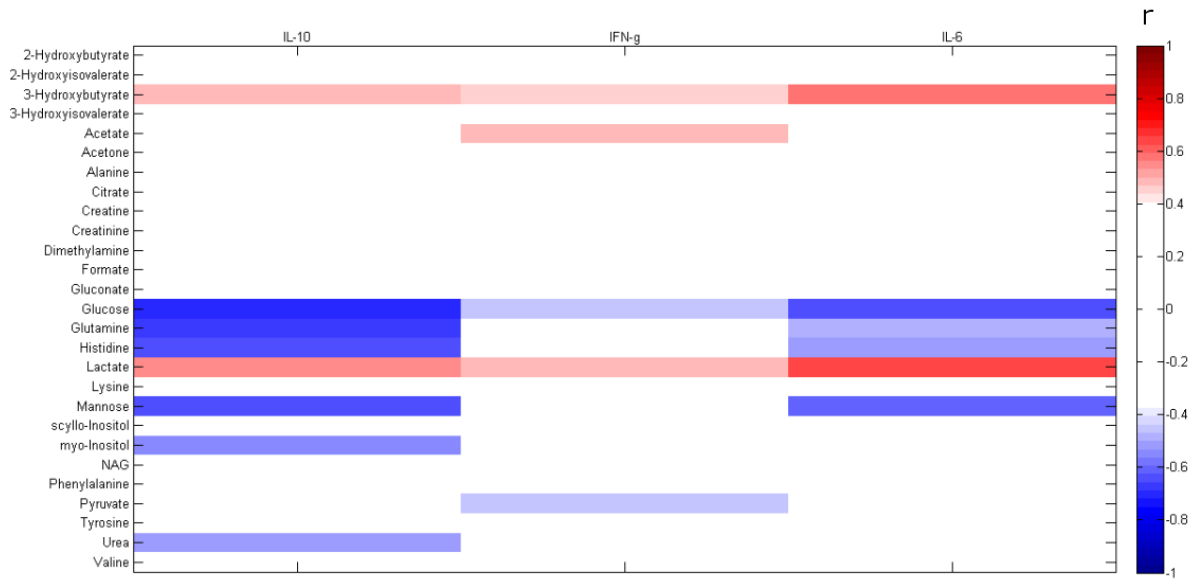


Figure 5.13: IL-10 and IL-6 Display Similar Correlation Patterns with CSF Metabolites

2D Pearson-based correlations performed between integrals of ^{1}H NMR metabolite peaks and cytokine concentrations, measured in CSF of HAT patients (n=31). Presented correlations were all statistically significant (FDR corrected p<0.05), with positive correlations in red and negative correlations in blue. Abbreviations as described previously.

brucei infection and to investigate potential metabolic associations with host immunity and clinical symptoms, resulting in the discovery of several disease-associated biomarkers with diagnostic potential.

5.6.1 Strong HAT-associated Differences in Plasma Linked to Changes in Lipid and Amino Acid Metabolism

HAT was associated with dominant variations in plasma metabolic phenotypes profiles, observed using both ^{1}H NMR spectroscopy and UPLC-MS, though no significant differences were determined between early and late stage patients. In total, 16 metabolites were discovered as significantly discriminatory for HAT *via* NMR, which largely consisted of lipids and amino acids, followed by a further 37 lipid markers determined *via* UPLC-MS. Both sets of markers were capable of correctly identifying HAT samples from controls with high accuracy (up to 93.75% sensitivity and 95.45% specificity). Correlation analyses between metabolites measured across these two platforms showed apparent links, particularly amongst the lipids, yet none of these were statistically significant, most likely due to lack of statistical power.

The single most discriminatory metabolite that was detected by NMR was phenylalanine, which displayed the largest significant percentage difference in plasma concentration between HAT and controls (higher in patients by 51.2%). Elevated phenylalanine in the blood has been known to

be associated with inflammation and infection[279], and has been recorded in patients with enteric fever[280] and malaria.[281] Authors from the malaria study have suggested that raised phenaylalaninaemia may be linked with neurological effects seen in cerebral malaria, as high levels of phenylalanine are known to be neurotoxic.[282] Interestingly, plasma phenylalanine was the only metabolite who's concentration in patient plasma samples was also directly correlated with its concentrations in CSF, implicating its neurological involvement in HAT as well.

Other changes in amino acids included significantly lower concentrations of histidine and alanine in HAT patients, compared with controls. Reduced histidine concentration have been shown to be correlated with inflammation and oxidative stress[283,284], as well as malarial infection in humans.[285] Furthermore, plasma concentrations of *N*-acetyl groups of glycoproteins were elevated in patients. The majority of the observed NMR *N*-acetyl-glycoprotein signals have been shown to belong to α1 acid glycoprotein[286,287], an inflammatory marker known to be linked with the acute-phase response. This result is similar to data generated from murine infection with *T. b. brucei*, where plasma acetylated glycoproteins (largely O-acetyls) were also significantly elevated post-infection.[92] In addition, increased plasma creatinine concentrations were significantly higher in HAT patients and have been observed in experimental *T. brucei gambiense* infection in rats.[288] Increased creatinine may indicate renal dysfunction in patients, a clinical feature often observed in HAT.[58] Collectively, the presented findings raise the implication that metabolic changes due to *Trypanosoma* infection induce a systemic inflammatory response in the host, a result that is consistent with clinical evidence.[66]

The majority of HAT-associated differences in NMR plasma spectra involved variation in lipid composition between the two groups. Lipids are largely known for their function as energy reserves and as major constituents of membranes, yet they also represent highly biologically active metabolites. They are involved in signalling and a range of inflammatory processes, from prostaglandins and leukotrienes synthesis, to acting as ligands for nuclear receptors (e.g. peroxisome-proliferator-activated receptors, PPAR's), which in turn selectively activate or repress genes that regulate fatty-acid oxidation, adipocyte differentiation, and glucose metabolism.[289,290] In this study, plasma concentrations of the majority of discriminatory lipids were significantly lower in patients than controls. NMR results indicated that these decreased lipids contained unsaturated bonds (CH=CH and C=CCH2C=C), CH_3 and CH_2CH_2CO moieties, in addition to decreases in albumin lysyl groups and high-density lipoprotein (HDL) cholesterol. Disturbances in lipid metabolism have been observed across a range of infections, primarily assayed using lipid-specific biochemical serum tests (as opposed to comprehensive metabolic profiling experiments). Significant decreases in

total cholesterol and HDL have been reported in malaria[291], enteric fever[292], sepsis[293], visceral leishmaniasis[294], and HAT *gambiense*.[295] Reduced lipoprotein CH_3 moieties were also reported in the murine model of trypanosomiasis infected with *T. b. brucei*[92], re-asserting the rationale for further assessment into the different of lipid species and their possible roles.

Subsequent examination of plasma lipid profiles *via* UPLC-MS showed that HAT was associated with significant reductions in polyunsaturated phosphatidylcholines (e.g. PC 38:3, 38:5, and 40:6), whilst PC molecules with 2 or fewer unsaturated bonds were increased. This phenotype presents a more complex result than was previously seen in mice infected with *T. b. brucei*, where authors showed a general reduction in plasma PC concentrations, as measured by NMR.[92] Levels of numerous lysophosphatidylcholines (e.g. LysoPC 16:0, 18:0, 18:2) were also significantly lower in HAT patients. Despite the fact that *Trypanosoma brucei ssp* are known to be capable of producing many of their lipids *de novo*, evidence suggests that not only do they tend to scavenge lipid precursors from the host, but that lipid uptake is also an absolute requirement for optimal growth.[296,297] These particularly include fatty acids and lyso-phospholipids, that become assembled to create phospholipids, the most dominant constituents of *T. brucei ssp* lipids, found in their membranes.[298] Additionally, bloodstream forms of *T. b. rhodesiense* contain a high proportion of polyunsaturated fatty acyl chains within their lipids[299,300], giving rise to the suggestion that decreased polyunsaturated plasma lipids in HAT patients may be attributed to augmented parasitic uptake.

A range of sphingomyelin species were also significantly reduced in HAT (including SM d40:1, d41:1 and d42:1), alongside an increase in ceramide plasma levels. Sphingolipid turnover is regulated by the balance between enzymatic synthesis and degradation, the latter resulting in the generation of ceramide as a major break-down product. Triggers including oxidative stress, infection and inflammatory cytokines all can activate sphingomyelinase enzymes to increase ceramide production. In turn, augmented levels of ceramide have been associated with increased cell autophagy and apoptosis (especially in endothelial cells which results in increased vascular permeability), further pro-inflammatory cytokine and chemokine synthesis, and other metabolic disorders such as obesity and insulin resistance.[301] Unlike the majority of the plasma sphingomyelins measured in this study, levels of SM d34:1 were higher in patients than controls. This particular species not only represented one of top 5 discriminatory markers measured *via* UPLC-MS but was also significantly correlated with swollen lymph nodes in patients. Thus, these results emphasise that the disturbances in sphingolipid metabolism, imposed by *T. b. rhodesiense* infection, are associated with widespread inflammatory effects in HAT.

Hypertriglyceridaemia was another key observation in plasma from HAT patients, with the largest significant increases in TG 51:2 and 53:2. Elevated plasma/serum triglycerides has been reported in *in vivo* models of *T. brucei spp* infections[302,303] and in clinical samples from HAT patients infected with *T. b. gambiense*.[295] However, as hypertriglyceridaemia (often combined with increases in circulating free fatty acids) appears to be a common trait to a range of human infectious diseases (e.g. enteric fever[292], sepsis[293] and malaria[291]), further identification of the different triglyceride species elevated in the various infections would be required to optimally assess disease specificity.

5.6.2 Disease-related Changes in CSF Provide Insights into Neuropathology of HAT

The metabolic changes associated with HAT in CSF samples were less pronounced than plasma samples. Only 5 metabolites, as detected by ^{1}H NMR, were significantly different between HAT patients and control, which collectively only displayed a sensitivity of 60.8% for HAT identification. Out of these, only histidine (reduced in patients) also represented a discriminatory metabolite for HAT in plasma. The remaining markers included glucose, glutamine, and *scyllo*-inositol, all lowered in HAT, and 2-hydroxybutyrate (2HB), which was increased. Reductions in glucose (and glutamine) were also been measured in plasma of *T. b. brucei* murine infection model, which authors have attributed towards increased parasitic utilisation for energy requirements.[92] Furthermore, HAT patients often present with hypoglycaemia, which itself is also an adverse complication of one the drugs used to treat HAT.[58,68] Limited literature exists on the function of *scyllo*-inositol, which is predominantly found in the brain along with inositol stereoisomers, (particularly *myo*-inositol). Unlike *myo*-inositol, *scyllo*-inositol does not appear to be involved in maintaining golgi body trafficking and cellular membrane integrity and signalling, as it has not been shown to be incorporated into glycerophospholipids (as phosphatidylinositol).[304,305] Another key role of inositols are their putative cellular protective functions by acting as osmolytes[305,306], therefore a reduction seen in the CSF may indicate an osmotic imbalance in the brain of HAT patients.

2HB is derived from 2-oxobutanoate, which in turn is generated primarily in the liver from both threonine catabolism and glutathione synthesis. Recent evidence suggests that increased oxidative stress and elevated lipid oxidation results in augmented 2HB production to meet the high demand for glutathione, the main anti-oxidant in mammalian cells.[307] Inflammatory responses are known to induce oxidative stress.[308] As a result, elevated 2HB concentration in CSF may act as a marker of the sustained inflammation observed in HAT, potentially neurological inflammation present at

late stage.

No significant differences in CSF metabolite levels were observed between the diagnostic stages, though this may partly also be due to lack of statistical power. Nevertheless, observed trends included decreased mannose concentration and increased levels of 3HB and lactate in late stage patients, compared with early stage. As neurological symptoms are typically limited to the late stage of disease (once the parasite has crossed the blood-brain barrier), it did thus not come unexpected that this set of metabolites was also involved in the differentiation between patients with and without neurological symptoms, e.g impaired gait.

5.6.3 Inflammation Observed in HAT Linked with Altered Metabolism in Host

Inflammatory-associated metabolic changes were accompanied with significantly higher levels of both inflammatory cytokine IFNγ and regulatory cytokine IL-10 in plasma of HAT patients, as well as apparent increases in TNFα and IL-6. These cytokine profiles are in agreement with several clinical studies[79,80] and mouse models.[82] IL-10 was the most prominent cytokine detected in both plasma and CSF samples and although no control cytokine results was available for CSF, evidence has shown that HAT is associated with augmented IL-10 in CSF of both human studies[78,81] and mouse models.[82] Increased pro-inflammatory cytokines, including TNFα, IFNγ and IL-6, have also been measured in brain tissue from mice infected with *T. b. brucei*, thought to be primarily produced by activated atrocytes within the central nervous system.[83,84] Furthermore, the reduction of IL-10 and IL-6 in plasma from in late stage patients, compared with early stage HAT, yet increase of IL-10 and IL-6 in CSF, has also been reported[78,79,81], highlighting the disparities between systemic and local immune responses in the host. The imbalance of pro- and counter-inflammatory cytokines, as observed in this study, has been shown to be strongly linked disease outcome, particularly neuropathogenesis, where inflammation-induced increased vascular permeability is thought to promote parasitic invasion into the brain.[309]

Immune-metabolic associations between cytokine and metabolite concentrations were observed in both plasma and CSF, though these were much more prominent in the latter. In plasma, the majority of these links involved positive correlations between IL-6 and metabolites that were unrelated with HAT discrimination (e.g. acetate and pyruvate) between HAT and controls. Whilst trends were observed between lipid species and cytokines (particularly amongst the lysoPC's), none of these correlations were significant, and would thus need to be repeated with larger sample groups for confirmation. In contrast, strong links were observed in the CSF of HAT patients. IL-10, IL-6 and IFNγ were all positively correlated with 3HB and lactate and negatively linked

with glucose, signifying the crucial role of energy metabolism in immune response. Additionally, mannose, glutamine and histidine were negatively linked with IL-10 and IL-6. These results imply that the metabolites, which were differentially expressed between HAT patients and controls, were also connected with neuronal immunopathology.

5.6.4 Study Limitations

The strongest limitation to this study was the modest and incomplete number of samples available for both metabolic and cytokine investigations. Cytokine measurements had been performed well in advance of the presented metabolic profiling work, on a cohort of samples where only a subset had also been available for any additional experimental analyses. The reduction in the dataset was particularly evident for the plasma samples analysed *via* UPLC-MS, which consisted of 16 HAT patients, of which only 8 had also undergone cytokine samples, thus severely restricting the statistical power required to determine significant immune-metabolic correlations in this dataset.

Despite the limited number of samples, variation due to confounding factors such as gender and age, were minimal in this study, indicating a relatively homogeneous population, from a metabolic point of view, which facilitated the detection of disease associated imprints on the host. Ideally, additional measures such as patient weight and body mass index (BMI) would have also aided in the identification of any further metabolic variances within the studied population, and their record would be highly recommended for future investigations. An important consideration is that one cannot exclude the fact that HAT patients (especially within the late stage) may have reduced dietary intake or appetite, due their illness, and that the decreased nutrient levels, e.g. in glucose, histidine, glucosamine in CSF, were merely reflection of reduced intake. However, the changes appeared to target only certain metabolites, since decreases in other common dietary biomarkers, such as carnitines or trimethyl-*N*-oxide[310,311], were not observed, suggesting that the differences observed in this study are disease linked. Nevertheless, a record of patient's dietary intake (in addition to BMI/weight) would prove highly useful in future studies, for confirmation of findings.

The fact that little demographic subject information was available for the CSF control samples also casts a level of uncertainty on the current interpretation of the presented results. One cannot be sure that the differences in CSF metabolite concentrations between the HAT samples and controls were not also due to confounders such as subject age, gender, sample collection site, and/or sample storage time. Consequently, statistical correlations between metabolite concentrations in HAT patient plasma samples (with known demographics) and CSF samples were performed, as an additional method to aid the discovery of CSF metabolites linked with disease (i.e. phenylalanine).

There was also a lack of comprehensive parasitological assessment in this study. Whilst previous studies have shown that HAT patients from the analysed geographical sites are all infected by the *T. b. rhodesiense* with minimal genetic variation[79], this ideally should be have been confirmed with the current sample set. In addition to testing for *Trypanosoma*, patients were also screened for malaria, filariasis and schistosomiasis, yet co-infection from other agents (e.g. hookworm, or *Salmonella enterica ssp.*, which cause enteric fever) cannot be excluded. Extended screening of different infectious pathogens on a larger cohort, followed by comparative analyses between patients that are single infected with *Trypanosoma* and patients that are co-infected with other agents, are required for optimal identification of effects on host metabolism that are truly specific to trypanosomes.

Lastly, the presence of artefactual and/or contaminant ^{1}H NMR spectral peaks that required removal may have resulted in the loss of endogenous metabolic information within the excluded regions. This was particularly the case for EDTA, which produced large artefactual peaks within ^{1}H NMR spectra[278] that could have restricted metabolic interpretation. For any further metabolic profiling studies, it would thus be more appropriate to use blood collection tubes that are EDTA free, e.g. without any additives, or using lithium heparin tubes for plasma collection, following published recommendations of practice.[108]

5.6.5 Conclusions and Future Work

In conclusion, the presented results have uniquely shown the successfully application of a comprehensive metabolic profiling approach for the investigation of HAT, with the determination of set of disease-associated metabolites with diagnostic potential, predominantly in plasma. Such changes included perturbations in lipid and amino acid metabolism, linked with systemic inflammation. Changes in the combined panel of these discriminatory metabolites has not yet been reported and may prove diagnostically effective in differentiating HAT from other endemics diseases. Furthermore, whilst a selection of significantly changed metabolites were also associated with mice infected with *T. b. brucei*, important differences were observed from mouse model, highlighting the necessity of validating experimental animal study data in human studies. The statistical models generated based from both NMR- and MS-based data were robust, thus the use of these technologies in an exploratory capacity is clear. However, it is obvious that neither NMR spectroscopy nor UPLC-MS are appropriate platforms for point-of-care in developing countries. Instead, the use of their technology lies in characterisation of the metabolic imprints of a parasitic infection, as shown in this study, with subsequent development of simpler clinical assays, e.g. based on biochemical dipstick

tests, on the strongest discriminatory biomarkers.

Immunological investigations have shown significant differences in inflammatory and regulatory cytokine concentrations between patients and controls, and between the two diagnostic stages of HAT. Yet the novelty of this study is not only the detection of additional metabolic differences between the stages but also the discovery of strong immune-metabolic correlations between energy-associated metabolites and host inflammatory responses within the brain of HAT patients, which appear to contribute towards neuropathogenesis of HAT.

As plasma lipids proved the most valuable candidates for HAT identification, further investigation of inflammatory lipid mediators such as prostaglandins and leukotrienes would also be of particular interest, to better understand pathology and provide key insights into links between lipid metabolism and immune responses in the host. Future examinations focused on lipid sub-species identification, e.g. with the use of MS/MS and lipid reference standards, and characterisation would also be required to help isolate candidates with most clinical diagnostic potential, to ultimately assist in the detection and management of *T. b. rhodesiense* HAT.

Chapter 6

General Discussion and Concluding Remarks

6.1 Summary of Findings 183
6.2 General Considerations and Challenges 188
6.3 Recommendations for Future Work 190
6.4 Conclusion 191

6.1 Summary of Findings

The overall aims of this PhD project were to investigate parasitic trypanosomatid infections, using a metabolic profiling approach, to explore possible metabolic interactions with other host response measures, in order to describe potential pathological mechanisms that underlie infectious disease. Using a combination of different molecular techniques and statistical tools, these aims have been successfully implemented to characterise experimental *L. major* infection in two mouse strains, as well as *T. b. rhodesiense* infection in humans. Significant infection-related changes in host metabolism were identified for both parasite species, which were predominately associated with inflammatory immune responses in the host, that appeared to negatively impact disease outcome.

6.1.1 Systemic Characterisation of Murine Strains

An in-depth systemic characterisation of BALB/c and C57BL/6 mice was achieved using a combination of ^{1}H NMR spectroscopy for metabolic analyses of multiple biofluids and tissues, MSD technology for plasma cytokine measurements (and additional haemocrit observations) as markers of host

immune/inflammatory status, and faecal microbial abundance *via* next-generation pyroseuquencing, as metrics for gut microbial diversity, as gut microbiota have been shown to impact parasitic infection responses *in vivo*. Each of these measures displayed significant differences between the two strains, with the urinary metabolic phenotype being the most discriminatory. Despite identical diet for both mouse strains, ^{1}H NMR results revealed inherent strain-associated differences in their glucose and lipid metabolism across numerous compartments, i.e. in plasma, urine, faeces, spleen, kidney and the hippocampus region of the brain. Such differences included variations in concentrations of lipid methyl and methylene moeities, and (triglyceride component) glycerol, which were all higher in BALB/c mice. C57BL/6 mice, on the other hand, had significantly higher levels of glucose, glycolytic and/or TCA intermediates e.g. pyruvate and citrate, as well as increased urinary and plasma bile acids. Such disparities in plasma lipid and glucose concentrations are in agreement with literature[205] and animal suppliers' information.[206–209] These inherent strain-related differences in metabolite abundances may signify a natural preference of C57BL/6 mice for the use of glucose as their main source of energy, whilst BALB/c mice also heavily depend on lipid catabolism.

Furthermore, strain-related divergences were observed in the gut microbial populations, highlighted by numerous differences in the faecal and urine concentrations of microbial-associated metabolites (phenyl derivatives in particular), and in relative faecal bacterial family abundances. *Firmicutes*, including bacterial species within the *Lachnospiraceae* family, were significantly higher in BALB/c mice, whilst C57BL/6 mice had augmented levels of *Erysipelotrichiaccae*.

This study also highlighted the presence of background inflammatory processes in BALB/c mice, compared to C57BL/6 mice, reflected by higher plasma concentrations of pro-inflammatory cytokines IL-1β and CXCL1, and a higher haematocrit. It is well known that this particular mouse strain is more susceptible to infection with a range of pathogens (e.g. *Leishmania* , *Listeria* and *Trypanosoma spp*), resulting in more aggravated disease than other strains, such as C57BL/6 mice.[203] It is thus plausible that the apparent ongoing inflammation in BALB/c mice could cause a continuous drain on the immune system within the host by redirecting resources, and thus impede the host from mounting an appropriate immune response towards infectious agents.

6.1.2 Sytemic Characterisation of Contrasting Murine Models of CL

Metabolic characterisation of biofluids from BALB/c mice infected with *L. major* , demonstrated significant parasite-associated changes of host metabolism in plasma and urine, already from the first few days of infection. Changes in plasma included a depletion of glucose and increased lipid

breakdown, reflected by a reduction of several lipid moieties and increased levels of ketone body, 3-hydroxybutyrate. Alterations in the urinary metabolic profiles included a decrease in hippurate concentration and an increase in phenylacetylglycine concentration, both of which appear to be commonly associated with parasitic infection since they have also been observed in several other parasitic infections in murine hosts.[92,93,150,257,258] Additionally, alanine and succinate, two major end-products of *Leishmania* metabolism[34,260], were increased after infection (in plasma and urine, respectively), raising the possibility that these metabolites could have been parasite-derived.

As for control mice, extensive metabolic phenotyping of biofluids and tissues, and cytokine and faecal microbial analyses were also performed for *L. major* infected animals. Infection was investigated and compared between two mouse models of CL, i.e. in non-healing BALB/c mice and in the self-healing C57BL/6 strain, in order to assess the various contributions of host response measures in relation to differing disease outcomes. Comparisons between the two mouse models revealed multiple significant divergences across the biological measures applied. Metabolic differences between the models were most prominent in the urine and plasma profiles, which included several metabolites that were also discriminatory for *L. major* infection, e.g. including augmented alanine and succinate and reduced urinary hippurate in non-healing mice compared with self-healing mice. The inherent differences in glucose and lipid metabolism between these two strains were even more pronounced during infection, presumably to deal with the increased energy demand exerted by host immunity to respond to parasitic infection.

Similar to the findings from uninfected mice, plasma cytokine measurements and haematocrit results in infected mice revealed a pro-inflammatory phenotype in the non-healing strain, which had significantly higher levels of IL-1β. Recent *Leishmania -in vivo* studies have shown that increased IL-1β function resulted in disease exacerbation, linked with increased levels of other inflammatory cytokines and suppressed T cell responses.[262] Consequently, evidence strongly suggests that (IL-1β-associated) inflammatory processes within the non-healing BALB/c strain hinders the generation of an effective immune response against *Leishmania*, already in the early stages of infection, actively promoting disease propagation.

Infection resulted in temporal changes in faecal microbial populations in self-healing C57BL/6 mice, whilst those in the non-healing strain remained relatively homoeostatic. Results demonstrated that the two most dominant bacterial classes found in this study, *Clostridia* (higher in non-healing mice) and *Gammaproteobacteria* (higher in self-healing mice) differed systematically between two mouse models, particularly evident towards the end of the study. They were also inversely correlated with each other, a relationship that has also been reported in rats undergoing weight loss following

bariatric surgery.[263] However, this has not been documented in infection models and determining whether this finding has a significant impact on infection outcome would thus require further investigation.

In addition to demonstrating differences in host response measures, for each biological matrix between the infection models, this study also uniquely applied a correlation network approach to analyse and visualise statistical connections between the different response measures. Such analyses revealed remarkable divergences in the number of correlations observed between the two mouse models and between sampling time-points. Notably, the self-healing model displayed a high degree of both inter- and intra-matrix correlations, already two days post-infection. These associations were organised into functional network hubs — namely acetate, citrate and alanine — that appeared to orchestrate a systemic response towards *Leishmania*. Non-healing mice, on the other hand, only displayed a high degree of associations from day 6 post-infection onwards, which unlike C57BL/6, appeared to be structured more randomly.

Experimental evidence has shown that parasitic clearance of *Leishmania* in C57BL/6 mice is linked to the recruitment of a diverse repetoire of leukocytes to the site of infection already from one day after infection, whereas recruitment in infected BALB/c mice is restricted to polymorphonuclear cells.[270] Combined, these results implicate the necessity for rapid communication and interaction across multiple host compartments to generate an effective immune response against *Leishmania* parasites, as observed in the self-healing model of CL. In contrast, the delay and disorganised interactions between metabolites, cytokines and microbiota observed in non-healing strain, possibly due to excessive inflammatory processes, results in an inefficient response to the infection, thus allowing parasitic survival.

6.1.3 Metabolic and Immune Characterisation of HAT

Despite the multitude of inherent and external factors that contribute towards the high variability observed in human metabolic phenotypes[98], this study represents one of the few that has been able to successfully implement metabolic profiling of human parasitic infections and determine disease-associated markers with diagnostic potential. Metabolic investigations were performed on plasma and CSF samples from Ugandan HAT *rhodesiense* patients and healthy endemic controls. Using ^{1}H NMR spectroscopy, striking perturbations, particularly in lipid and amino acid metabolism, were observed in plasma from HAT patients.

A total of 16 significantly altered metabolites were found to be discriminatory for HAT, with a collective specificity of 96.7% and sensitivity of 84.4% (in contrast to standard microscopic as-

sessment where sensitivity can range between 70–80%).[58] Metabolites whose concentrations were augmented in HAT included phenylalanine, *N*-acetylated glycoprotein, and lipid glyceryl in HAT patients. HAT was also associated with reductions in levels of histidine, cholesterol, and a range of lipid signals (e.g. methyl [CH_3] and ketene [CH_2CO] groups, and unsaturated fatty acid chains containing [$C{=}CCH_2C{=}C$] and [CH=CH] moieties).

Extended profiling of plasma lipids *via* UPLC-MS revealed an additional 37 HAT-characteristic metabolic markers, with a collective specificity of 93.8% and sensitivity of 85.7%. Significant differences between patients and controls included higher levels of a variety of triglyceride species and ceramide in HAT patients, alongside lowered levels in of a range lysophosphatidylcholines, polyunsaturated phosphatidylcholines and sphingomyelines.

Evidence from murine trypanosomiasis model demonstrated that certain infection-related metabolic features in plasma that were shared with this study, namely reduced lipid CH_3 and elevated acetylated glycoproteins.[92] However, the majority of the changes measured in the mouse model were not translated to humans, confirming the need to validate experimental findings in animal models with human samples for clinical relevance.

The general feature of HAT-associated metabolic changes in plasma was the link with inflammation. Many of the discriminatory metabolites were found to be directly correlated with pro-inflammatory cytokines, such as IL-6, or with general inflammatory clinical symptoms e.g. swollen lymph nodes, in this study. Many more have also been reported by others to be general markers of infection and/inflammation, e.g. hypertriglyceridaemia[293,302] and increased phenylalanine[279], and/or have been shown to be actively involved in inflammatory signalling, such as increased plasma concentrations of ceramide, which induces the synthesis of pro-inflammatory cytokines and chemokines, and can increase vascular permeability.[301]

The metabolic changes associated with HAT in CSF samples were less pronounced than in plasma samples; only 5 significantly discriminatory metabolite markers were determined *via* ^{1}H NMR. Whilst these metabolites may not represent optimal clinical markers of disease (collectively their sensitivity only reached 60.8%), they still gave valuable insights into possible mechanisms of neuropathology within HAT. HAT-associated differences highlighted an increased demand for energy, reflected by parasitic-induced depletion of glucose and glucosamine, possible osmolytic imbalances within the brain due to decreased *scyllo*-inositol[305,306], and potential increase in inflammation-induced oxidative stress, signified by increased concentrations of 2-oxobutanoate (produced as a by-product of glutathione synthesis).[307]

Unfortunately, no significant metabolic differences between early and late stage HAT were

determined in either plasma or CSF samples, though an apparent increase in lactate concentrations was observed in both biofluids. This may indicate increased glycolysis in the late stage of disease, most likely within host cells, as *Trypanosoma* species lack lactate dehydrogenase to produce lactate (and release pyruvate as glycolytic end-product instead).[32]

6.2 General Considerations and Challenges

Metabolic profiling has shown great capacity for giving a global perspective of changes in metabolism, which in the case for parasitic infection, involves the complex metabolic exchange between the parasites and the host. Much of the metabolic aspects presented in this thesis have been focused on the investigations of biofluids, such as plasma and urine. Whilst these biofluids represent samples of most practical clinical value (e.g. due to their simple, plentiful and relatively non-invasive method of collection), one should be mindful that any metabolic changes observed in these biological compartments reflect changes of the host organism, as a whole. It is thus extremely difficult to ascertain the exact origin of these changes in metabolites, whether it is determining the responsible organ and/or cell type in the host, or differentiating between host and parasite molecules. Additional studies that are specifically designed to track involved metabolic pathways, e.g. through the use of (radio- or fluorescently-) labelled metabolic substrates, would be required to help answer these questions, though this will prove particularly challenging when examining within complex *in vivo* systems.

Another important consideration for metabolic profiling is that this approach represents a hypothesis-generating method, with the purpose being the maximal capture of biological information.[88] Consequently, results produced with this method are highly descriptive in nature, including the disclosure of a list of discriminatory metabolites of interest, as has been the case in this PhD project. Whilst the validity of such markers can be demonstrated through statistical means in the presented studies, they still need to be confirmed in additional experiments, e.g. with the use of targeted assays specific for the markers in question, in order determine their robustness across experimental conditions.

The associations presented in this project, e.g. between metabolites and disease status or other host response measures, are based on statistical linkages, e.g. using correlation analyses. However, one cannot conclude neither the sequence order between measures in such associations, nor determine whether the associated measure were a cause or an effect of each other. Targeting specific metabolites and/or metabolic pathways that were discriminatory either for infection or a particular infection model (e.g. self-healing phenotype), through use of enzyme inhibitors/agonists

or dietary interventions, could help further elucidate their involvement in either disease propagation or amelioration.

It is also noteworthy that the gut microbial analyses were based on next-generation pyrosequencing of 16S ribosomal RNA genes. There are many benefits of this technology, such as its high sensitivity, accuracy and high-throughput.[183] This method relies on the sequencing of approximately 500 base pairs of 16S ribosomal DNA, which has been shown to be sufficient in reliably identifying most known microbial species but not all.[1][312] Complete microbial genome-sequencing, whilst more expensive and time-intensive, is a more comprehensive and unbiased approach for microbial screening (especially for describing new species), and would have provided a perspective that is more representative of the microbial population.[312] Another consideration is that whilst *Clostridia* has been shown to represent one of the more common gut microbial classes, *Gammaproteobacteria* are usually reported in high proportions only following an insult, e.g. experimental bariatric surgery[263] or infection-induced inflammation.[313] Instead, *Bacteroidia* are typically reported as the most dominant bacterial class, yet in this study they were only measured at low levels (usually <5%). It is unclear whether the atypical distribution of gut microbiota observed in this study represents an artefactual error, a possible enteropathogenic bacterial contamination in some of these mice, or simply natural variation in microbial population. No significant changes in animal weight or haematocrit were measured for a given strain, implying that the effects of any possible bacterial infection were subtle, if present.

One of the major challenges I faced in this PhD was actually the hallmark of this project, namely the quantity and complexity of acquired data. This was particularly the case for the murine *Leishmania* infection study, which consisted of metabolic spectroscopic from multiple biofluids and tissues, faecal microbial results and cytokine data, each from four different test groups, over four different time-points. As a result, different strategies of analyses could have been performed (and were in fact attempted), depending on the research question. In this study, the main aim was to characterise the difference between infection-related responses in self-healing vs. non-healing CL. Consequently, the results focused on comparing control-strain adjusted (*via* subtraction) data from infected strains, as was presented in Chapter 4. Yet much more latent information could have been presented from the acquired data, such as extensive comparisons between metabolic phenotypes of infected and non-infected samples for both murine models, including tissue samples, or alternative statistical methods for control-strain adjustment. Much of these analyses were in fact initially performed but finally aborted and omitted from the thesis, to allow for a more focused piece of

[1] For some, bacterial species, sequencing of the entire 1,500 bp 16S gene is needed.

work.

6.3 Recommendations for Future Work

The multiple infection-related metabolic changes and interactions with host immune and microbial responses, discovered in this PhD project, have opened even more avenues into the unknown. As has been highlighted previously, the main findings, such as the list of discriminatory markers, would first need to be confirmed in additional studies, ideally using a larger number of animals/patients. Additionally, investigations of *Leishmania* infection have been limited to mouse models for this PhD. Future characterisation of human cutaneous leishmanisis is of utmost importance, in order to determine the translatability of the mouse results, especially considering the current divergences that were observed between human and mouse *T. brucei* infection.

As the infective species is known to have a dramatic effect on disease outcome, for both leishmaniasis and HAT, it would be highly interesting to compare the current work with metabolic profiles of mice infected with at least one of the other clinically relevant *Leishmania* species, and with *T. b. gambiense* infection in humans. Eventhough rapid and accurate diagnostic screening methods already exist for HAT *gambiense*, metabolic phenotyping of plasma and CSF samples from *T. b. gambiense* infected individuals would still be beneficial in aiding the discrimination between HAT-associated markers that were either specific for *T. b. gambiense*/ *T. b. rhodesiense*, or common to both. Similarly, screening of murine infection models (and ultimately, human infections) with other clinically relevant *Leishmania* species, could aid the discovery of either species-specific or pan-leishmanial disease markers. Lastly, comparing the response of BALB/c and C57BL/6 mouse strains to other infections would also be interesting to study, to determine whether systematic differences are also observed elsewhere.

In the pursuit for novel diagnostic markers, the next (and by no means trivial) step would be the development of a clinical assay that incorporates the detection of the presented set of disease-discriminatory markers and is relatively simple, high-throughput and cost-effective, e.g. in the form of a dipstick- or lateral-flow-based immunoassay.[314] Subsequent validation of these markers/newly developed assays in large-scale longitudinal studies, comparing with the currently standard diagnostic methods, would finally be required to confirm the benefit of their implementation in the clinic.[314]

Additional exploratory ventures from this PhD would include further investigations into the mechanisms of action of the metabolites of interest, e.g. the network hubs, acetate and citrate, observed in the self-healing CL model. Preliminary investigations with these metabolites could

involve *in vitro* supplementations studies, by adding the aforementioned metabolites into culture media containing *L. major* cultures, to explore possible direct leishmanicidal effects, or with infected macrophage cultures, for possible indirect toxic effects. Such pilot studies have already been commenced in-house, using picolinic acid, an indirect stimulator of macrophage nitric oxide production, which demonstrated decreased parasitic burden and altered metabolic phenotypes *in vitro*.[271,272]

Hallmarks of human *T. b. rhodesiense* infection, on the other hand, included alterations in plasma lipid species, e.g. depletions of lysophosphatidylcholines and polyunsaturated phosphatidylcholines. Extended characterisation of these different lipid species, with the use of lipid reference standards and additional MS-MS are recommended to specify the exact composition of these species that were altered in HAT. Once identified, additional *in vitro* investigations of trypanosome viability in lipid-restricted culture media, may also help to shed light on which lipid species are most critical for parasite growth, and thus represent possible novel therapeutic targets.

6.4 Conclusion

Parasitic trypanosomatid infections cause several neglected tropical diseases in humans, including cutaneous leishmaniasis and human African trypanosomiasis. Both diseases remain poorly managed, with suboptimal diagnosis and dated, limited, and often toxic drugs, and would thus greatly benefit from the discovery of novel diagnostic and therapeutic targets through biomedical research. Using a combination of advanced metabolic profiling technology and multivariate and univariate data analyses, the presented findings provide the first comprehensive descriptions of systemic metabolic changes in the host, associated with experimental *L. major* infection in two murine models of cutaneous leishmaniasis, and with *T. b. rhodesiense* infection in human African trypanosomiasis.

Multiple statistical interactions between metabolites and other host response measures were discovered, particularly with pro-inflammatory cytokines, emphasising the connection between altered host metabolism and inflammation during infectious disease. Correlation network analysis was used as a statistical method to integrate and co-analyse metabolic, immunological and microbial data, which was presented in two different formats (as a wheel or network map), to facilitate interpretation of their associations. Such analysis uniquely revealed striking differences in the degree of interaction between host responses in the self-healing and the non-healing *L. major* infection models, an approach that can also be applied to investigate other diseases and/or physiological states. A highly connected and functionally organised systems response was generated shortly after infection in the self-healing CL model, linked to effective elimination of *Leishmania* parasites, yet the

non-healing strain displayed a delayed and disordered connections, thought to contribute towards an ineffective response.

As well as giving a global insight into underlying host-parasitic metabolic processes that may contribute to pathology, this PhD project has shown the remarkable exploratory capacity of metabolic profiling for the discovery of infection-discriminatory markers. A range of significantly differentially expressed metabolites were determined for both *L. major* and *T. b. rhodesiense* infections. These included increased concentrations of putative parasite-derived metabolic end-products, general markers of inflammation, and differences in a range of plasma lipid species, which have the potential be exploited as novel disease biomarkers. Providing more research into these markers is undertaken, only time will reveal their translatability and efficacy as new clinical diagnostic tools. The ultimate aim would be to develop a miniaturised test that can detect a panel of such biomakers with high accuracy yet be relatively inexpensive, to have a significant impact towards improving the management of cutaneous leishmaniasis and HAT, especially for use in the field.

Chapter 7

Bibliography

[1] World Health Organization. *Neglected Tropical Diseases*. URL: `http://www.who.int/neglected_diseases/diseases/en/` (visited on 04/10/2015).

[2] Peter J Hotez et al. 'The Global Burden of Disease Study 2010: Interpretation and Implications for the Neglected Tropical Diseases'. In: *PLoS Neglected Tropical Diseases* 8.7 (2014).

[3] Peter J Hotez, Aruna Kamath et al. 'Neglected tropical diseases in sub-Saharan Africa: review of their prevalence, distribution, and disease burden'. In: *PLoS Neglected Tropical Diseases* 3.8 (2009), e412.

[4] Keith Vickerman. 'The evolutionary expansion of the trypanosomatid flagellates'. In: *International Journal for Parasitology* 24.8 (1994), pp. 1317–1331.

[5] Theresa A Shapiro and Paul T Englund. 'The structure and replication of kinetoplast DNA'. In: *Annual Reviews in Microbiology* 49.1 (1995), pp. 117–143.

[6] Paul AM Michels et al. 'Metabolic functions of glycosomes in trypanosomatids'. In: *Biochimica et Biophysica Acta (BBA)-Molecular Cell Research* 1763.12 (2006), pp. 1463–1477.

[7] Richard John Wheeler, Eva Gluenz and Keith Gull. 'The Limits on Trypanosomatid Morphological Diversity'. In: *PLoS ONE* 8.11 (2013).

[8] Christopher JL Murray et al. 'Disability-adjusted life years (DALYs) for 291 diseases and injuries in 21 regions, 1990–2010: a systematic analysis for the Global Burden of Disease Study 2010'. In: *The Lancet* 380.9859 (2013), pp. 2197–2223.

[9] P Desjeux. 'Leishmaniasis: current situation and new perspectives'. In: *Comparative Immunology, Microbiology and Infectious Diseases* 27.5 (2004), pp. 305–318.

[10] Richard Reithinger et al. 'Cutaneous leishmaniasis'. In: *The Lancet Infectious Diseases* 7.9 (2007), pp. 581–596.

[11] François Chappuis et al. 'Visceral leishmaniasis: what are the needs for diagnosis, treatment and control?' In: *Nature Reviews Microbiology* 5.11 (2007), pp. 873–882.

[12] Piero Olliaro et al. 'Methodology of clinical trials aimed at assessing interventions for cutaneous leishmaniasis'. In: *PLoS Neglected Tropical Diseases* 7.3 (2013), e2130.

[13] Tim Van Assche et al. 'Leishmania–macrophage interactions: insights into the redox biology'. In: *Free Radical Biology and Medicine* 51.2 (2011), pp. 337–351.

[14] Simon L Croft and P Olliaro. 'Leishmaniasis chemotherapy—challenges and opportunities'. In: *Clinical Microbiology and Infection* 17.10 (2011), pp. 1478–1483.

[15] Rajiv Kumar and Christian Engwerda. 'Vaccines to prevent leishmaniasis'. In: *Clinical & Translational Immunology* 3.3 (2014), e13.

[16] Richard Reithinger and Jean-Claude Dujardin. 'Molecular diagnosis of leishmaniasis: current status and future applications'. In: *Journal of Clinical Microbiology* 45.1 (2007), pp. 21–25.

[17] Kristen A Weigle et al. 'Leishmanin skin test standardization and evaluation of safety, dose, storage, longevity of reaction and sensitization.' In: *The American Journal of Tropical Medicine and Hygiene* 44.3 (1991), pp. 260–271.

[18] Esther Bensoussan et al. 'Comparison of PCR assays for diagnosis of cutaneous leishmaniasis'. In: *Journal of Clinical Microbiology* 44.4 (2006), pp. 1435–1439.

[19] Henry W Murray et al. 'Advances in leishmaniasis'. In: *The Lancet* 366.9496 (2005), pp. 1561–1577.

[20] Susan Wyllie, Mark L Cunningham and Alan H Fairlamb. 'Dual action of antimonial drugs on thiol redox metabolism in the human pathogen Leishmania donovani'. In: *Journal of Biological Chemistry* 279.38 (2004), pp. 39925–39932.

[21] Nisha Singh, Manish Kumar and Rakesh Kumar Singh. 'Leishmaniasis: current status of available drugs and new potential drug targets'. In: *Asian Pacific Journal of Tropical Medicine* 5.6 (2012), pp. 485–497.

[22] Tatiana S Tiuman et al. 'Recent advances in leishmaniasis treatment'. In: *International Journal of Infectious Diseases* 15.8 (2011), e525–e532.

[23] Simon L Croft and Graham H Coombs. 'Leishmaniasis–current chemotherapy and recent advances in the search for novel drugs'. In: *Trends in Parasitology* 19.11 (2003), pp. 502–508.

[24] Richard W Ashford. 'Leishmaniasis reservoirs and their significance in control'. In: *Clinics in Dermatology* 14.5 (1996), pp. 523–532.

[25] Paul Kaye and Phillip Scott. 'Leishmaniasis: complexity at the host–pathogen interface'. In: *Nature Reviews Microbiology* 9.8 (2011), pp. 604–615.

[26] Thomas Naderer and Malcolm J McConville. 'The Leishmania–macrophage interaction: a metabolic perspective'. In: *Cellular microbiology* 10.2 (2008), pp. 301–308.

[27] Gerald F Späth et al. 'The role (s) of lipophosphoglycan (LPG) in the establishment of Leishmania major infections in mammalian hosts'. In: *Proceedings of the National Academy of Sciences* 100.16 (2003), pp. 9536–9541.

[28] Salvatore J Turco and Albert Descoteaux. 'The lipophosphoglycan of Leishmania parasites'. In: *Annual Reviews in Microbiology* 46.1 (1992), pp. 65–92.

[29] Efrat Barak et al. 'Differentiation of Leishmania donovani in host-free system: analysis of signal perception and response'. In: *Molecular and Biochemical Parasitology* 141.1 (2005), pp. 99–108.

[30] Alessio Bocedi et al. 'Trypanothione efficiently intercepts nitric oxide as a harmless iron complex in trypanosomatid parasites'. In: *The FASEB Journal* 24.4 (2010), pp. 1035–1042.

[31] David Sacks and Nancy Noben-Trauth. 'The immunology of susceptibility and resistance to Leishmania major in mice'. In: *Nature Reviews Immunology* 2.11 (2002), pp. 845–858.

[32] Aloysius GM Tielens and Jaap J van Hellemond. 'Surprising variety in energy metabolism within Trypanosomatidae'. In: *Trends in Parasitology* 25.10 (2009), pp. 482–490.

[33] Eleanor C Saunders et al. 'Central carbon metabolism of Leishmania parasites'. In: *Parasitology* 137.09 (2010), pp. 1303–1313.

[34] Fred R Opperdoes and Graham H Coombs. 'Metabolism of Leishmania: proven and predicted'. In: *Trends in Parasitology* 23.4 (2007), pp. 149–158.

[35] Doron Rosenzweig et al. 'Retooling Leishmania metabolism: from sand fly gut to human macrophage'. In: *The FASEB journal* 22.2 (2008), pp. 590–602.

[36] Thomas Naderer et al. 'Virulence of Leishmania major in macrophages and mice requires the gluconeogenic enzyme fructose-1, 6-bisphosphatase'. In: *Proceedings of the National Academy of Sciences* 103.14 (2006), pp. 5502–5507.

[37] Phillip Scott. 'IFN-gamma modulates the early development of Th1 and Th2 responses in a murine model of cutaneous leishmaniasis.' In: *The Journal of Immunology* 147.9 (1991), pp. 3149–3155.

[38] FP Heinzel et al. 'Recombinant interleukin 12 cures mice infected with Leishmania major.' In: *The Journal of Experimental Medicine* 177.5 (1993), pp. 1505–1509.

[39] JP Sypek et al. 'Resolution of cutaneous leishmaniasis: interleukin 12 initiates a protective T helper type 1 immune response.' In: *The Journal of Experimental Medicine* 177.6 (1993), pp. 1797–1802.

[40] Audrey Y Park, Brian D Hondowicz and Phillip Scott. 'IL-12 is required to maintain a Th1 response during Leishmania major infection'. In: *The Journal of Immunology* 165.2 (2000), pp. 896–902.

[41] Kazuo Yamakami et al. 'A single intradermal administration of soluble leishmanial antigen and plasmid expressing interleukin-12 protects BALB/c mice from Leishmaniamajor infection'. In: *Parasitology International* 50.2 (2001), pp. 81–91.

[42] Virginia Iniesta et al. 'Arginase I induction during Leishmania major infection mediates the development of disease'. In: *Infection and Immunity* 73.9 (2005), pp. 6085–6090.

[43] Pascale Kropf et al. 'Arginase and polyamine synthesis are key factors in the regulation of experimental leishmaniasis in vivo'. In: *The FASEB journal* 19.8 (2005), pp. 1000–1002.

[44] Sabrina D Lamour et al. 'Metabolic characterization of Leishmania major infection in activated and nonactivated macrophages.' In: *Journal of Proteome Research* 11.8 (2012), pp. 4211–4222.

[45] Evelyne Mougneau, Franck Bihl and Nicolas Glaichenhaus. 'Cell biology and immunology of Leishmania'. In: *Immunological Reviews* 240.1 (2011), pp. 286–296.

[46] Liang Zhou, Mark MW Chong and Dan R Littman. 'Plasticity of CD4+ T cell lineage differentiation'. In: *Immunity* 30.5 (2009), pp. 646–655.

[47] John J O'Shea and William E Paul. 'Mechanisms underlying lineage commitment and plasticity of helper CD4+ T cells'. In: *Science* 327.5969 (2010), pp. 1098–1102.

[48] Tanya Scharton-Kersten and Phillip Scott. 'The role of the innate immune response in Th1 cell development following Leishmania major infection.' In: *Journal of Leukocyte Biology* 57.4 (1995), pp. 515–522.

[49] Florian Woelbing et al. 'Uptake of Leishmania major by dendritic cells is mediated by Fcγ receptors and facilitates acquisition of protective immunity'. In: *The Journal of Experimental Medicine* 203.1 (2006), p. 177.

[50] Nathan C Peters et al. 'In vivo imaging reveals an essential role for neutrophils in leishmaniasis transmitted by sand flies'. In: *Science* 321.5891 (2008), pp. 970–974.

[51] C Leandro et al. 'Cell mediated immunity and specific IgG1 and IgG2 antibody response in natural and experimental canine leishmaniosis'. In: *Veterinary Immunology and Immunopathology* 79.3 (2001), pp. 273–284.

[52] Anirban Ghosh, Wen Wei Zhang and Greg Matlashewski. 'Immunization with A2 protein results in a mixed Th1/Th2 and a humoral response which protects mice against Leishmania donovani infections'. In: *Vaccine* 20.1 (2001), pp. 59–66.

[53] Yasmine Belkaid et al. 'CD4+ CD25+ regulatory T cells control Leishmania major persistence and immunity'. In: *Nature* 420.6915 (2002), pp. 502–507.

[54] MR Oliveira et al. 'Germ-free mice produce high levels of interferon-gamma in response to infection with Leishmania major but fail to heal lesions'. In: *Parasitology* 131.04 (2005), pp. 477–488.

[55] C Pirmez et al. 'Cytokine patterns in the pathogenesis of human leishmaniasis.' In: *Journal of Clinical Investigation* 91.4 (1993), p. 1390.

[56] Kai Zwingenberger et al. 'Determinants of the immune response in visceral leishmaniasis: evidence for predominance of endogenous interleukin 4 over interferon-γ production'. In: *Clinical Immunology and Immunopathology* 57.2 (1990), pp. 242–249.

[57] Hashim W Ghalib et al. 'Interleukin 10 production correlates with pathology in human Leishmania donovani infections.' In: *Journal of Clinical Investigation* 92.1 (1993), p. 324.

[58] Reto Brun et al. 'Human african trypanosomiasis'. In: *The Lancet* 375.9709 (2010), pp. 148–159.

[59] Florencia La Greca and Stefan Magez. 'Vaccination against trypanosomiasis: can it be done or is the trypanosome truly the ultimate immune destroyer and escape artist?' In: *Human Vaccines* 7.11 (2011), pp. 1225–1233.

[60] D Malvy and François Chappuis. 'Sleeping sickness'. In: *Clinical Microbiology and Infection* 17.7 (2011), pp. 986–995.

[61] World Health Organization et al. 'Accelerating work to overcome the global impact of neglected tropical diseases: a roadmap for implementation: executive summary'. In: (2012).

[62] PP Simarro et al. 'Diversity of human African trypanosomiasis epidemiological settings requires fine-tuning control strategies to facilitate disease elimination.' In: *Research & Reports in Tropical Medicine* 4 (2013).

[63] World Health Organization et al. 'Control and surveillance of human African trypanosomiasis.' In: *World Health Organization technical report series* 984 (2013), p. 1.

[64] Dominique Legros et al. 'Treatment of human African trypanosomiasis—present situation and needs for research and development'. In: *The Lancet Infectious Diseases* 2.7 (2002), pp. 437–440.

[65] François Chappuis et al. 'Options for field diagnosis of human African trypanosomiasis'. In: *Clinical Microbiology Reviews* 18.1 (2005), pp. 133–146.

[66] JM Sternberg. 'Human African trypanosomiasis: clinical presentation and immune response'. In: *Parasite Immunology* 26.11-12 (2004), pp. 469–476.

[67] World Health Organization et al. 'Control and surveillance of African trypanosomiasis: report of a WHO expert committee'. In: (1998).

[68] Michael P Barrett et al. 'Human African trypanosomiasis: pharmacological re-engagement with a neglected disease'. In: *British Journal of Pharmacology* 152.8 (2007), pp. 1155–1171.

[69] Jose R Franco et al. 'Monitoring the use of nifurtimox-eflornithine combination therapy (NECT) in the treatment of second stage gambiense human African trypanosomiasis'. In: *Journal of Research and Reports in Tropical Medicine* 3 (2012), pp. 93–101.

[70] Monika Iten et al. 'Alterations in ornithine decarboxylase characteristics account for tolerance of Trypanosoma brucei rhodesiense to D, L-alpha-difluoromethylornithine.' In: *Antimicrobial Agents and Chemotherapy* 41.9 (1997), pp. 1922–1925.

[71] Katelyn Fenn and Keith R Matthews. 'The cell biology of Trypanosoma brucei differentiation'. In: *Current Opinion in Microbiology* 10.6 (2007), pp. 539–546.

[72] Gerasimos Langousis and Kent L Hill. 'Motility and more: the flagellum of Trypanosoma brucei'. In: *Nature Reviews Microbiology* 12.7 (2014), pp. 505–518.

[73] Keith R Matthews. 'The developmental cell biology of Trypanosoma brucei'. In: *Journal of Cell Science* 118.2 (2005), pp. 283–290.

[74] Frédéric Bringaud, Loïc Rivière and Virginie Coustou. 'Energy metabolism of trypanosomatids: adaptation to available carbon sources'. In: *Molecular and Biochemical Parasitology* 149.1 (2006), pp. 1–9.

[75] Susanne Diehl et al. 'Analysis of stage-specific gene expression in the bloodstream and the procyclic form of Trypanosoma brucei using a genomic DNA-microarray'. In: *Molecular and Biochemical Parasitology* 123.2 (2002), pp. 115–123.

[76] Matthew Berriman et al. 'The genome of the African trypanosome Trypanosoma brucei'. In: *Science* 309.5733 (2005), pp. 416–422.

[77] Philippe Vincendeau and Bernard Bouteille. 'Immunology and immunopathology of African trypanosomiasis'. In: *Anais da Academia Brasileira de Ciências* 78.4 (2006), pp. 645–665.

[78] Lorna MacLean, Martin Odiit and Jeremy M Sternberg. 'Nitric oxide and cytokine synthesis in human African trypanosomiasis'. In: *Journal of Infectious Diseases* 184.8 (2001), pp. 1086–1090.

[79] Lorna MacLean et al. 'Severity of human African trypanosomiasis in East Africa is associated with geographic location, parasite genotype, and host inflammatory cytokine response profile'. In: *Infection and Immunity* 72.12 (2004), pp. 7040–7044.

[80] David Courtin et al. 'Comparison of cytokine plasma levels in human African trypanosomiasis'. In: *Tropical Medicine & International Health* 11.5 (2006), pp. 647–653.

[81] Alexandre Hainard et al. 'A combined CXCL10, CXCL8 and H-FABP panel for the staging of human African trypanosomiasis patients'. In: *PLoS Neglected Tropical Diseases* 3.6 (2009), e459–e459.

[82] Jeremy M Sternberg et al. 'Meningoencephalitic African trypanosomiasis: brain IL-10 and IL-6 are associated with protection from neuro-inflammatory pathology'. In: *Journal of Neuroimmunology* 167.1 (2005), pp. 81–89.

[83] CA Hunter et al. 'Immunopathology of experimental African sleeping sickness: detection of cytokine mRNA in the brains of Trypanosoma brucei brucei-infected mice.' In: *Infection and Immunity* 59.12 (1991), pp. 4636–4640.

[84] Chris A Hunter et al. 'Astrocyte activation correlates with cytokine production in central nervous system of Trypanosoma brucei brucei-infected mice.' In: *Laboratory Investigation; a journal of technical methods and pathology* 67.5 (1992), pp. 635–642.

[85] JK Nicholson, JC Lindon and E Holmes. 'Metabonomics: understanding the metabolic responses of living systems to pathophysiological stimuli via multivariate statistical analysis of biological NMR spectroscopic data'. In: *Xenobiotica* 29.11 (1999), pp. 1181–1189.

[86] John C Lindon, Elaine Holmes and Jeremy K Nicholson. 'Metabonomics techniques and applications to pharmaceutical research & development'. In: *Pharmaceutical Research* 23.6 (2006), pp. 1075–1088.

[87] Warwick B Dunn and David I Ellis. 'Metabolomics: current analytical platforms and methodologies'. In: *TrAC Trends in Analytical Chemistry* 24.4 (2005), pp. 285–294.

[88] Warwick B Dunn et al. 'Systems level studies of mammalian metabolomes: the roles of mass spectrometry and nuclear magnetic resonance spectroscopy'. In: *Chemical Society Reviews* 40.1 (2011), pp. 387–426.

[89] Rongwei Teng et al. 'Metabolite profiling of the intraerythrocytic malaria parasite Plasmodium falciparum by 1H NMR spectroscopy'. In: *NMR in Biomedicine* 22.3 (2009), pp. 292–302.

[90] Kellen L Olszewski et al. 'Host-parasite interactions revealed by Plasmodium falciparum metabolomics'. In: *Cell Host & Microbe* 5.2 (2009), pp. 191–199.

[91] Yulan Wang et al. 'Metabonomic investigations in mice infected with Schistosoma mansoni: an approach for biomarker identification'. In: *Proceedings of the National Academy of Sciences* 101.34 (2004), pp. 12676–12681.

[92] Yulan Wang et al. 'Global metabolic responses of mice to Trypanosoma brucei brucei infection'. In: *Proceedings of the National Academy of Sciences* 105.16 (2008), pp. 6127–6132.

[93] Jasmina Saric et al. 'Metabolic profiling of an Echinostoma caproni infection in the mouse for biomarker discovery'. In: *PLoS Neglected Tropical Diseases* 2.7 (2008), e254.

[94] Lucienne Tritten et al. 'Metabolic profiling framework for discovery of candidate diagnostic markers of malaria'. In: *Scientific reports* 3 (2013).

[95] Lucienne Tritten et al. 'Comparing systemic metabolic responses in mice to single or dual infection with Plasmodium berghei and Heligmosomoides bakeri'. In: *Molecular BioSystems* 10.9 (2014), pp. 2358–2367.

[96] A Kamleh et al. 'Metabolomic profiling using Orbitrap Fourier transform mass spectrometry with hydrophilic interaction chromatography: a method with wide applicability to analysis

of biomolecules'. In: *Rapid Communications in Mass Spectrometry* 22.12 (2008), pp. 1912–1918.

[97] Elaine Holmes et al. 'Human metabolic phenotype diversity and its association with diet and blood pressure'. In: *Nature* 453.7193 (2008), pp. 396–400.

[98] Cinzia Stella et al. 'Susceptibility of human metabolic phenotypes to dietary modulation'. In: *Journal of Proteome Research* 5.10 (2006), pp. 2780–2788.

[99] Tiziana Pacchiarotta, André M Deelder and Oleg A Mayboroda. 'Metabolomic investigations of human infections'. In: *Bioanalysis* 4.8 (2012), pp. 919–925.

[100] Judith R Denery et al. 'Metabolomics-based discovery of diagnostic biomarkers for onchocerciasis'. In: *PLoS Neglected Tropical Diseases* 4.10 (2010), e834.

[101] Burton H Singer et al. 'Exploiting the potential of metabonomics in large population studies: three venues'. In: *The Handbook of Metabonomics and Metabolomics* (2007), pp. 289–325.

[102] Crina IA Balog et al. 'Metabonomic investigation of human Schistosoma mansoni infection'. In: *Molecular Biosystems* 7.5 (2011), pp. 1473–1480.

[103] Sabrina D Lamour et al. 'Changes in metabolic phenotypes of Plasmodium falciparum in vitro cultures during gametocyte development'. In: *Malaria Journal* 13.1 (2014), p. 468.

[104] Harald Günther. *NMR spectroscopy: basic principles, concepts and applications in chemistry*. John Wiley & Sons, 2013.

[105] Peter J Hore. *Nuclear Magnetic Resonance (Oxford Chemistry Primers)*. Oxford, UK: Oxford University Press Inc., May 1995.

[106] Hector C Keun et al. 'Analytical reproducibility in 1H NMR-based metabonomic urinalysis'. In: *Chemical Research in Toxicology* 15.11 (2002), pp. 1380–1386.

[107] Marc-Emmanuel Dumas et al. 'Assessment of analytical reproducibility of 1H NMR spectroscopy based metabonomics for large-scale epidemiological research: the INTERMAP Study'. In: *Analytical Chemistry* 78.7 (2006), pp. 2199–2208.

[108] Olaf Beckonert et al. 'Metabolic profiling, metabolomic and metabonomic procedures for NMR spectroscopy of urine, plasma, serum and tissue extracts'. In: *Nature Protocols* 2.11 (2007), pp. 2692–2703.

[109] Olaf Beckonert et al. 'High-resolution magic-angle-spinning NMR spectroscopy for metabolic profiling of intact tissues'. In: *Nature Protocols* 5.6 (2010), pp. 1019–1032.

[110] NJ Waters et al. 'High-resolution magic angle spinning 1 H NMR spectroscopy of intact liver and kidney: optimization of sample preparation procedures and biochemical stability of tissue during spectral acquisition'. In: *Analytical Biochemistry* 282.1 (2000), pp. 16–23.

[111] Saul Meiboom and David Gill. 'Modified spin-echo method for measuring nuclear relaxation times'. In: *Review of Scientific Instruments* 29.8 (1958), pp. 688–691.

[112] Kirill A Veselkov et al. 'Recursive segment-wise peak alignment of biological 1H NMR spectra for improved metabolic biomarker recovery'. In: *Analytical Chemistry* 81.1 (2008), pp. 56–66.

[113] Frank Dieterle et al. 'Probabilistic quotient normalization as robust method to account for dilution of complex biological mixtures. Application in 1H NMR metabonomics'. In: *Analytical Chemistry* 78.13 (2006), pp. 4281–4290.

[114] PJD Foxall et al. 'Analysis of biological fluids using 600 MHz proton NMR spectroscopy: application of homonuclear two-dimensional J-resolved spectroscopy to urine and blood plasma for spectral simplification and assignment'. In: *Journal of Pharmaceutical and Biomedical Analysis* 11.1 (1993), pp. 21–31.

[115] Nikolaos Psychogios et al. 'The human serum metabolome'. In: *PLoS ONE* 6.2 (2011), e16957.

[116] Katja Dettmer, Pavel A Aronov and Bruce D Hammock. 'Mass spectrometry-based metabolomics'. In: *Mass Spectrometry Reviews* 26.1 (2007), pp. 51–78.

[117] James J Pitt. 'Principles and applications of liquid chromatography-mass spectrometry in clinical biochemistry'. In: *The Clinical Biochemist Reviews* 30.1 (2009), p. 19.

[118] WMA Niessen. 'Progress in liquid chromatography–mass spectrometry instrumentation and its impact on high-throughput screening'. In: *Journal of Chromatography A* 1000.1 (2003), pp. 413–436.

[119] Walter A Korfmacher. 'Foundation review: Principles and applications of LC-MS in new drug discovery'. In: *Drug Discovery Today* 10.20 (2005), pp. 1357–1367.

[120] Giorgis Isaac, Stephen McDonald and Giuseppe Astarita. 'Lipid separation using UPLC with charged surface hybrid technology'. In: *Waters Application Note* (2011).

[121] Panagiotis A Vorkas et al. 'Untargeted UPLC-MS profiling pipeline to expand tissue metabolome coverage: Application to cardiovascular disease'. In: *Analytical Chemistry* 87.8 (2015), pp. 4184–4193.

[122] Elizabeth J Want et al. 'Global metabolic profiling procedures for urine using UPLC–MS'. In: *Nature Protocols* 5.6 (2010), pp. 1005–1018.

[123] Elizabeth J Want et al. 'Global metabolic profiling of animal and human tissues via UPLC-MS'. In: *Nature Protocols* 8.1 (2013), pp. 17–32.

[124] Jun Feng Xiao, Bin Zhou and Habtom W Ressom. 'Metabolite identification and quantitation in LC-MS/MS-based metabolomics'. In: *TrAC Trends in Analytical Chemistry* 32 (2012), pp. 1–14.

[125] Anas El-Aneed, Aljandro Cohen and Joseph Banoub. 'Mass spectrometry, review of the basics: electrospray, MALDI, and commonly used mass analyzers'. In: *Applied Spectroscopy Reviews* 44.3 (2009), pp. 210–230.

[126] Thomas M Annesley. 'Ion suppression in mass spectrometry'. In: *Clinical Chemistry* 49.7 (2003), pp. 1041–1044.

[127] Scott J Geromanos et al. 'The detection, correlation, and comparison of peptide precursor and product ions from data independent LC-MS with data dependant LC-MS/MS'. In: *Proteomics* 9.6 (2009), pp. 1683–1695.

[128] Magali H Sarafian et al. 'Objective Set of Criteria for Optimization of Sample Preparation Procedures for Ultra-High Throughput Untargeted Blood Plasma Lipid Profiling by Ultra Performance Liquid Chromatography–Mass Spectrometry'. In: *Analytical Chemistry* 86.12 (2014), pp. 5766–5774.

[129] Colin A. Smith et al. 'XCMS: Processing Mass Spectrometry Data for Metabolite Profiling Using Nonlinear Peak Alignment, Matching, and Identification'. In: *Analytical Chemistry* 78.3 (2006), pp. 779–787. URL: `http://dx.doi.org/10.1021/ac051437y`.

[130] Robert A van den Berg et al. 'Centering, scaling, and transformations: improving the biological information content of metabolomics data'. In: *BMC Genomics* 7.1 (2006), p. 142.

[131] Svante Wold, Kim Esbensen and Paul Geladi. 'Principal component analysis'. In: *Chemometrics and Intelligent Laboratory Systems* 2.1 (1987), pp. 37–52.

[132] Harold Hotelling. 'Analysis of a complex of statistical variables into principal components.' In: *Journal of Educational Psychology* 24.6 (1933), p. 417.

[133] R Stoyanova and TR Brown. 'NMR spectral quantitation by principal component analysis'. In: *NMR in Biomedicine* 14.4 (2001), pp. 271–277.

[134] Lydia E Kavraki. 'Dimensionality Reduction Methods for Molecular Motion'. In: *Geometric Methods in Structural Computational Biology*. 10th ed. Online e-book. Texas, USA: Rice University, Openstax Connexions, July 2003. URL: http://cnx.org/contents/02ff5dd2-fe30-4bf5-8e2a-83b5c3dc0333@10/Dimensionality_Reduction_Metho.

[135] Svante Wold et al. 'The collinearity problem in linear regression. The partial least squares (PLS) approach to generalized inverses'. In: *SIAM Journal on Scientific and Statistical Computing* 5.3 (1984), pp. 735–743.

[136] Svante Wold, Michael Sjöström and Lennart Eriksson. 'PLS-regression: a basic tool of chemometrics'. In: *Chemometrics and Intelligent Laboratory Systems* 58.2 (2001), pp. 109–130.

[137] Lennart Eriksson et al. 'Using chemometrics for navigating in the large data sets of genomics, proteomics, and metabonomics (gpm)'. In: *Analytical and Bioanalytical Chemistry* 380.3 (2004), pp. 419–429.

[138] Svante Wold et al. 'Orthogonal signal correction of near-infrared spectra'. In: *Chemometrics and Intelligent Laboratory Systems* 44.1 (1998), pp. 175–185.

[139] Max Bylesjö et al. 'OPLS discriminant analysis: combining the strengths of PLS-DA and SIMCA classification'. In: *Journal of Chemometrics* 20.8-10 (2006), pp. 341–351.

[140] Olivier Cloarec et al. 'Evaluation of the orthogonal projection on latent structure model limitations caused by chemical shift variability and improved visualization of biomarker changes in 1H NMR spectroscopic metabonomic studies'. In: *Analytical Chemistry* 77.2 (2005), pp. 517–526.

[141] Umetrics. *SIMCA –P and Multivariate Analysis Frequently Asked Questions*. Version 1.01. URL: http://www.umetrics.com/sites/default/files/kb/multivariate_faq.pdf (visited on 29/04/2015).

[142] David S Wishart et al. 'HMDB: the human metabolome database'. In: *Nucleic Acids Research* 35.suppl 1 (2007), pp. D521–D526.

[143] David S Wishart et al. 'HMDB: a knowledgebase for the human metabolome'. In: *Nucleic Acids Research* 37.suppl 1 (2009), pp. D603–D610.

[144] David S Wishart et al. 'HMDB 3.0—the human metabolome database in 2013'. In: *Nucleic Acids Research* (2012), gks1065.

[145] Jeremy K Nicholson et al. '750 MHz 1H and 1H-13C NMR spectroscopy of human blood plasma'. In: *Analytical Chemistry* 67.5 (1995), pp. 793–811.

[146] Brian C Sweatman et al. '600 MHz 1 H-NMR spectroscopy of human cerebrospinal fluid: effects of sample manipulation and assignment of resonances'. In: *Journal of Pharmaceutical and Biomedical Analysis* 11.8 (1993), pp. 651–664.

[147] David S Wishart et al. 'The human cerebrospinal fluid metabolome'. In: *Journal of Chromatography B* 871.2 (2008), pp. 164–173.

[148] Marc-Emmanuel Dumas et al. 'Metabolic profiling reveals a contribution of gut microbiota to fatty liver phenotype in insulin-resistant mice'. In: *Proceedings of the National Academy of Sciences* 103.33 (2006), pp. 12511–12516.

[149] Francois-Pierre J Martin et al. 'Probiotic modulation of symbiotic gut microbial–host metabolic interactions in a humanized microbiome mouse model'. In: *Molecular Systems Biology* 4.1 (2008), p. 157.

[150] Jia V Li et al. 'Chemometric analysis of biofluids from mice experimentally infected with Schistosoma mansoni'. In: *Parasites & Vectors* 4 (2011), p. 179.

[151] Jasmina Saric et al. 'Species variation in the fecal metabolome gives insight into differential gastrointestinal function'. In: *Journal of Proteome Research* 7.01 (2007), pp. 352–360.

[152] Varanavasi Govindaraju, Karl Young and Andrew A Maudsley. 'Proton NMR chemical shifts and coupling constants for brain metabolites'. In: *NMR in Biomedicine* 13.3 (2000), pp. 129–153.

[153] François-Pierre J Martin et al. 'A top-down systems biology view of microbiome-mammalian metabolic interactions in a mouse model'. In: *Molecular Systems Biology* 3.1 (2007), p. 112.

[154] Sandrine P Claus et al. 'Systemic multicompartmental effects of the gut microbiome on mouse metabolic phenotypes'. In: *Molecular Systems Biology* 4.1 (2008), p. 219.

[155] Jia V Li et al. 'Metabolic profiling of a Schistosoma mansoni infection in mouse tissues using magic angle spinning-nuclear magnetic resonance spectroscopy'. In: *International Journal for Parasitology* 39.5 (2009), pp. 547–558.

[156] Olivier Cloarec et al. 'Statistical total correlation spectroscopy: an exploratory approach for latent biomarker identification from metabolic 1H NMR data sets'. In: *Analytical Chemistry* 77.5 (2005), pp. 1282–1289.

[157] Susanne Wiklund et al. 'Visualization of GC/TOF-MS-based metabolomics data for identification of biochemically interesting compounds using OPLS class models'. In: *Analytical Chemistry* 80.1 (2008), pp. 115–122.

[158] Panagiotis A Vorkas et al. 'Metabolic phenotyping of atherosclerotic plaques reveals latent associations between free cholesterol and ceramide metabolism in atherogenesis'. In: *Journal of Proteome Research* 14.3 (2015), pp. 1389–1399.

[159] Koen Sandra et al. 'Comprehensive blood plasma lipidomics by liquid chromatography/quadrupole time-of-flight mass spectrometry'. In: *Journal of Chromatography A* 1217.25 (2010), pp. 4087–4099.

[160] Gerhard Liebisch et al. 'Shorthand notation for lipid structures derived from mass spectrometry'. In: *Journal of Lipid Research* (2013), jlr–M033506.

[161] Ken-ichi Arai et al. 'Cytokines: coordinators of immune and inflammatory responses'. In: *Annual Review of Biochemistry* 59.1 (1990), pp. 783–836.

[162] Charles A Janeway et al. *Immunobiology: the immune system in health and disease*. Vol. 2. Churchill Livingstone, 2001.

[163] R Thorpe et al. 'Detection and measurement of cytokines'. In: *Blood Reviews* 6.3 (1992), pp. 133–148.

[164] Anthony Meager. 'Measurement of cytokines by bioassays: theory and application'. In: *Methods* 38.4 (2006), pp. 237–252.

[165] Joseph G Cannon et al. 'Measuring circulating cytokines'. In: *Journal of Applied Physiology* 75.4 (1993), pp. 1897–1902.

[166] Sean X Leng et al. 'ELISA and multiplex technologies for cytokine measurement in inflammation and aging research'. In: *The Journals of Gerontology Series A: Biological Sciences and Medical Sciences* 63.8 (2008), pp. 879–884.

[167] Meso Scale Diagnostics Llc. *Electrochemiluminescence*. URL: `https://www.mesoscale.com/en/technical_resources/our_technology/ecl/` (visited on 08/05/2015).

[168] Lorna MacLean et al. 'Stage progression and neurological symptoms in Trypanosoma brucei rhodesiense sleeping sickness: role of the CNS inflammatory response'. In: *PLoS Neglected Tropical Diseases* 6.10 (2012), e1857.

[169] Andrew L Kau et al. 'Human nutrition, the gut microbiome and the immune system'. In: *Nature* 474.7351 (2011), pp. 327–336.

[170] Jacoline Gerritsen et al. 'Intestinal microbiota in human health and disease: the impact of probiotics'. In: *Genes & Nutrition* 6.3 (2011), pp. 209–240.

[171] Jeremy K Nicholson et al. 'Host-gut microbiota metabolic interactions'. In: *Science* 336.6086 (2012), pp. 1262–1267.

[172] Seth T Walk et al. 'Alteration of the murine gut microbiota during infection with the parasitic helminth Heligmosomoides polygyrus'. In: *Inflammatory Bowel Diseases* 16.11 (2010), pp. 1841–1849.

[173] Scot E Dowd et al. 'Polymicrobial nature of chronic diabetic foot ulcer biofilm infections determined using bacterial tag encoded FLX amplicon pyrosequencing (bTEFAP)'. In: *PLoS ONE* 3.10 (2008), e3326.

[174] Randall D Wolcott et al. 'Evaluation of the bacterial diversity among and within individual venous leg ulcers using bacterial tag-encoded FLX and titanium amplicon pyrosequencing and metagenomic approaches'. In: *BMC Microbiology* 9.1 (2009), p. 226.

[175] Viktoria Gontcharova et al. 'Black box chimera check (B2C2): a windows-based software for batch depletion of chimeras from bacterial 16S rRNA gene datasets'. In: *The Open Microbiology Journal* 4 (2010), p. 47.

[176] QIAGEN. *QIAamp ®DNA Stool Handbook*. Apr. 2010. URL: `https://www.qiagen.com/gb/resources/resourcedetail?id=c8fe97e7-78cc-4275-bbac-72c9b7c3de38&lang=en`.

[177] Maria Grazia Paglia and Paolo Visca. 'An improved PCR-based method for detection and differentiation of Entamoeba histolytica and Entamoeba dispar in formalin-fixed stools'. In: *Acta Tropica* 92.3 (2004), pp. 273–277.

[178] Philippe Desjardins and Deborah Conklin. 'NanoDrop microvolume quantitation of nucleic acids'. In: *Journal of Visualized Experiments: JoVE* 45 (2010).

[179] Enzoklop (own work) via Wikimedia Commons - Wikepedia. Version 5 April 2014. URL: `http://commons.wikimedia.org/wiki/File:Polymerase_chain_reaction.svg` (visited on 27/05/2015).

[180] Henry A Erlich. 'Polymerase chain reaction'. In: *Journal of Clinical Immunology* 9.6 (1989), pp. 437–447.

[181] Debbie A Lewis et al. 'The human urinary microbiome; bacterial DNA in voided urine of asymptomatic adults'. In: *Frontiers in Cellular and Infection Microbiology* 3 (2013).

[182] Mostafa Ronaghi, Mathias Uhlén and Pål Nyrén. 'A sequencing method based on real-time pyrophosphate'. In: *Science* 281.5375 (1998), pp. 363–365.

[183] Marcus Droege and Brendon Hill. 'The Genome Sequencer FLX^{TM} System—Longer reads, more applications, straight forward bioinformatics and more complete data sets'. In: *Journal of Biotechnology* 136.1 (2008), pp. 3–10.

[184] Research and Testing Laboratory. *Data Analysis Methodology.* (Original Version used during PhD project dated 26/10/2012). July 2014. URL: `http://www.researchandtesting.com/docs/Data_Analysis_Methodology.pdf`.

[185] William H Kruskal and W Allen Wallis. 'Use of ranks in one-criterion variance analysis'. In: *Journal of the American Statistical Association* 47.260 (1952), pp. 583–621.

[186] Henry B Mann and Donald R Whitney. 'On a test of whether one of two random variables is stochastically larger than the other'. In: *The Annals of Mathematical Statistics* (1947), pp. 50–60.

[187] Milton Friedman. 'The use of ranks to avoid the assumption of normality implicit in the analysis of variance'. In: *Journal of the American Statistical Association* 32.200 (1937), pp. 675–701.

[188] Frank Wilcoxon. 'Individual comparisons by ranking methods'. In: *Biometrics Bulletin* (1945), pp. 80–83.

[189] Olive Jean Dunn. 'Multiple comparisons among means'. In: *Journal of the American Statistical Association* 56.293 (1961), pp. 52–64.

[190] Yosef Hochberg and Yoav Benjamini. 'More powerful procedures for multiple significance testing'. In: *Statistics in Medicine* 9.7 (1990), pp. 811–818.

[191] Yoav Benjamini and Yosef Hochberg. 'Controlling the false discovery rate: a practical and powerful approach to multiple testing'. In: *Journal of the Royal Statistical Society. Series B (Methodological)* (1995), pp. 289–300.

[192] John D Storey and Robert Tibshirani. 'Statistical significance for genomewide studies'. In: *Proceedings of the National Academy of Sciences* 100.16 (2003), pp. 9440–9445.

[193] Jan Hauke and Tomasz Kossowski. 'Comparison of values of Pearson's and Spearman's correlation coefficients on the same sets of data'. In: *Quaestiones Geographicae* 30.2 (2011), pp. 87–93.

[194] Derek J Crockford et al. 'Statistical heterospectroscopy, an approach to the integrated analysis of NMR and UPLC-MS data sets: application in metabonomic toxicology studies'. In: *Analytical Chemistry* 78.2 (2006), pp. 363–371.

[195] Jasmina Saric et al. 'Systems parasitology: effects of Fasciola hepatica on the neurochemical profile in the rat brain'. In: *Molecular Systems Biology* 6.1 (2010), p. 396.

[196] Joram M Posma et al. 'MetaboNetworks, an interactive Matlab-based toolbox for creating, customizing and exploring sub-networks from KEGG'. In: *Bioinformatics* (2013), btt612.

[197] Sabrina D Lamour et al. 'Metabolic, immune, and gut microbial signals mount a systems response to Leishmania major infection'. In: *Journal of Proteome Research* 14.1 (2014), pp. 318–329.

[198] Simon Festing and Robin Wilkinson. 'The ethics of animal research'. In: *EMBO Reports* 8.6 (2007), pp. 526–530.

[199] Pandora Pound et al. 'Where is the evidence that animal research benefits humans?' In: *BMJ* 328.7438 (2004), pp. 514–517.

[200] The Jackson Laboratory. *Strain Name: C57BL/6J Stock Number: 000664*. Version 6.10. 2014. URL: `http://jaxmice.jax.org/strain/000664.html` (visited on 08/07/2015).

[201] Asif T Chinwalla et al. 'Initial sequencing and comparative analysis of the mouse genome'. In: *Nature* 420.6915 (2002), pp. 520–562.

[202] Charles River Laboratories International Inc. *The C57BL/6 Mouse White Paper: The Role of the C57BL/6N Mouse in the Creation of Future Genetically Engineered Models*. 2013. URL: `http://www.criver.com/files/pdfs/rms/c57bl6/rm_rm_r_c57bl6_white_paper.aspx` (visited on 08/07/2015).

[203] The Jackson Laboratory. *Strain Name: BALB/cJ Stock Number: 000651*. Version 6.10. 2014. URL: `http://jaxmice.jax.org/strain/000651.html` (visited on 08/07/2015).

[204] Mary E Bollard et al. 'NMR-based metabonomic approaches for evaluating physiological influences on biofluid composition'. In: *NMR in Biomedicine* 18.3 (2005), pp. 143–162.

[205] Marie-France Champy et al. 'Genetic background determines metabolic phenotypes in the mouse'. In: *Mammalian Genome* 19.5 (2008), pp. 318–331.

[206] The Jackson Laboratory. *Physiological Data Summary - BALB/cJ (000651)*. 2007. URL: `http://jaxmice.jax.org/support/phenotyping/BALBcJdata000651.pdf` (visited on 08/07/2015).

[207] The Jackson Laboratory. *Physiological Data Summary – C57BL/6J (000664)*. 2013. URL: `http://jaxmice.jax.org/support/phenotyping/B6data000664.pdf` (visited on 08/07/2015).

[208] Charles River Laboratories International Inc. *BALB/C Mouse Hematology*. 2012. URL: http://www.criver.com/files/pdfs/rms/balbc/rm_rm_r_balb-c_mouse_clinical_pathology_data.aspx (visited on 08/07/2015).

[209] Charles River Laboratories International Inc. *C57BL/6 Mouse Hematology*. 2012. URL: http://www.criver.com/files/pdfs/rms/c57bl6/rm_rm_r_c57bl6_mouse_clinical_pathology_data.aspx (visited on 08/07/2015).

[210] Claire L Gavaghan et al. 'An NMR-based metabonomic approach to investigate the biochemical consequences of genetic strain differences: application to the C57BL10J and Alpk: ApfCD mouse'. In: *FEBS letters* 484.3 (2000), pp. 169–174.

[211] Robert Plumb et al. 'Metabonomic analysis of mouse urine by liquid-chromatography-time of flight mass spectrometry (LC-TOFMS): detection of strain, diurnal and gender differences'. In: *Analyst* 128.7 (2003), pp. 819–823.

[212] Jia V Li et al. 'Metabonomic investigations of age-and batch-related variations in female NMRI mice using proton nuclear magnetic resonance spectroscopy'. In: *Molecular BioSystems* 9.12 (2013), pp. 3155–3165.

[213] Andrew Craig et al. 'Systems toxicology: integrated genomic, proteomic and metabonomic analysis of methapyrilene induced hepatotoxicity in the rat'. In: *Journal of Proteome Research* 5.7 (2006), pp. 1586–1601.

[214] Alison Waldram et al. 'Top-down systems biology modeling of host metabotype- microbiome associations in obese rodents'. In: *Journal of Proteome Research* 8.5 (2009), pp. 2361–2375.

[215] Donovan H Parks and Robert G Beiko. 'Identifying biologically relevant differences between metagenomic communities'. In: *Bioinformatics* 26.6 (2010), pp. 715–721.

[216] James Robert White, Niranjan Nagarajan and Mihai Pop. 'Statistical methods for detecting differentially abundant features in clinical metagenomic samples'. In: *PLoS Computational Biology* 5.4 (2009), e1000352.

[217] K Janne et al. 'Hierarchical principal component analysis (PCA) and projection to latent structure (PLS) technique on spectroscopic data as a data pretreatment for calibration'. In: *Journal of Chemometrics* 15.4 (2001), pp. 203–213.

[218] Peter J Havel. 'Peripheral signals conveying metabolic information to the brain: short-term and long-term regulation of food intake and energy homeostasis'. In: *Experimental Biology and Medicine* 226.11 (2001), pp. 963–977.

[219] Trairak Pisitkun, Rose Johnstone and Mark A Knepper. 'Discovery of urinary biomarkers'. In: *Molecular & Cellular Proteomics* 5.10 (2006), pp. 1760–1771.

[220] HL Kornberg and HA Krebs. 'Synthesis of cell constituents from C2-units by a modified tricarboxylic acid cycle.' In: *Nature* 179.4568 (1957), p. 988.

[221] Steven H Zeisel. 'Choline: Critical Role During Fetal Development and Dietary Requirements in Adults'. In: *Annual Review of Nutrition* 26 (2006), pp. 229–250.

[222] AE Harper, RHr Miller and KP Block. 'Branched-chain amino acid metabolism'. In: *Annual Review of Nutrition* 4.1 (1984), pp. 409–454.

[223] Melissa K Friswell et al. 'Site and Strain-Specific Variation in Gut Microbiota Profiles and Metabolism in Experimental Mice'. In: *PLoS ONE* 5.1 (2010).

[224] Majbritt Ravn Hufeldt et al. 'Variation in the gut microbiota of laboratory mice is related to both genetic and environmental factors'. In: *Comparative Medicine* 60.5 (2010), p. 336.

[225] James H Campbell et al. 'Host genetic and environmental effects on mouse intestinal microbiota'. In: *The ISME Journal* 6.11 (2012), pp. 2033–2044.

[226] Melanie D Spencer et al. 'Association between composition of the human gastrointestinal microbiome and development of fatty liver with choline deficiency'. In: *Gastroenterology* 140.3 (2011), pp. 976–986.

[227] Joseph P Zackular et al. 'The gut microbiome modulates colon tumorigenesis'. In: *MBio* 4.6 (2013), e00692–13.

[228] Nadja Larsen et al. 'Gut microbiota in human adults with type 2 diabetes differs from non-diabetic adults'. In: *PLoS ONE* 5.2 (2010), e9085.

[229] Jan S Suchodolski et al. 'The Fecal Microbiome in Dogs with Acute Diarrhea and Idiopathic Inflammatory Bowel Disease'. In: *PLoS ONE* 7.12 (2012).

[230] RM Locksley et al. 'Murine cutaneous leishmaniasis: susceptibility correlates with differential expansion of helper T-cell subsets'. In: *Annales de l'Institut Pasteur/Immunologie*. Vol. 138. 5. Elsevier. 1987, pp. 744–749.

[231] Ingo B Autenrieth et al. 'Immune responses to Yersinia enterocolitica in susceptible BALB/c and resistant C57BL/6 mice: an essential role for gamma interferon.' In: *Infection and Immunity* 62.6 (1994), pp. 2590–2599.

[232] Xi Yang, Kent T HayGlass and Robert C Brunham. 'Genetically determined differences in IL-10 and IFN-gamma responses correlate with clearance of Chlamydia trachomatis mouse pneumonitis infection.' In: *The Journal of Immunology* 156.11 (1996), pp. 4338–4344.

[233] Tie Liu et al. 'Differences in interleukin-12 and-15 production by dendritic cells at the early stage of Listeria monocytogenes infection between BALB/c and C57 BL/6 mice'. In: *Cellular Immunology* 202.1 (2000), pp. 31–40.

[234] Kim E Schmidt et al. 'Induction of pro-inflammatory mediators in Plasmodium berghei infected BALB/c mice breaks blood-brain-barrier and leads to cerebral malaria in an IL-12 dependent manner'. In: *Microbes and Infection* 13.10 (2011), pp. 828–836.

[235] E Roggero et al. 'Differential susceptibility to acute Trypanosoma cruzi infection in BALB/c and C57BL/6 mice is not associated with a distinct parasite load but cytokine abnormalities'. In: *Clinical & Experimental Immunology* 128.3 (2002), pp. 421–428.

[236] Stephen John Hopkins. 'The pathophysiological role of cytokines'. In: *Legal Medicine* 5 (2003), S45–S57.

[237] Charles A Dinarello. 'Biologic basis for interleukin-1 in disease'. In: *Blood* 87.6 (1996), pp. 2095–2147.

[238] Luke A O'Neill and Catherine Greene. 'Signal transduction pathways activated by the IL-1 receptor family: ancient signaling machinery in mammals, insects, and plants.' In: *Journal of Leukocyte Biology* 63.6 (1998), pp. 650–657.

[239] Christoph Schumacher et al. 'High-and low-affinity binding of GRO alpha and neutrophil-activating peptide 2 to interleukin 8 receptors on human neutrophils'. In: *Proceedings of the National Academy of Sciences* 89.21 (1992), pp. 10542–10546.

[240] Michael D Sadick et al. 'Cure of murine leishmaniasis with anti-interleukin 4 monoclonal antibody. Evidence for a T cell-dependent, interferon gamma-independent mechanism.' In: *The Journal of Experimental Medicine* 171.1 (1990), pp. 115–127.

[241] René Chatelain, Kari Varkila and Robert L Coffman. 'IL-4 induces a Th2 response in Leishmania major-infected mice.' In: *The Journal of Immunology* 148.4 (1992), pp. 1182–1187.

[242] Manfred Kopf et al. 'IL-4-deficient Balb/c mice resist infection with Leishmania major.' In: *The Journal of Experimental Medicine* 184.3 (1996), pp. 1127–1136.

[243] Jasmina Saric et al. 'Integrated cytokine and metabolic analysis of pathological responses to parasite exposure in rodents'. In: *Journal of Proteome Research* 9.5 (2010), pp. 2255–2264.

[244] Benjamin Moreno et al. '31P NMR spectroscopy of Trypanosoma brucei, Trypanosoma cruzi, and Leishmania major Evidence for high levels of condensed inorganic phosphates'. In: *Journal of Biological Chemistry* 275.37 (2000), pp. 28356–28362.

[245] Reimmel K Adosraku et al. 'Proton NMR lipid profile of Leishmania donovani promastigotes'. In: *Molecular and Biochemical Parasitology* 62.2 (1993), pp. 251–262.

[246] Ruben t'Kindt et al. 'Towards an unbiased metabolic profiling of protozoan parasites: optimisation of a Leishmania sampling protocol for HILIC-orbitrap analysis'. In: *Analytical and Bioanalytical Chemistry* 398.5 (2010), pp. 2059–2069.

[247] Gisele AB Canuto et al. 'CE-ESI-MS metabolic fingerprinting of Leishmania resistance to antimony treatment'. In: *Electrophoresis* 33.12 (2012), pp. 1901–1910.

[248] R Da Silva and DL Sacks. 'Metacyclogenesis is a major determinant of Leishmania promastigote virulence and attenuation.' In: *Infection and Immunity* 55.11 (1987), pp. 2802–2806.

[249] QIAGEN. *DNeasy ®Blood & Tissue Handbook*. July 2006. URL: `https://www.qiagen.com/gb/resources/resourcedetail?id=6b09dfb8-6319-464d-996c-79e8c7045a50&lang=en` (visited on 29/07/2015).

[250] Applied BioSystems. *Fast SYBR® Green Master Mix Protocol*. July 2010. URL: `https://tools.lifetechnologies.com/content/sfs/manuals/cms_046776.pdf` (visited on 29/07/2015).

[251] Emilie de La Llave et al. 'A combined luciferase imaging and reverse transcription polymerase chain reaction assay for the study of Leishmania amastigote burden and correlated mouse tissue transcript fluctuations'. In: *Cellular Microbiology* 13.1 (2011), pp. 81–91.

[252] Manuela da Silva Solcà et al. 'Qualitative and quantitative polymerase chain reaction (PCR) for detection of Leishmania in spleen samples from naturally infected dogs'. In: *Veterinary Parasitology* 184.2 (2012), pp. 133–140.

[253] Jan Hellemans et al. 'qBase relative quantification framework and software for management and automated analysis of real-time quantitative PCR data'. In: *Genome Biology* 8.2 (2007), R19.

[254] Michael W Pfaffl. 'A new mathematical model for relative quantification in real-time RT–PCR'. In: *Nucleic Acids Research* 29.9 (2001), e45–e45.

[255] AAM Morris. 'Cerebral ketone body metabolism'. In: *Journal of Inherited Metabolic Disease* 28.2 (2005), pp. 109–121.

[256] Imen Rabhi et al. 'Transcriptomic Signature of Leishmania Infected Mice Macrophages: A Metabolic Point of View'. In: *PLoS Neglected Tropical Diseases* 6.8 (2012).

[257] Jia V Li et al. 'Metabonomic investigation of single and multiple strain Trypanosoma brucei brucei infections'. In: *The American Journal of Tropical Medicine and Hhygiene* 84.1 (2011), pp. 91–98.

[258] Jia V Li et al. 'Global metabolic responses of NMRI mice to an experimental Plasmodium berghei infection'. In: *Journal of Proteome Research* 7.9 (2008), pp. 3948–3956.

[259] Laetitia Dou et al. 'The uremic solutes p-cresol and indoxyl sulfate inhibit endothelial proliferation and wound repair'. In: *Kidney International* 65.2 (2004), pp. 442–451.

[260] Malcolm J McConville et al. 'Living in a phagolysosome; metabolism of Leishmania amastigotes'. In: *Trends in Parasitology* 23.8 (2007), pp. 368–375.

[261] Beatriz León, María López-Bravo and Carlos Ardavín. 'Monocyte-derived dendritic cells formed at the infection site control the induction of protective T helper 1 responses against Leishmania'. In: *Immunity* 26.4 (2007), pp. 519–531.

[262] Elena Voronov et al. 'IL-1-induced inflammation promotes development of leishmaniasis in susceptible BALB/c mice'. In: *International Immunology* 22.4 (2010), pp. 245–257.

[263] Jia V Li et al. 'Metabolic surgery profoundly influences gut microbial–host metabolic cross-talk'. In: *Gut* 60.9 (2011), pp. 1214–1223.

[264] Oksana Lukjancenko, David W Ussery and Trudy M Wassenaar. 'Comparative genomics of Bifidobacterium, Lactobacillus and related probiotic genera'. In: *Microbial Ecology* 63.3 (2012), pp. 651–673.

[265] Cyrille Hoarau et al. 'Supernatant of Bifidobacterium breve induces dendritic cell maturation, activation, and survival through a Toll-like receptor 2 pathway'. In: *Journal of Allergy and Clinical Immunology* 117.3 (2006), pp. 696–702.

[266] Seong Gyu Jeon et al. 'Probiotic Bifidobacterium breve induces IL-10-producing Tr1 cells in the colon'. In: *PLoS Pathogens* 8.5 (2012), e1002714.

[267] Wenjun Ouyang et al. 'Regulation and functions of the IL-10 family of cytokines in inflammation and disease'. In: *Annual Review of Immunology* 29 (2011), pp. 71–109.

[268] E Hijova and A Chmelarova. 'Short chain fatty acids and colonic health.' In: *Bratislavske lekarske listy - Bratislava Medical Journal* 108.8 (2006), pp. 354–358.

[269] Philip Felig et al. 'Alanine: key role in gluconeogenesis'. In: *Science* 167.3920 (1970), pp. 1003–1004.

[270] WJ Beil et al. 'Differences in the onset of the inflammatory response to cutaneous leishmaniasis in resistant and susceptible mice.' In: *Journal of Leukocyte Biology* 52.2 (1992), pp. 135–142.

[271] Aysha Rashid. 'Effects of picolinic acid on the macrophage metabolome in in vitro Leishmania infection'. BSc Thesis. Imperial College London, 2011.

[272] Peter M Orchard. 'Metabolic profiling of Leishmania parasites and their host cells'. BSc Thesis. Imperial College London, 2011.

[273] Stephanie Chrysanthou. 'In vitro metabolic profiling investigations on therapeutic supplements and virulence factors for Leishmania major'. MRes Thesis. Imperial College London, 2013.

[274] Jeremy M Sternberg et al. 'Evaluation of the Diagnostic Accuracy of Prototype Rapid Tests for Human African Trypanosomiasis'. In: *PLoS Neglected Tropical Diseases* 8.12 (2014), e3373.

[275] Sabrina D Lamour et al. 'Discovery of Infection Associated Metabolic Markers in Human African Trypanosomiasis'. In: *PLoS Neglected Tropical Diseases* 9.10 (2015), pp. 1–17.

[276] Lorna M MacLean et al. 'Focus-specific clinical profiles in human African Trypanosomiasis caused by Trypanosoma brucei rhodesiense'. In: *PLoS Neglected Tropical Diseases* 4.12 (2010), e906.

[277] Juan Carlos Oliveros. *Venny An interactive tool for comparing lists with Venn Diagrams. 2007–2015.* URL: `http://bioinfogp.cnb.csic.es/tools/venny/index.html`.

[278] Richard H Barton et al. 'The influence of EDTA and citrate anticoagulant addition to human plasma on information recovery from NMR-based metabolic profiling studies'. In: *Molecular BioSystems* 6.1 (2010), pp. 215–224.

[279] RW Wannemacher et al. 'The significance and mechanism of an increased serum phenylalanine-tyrosine ratio during infection.' In: *The American Journal of Clinical Nutrition* 29.9 (1976), pp. 997–1006.

[280] Elin Näsström et al. 'Salmonella Typhi and Salmonella Paratyphi A elaborate distinct systemic metabolite signatures during enteric fever'. In: *ELife* 3 (2014), e03100.

[281] Bert K Lopansri et al. 'Elevated plasma phenylalanine in severe malaria and implications for pathophysiology of neurological complications'. In: *Infection and Immunity* 74.6 (2006), pp. 3355–3359.

[282] Seymour Kaufman. 'An evaluation of the possible neurotoxicity of metabolites of phenylalanine'. In: *The Journal of Pediatrics* 114.5 (1989), pp. 895–900.

[283] Makoto Watanabe et al. 'Consequences of low plasma histidine in chronic kidney disease patients: associations with inflammation, oxidative stress, and mortality'. In: *The American Journal of Clinical Nutrition* 87.6 (2008), pp. 1860–1866.

[284] Yu-Cun Niu et al. 'Histidine and arginine are associated with inflammation and oxidative stress in obese women'. In: *British Journal of Nutrition* 108.01 (2012), pp. 57–61.

[285] CO Enwonwu et al. 'Increased plasma levels of histidine and histamine in falciparum malaria: relevance to severity of infection'. In: *Journal of Neural Transmission* 107.11 (2000), pp. 1273–1287.

[286] Jimmy D Bell et al. 'Assignment of resonances for 'acute-phase'glycoproteins in high resolution proton NMR spectra of human blood plasma'. In: *FEBS Letters* 215.2 (1987), pp. 311–315.

[287] Krista Fischer et al. 'Biomarker profiling by nuclear magnetic resonance spectroscopy for the prediction of all-cause mortality: an observational study of 17,345 persons'. In: *PLoS Medicine* 11.2 (2014), e1001606.

[288] Ismaila A Umar et al. 'The Role of Vitamin C Administration in Alleviation of Organ Damage in Rats Infected with Trypanosoma brucei.' In: *Journal of Clinical Biochemistry and Nutrition* 28.1 (2000), pp. 1–7.

[289] Philip C Calder. 'Polyunsaturated fatty acids, inflammation, and immunity'. In: *Lipids* 36.9 (2001), pp. 1007–1024.

[290] Steven J Bensinger and Peter Tontonoz. 'Integration of metabolism and inflammation by lipid-activated nuclear receptors'. In: *Nature* 454.7203 (2008), pp. 470–477.

[291] Benjamin J Visser et al. 'Serum lipids and lipoproteins in malaria-a systematic review and meta-analysis'. In: *Malaria Journal* 12.1 (2013), p. 442.

[292] SN Khosla, N Goyle and RK Seth. 'Lipid profile in enteric fever.' In: *The Journal of the Association of Physicians of India* 39.3 (1991), pp. 260–262.

[293] Concepclon Alvarez and A Ramos. 'Lipids, lipoproteins, and apoproteins in serum during infection.' In: *Clinical Chemistry* 32.1 (1986), pp. 142–145.

[294] Neci Matos Soares et al. 'Plasma lipoproteins in visceral leishmaniasis and their effect on Leishmania-infected macrophages'. In: *Parasite immunology* 32.4 (2010), pp. 259–266.

[295] G Huet et al. 'Serum lipid and lipoprotein abnormalities in human African trypanosomiasis'. In: *Transactions of the Royal Society of Tropical Medicine and Hygiene* 84.6 (1990), pp. 792–794.

[296] Samuel Black and Veerle Vandeweerd. 'Serum lipoproteins are required for multiplication of Trypanosoma brucei brucei under axenic culture conditions'. In: *Molecular and Biochemical Parasitology* 37.1 (1989), pp. 65–72.

[297] Isabelle Coppens, Thierry Levade and Pierre J Courtoy. 'Host plasma low density lipoprotein particles as an essential source of lipids for the bloodstream forms of Trypanosoma brucei'. In: *Journal of Biological Chemistry* 270.11 (1995), pp. 5736–5741.

[298] Gregory S Richmond et al. 'Lipidomic analysis of bloodstream and procyclic form Trypanosoma brucei'. In: *Parasitology* 137.09 (2010), pp. 1357–1392.

[299] Honor Dixon and J Williamson. 'The lipid composition of blood and culture forms of Trypanosoma lewisi and Trypanosoma rhodesiense compared with that of their environment'. In: *Comparative Biochemistry and Physiology* 33.1 (1970), pp. 111–128.

[300] Terry K Smith and Peter Bütikofer. 'Lipid metabolism in Trypanosoma brucei'. In: *Molecular and Biochemical Parasitology* 172.2 (2010), pp. 66–79.

[301] Michael Maceyka and Sarah Spiegel. 'Sphingolipid metabolites in inflammatory disease'. In: *Nature* 510.7503 (2014), pp. 58–67.

[302] Carol A Rouzer and Anthony Cerami. 'Hypertriglyceridemia associated with Trypanosoma brucei brucei infection in rabbits: role of defective triglyceride removal'. In: *Molecular and Biochemical Parasitology* 2.1 (1980), pp. 31–38.

[303] Yoshio Nakamura. 'Alterations of serum lipid, lipoprotein and inflammatory cytokine profiles of rabbits infected with Trypanosoma brucei brucei'. In: *Veterinary parasitology* 80.2 (1998), pp. 117–125.

[304] Daniela Fenili et al. 'Properties of scyllo–inositol as a therapeutic treatment of AD-like pathology'. In: *Journal of Molecular Medicine* 85.6 (2007), pp. 603–611.

[305] Robert H Michell. 'Inositol derivatives: evolution and functions'. In: *Nature Reviews Molecular Cell Biology* 9.2 (2008), pp. 151–161.

[306] Maurice B Burg and Joan D Ferraris. 'Intracellular organic osmolytes: function and regulation'. In: *Journal of Biological Chemistry* 283.12 (2008), pp. 7309–7313.

[307] Walter E Gall et al. 'α-Hydroxybutyrate Is an Early Biomarker of Insulin Resistance and Glucose Intolerance in a Nondiabetic Population'. In: *PLoS ONE* 5.5 (2010).

[308] Alessandro Federico et al. 'Chronic inflammation and oxidative stress in human carcinogenesis'. In: *International Journal of Cancer* 121.11 (2007), pp. 2381–2386.

[309] B Enanga et al. 'Sleeping sickness and the brain'. In: *Cellular and Molecular Life Sciences CMLS* 59.5 (2002), pp. 845–858.

[310] David S Wishart. 'Metabolomics: applications to food science and nutrition research'. In: *Trends in Food Science & Technology* 19.9 (2008), pp. 482–493.

[311] Aifric O'Sullivan, Michael J Gibney and Lorraine Brennan. 'Dietary intake patterns are reflected in metabolomic profiles: potential role in dietary assessment studies'. In: *The American Journal of Clinical Nutrition* 93.2 (2011), pp. 314–321.

[312] José F Siqueira Jr, Ashraf F Fouad and Isabela N Rôças. 'Pyrosequencing as a tool for better understanding of human microbiomes'. In: *Journal of Oral Microbiology* 4 (2012).

[313] Claudia Lupp et al. 'Host-mediated inflammation disrupts the intestinal microbiota and promotes the overgrowth of Enterobacteriaceae'. In: *Cell Host & Microbe* 2.2 (2007), pp. 119–129.

[314] Elisabeth Drucker and Kurt Krapfenbauer. 'Pitfalls and limitations in translation from biomarker discovery to clinical utility in predictive and personalised medicine'. In: *The EPMA Journal* 4.1 (2013), pp. 1–10.

Appendix A

Appendix

List of Appendix Figures . . . 219
List of Appendix Tables . . . 220
Appendix Figures and Tables . . . 221
Publications . . . 245
Permissions for Third Party Copyright Works . . . 309

List of Appendix Figures

A.1 Faecal Clostridia and Gammaproteobacteria Abundances Demonstrate Outliers within Dataset . . . 227
A.2 Differential Faecal Family OTU Data in Mouse Strains Using Bonferroni Multiple Test Correction . . . 230
A.3 Strain-related Variation Dominates Infection-associated Changes . . . 231
A.4 Cage-related Variation Observed in Microbial Data . . . 236
A.5 Correlation Networks of Host Metabolic, Cytokine and Microbial Markers in CL Models Prior to Infection . . . 237
A.6 No Dominant Metabolic Variation in Plasma ^{1}H NMR Phenotypes Associated with HAT Diagnostic Stage . . . 239
A.7 Correlations between Plasma ^{1}H NMR and UPLC-MS Markers after FDR Correction 240
A.8 Correlations between Plasma Cytokines and UPLC-MS Markers prior to FDR Correction . . . 243
A.9 Metabolic Changes in CSF Associated with Impaired Gait . . . 244

List of Appendix Tables

A.1 List of ^{1}H NMR Metabolite Assignments Measured in Mouse Biofluids 221
A.2 List of ^{1}H NMR Metabolite Assignments Measured in Mouse Tissues 223
A.3 List of Faecal Bacteria Detected in BALB/c and C57BL/6 Mice 228
A.4 Physiological and Pathological Measures of CL Models 231
A.5 CL-Model Discriminatory Metabolites in Biofluids per Time-point 232
A.6 CL-Model Discriminatory Metabolites per Tissue 234
A.7 List of ^{1}H NMR Metabolite Assignments Measured in Human Biofluids 238
A.8 Assignment of HAT Discriminatory Lipid Species Measured in Plasma via UPLC-MS 241
A.9 Accuracy of the Collection of CSF NMR Discriminatory Markers for HAT 243
A.10 Summary of Permission for Third Party Copyright Works 309

Appendix Figures and Tables

Table A.1: List of ^{1}H NMR Metabolite Assignments Measured in Mouse Biofluids

Chemical shift (ppm)	Peak Multiplicity	Molecule	Biofluid	Method of Identification
0.68	m	cholesterol	p	database, lit, STOCSY
0.70	m	bile acids	f	database, lit, STOCSY
0.84	m	lipid CH_3	f, p	database, lit, STOCSY
0.90	t	3-methyl-2-oxovalerate*	u	chenomx (partial fit), STOCSY
0.90	t	butyrate	f	chenomx, database, lit, STOCSY
0.92	t	isoleucine	p	chenomx, database, lit, STOCSY
0.93	t	2-oxovalerate	u	chenomx, lit, STOCSY
0.94	d	2-oxoisocaproate	u	chenomx, database, lit, STOCSY
0.95	d	leucine	p	chenomx, database, lit, STOCSY
0.96	d	leucine	p	chenomx, database, lit, STOCSY
0.98	d	valine	p	chenomx, database, lit, STOCSY
1.00	d	isoleucine	p	chenomx, database, lit, STOCSY
1.04	d	valine	p	chenomx, database, lit, STOCSY
1.06	t	propionate	f	chenomx, database, lit, STOCSY
1.07	d	UK 1.07(d)	p	–
1.11	d	3-methyl-2-oxovalerate*	u	chenomx (partial fit), STOCSY
1.13	d	2-oxoisovalerate*	u	lit (partial fit), HMDB
1.19	d	3-hydroxybutyrate	p	chenomx, database, lit, STOCSY
1.26	s	3-hydroxyisovalerate*	f	chenomx, lit
1.26	m	lipid CH_2	p	database, lit, STOCSY
1.30	m	lipid CH_2	f	database, lit, STOCSY
1.33	d	lactate	f, p, u	chenomx, database, lit, STOCSY
1.43	m	isoleucine	p	chenomx, database, lit, STOCSY
1.43	d	UK 1.43(d)	p	–
1.44	m	lysine	p	chenomx, database, lit, STOCSY
1.46	m	3-methyl-2-oxovalerate*	u	chenomx (partial fit), STOCSY
1.48	d	alanine	f, p	chenomx, database, lit, STOCSY
1.50	m	lysine	p	chenomx, database, lit, STOCSY
1.56	m	butyrate	f	chenomx, database, lit, STOCSY
1.57	m	lipid CH_2CH_2CO	p	database, lit, STOCSY
1.62	m	2-oxovalerate	u	chenomx, lit, STOCSY
1.65	m	5-aminovalerate	f	database, lit, STOCSY
1.69	m	3-methyl-2-oxovalerate*	u	chenomx (partial fit), STOCSY
1.71	m	leucine	p	chenomx, database, lit, STOCSY
1.72	m	lysine	p	chenomx, database, lit, STOCSY
1.83	m	2-oxoadipate	u	database, lit, STOCSY
1.88	m	lysine	p	chenomx, database, lit, STOCSY
1.88	dd	UK 1.88(dd)	u	–
1.92	s	acetate	f, p, u	chenomx, database, lit
2.00	m	lipid $CH_2C{=}C$	p	database, lit, STOCSY
2.04	m	*N*-acetyl glycoproteins	u	database, lit
2.05	m	glutamate	f, p	chenomx, database, lit, STOCSY
2.06	m	*N*-acetyl glycoproteins	f, u	database, lit
2.11	m	glutamate	f, p	chenomx, database, lit, STOCSY
2.13	m	glutamine	p	chenomx, database, lit, STOCSY
2.16	t	butyrate	f	chenomx, database, lit, STOCSY
2.18	q	propionate	f	chenomx, database, lit, STOCSY
2.22	d	2-oxoadipate	u	database, lit, STOCSY
2.22	s	acetoacetate	p	chenomx (partial fit), database, lit
2.24	t	5-aminovalerate	f	database, lit, STOCSY
2.24	m	lipid CH_2CO	p	chenomx, database, lit, STOCSY
2.32	m	3-hydroxybutyrate	p	chenomx, database, lit, STOCSY
2.35	m	glutamate	f, p	chenomx, database, lit, STOCSY
2.36	s	pyuvate	p	chenomx, database, lit
2.38	t	ureidopropanoate	u	lit, STOCSY
2.41	m	3-hydroxybutyrate	p	chenomx, database, lit, STOCSY
2.41	s	succinate	f, p, u	chenomx, database, lit
2.43	m	glutamine	p	chenomx, database, lit, STOCSY
2.45	t	2-oxoglutarate	u	chenomx, database, lit, STOCSY
2.48	t	3-hydroxyphenyl propioniate	f	lit, STOCSY
2.51	m	UK 2.51(d)	u	STOCSY**
2.52	d	citrate	p, u	chenomx, database, lit, STOCSY
2.61	s	methylamine	u	chenomx, database, lit
2.62	d	2-oxoisocaproate	u	chenomx, database, lit, STOCSY
2.68	d	citrate	p, u	chenomx, database, lit, STOCSY
2.70	t	2-oxovalerate	u	chenomx, lit

Continued on next page

Table A.1: (continued)

Chemical shift (ppm)	Peak Multiplicity	Molecule	Biofluid	Method of Identification
2.72	s	dimethylamine	f	chenomx, database, lit
2.74	m	lipid C=CCH_2C=C	p	database, lit, STOCSY
2.76	m	glutamate	f, p	chenomx, database, lit
2.78	t	2-oxoadipate	u	database, lit, STOCSY
2.84	t	3-hydroxyphenyl propioniate	f	lit, STOCSY
2.88	s	trimethylamine	u	chenomx, database, lit
2.91	t	UK 2.91(t)	u	STOCSY**
3.00	t	2-oxoglutarate	u	chenomx, database, lit, STOCSY
3.02	t	5-aminovalerate	f	database, lit, STOCSY
3.03	t	lysine	p	chenomx, database, lit, STOCSY
3.04	s	creatine	p, u	chenomx, database, lit, STOCSY
3.05	s	creatinine	u	chenomx, database, lit, STOCSY
3.12	d	cis-aconitate	u	chenomx, lit
3.16	s	UK 3.16(s)	u	–
3.16	t?	UK 3.16(t?)	f	STOCSY***
3.17	s	UK 3.17(s)	u	–
3.20	s	choline	p	chenomx, database, lit, STOCSY
3.21	a	phosphocholine	p	chenomx, database, lit, STOCSY
3.22	s	glycerophosphocholine	p	chenomx, database, lit, STOCSY
3.24	dd	β-glucose	p	chenomx, database, lit, STOCSY
3.27	t	taurine	f, u	chenomx, database, lit, STOCSY
3.27	s	trimethylamine-*N*-oxide	u	chenomx, database, lit, STOCSY
3.30	t	ureidopropanoate	u	lit, STOCSY
3.36	s	scyllo-inositol	p, u	chenomx, database, lit
3.40	t	β-glucose	p	chenomx, database, lit, STOCSY
3.42	t	α-glucose	p	chenomx, database, lit, STOCSY
3.43	t	taurine	f, u	chenomx, database, lit, STOCSY
3.45	s	3-hydroxyphenylacetate	f	chenomx, HMDB, STOCSY
3.46	dd	β-glucose	p	chenomx, database, lit, STOCSY
3.49	t	β-glucose	p	chenomx, database, lit, STOCSY
3.51	m	choline	p	chenomx, database, lit, STOCSY
3.55	s	glycine	u	chenomx, database, lit
3.56	dd	glycerol	p	chenomx, database, lit, STOCSY
3.56	s	glycine	p	chenomx, database, lit
3.57	dd	α-glucose	p	chenomx, database, lit, STOCSY
3.60	m	phosphocholine	p	chenomx, database, lit, STOCSY
3.61	m	glycerophosphocholine	p	chenomx, database, lit, STOCSY
3.62	d	valine	p	chenomx, database, lit
3.64	dd	glycerol	p	chenomx, database, lit, STOCSY
3.65	dd	glycerol	p	chenomx, database, lit, STOCSY
3.65	m	UK 3.65(m)	f	–
3.67	m	glycerophosphocholine	p	chenomx, database, lit, STOCSY
3.67	d	isoleucine	p	chenomx, database, lit
3.68	s	phenylacetylglycine	u	chenomx, database, lit, STOCSY
3.71	t	α-glucose	p	chenomx, database, lit, STOCSY
3.72	dd	β-glucose	p	chenomx, database, lit, STOCSY
3.74	t	lysine	p	chenomx, database, lit, STOCSY
3.75	t	glutamine	p	chenomx, database, lit
3.76	m	α-glucose	p	chenomx, database, lit, STOCSY
3.76	d	phenylacetylglycine	u	chenomx, database, lit, STOCSY
3.77	m	glutamate	f	chenomx, database, lit, STOCSY
3.78	q	alanine	f, p	chenomx, database, lit, STOCSY
3.79	m	glycerol	p	chenomx, database, lit, STOCSY
3.80	s	guanidoacetate	u	chenomx, database, lit
3.82	dd	α-glucose	p	chenomx, database, lit, STOCSY
3.83	m	α-glucose	p	chenomx, database, lit, STOCSY
3.90	dd	β-glucose	p	chenomx, database, lit, STOCSY
3.93	s	creatine	p, u	chenomx, database, lit, STOCSY
3.96	s	glycolate	u	chenomx, lit
3.96	t	UK 3.96(t)	f	STOCSY***
3.97	d	hippurate	u	chenomx, database, lit, STOCSY
4.05	s	creatinine	u	chenomx, database, lit, STOCSY
4.06	m	choline	p	chenomx, database, lit, STOCSY
4.12	q	lactate	f, p, u	chenomx, database, lit, STOCSY
4.15	m	3-hydroxybutyrate	p	chenomx, database, lit
4.17	m	phosphocholine	p	chenomx, database, lit
4.33	m	glycerophosphocholine	p	chenomx, database, lit, STOCSY
4.48	s	1-methylnicotinamide	u	chenomx, HMDB, STOCSY
4.64	dd	β-glucose	p	chenomx, database, lit, STOCSY
5.23	d	α-glucose	p	chenomx, database, lit, STOCSY

Continued on next page

Table A.1: (continued)

Chemical shift (ppm)	Peak Multiplicity	Molecule	Biofluid	Method of Identification
5.28	m	lipid CH=CH unsaturated	p	database, lit, STOCSY
5.39	s	allantoin	p, u	chenomx, database, lit
6.53	s	fumarate	u	chenomx, database, lit
6.76	m	3-hydroxyphenylacetate	f	chenomx, HMDB, STOCSY
6.81	m	3-hydroxyphenylacetate	f	chenomx, HMDB, STOCSY
6.87	d	3-hydroxyphenylacetate	f	chenomx, HMDB, STOCSY
6.91	d?	o-cresol*	u	chenomx (partial fit), STOCSY
6.91	d	tyramine*	u	chenomx (partial fit), STOCSY
7.14	t?	o-cresol*	u	chenomx (partial fit), STOCSY
7.16	m	UK 7.16(m)	u	STOCSY**
7.20	t	indoxylsulphate	u	chenomx, database, lit, STOCSY
7.20	d	tyramine*	u	chenomx (partial fit), STOCSY
7.21	m	UK 7.21(m)	u	STOCSY**
7.25	t	3-hydroxyphenylacetate	f	chenomx, HMDB, STOCSY
7.27	t	indoxylsulphate	u	chenomx, database, lit, STOCSY
7.31	m	phenylacetate*	u	chenomx (partial fit), HMDB, STOCSY
7.36	s	indoxylsulphate	u	chenomx, database, lit, STOCSY
7.36	t	phenylacetylglycine	u	chenomx, database, lit, STOCSY
7.38	m	phenylacetate*	u	chenomx (partial fit), HMDB, STOCSY
7.39	m	UK 7.39(m)	u	STOCSY**
7.43	t	phenylacetylglycine	u	chenomx, database, lit, STOCSY
7.51	d	indoxylsulphate	u	chenomx, database, lit, STOCSY
7.56	t	hippurate	u	chenomx, database, lit, STOCSY
7.64	t	hippurate	u	chenomx, database, lit, STOCSY
7.68	s	UK 7.68(s)	f, u	–
7.71	d	indoxylsulphate	u	chenomx, database, lit, STOCSY
7.84	d	hippurate	u	chenomx, database, lit, STOCSY
7.87	s	UK 7.87(s)	f	–
7.97	m	phenylacetylglycine	u	chenomx, database, lit, STOCSY
8.18	t	1-methylnicotinamide	u	chenomx, HMDB, STOCSY
8.46	s	formate	f, p	chenomx, database, lit
8.90	d	1-methylnicotinamide	u	chenomx, HMDB, STOCSY
8.96	d	1-methylnicotinamide	u	chenomx, HMDB, STOCSY
9.26	s	1-methylnicotinamide	u	chenomx, HMDB, STOCSY

Key: *, tentative assignments. **, strong positive correlation between 3.16(t?) and 3.96(t); ***, strong positive correlation between 2.51(d), 2.91(t), 7.16(m), 7.21(m), and 7.39(t); (d), doublet; (dd), doublet of doublets; (m), multiplet; (s), singlet; (t), triplet; f, faeces; lit, literature (see Methods Section 2.5.2 for list of references); HMDB, the Human Metabolome Database[142–144]; p, plasma; STOCSY, Statistical total correlation spectroscopy; u, urine; UK, unknown metabolites

Table A.2: List of ^{1}H NMR Metabolite Assignments Measured in Mouse Tissues

Chemical shift (ppm)	Peak Multiplicity	Molecule	Tissue	Method of Identification
0.68	m	bile acids	il, li	database, lit, STOCSY
0.86	m	lipid CH_3	b, il, k, li, lu, ly, t	database, lit, STOCSY
0.94	t	isoleucine	b, h, il, k, li, lu, ly, s, t	chenomx, database, lit, STOCSY
0.95	d	leucine	b, h, il, k, li, lu, ly, s, t	chenomx, database, lit, STOCSY
0.96	d	leucine	b, h, il, k, li, lu, ly, s, t	chenomx, database, lit, STOCSY
1.00	d	valine	b, h, il, k, li, lu, ly, s, t	chenomx, database, lit, STOCSY
1.01	d	isoleucine	b, h, il, k, li, lu, ly, s, t	chenomx, database, lit, STOCSY
1.05	d	valine	b, h, il, k, li, lu, ly, s, t	chenomx, database, lit, STOCSY
1.19	d	3-hydroxybutyrate	b, h, k, li, lu, ly, s, t	chenomx, database, lit, STOCSY
1.26	s	3-hydroxyisovalerate	h, k, lu, t	chenomx, lit
1.26	m	lipid CH_2	b, il, k, li, lu, ly, t	database, lit, STOCSY
1.33	d	lactate	b, h, il, k, li, lu, ly, s, t	chenomx, database, lit, STOCSY
1.43	d	UK 1.43(d)	b, h, k	–
1.46	m	isoleucine	b, h, il, li, lu, ly, s, t	chenomx, database, lit, STOCSY
1.49	d	alanine	b, h, il, k, li, lu, ly, s, t	chenomx, database, lit, STOCSY
1.58	m	lipid CH_2CH_2CO	k, ly	database, lit, STOCSY
1.71	m	leucine	b, h, il, k, li, lu, ly, s, t	chenomx, database, lit, STOCSY
1.72	m	lysine	ly	chenomx, database, lit, STOCSY

Continued on next page

Table A.2: (continued)

Chemical shift (ppm)	Peak Multiplicity	Molecule	Tissue	Method of Identification
1.91	m	4-aminobutyrate	b	chenomx, lit, STOCSY
1.91	m	lysine	ly	chenomx, database, lit, STOCSY
1.92	s	acetate	b, h, il, k, li, lu, ly, s, t	chenomx, database, lit
1.98	m	isoleucine	b, h, il, li, lu, ly, s, t	chenomx, database, lit, STOCSY
2.00	m	lipid $CH_2C{=}C$	k, ly	database, lit, STOCSY
2.03	s	*N*-acetylaspartate	b, t	chenomx, database, lit, STOCSY
2.07	m	glutamate	b, h, k, li, lu, ly, s, t	chenomx, database, lit, STOCSY
2.14	m	glutamine	b, h, k, li, lu, s, t	chenomx, database, lit, STOCSY
2.16	m	glutathione	b, h, il, li, s, t	chenomx, lit, STOCSY
2.26	m	lipid CH_2CO	ly	database, lit, STOCSY
2.28	m	valine	b, h, il, k, li, lu, ly, s, t	chenomx, database, lit
2.30	t	4-aminobutyrate	b	chenomx, lit, STOCSY
2.32	m	3-hydroxybutyrate	b, h, k, li, lu, ly, s, t	chenomx, database, lit, STOCSY
2.35	m	glutamate	b, h, k, li, lu, ly, s, t	chenomx, database, lit, STOCSY
2.39	m	3-hydroxybutyrate	b, h, k, li, lu, ly, s, t	chenomx, database, lit, STOCSY
2.41	s	succinate	b, h, il, k, li, lu, s, t	chenomx, database, lit
2.46	m	glutamine	b, h, k, li, lu, s, t	chenomx, database, lit, STOCSY
2.51	dd	*N*-acetylaspartate	b, t	chenomx, database, lit, STOCSY
2.52	d	citrate	il, k, lu	chenomx, database, lit, STOCSY
2.52	m	glutathione	b, h, il, li, s, t	chenomx, lit, STOCSY
2.56	m	glutathione	b, h, il, li, s, t	chenomx, lit, STOCSY
2.68	d	citrate	il, k, lu	chenomx, database, lit, STOCSY
2.69	dd	aspartate	b, h, k, li, lu, ly, s, t	chenomx, lit, STOCSY
2.70	dd	*N*-acetylaspartate	b, t	chenomx, database, lit, STOCSY
2.72	s	dimethylamine	il, k, li	chenomx, databse, lit
2.74	s	sarcosine	lu	chenomx**, HMDB, lit, STOCSY
2.77	m	lipid $C{=}CCH_2C{=}C$	ly	database, lit, STOCSY
2.82	dd	aspartate	b, h, k, li, lu, ly, s, t	chenomx, database, lit, STOCSY
2.94	dd	glutathione	b, h, il, li, s, t	chenomx, lit, STOCSY
2.98	dd	glutathione	b, h, il, li, s, t	chenomx, lit, STOCSY
3.01	t	4-aminobutyrate	b	chenomx, lit, STOCSY
3.03	t	lysine	ly	chenomx, database, lit, STOCSY
3.04	s	creatine	b, h, il, k, li, lu, ly, s	chenomx, database, lit, STOCSY
3.11	s	malonate	lu	chenomx, HMDB, lit
3.15	t	ethanolamine	b, il, k, lu, ly, s, t	chenomx, database, lit, STOCSY
3.21	s	choline	b, h, il, k, li, lu, ly, s, t	chenomx, database, lit, STOCSY
3.22	s	phosphocholine	b, h, il, k, li, lu, ly, s, t	chenomx, database, lit, STOCSY
3.23	s	glycerophosphocholine	b, h, il, k, li, lu, ly, s, t	chenomx, database, lit, STOCSY
3.23	t	phosphoethanolamine	ly, s, t	chenomx, database, lit, STOCSY
3.26	dd	β-glucose	li	chenomx, database, lit, STOCSY
3.27	s	betaine	il, k, lu, ly, s, t	chenomx, database, lit, STOCSY
3.28	t	taurine	b, h, il, k, lu, ly, s	chenomx, database, lit, STOCSY
3.30	t	myo-inositol	b, k, s	chenomx, database, lit, STOCSY
3.36	s	scyllo-inositol	b, k, il, lu, ly	chenomx, database, lit
3.39	m	mannose	li	chenomx, HMDB, STOCSY
3.42	t	taurine	b, h, il, k, lu, ly, s	chenomx, database, lit, STOCSY
3.43	t	α-glucose	li	chenomx, database, lit, STOCSY
3.45	m	glycogen	li	HMDB, lit, STOCSY
3.46	dd	β-glucose	li	chenomx, database, lit, STOCSY
3.48	d	UK 3.48(d)	b, il, k	–
3.50	t	β-glucose	li	chenomx, database, lit, STOCSY
3.52	m	choline	b, h, il, k, li, lu, ly, s, t	chenomx, database, lit, STOCSY
3.54	dd	myo-inositol	b, k, s	chenomx, database, lit, STOCSY
3.55	dd	α-glucose	li	chenomx, database, lit, STOCSY
3.56	dd	glycerol	b, h, il, ly, s	chenomx, database, lit, STOCSY
3.57	s	glycine	b, h, k, li, lu, ly, s	chenomx, database, lit
3.57	t	mannose	li	chenomx, HMDB, STOCSY
3.60	m	phosphocholine	b, h, il, k, li, lu, ly, s, t	chenomx, database, lit, STOCSY
3.62	d	valine	b, h, il, k, li, lu, ly, s, t	chenomx, database, lit, STOCSY
3.63	t	myo-inositol	b, k, s	chenomx, database, lit, STOCSY
3.64	dd	glycerol	b, h, il, ly, s	chenomx, database, lit, STOCSY
3.64	m	mannose	li	chenomx, HMDB, STOCSY
3.65	dd	glycerol	b, h, il, ly, s	chenomx, database, lit, STOCSY
3.65	m	glycogen	li	HMDB, lit, STOCSY
3.68	m	glycerophosphocholine	b, h, il, k, li, lu, ly, s, t	chenomx, database, lit, STOCSY
3.72	t	α-glucose	li	chenomx, database, lit, STOCSY
3.72	t	leucine	b, h, il, k, li, lu, ly, s, t	chenomx, database, lit
3.73	m	mannose	li	chenomx, HMDB, STOCSY
3.74	dd	β-glucose	li	chenomx, database, lit, STOCSY
3.74	t	lysine	ly	chenomx, database, lit, STOCSY

Continued on next page

Table A.2: (continued)

Chemical shift (ppm)	Peak Multiplicity	Molecule	Tissue	Method of Identification
3.76	m	glutamate	b, h, k, li, lu, ly, s, t	chenomx, database, lit, STOCSY
3.77	m	glutamine	b, h, k, li, lu, s, t	chenomx, database, lit, STOCSY
3.77	m	UDP-glucose	li, s	chenomx, HMDB, STOCSY
3.78	q	alanine	b, h, il, k, li, lu, ly, s, t	chenomx, database, lit, STOCSY
3.78	m	α-glucose	li	chenomx, database, lit, STOCSY
3.78	m	glutathione	b, h, il, li, s, t	chenomx, lit, STOCSY
3.79	m	glycerol	b, h, il, ly, s	chenomx, database, lit
3.82	dd	α-glucose	li	chenomx, database, lit, STOCSY
3.83	t	ethanolamine	b, il, k, lu, ly, s, t	chenomx, database, lit, STOCSY
3.84	m	α-glucose	li	chenomx, database, lit, STOCSY
3.84	dd	inosine	b, h, il, k, li, ly, s	chenomx, database, lit, STOCSY
3.85	m	glycogen	li	HMDB, lit, STOCSY
3.85	m	mannose	li	chenomx, HMDB, STOCSY
3.89	m	mannose	li	chenomx, HMDB, STOCSY
3.90	dd	β-glucose	li	chenomx, database, lit, STOCSY
3.90	m	UDP-glucose	li, s	chenomx, HMDB, STOCSY
3.91	dd	aspartate	b, h, il, lu, ly, ly, s, t	chenomx, database, lit, STOCSY
3.91	s	betaine	il, k, lu, ly, s, t	chenomx, database, lit, STOCSY
3.94	s	creatine	b, h, il, k, li, lu, ly, s	chenomx, database, lit, STOCSY
3.95	dd	mannose	li	chenomx, HMDB, STOCSY
3.97	m	glycogen	li	HMDB, lit, STOCSY
3.98	m	mannose	li	chenomx, HMDB, STOCSY
3.98	q	phosphoethanolamine	ly, s, t	chenomx, database, lit, STOCSY
3.99	q	phenylalanine	il, li, ly	chenomx, database, lit, STOCSY
4.02	m	AMP*	b, h, il, k, li, lu, s, t	chenomx, HMDB, STOCSY
4.06	m	choline	b, h, il, k, li, lu, ly, s, t	chenomx, database, lit
4.08	t	myo-inositol	b, s	chenomx, database, lit, STOCSY
4.12	q	lactate	b, h, il, k, li, lu, ly, s, t	chenomx, database, lit, STOCSY
4.15	m	uridine	k, li	chenomx, database, lit, STOCSY
4.17	m	phosphocholine	b, h, il, k, li, lu, ly, s, t	chenomx, database, lit, STOCSY
4.22	m	ADP*	b, h, lu, s, t	chenomx, lit
4.22	m	UDP-glucose	li, s	chenomx, HMDB, STOCSY
4.24	m	ATP*	b, h, il, li	chenomx, lit
4.24	t	uridine	k, li	chenomx, database, lit, STOCSY
4.28	q	inosine	b, h, il, k, li, ly, s	chenomx, database, lit
4.30	q	adenosine	b	chenomx, database, lit
4.33	m	glycerophosphocholine	b, h, il, k, li, lu, ly, s, t	chenomx, database, lit, STOCSY
4.36	t	uridine	k, li	chenomx, database, lit, STOCSY
4.37	m	AMP*	b, h, il, k, li, lu, s, t	chenomx, HMDB, STOCSY
4.38	m	ADP*	b, h, lu, s, t	chenomx, lit, STOCSY
4.38	m	UDP-glucose	li, s	chenomx, HMDB, STOCSY
4.39	m	*N*-acetylaspartate	b, t	chenomx, database, lit, STOCSY
4.40	t	inosine	b, h, il, k, li, ly, s	chenomx, database, lit, STOCSY
4.45	t	adenosine	b	chenomx, lit
4.46	t	inosine	b, h, il, k, li, ly, s	chenomx, database, lit, STOCSY
4.51	t	AMP*	b, h, il, k, li, lu, s, t	chenomx, HMDB, STOCSY
4.52	m	NAD^+*	b	chenomx**, HMDB, STOCSY
4.64	dd	β-glucose	li	chenomx, database, lit, STOCSY
5.19	d	mannose	li	chenomx, HMDB, STOCSY
5.23	m	lipid, glyceryl	ly	database, lit, STOCSY
5.25	d	α-glucose	li	chenomx, database, lit, STOCSY
5.31	m	lipid CH=CH unsaturated	k, ly	database, lit, STOCSY
5.42	m	glycogen	li	HMDB, lit, STOCSY
5.62	m	UDP-glucose	li, s	chenomx, HMDB, STOCSY
5.91	d	cytidine*	ly	chenomx**, HMDB, STOCSY
5.91	s	uridine	k, li	chenomx, database, lit, STOCSY
5.92	d	uridine	k, li	chenomx, database, lit, STOCSY
5.99	m	UDP-glucose	li, s	chenomx, HMDB, STOCSY
6.04	d	NAD^+*	b	chenomx**, HMDB, STOCSY
6.05	d	cytidine*	ly	chenomx**, HMDB, STOCSY
6.09	d	adenosine	b	chenomx, lit, STOCSY
6.11	d	inosine	b, h, il, k, li, ly, s	chenomx, database, lit, STOCSY
6.14	d	ADP*	b, h, lu, s, t	chenomx, lit, STOCSY
6.15	d	AMP*	b, h, il, k, li, lu, s, t	chenomx, HMDB, STOCSY
6.15	d	ATP*	b, h, il, li,	chenomx, lit, STOCSY
6.53	s	fumarate	h, h, il, k, li, lu, ly, s, t	chenomx, database, lit, STOCSY
6.91	d	tyrosine	b, il, k, li, ly	chenomx**, database, lit, STOCSY
7.08	s	histidine*	b	chenomx**, database, lit, STOCSY
7.13	s	histamine*	il	chenomx**, database, STOCSY

Continued on next page

Table A.2: (continued)

Chemical shift (ppm)	Peak Multiplicity	Molecule	Tissue	Method of Identification
7.20	d	tyrosine	b, il, k, li, ly	chenomx**, database, lit, STOCSY
7.33	m	phenylalanine	il, li, ly	chenomx, database, lit, STOCSY
7.38	m	phenylalanine	il, li, ly	chenomx, database, lit, STOCSY
7.43	m	phenylalanine	il, li, ly	chenomx, database, lit, STOCSY
7.60	q	nicotinamide	b, h, il, k, li, lu, ly, s	chenomx, database, STOCSY
7.68	s	UK 7.68(s)	h, il, li, lu	–
7.85	d	cytidine*	ly	chenomx**, HMDB, STOCSY
7.88	d	uridine	k, li	chenomx, database, lit, STOCSY
7.92	s	xanthine	li	chenomx, HMDB, lit
7.94	d	*N*-acetylaspartate	b, t	chenomx, database, lit, STOCSY
7.96	d	UDP-glucose	li, s	chenomx, HMDB, STOCSY
7.98	s	histamine*	il	chenomx**, database, STOCSY
8.08		histidine*	b	chenomx**, database, lit, STOCSY
8.15	s	NAD^+*	b	chenomx**, HMDB, STOCSY
8.18	m	NAD^+*	b	chenomx**, HMDB, STOCSY
8.21	s	hypoxanthine	b, il, s, t, lu, ly	chenomx, database, STOCSY
8.22	s	hypoxanthine	b, il, s, t, lu, ly	chenomx, database, STOCSY
8.25	s	inosine	b, h, il, k, li, ly, s	chenomx, database, STOCSY
8.25	dd	nicotinamide	b, h, il, k, li, lu, ly, s	chenomx, database, STOCSY
8.26	s	adenosine	b	chenomx, lit, STOCSY
8.28	s	AMP*	b, h, il, k, li, lu, s, t	chenomx, HMDB, STOCSY
8.36	s	adenosine	b	chenomx, lit, STOCSY
8.36	s	inosine	b, h, il, k, li, ly, s	chenomx, database, lit, STOCSY
8.43	s	NAD^+*	b	chenomx**, HMDB, STOCSY
8.46	s	formate	h, il, k, li, lu, ly, s, t	chenomx, database, lit
8.54	s	ATP*	b, h, il, li	chenomx, lit, STOCSY
8.58	s	ADP*	b, h, lu, s, t	chenomx, lit, STOCSY
8.61	s	AMP*	b, h, il, k, li, lu, s, t	chenomx, HMDB, STOCSY
8.72	dd	nicotinamide	b, h, il, k, li, lu, ly, s	chenomx, database, STOCSY
8.84	d	NAD^+*	b	chenomx**, HMDB, STOCSY
8.95	s	nicotinamide	b, h, il, k, li, lu, ly, s	chenomx, database, STOCSY
9.14	d	NAD^+*	b	chenomx**, HMDB, STOCSY
9.24	m	UK 9.24(m)	ly	STOCSY***
9.35	s	NAD^+*	b	chenomx**, HMDB, STOCSY
9.65	m	UK 9.65(m)	ly	STOCSY***

*Tentative assignments. **Partial match with chenomx peaks. ***Strong positive correlation 9.24(m) and 9.65(m).
Key as described in Table A.1 with the following additions: ADP, adenosine diphosphate; AMP, adenosine monophosphate; ATP, adenosine triphosphate; b, brain; h, heart; i, ileum; k, kidney; li, liver; lu; lung; ly, lymph nodes; NAD^+, nicotinamide adenine dinucleotide (oxidised ion); s, spleen; t, thymus

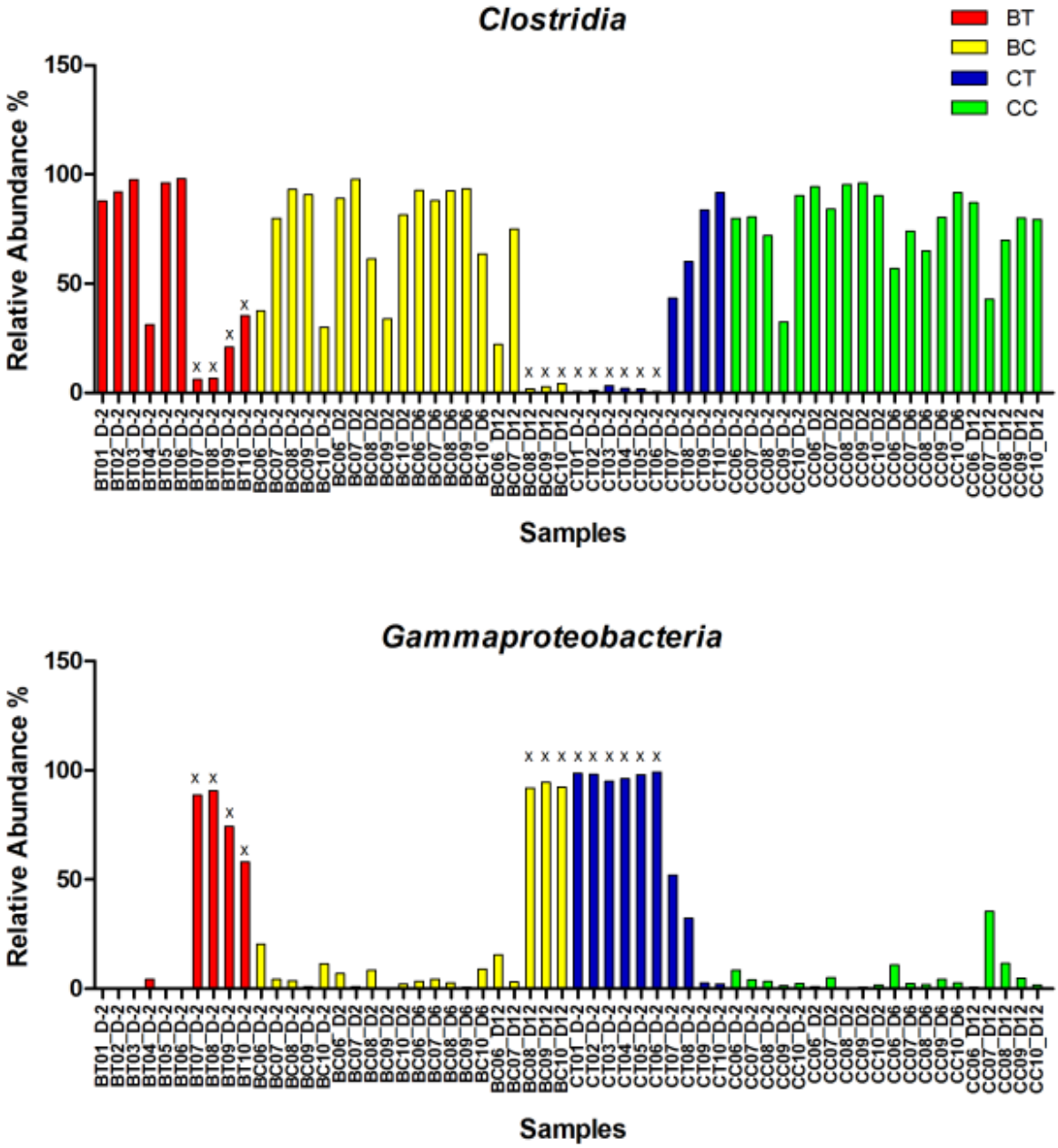


Figure A.1: Faecal Clostridia and Gammaproteobacteria Abundances Demonstrate Outliers within Dataset

Relative percentage abundance of *Clostridia* and *Gammaproteobacteria* classes are plotted for each sample. Sample outliers that displayed unusually high *Gammaproteobacteria* coupled with low *Clostridia* are marked with an 'x'. Abbreviations: BC, control BALB/c (BC); BT, pre-infected BALB/c (day -2); CC, control C57BL/6; CT, pre-infected C57BL/6 (day -2); D; day.

Table A.3: List of Faecal Bacteria Detected in BALB/c and C57BL/6 Mice

Phylum	**Class**	**Order**	**Family**	**Genus**
Actinobacteria	*Actinobacteria*	*Actinomycetales*	*Dermacoccaceae*	*Dermacoccus*
Actinobacteria	*Actinobacteria*	*Actinomycetales*	*Microbacteriaceae*	*Curtobacterium, Microbacterium*
Actinobacteria	*Actinobacteria*	*Actinomycetales*	*Propionibacteriaceae*	*Propionibacterium*
Firmicutes	*Bacilli*	*Bacillales*	*Bacillaceae*	*Bacillus*
Firmicutes	*Bacilli*	*Bacillales*	*Paenibacillaceae*	*Paenibacillus*
Firmicutes	*Bacilli*	*Bacillales*	*Staphylococcaceae*	*Staphylococcus*
Firmicutes	*Clostridia*	*Clostridiales*	*Clostridiales Incertae Sedis XI*	*Peptoniphilus, Tepidimicrobium*
Firmicutes	*Clostridia*	*Clostridiales*	*Lachnospiraceae*	*Clostridium* cluster XlVb, *Dorea, Lachnospiracea incertae sedis,* unclassified (x 51)
Firmicutes	*Clostridia*	*Clostridiales*	*Peptostreptococcaceae*	*Clostridium* cluster XI
Firmicutes	*Clostridia*	*Clostridiales*	*Ruminococcaceae*	*Anaerotruncus, Butyricicoccus, Oscillibacter,* unclassified (x 2)
Firmicutes	*Clostridia**	–	–	–
Firmicutes	*Erysipelotrichia*	*Erysipelotrichales*	*Erysipelotrichaceae*	*Turicibacter*
*Firmicutes**	–	–	–	–
Proteobacteria	*Alphaproteobacteria*	*Caulobacterales*	*Caulobacteraceae*	*Brevundimonas*
Proteobacteria	*Alphaproteobacteria*	*Rhizobiales*	*Methylobacteriaceae*	*Methylobacterium*
Proteobacteria	*Alphaproteobacteria*	*Rhodobacterales*	*Rhodobacteraceae*	Paracoccus, unclassified
Proteobacteria	*Alphaproteobacteria*	*Rhodospirillales*	*Acetobacteraceae*	unclassified (x 2)
Proteobacteria	*Alphaproteobacteria*	*Sphingomonadales*	*Sphingomonadaceae*	*Sphingomonas*
Proteobacteria	*Alphaproteobacteria**	–	–	–
Proteobacteria	*Betaproteobacteria*	*Burkholderiales*	*Alcaligenaceae*	*Achromobacter*
Proteobacteria	*Betaproteobacteria*	*Burkholderiales*	*Burkholderiaceae*	*Ralstonia*
Proteobacteria	*Betaproteobacteria*	*Burkholderiales*	*Comamonadaceae*	*Acidovorax, Schlegelella*
Proteobacteria	*Betaproteobacteria*	*Burkholderiales*	*Oxalobacteraceae*	*Massilia*
Proteobacteria	*Betaproteobacteria*	*Neisseriales*	*Neisseriaceae*	*Aquaspirillum*
Proteobacteria	*Gammaproteobacteria*	*Enterobacteriales*	*Enterobacteriaceae*	*Escherichia/Shigella, Serratia,* unclassified (x 2)
Proteobacteria	*Gammaproteobacteria*	*Pseudomonadales*	*Pseudomonadaceae*	*Pseudomonas*
Proteobacteria	*Gammaproteobacteria*	*Xanthomonadales*	*Xanthomonadaceae*	*Stenotrophomonas*
*Proteobacteria**	–	–	–	–
Otu (rare & singletons)	–	–	–	–
Otu00016	–	–	–	–
Otu00018	–	–	–	–
Otu00031	–	–	–	–
Otu00032	–	–	–	–
Otu00036	–	–	–	–
Otu00042	–	–	–	–
Otu00046	–	–	–	–
Otu00050	–	–	–	–
Otu00053	–	–	–	–
Otu00065	–	–	–	–
Otu00069	–	–	–	–
Otu00081	–	–	–	–
Otu00082	–	–	–	–

Continued on next page

Table A.3: (continued)

Phylum	Class	Order	Family	Genus
Otu00091	–	–	–	–
Otu00099	–	–	–	–
Otu00103	–	–	–	–
Otu00115	–	–	–	–
Otu00116	–	–	–	–
Otu00142	–	–	–	–
Otu00148	–	–	–	–
Otu00224	–	–	–	–
Otu00233	–	–	–	–
Otu00235	–	–	–	–
Otu00243	–	–	–	–
Otu00248	–	–	–	–
Otu00249	–	–	–	–
Otu00346	–	–	–	–

List of microbial OTU's (as determined using MORTHUR analyses) that were detected in the faeces of the control and pre-infected mice, of both BALB/c and C57BL/6 strains. OTU's are labelled to their genus level, though total OTU counts at the family level were used as loading variables for final PCA analyses. The number of OTU's unclassified at the genus level are shown in brackets. OTU's which were only present in one animal (singleton) or only 1 count was detected (rare) were pooled for analysis. *OTU's were resolved only to specified taxonomic level.

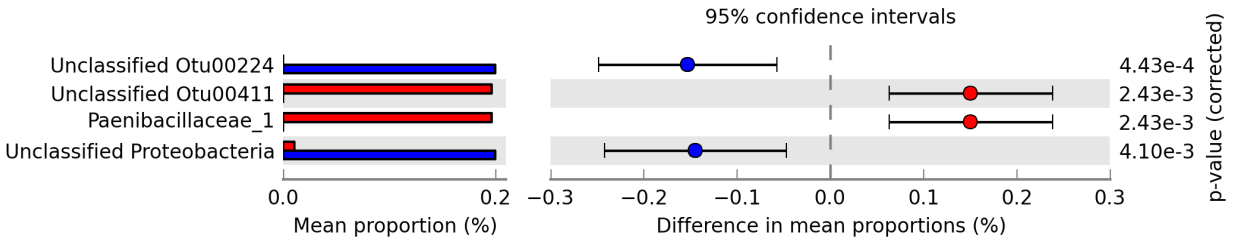


Figure A.2: Differential Faecal Family OTU Data in Mouse Strains Using Bonferroni Multiple Test Correction

Extended error plot showing average percentage relative abundance (proportion) of discriminatory faecal microbial OTU's on the left plot, and average difference between the group means and their confidence intervals on the right, based on OTU count data resolved at the family taxonomic level. Significant differences between BALB/c mouse strain (red bars) and C57BL/6 strains (blue bars) were tested using White's non-parametric test with Bonferroni correction. Corrected p-values shown in the plot.

Table A.4: Physiological and Pathological Measures of CL Models

	Time-point	Group Median (range) Infected BALB/c	Infected C57BL/6	P-value Difference between strains	Difference over time	Difference between infection status[a]
Weight (g)	Day -2	18.2 (17.2–19.4)	18.7 (17.3–19.5)	NS	NS	NS
	Day 2	18.7 (17.6–20.0)	18.3 (17.2–19.5)	NS	NS	NS
	Day 6	19.1 (17.8–20.0)	18.5 (17.1–19.6)	NS	NS	NS
	Day 12	–	–	–	–	–
PCV (%)	Day -2	62.5 (56.9–66.1)	54.0 (47.5–58.7)	0.001 (BT > CT)	NS	NS
	Day 2	59.9 (58.7–65.0)	53.2 (50.00–54.7)	0.001 (BT > CT)	NS	NS
	Day 6	61.9 (60.0–64.5)	53.2 (46.0–55.6)	0.001 (BT > CT)	NS	0.044 (BT > BC)
	Day 12	60.0 (58.1–65.1)	51.2 (43.8–54.8)	0.001 (BT > CT)	NS	NS
Lesion	Day of Onset	7.5 (5–10)	7.9 (5–11)	NS	–	–
	Max Area (mm^2)	20.22 (3.93–28.27)	11.78 (9.42–25.13)	NS	–	–

[a] See Table 3.1 for results from control mice. Key: –, not tested/applicable; BT > CT, levels in infected BALC/c (BT) greater than infected C57BL/6 (CT); NS, not statistically significant; PCV, packed cell volume.

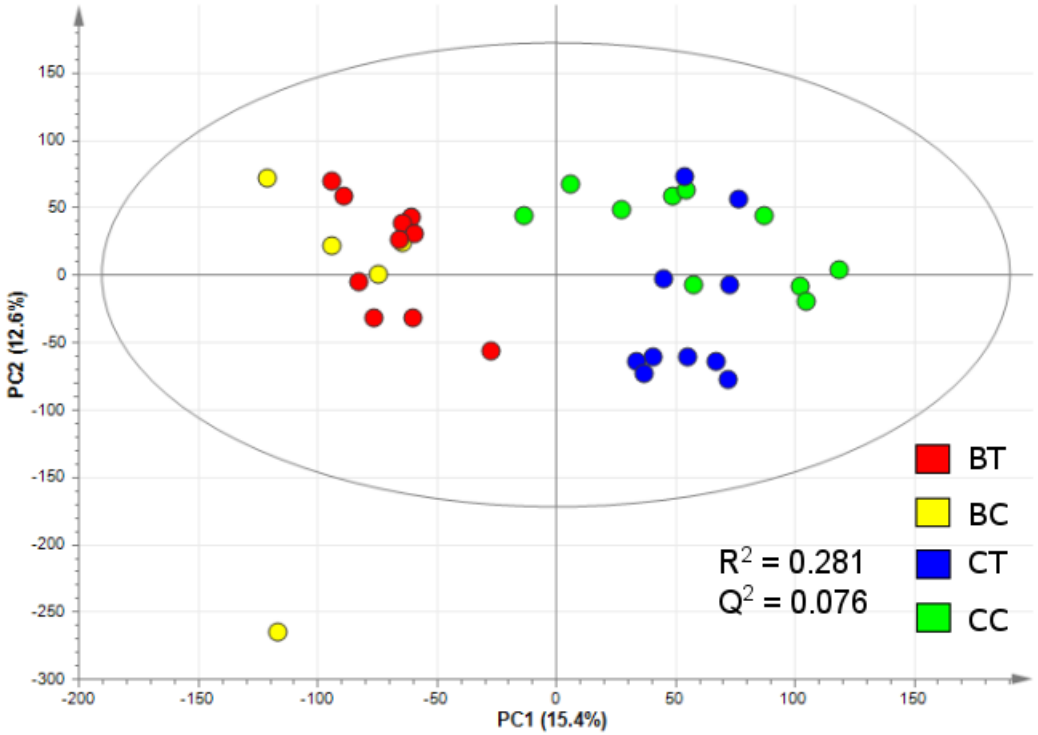


Figure A.3: Strain-related Variation Dominates Infection-associated Changes

Scores plot of PCA model based plasma ^{1}H NMR spectral data from all four groups on D6. Abbreviations: BC, control BALB/c (BC); BT, infected BALB/c; CC, control C57BL/6; CT, infected C57BL/6.

Table A.5: CL-Model Discriminatory Metabolites in Biofluids per Time-point

Metabolite	Urine D-2	Urine D2	Urine D6	Urine Dterm	Plasma D2	Plasma D6	Plasma Dterm	Plasma CPMG D2	Plasma CPMG D6	Plasma CPMG Dterm	Faeces D-2	Faeces D2	Faeces D6	Faeces Dterm
2-Oxoadipate	C		B	C										
2-Oxoglutarate			B											
2-Oxoisocaproate	C		C											
2-Oxovalerate		B	C	B										
3-Hydroxybutyrate					B	B		B	B	B				
3-Hydroxyisovalerate*		C												
3-Methyl-2-Oxovalerate*	C	B		C										
Acetate				C	C				B		B	C		B
Alanine								B	B			B		
Bile Acid													B	C
Butyrate													C	B
Cholate*													B	C
cis-Aconitate		B		B										
Citrate			B	C				B						
Creatine	C		C		B			B	B					
Creatinine		B												
Dimethylamine		B												
Dimethylglycine		C		C										
Fumarate			B											
Glucose					B			B	B					
Glutamate													C	
Glutamine					B			B		B				
Glycerophosphocholine		C			C	C								
Glycocholate*													B	C
Glycine					C								B	
Hippurate	B	C	C											
Indoxylsulphate	B	C												
Lactate			B	C	C		C	C	B				B	
Lipid CH_2		B			C			C	B					
Lipid $CH_2C{=}C$					C									
Lipid CH_2CH_2CO					C									
Lipid CH_2CO					C									
Lipid CH_3					C			C	B				B	
scyllo-Inositol*								C		C				
Phenylacetylglycine		B		B										
Pyruvate					B			B						
Succinate			B											B
Taurine	C			C										
Trimethylamine				B								B		B
Trimethylamine-*N*-oxide													B	
Tyramine											B	B		B

Continued on next page

Table A.5: (continued)

Metabolite	Urine D-2	Urine D2	Urine D6	Urine Dterm	Plasma D2	Plasma D6	Plasma Dterm	Plasma CPMG D2	Plasma CPMG D6	Plasma CPMG Dterm	Faeces D-2	Faeces D2	Faeces D6	Faeces Dterm
Ureidopropanoate	C													
Valine								B	B					
UK 1.88(dd)		B	B											
UK 7.87(s)												B	B	

Table lists infection-associated metabolites that were discriminatory between the two CL models, as measured *via* O-PLS-DA. Key: *, tentative assignments; B, positively correlated with infected BALB/c strain; C, positively correlated with infected C57BL/6 strain; (dd), doublet of doublets; (s), singlet; UK, unknown metabolites.

Table A.6: CL-Model Discriminatory Metabolites per Tissue

Metabolite	Ileum	Liver	Spleen	Thymus	Lymph Nodes	Heart	Brain-hippo	Brain-cortex	Brain-cbell	Brain-stem
3-Hydroxybutyrate		B				C				
4-Aminobutyrate							B		C	
N-Aceytlaspartate								B	C	C
Acetate		B				C	B	B		B
Adenosine*							B			
ADP	C			B		B	B	B	B	
Alanine			B		B	C	B	B		B
AMP		B	B	C			C		C	C
Aspartate			B				C		C	B
ATP*		B	B					B		
Betaine			B							
Choline	B		B	B	B	C	B	B		B
Creatine	C		B	C			C	B	C	B
Creatinine								B		B
Dimethylamine			B							
Ethanolamine	B							B		
Fumarate	B									
Glutamate		B	B			C	B			B
Glutamine	B			C				B		
Glutathione	C	B	B							
Glycerol			B		B					
Glycerophosphocholine		B	B	C						
Glycine	B	B			B			B	C	B
Glycogen		C								
Histamine	B	B								
Hypoxanthine	B		B				B			B
Inosine	B	B			B	C	B		B	B
Lactate	C		B				B	B	B	B
Lipid C=CCH$_2$C=C	C									
Lipid CH$_2$	C			B	C	B	C			
Lipid CH2CH$_2$CO	C									
Lipid CH$_2$C=C	C									
Lipid CH$_2$O	C									
Lipid CH$_3$	C	B		B	C	B				
myo-Inositol			B				B	B		
NAD$^+$*							C			
Nicotinamide		B								
Phosphocholine	C	B	B							
Phosphoethanolamine			B	C						
Succinate			B	C		C				
Taurine		B	B		B		B	B		B
Tyrosine							C			

Continued on next page

Table A.6: (continued)

Metabolite	Ileum	Liver	Spleen	Thymus	Lymph nodes	Heart	Brain hippo	Brain cortex	Brain cbell	Brain stem
UDP-Glucose		B	B							
Valine						C				
UK 3.48(d)	B		B							
UK 7.68(s)				B			B			C

Table lists infection-associated metabolites that were discriminatory between the two CL models, as measured *via* O-PLS-DA. Kidney data omitted as no significant differences were found. Key: *, tentative assignments; ADP, adenosine diphosphate; ADP, adenosine diphosphate; AMP, adenosine monophosphate; ATP, adenosine triphosphate; ATP, adenosine triphosphate; B, positively correlated with infected BALB/c strain; C, positively correlated with infected C57BL/6 strain; cbell, cerebellum; (d), doublet; hippo, hippocampus; NAD^+, nicotinamide adenine dinucleotide (oxidised ion); (s), singlet; UDP, uridine diphospho; UK, unknown metabolites.

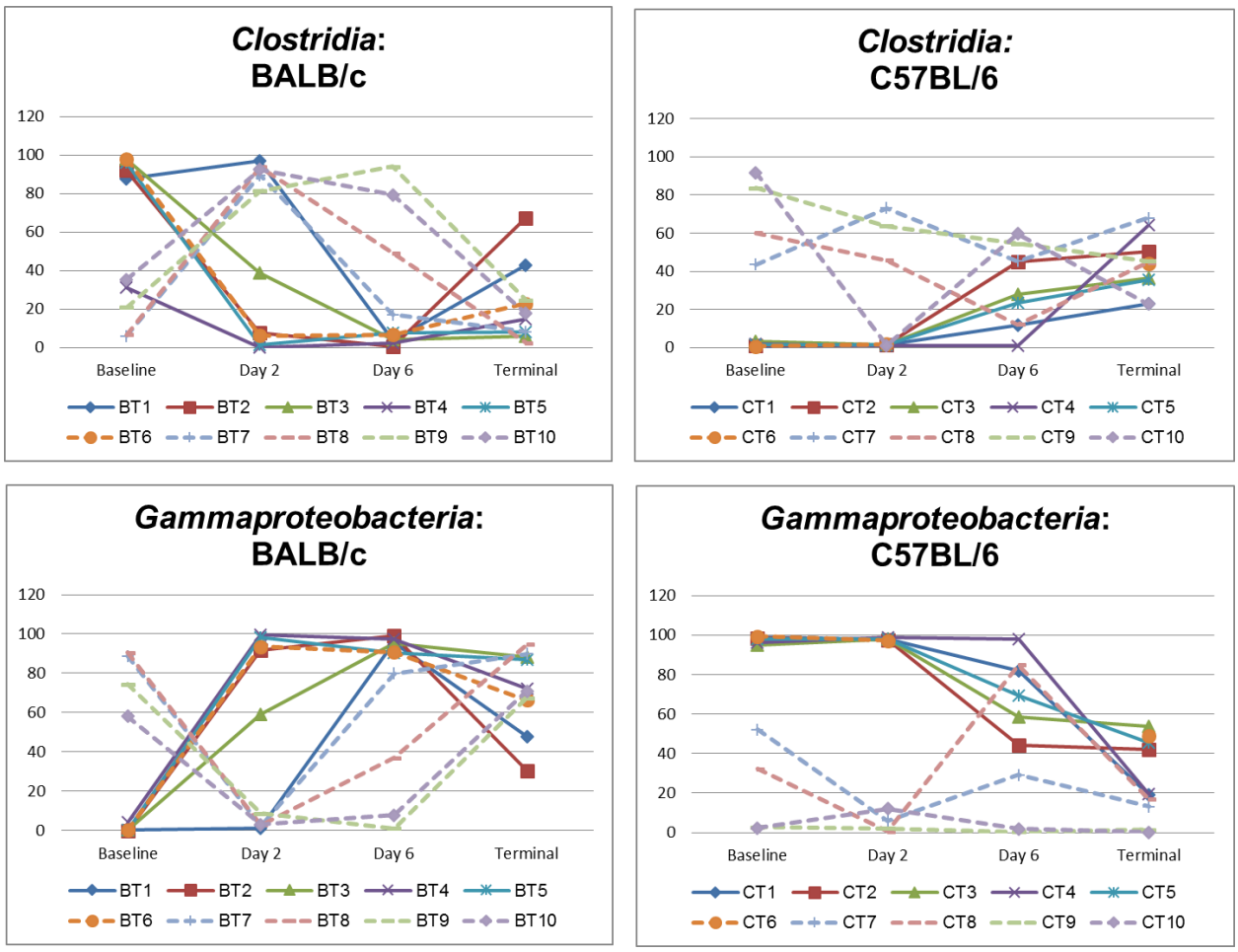


Figure A.4: Cage-related Variation Observed in Microbial Data

Graphs show individual time-trajectories of relative abundance of *Clostridia* and *Gammaproteobacteria* for each *L. major* infected mice in both CL models. For each strain, animals 1–5 (solid lines) were housed in one cage, whilst animals 6–10 (hashed lines) were housed together in a separate cage. Abbreviations: BT, infected BALB/c; CT, infected C57BL/6.

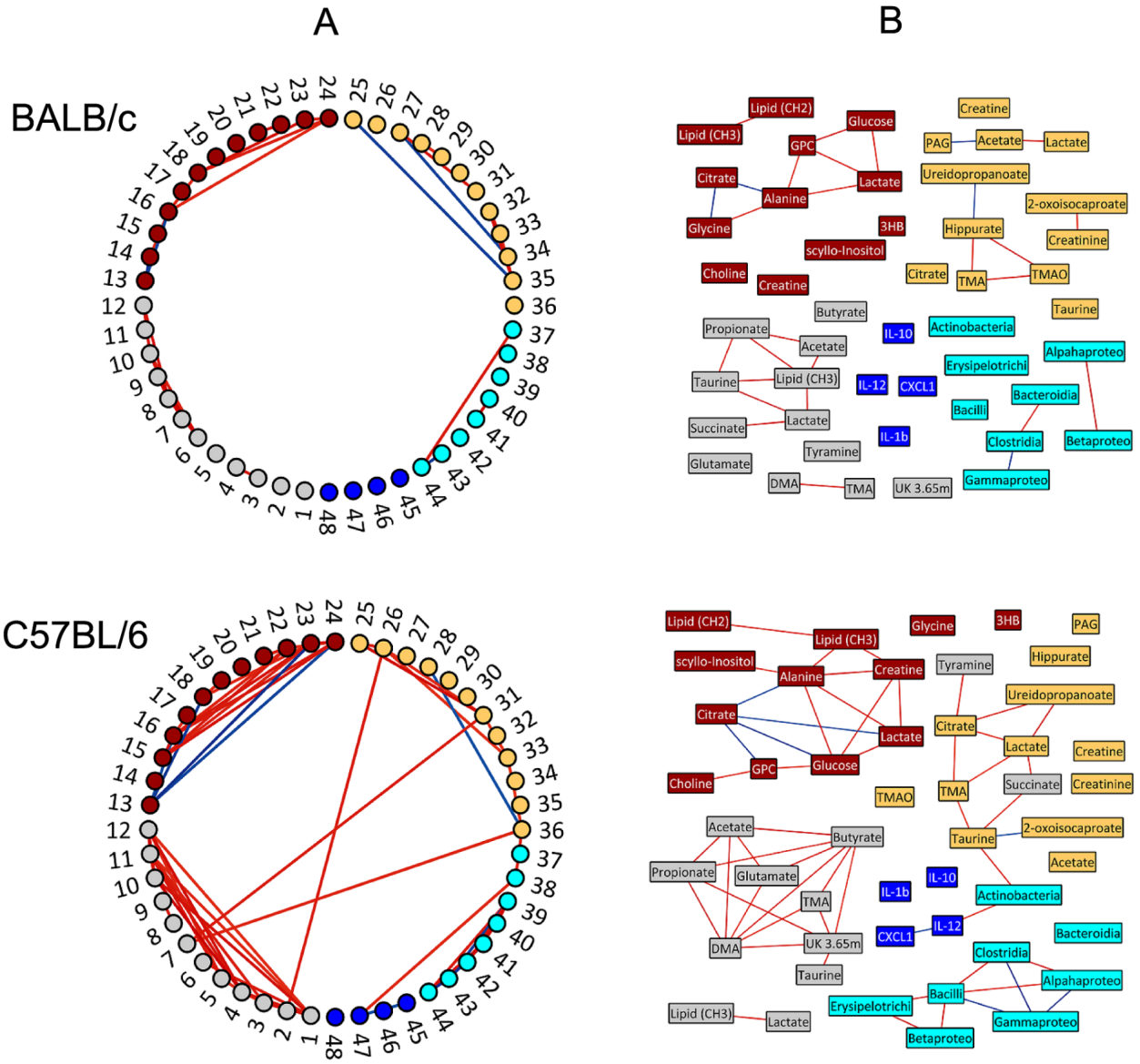


Figure A.5: Correlation Networks of Host Metabolic, Cytokine and Microbial Markers in CL Models Prior to Infection

Correlation network analyses in wheels format (A) and map format (B) based on pair-wise Pearson correlation between all variables for BALB/c and C57BL/6 mice, two days prior to infection. Significant positive and negative correlations between variables are shown as red lines and blue lines, respectively, in both formats. Each variable has a fixed position, with nodes along the circumference that are colour-coded according to biological matrix in the following order: faecal metabolites in grey (1–12); plasma metabolites in red (13–24); urinary metabolites in yellow (25–36); microbial classes in cyan (37–43), cytokines in dark blue (45–48). Complete list of variables as in Figure 4.8. Abbreviations as described previously. *Reprinted with permission from Lamour, et al. (2014)*[197] *Copyright 2014 American Chemical Society.*

Table A.7: List of ^{1}H NMR Metabolite Assignments Measured in Human Biofluids

Chemical shift (ppm)	Peak Multiplicity	Molecule	Biofluid	Method of Identification
0.68	m	cholesterol	p	database, lit, STOCSY
0.84	d	2-hydroxyisovalerate	c	chenomx, lit, STOCSY
0.86	m	lipid CH_3	p	database, lit, STOCSY
0.90	t	2-hydroxybutyrate	c	chenomx, lit, STOCSY
0.94	t	isoleucine	p	chenomx, database, lit, STOCSY
0.97	d	2-hydroxyisovalerate	c	chenomx, lit, STOCSY
0.97	d	leucine	p	chenomx, database, lit, STOCSY
0.98	d	leucine	p	chenomx, database, lit, STOCSY
1.01	d	valine	c, p	chenomx, database, lit, STOCSY
1.02	d	isoleucine	p	chenomx, database, lit, STOCSY
1.06	d	valine	p	chenomx, database, lit, STOCSY
1.19	t	ethanol[a]	c, p	chenomx, database, STOCSY
1.20	d	3-hydroxybutyrate	c, p	chenomx, database, lit, STOCSY
1.25	s	3-hydroxyisovalerate	c	chenomx, database
1.26	m	lipid CH_2	p	database, lit, STOCSY
1.34	d	lactate	c, p	chenomx, database, lit, STOCSY
1.45	d	unassigned	c, p	–
1.49	d	alanine	c, p	chenomx, database, lit, STOCSY
1.57	m	lipid CH_2CH_2CO	p	chenomx, lit, STOCSY
1.65	m	2-hydroxybutyrate	c	chenomx, lit, STOCSY
1.72	m	lysine	c, p	chenomx, database, lit, STOCSY
1.73	m	2-hydroxybutyrate	c	chenomx, lit, STOCSY
1.89	m	lysine	c, p	chenomx, database, lit, STOCSY
1.93	s	acetate	c, p	chenomx, database, lit, STOCSY
2.02	m	lipid $CH_2C{=}C$	p	lit, STOCSY
2.06	m	*N*-acetyl glycoprotein	c, p	database, lit, STOCSY
2.09	m	*N*-acetyl glycoprotein	c, p	database, lit, STOCSY
2.14	m	glutamine	c, p	chenomx, database, lit, STOCSY
2.15	s	paracetamol conjugate[a]	c	chenomx, lit, STOCSY
2.24	s	acetone	c	chenomx, lit
2.24	m	lipid CH_2CO	p	lit, STOCSY
2.25	s	acetoacetate	c, p	chenomx, database, lit, STOCSY
2.32	m	3-hydroxybutyrate	c, p	chenomx, database, lit, STOCSY
2.36	m	glutamate	p	chenomx, database, lit, STOCSY
2.37	s	pyuvate	c, p	chenomx, database, lit
2.42	m	3-hydroxybutyrate	c, p	chenomx, database, lit, STOCSY
2.46	m	glutamine	c, p	chenomx, database, lit, STOCSY
2.53	d	citrate	c, p	chenomx, database, lit, STOCSY
2.58	s	Ca-EDTA^{2-} [a]	p	lit, STOCSY
2.67	d	citrate	c, p	chenomx, database, lit, STOCSY
2.72	s	dimethylamine	c	chenomx, lit
2.72	s	Mg-EDTA^{2-} [a]	p	lit, STOCSY
2.74	m	lipid C=CCH2C=C	p	lit, STOCSY
2.92	m	lipid ε-CH_2 lysyl (albumin)	p	lit, STOCSY
2.98	m	lipid ε-CH_2 lysyl (albumin)	p	lit, STOCSY
3.01	m	lipid ε-CH_2 lysyl (albumin)	p	lit, STOCSY
3.03	t	*myo*-inositol	c	chenomx, database, lit, STOCSY
3.03	t	lysine	c	chenomx, database, lit, STOCSY
3.04	s	creatine	c, p	chenomx, database, lit, STOCSY
3.05	s	creatinine	c, p	chenomx, database, lit, STOCSY
3.15	m	Ca-EDTA^{2-} [a]	p	lit, STOCSY
3.22	s	glycerophosphocholine	p	chenomx, database, lit, STOCSY
3.23	s	H-EDTA^{3-} [a]	p	lit, STOCSY
3.26	dd	β-glucose	c, p	chenomx, database, lit, STOCSY
3.29	t	*myo*-inositol	c, p	chenomx, database, lit, STOCSY
3.38	s	*scyllo*-inositol/methanol[a]	c, p	chenomx, database, lit
3.42	t	β-glucose	c, p	chenomx, database, lit, STOCSY
3.43	t	α-glucose	c, p	chenomx, database, lit, STOCSY
3.46	s	acetoacetate	p	chenomx, database, lit, STOCSY
3.48	dd	β-glucose	c, p	chenomx, database, lit, STOCSY
3.51	t	β-glucose	c, p	chenomx, database, lit, STOCSY
3.56	dd	α-glucose	c, p	chenomx, database, lit, STOCSY
3.58	s	glycine	p	chenomx, database, lit
3.59	dd	glycerol	p	chenomx, database, lit, STOCSY
3.63	s	H-EDTA^{3-} [a]	p	lit, STOCSY
3.66	dd	glycerol	p	chenomx, database, lit, STOCSY
3.68	q	ethanol[a]	p	chenomx, database, STOCSY
3.72	t	α-glucose	c, p	chenomx, database, lit, STOCSY
3.73	dd	β-glucose	c, p	chenomx, database, lit, STOCSY
3.76	m	α-glucose	c, p	chenomx, database, lit, STOCSY

Continued on next page

Table A.7: (continued)

Chemical shift (ppm)	Peak Multiplicity	Molecule	Biofluid	Method of Identification
3.78	m	gluconate	c	chenomx, HMDB, STOCSY
3.84	dd	α-glucose	c, p	chenomx, database, lit, STOCSY
3.85	m	α-glucose	c, p	chenomx, database, lit, STOCSY
3.92	dd	β-glucose	c, p	chenomx, database, lit, STOCSY
3.94	s	creatine	c, p	chenomx, database, lit, STOCSY
4.00	m	phenylalanine	c, p	chenomx, database, lit, STOCSY
4.04	t	gluconate	c	chenomx, HMDB, STOCSY
4.07	s	creatinine	c, p	chenomx, database, lit, STOCSY
4.08	t	*myo*-inositol	c, p	chenomx, database, lit, STOCSY
4.13	q	lactate	c, p	chenomx, database, lit, STOCSY
4.15	d	gluconate	c	chenomx, HMDB, STOCSY
4.17	m	3-hydroxybutyrate	p	chenomx, database, lit, STOCSY
4.33	m	glycerophosphocholine	p	chenomx, database, lit, STOCSY
4.66	dd	β-glucose	c, p	chenomx, database, lit, STOCSY
4.91	d	mannose	c	chenomx, database, lit, STOCSY
5.19	d	mannose	c	chenomx, database, lit, STOCSY
5.20	m	lipid glyceryl	p	lit, STOCSY
5.25	d	α-glucose	c, p	chenomx, database, lit, STOCSY
5.30	m	lipid CH=CH unsaturated	p	database, lit, STOCSY
5.80	m	urea	c, p	chenomx, database, lit, STOCSY
6.92	s	paracetamol[a]	c	chenomx, lit, STOCSY
6.92	d	tyrosine	c, p	chenomx, database, lit, STOCSY
7.08	s	histidine	c, p	chenomx, database, lit, STOCSY
7.21	d	tyrosine	c, p	chenomx, database, lit, STOCSY
7.25	s	paracetamol[a]	c	chenomx, lit, STOCSY
7.35	d	phenylalanine	c, p	chenomx, database, lit, STOCSY
7.39	m	phenylalanine	c, p	chenomx, database, lit, STOCSY
7.44	m	phenylalanine	c, p	chenomx, database, lit, STOCSY
7.80	s	histidine	c, p	chenomx, database, lit, STOCSY
8.47	s	formate	c, p	chenomx, database, lit

[a] exogenous compound (drug metabolite, contaminant, or anticoagulant).
Key: (d), doublet; (dd), doublet of doublets; (m), multiplet; (s), singlet; (t), triplet; c, cerebral spinal fluid; EDTA, ethylenediaminetetraacetic acid; lit, literature[115,145–147]; HMDB, the Human Metabolome Database[142–144]; p, plasma; STOCSY, statistical total correlation spectroscopy.

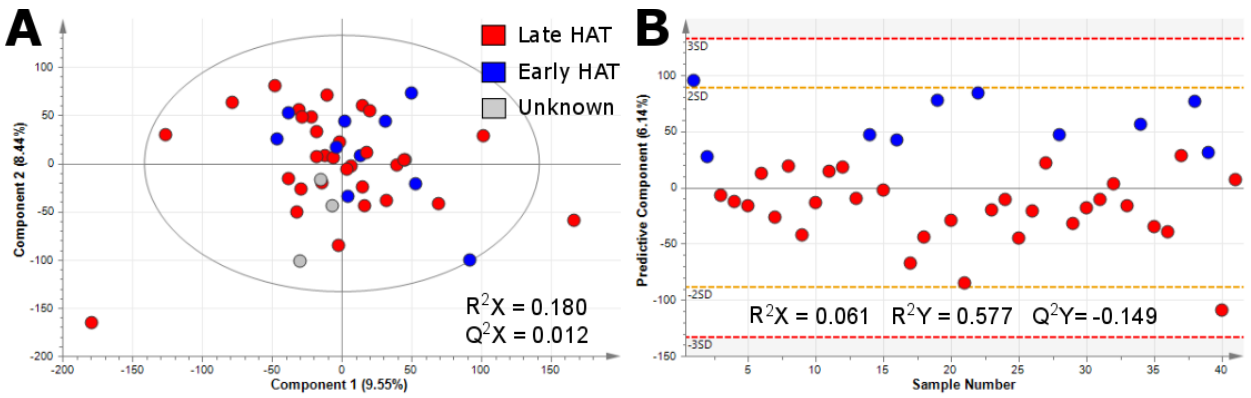


Figure A.6: No Dominant Metabolic Variation in Plasma ^{1}H NMR Phenotypes Associated with HAT Diagnostic Stage

PCA model (A) and O-PLS-DA model (B) score plots of plasma ^{1}H NMR spectra across HAT patients with early stage disease (n=10 in blue) vs. late stage (n=31 in red). Staging information was missing for 4 patients (in grey in PCA) which were omitted from O-PLS-DA. No significant differences between stages were shown (O-PLS-DA model not predictive with negative Q^2Y).

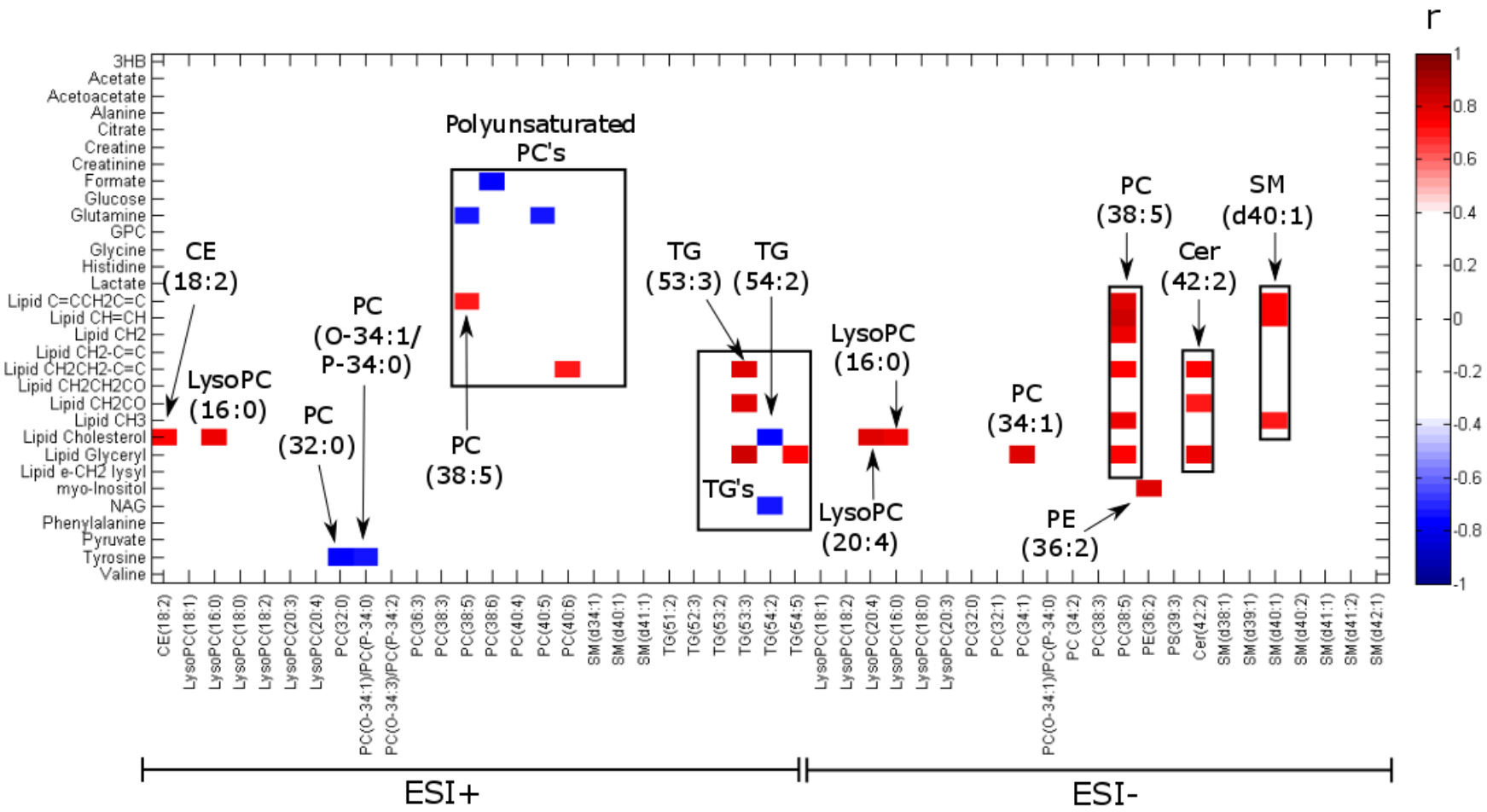


Figure A.7: Correlations between Plasma ^{1}H NMR and UPLC-MS Markers after FDR Correction

2D Pearson-based correlations performed between integrals of ^{1}H NMR metabolite peaks and lipid UPLC-MS features, measured in plasma of HAT patients (n=16). Positive correlations are shown in red and negative ones in blue. For all displayed associations, p<0.1, following Benjamini-Hochberg FDR correction.

Table A.8: Assignment of HAT Discriminatory Lipid Species Measured in Plasma via UPLC-MS

Lipid Class	Lipid Subclass	Composition	Formula	Observed Ion	RT (mins)	Predicted neutral mass	Measured m/z	Mass Accuracy Error (ppm)	P-value	% Difference
				ESI Positive Mode						
Sterol Lipids	CE	18:2	$C_{45}H_{76}O_2$	$[M+NH_4]^+$	15.70	648.585	666.621	3	1.12E-04	-53.1%
Glycerophospholipids	LysoPC	16:0	$C_{24}H_{50}NO_7P$	$[M+H]^+$	1.23	495.333	496.340	1	1.93E-05	-58.6%
Glycerophospholipids	LysoPC	18:1	$C_{26}H_{52}NO_7P$	$[M+H]^+$	1.29	521.348	522.355	0	1.98E-05	-66.5%
Glycerophospholipids	LysoPC	18:0	$C_{26}H_{54}NO_7P$	$[M+H]^+$	1.70	523.364	524.372	1	1.11E-04	-72.6%
Glycerophospholipids	LysoPC	18:2	$C_{26}H_{50}NO_7P$	$[M+H]^+$	1.05	519.333	520.338	3	3.28E-04	-80.5%
Glycerophospholipids	LysoPC	20:3	$C_{28}H_{52}NO_7P$	$[M+H]^+$	1.55	545.348	546.355	0	5.97E-05	-75.0%
Glycerophospholipids	LysoPC	20:4	$C_{28}H_{50}NO_7P$	$[M+H]^+$	1.03	543.333	544.340	0	4.36E-06	-75.6%
Glycerophospholipids	PC	32:0	$C_{40}H_{80}NO_8P$	$[M+H]^+$ *	7.69	733.562	734.572	3	2.49E-05	104.2%
Glycerophospholipids	PC	O-34:1/P-34:0	$C_{42}H_{84}NO_7P$	$[M+H]^+$	9.32	745.599	746.607	1	3.93E-04	86.2%
Glycerophospholipids	PC	O-34:3/P-34:2	$C_{42}H_{80}NO_7P$	$[M+H]^+$	7.24	741.567	742.576	1	8.43E-04	-52.8%
Glycerophospholipids	PC	36:3	$C_{44}H_{82}NO_8P$	$[M+H]^+$ **	6.89	783.578	785.591	1	6.56E-03	-33.9%
Glycerophospholipids	PC	38:3	$C_{46}H_{86}NO_8P$	$[M+H]^+$ **	9.16	811.609	813.622	5	1.93E-03	-43.0%
Glycerophospholipids	PC	38:5	$C_{46}H_{82}NO_8P$	$[M+H]^+$	6.24	807.578	808.587	2	1.71E-02	-40.9%
Glycerophospholipids	PC	38:6	$C_{46}H_{80}NO_8P$	$[M+H]^+$	6.15	805.562	806.573	4	1.63E-02	-33.3%
Glycerophospholipids	PC	40:4	$C_{48}H_{88}NO_8P$	$[M+H]^+$	9.95	837.625	838.635	3	3.25E-03	-47.9%
Glycerophospholipids	PC	40:5	$C_{48}H_{86}NO_8P$	$[M+H]^+$	8.22	835.609	836.619	3	1.13E-02	-39.2%
Glycerophospholipids	PC	40:6	$C_{48}H_{84}NO_8P$	$[M+H]^+$	7.53	833.594	834.603	2	2.94E-02	-32.7%
Sphingolipids	SM	d34:1	$C_{39}H_{79}N_2O_6P$	$[M+H]^+$	5.52	702.568	703.576	1	1.75E-02	96.1%
Sphingolipids	SM	d40:1	$C_{45}H_{91}N_2O_6P$	$[M+H]^+$	12.90	786.662	787.672	4	1.01E-03	-43.9%
Sphingolipids	SM	d41:1	$C_{46}H_{93}N_2O_6P$	$[M+H]^+$	13.22	800.677	801.687	3	1.06E-03	-58.6%
Glycerolipids	TG	51:2	$C_{54}H_{100}O_6$	$[M+NH_4]^+$	15.73	844.752	862.791	6	6.52E-05	98.0%
Glycerolipids	TG	52:3	$C_{55}H_{100}O_6$	$[M+NH_4]^+$	15.60	856.752	874.791	5	4.40E-03	45.4%
Glycerolipids	TG	53:2	$C_{56}H_{104}O_6$	$[M+NH_4]^+$	15.99	872.783	890.821	4	6.73E-04	86.1%
Glycerolipids	TG	53:3	$C_{56}H_{102}O_6$	$[M+NH_4]^+$	15.74	870.768	888.806	5	4.02E-04	77.5%
Glycerolipids	TG	54:2	$C_{57}H_{106}O_6$	$[M+NH_4]^+$	16.13	886.799	904.835	2	3.74E-04	70.3%
Glycerolipids	TG	54:5	$C_{57}H_{100}O_6$	$[M+NH_4]^+$	15.51	880.752	898.789	3	2.20E-03	68.2%
				ESI Negative Mode						
Sphingolipids	Cer	42:2	$C_{42}H_{81}NO_3$	$[M+FA-H]^-$	13.62	647.622	692.618	2	2.51E-03	54.3%
Glycerophospholipids	LysoPC	16:0	$C_{24}H_{50}NO_7P$	$[M+FA-H]^-$	1.23	495.333	540.329	1	5.42E-04	-60.1%
Glycerophospholipids	LysoPC	18:1	$C_{26}H_{52}NO_7P$	$[M+FA-H]^-$	1.30	521.348	566.344	4	1.00E-04	-62.5%
Glycerophospholipids	LysoPC	18:2	$C_{26}H_{50}NO_7P$	$[M+FA-H]^-$	1.05	519.333	564.330	1	3.18E-04	-78.2%
Glycerophospholipids	LysoPC	20:3	$C_{28}H_{52}NO_7P$	$[M+FA-H]^-$	1.14	545.348	590.346	0	7.00E-04	-77.9%
Glycerophospholipids	LysoPC	20:4	$C_{28}H_{50}NO_7P$	$[M+FA-H]^-$	1.01	543.333	588.331	0	2.62E-04	-73.1%
Glycerophospholipids	PC	32:0	$C_{40}H_{80}NO_8P$	$[M+FA-H]^-$	7.70	733.562	778.560	0	2.58E-03	111.7%
Glycerophospholipids	PC	32:1	$C_{40}H_{78}NO_8P$	$[M+FA-H]^-$	6.04	731.547	776.545	0	1.29E-04	62.5%
Glycerophospholipids	PC	O-34:1/P-34:0	$C_{42}H_{84}NO_7P$	$[M+FA-H]^-$	9.31	745.599	790.599	2	2.00E-03	84.3%
Glycerophospholipids	PC	34:1	$C_{42}H_{82}NO_8P$	$[M+FA-H]^-$	7.94	759.578	804.578	2	8.21E-03	35.1%

Continued on next page

Table A.8: (continued)

Lipid Class	Lipid Subclass	Composition	Formula	Observed Ion	RT (mins)	Predicted neutral mass	Measured m/z	Mass Accuracy Error (ppm)	P-value	% Difference
Glycerophospholipids	PC	34:2	$C_{43}H_{82}NO_{10}P$	$[M+FA-H]^-$ **	6.42	757.562	803.568	0	1.29E-04	40.2%
Glycerophospholipids	PC	38:3	$C_{46}H_{86}NO_8P$	$[M+FA-H]^-$	9.13	811.609	856.607	0	2.51E-03	-30.7%
Glycerophospholipids	PC	38:5	$C_{46}H_{82}NO_8P$	$[M+FA-H]^-$	6.27	807.578	852.576	0	1.29E-04	-28.8%
Glycerophospholipids	PE	36:2	$C_{41}H_{78}NO_8P$	$[M-H]^-$	9.16	743.547	742.541	2	2.51E-03	54.4%
Glycerophospholipids	PS	39:3	$C_{45}H_{82}NO_{10}P$	$[M+FA-H]^-$	7.94	827.568	872.567	1	4.42E-03	37.5%
Sphingolipids	SM	d38:1	$C_{43}H_{87}N_2O_6P$	$[M+FA-H]^-$	10.15	758.630	803.630	2	3.46E-03	-37.6%
Sphingolipids	SM	d39:1	$C_{44}H_{89}N_2O_6P$	$[M+FA-H]^-$	11.70	772.646	817.645	1	4.42E-03	-52.1%
Sphingolipids	SM	d40:1	$C_{45}H_{91}N_2O_6P$	$[M+FA-H]^-$	12.90	786.662	831.661	1	3.46E-03	-48.6%
Sphingolipids	SM	d40:2	$C_{45}H_{89}N_2O_6P$	$[M+FA-H]^-$	10.59	784.646	829.646	2	2.58E-03	-60.2%
Sphingolipids	SM	d41:1	$C_{46}H_{93}N_2O_6P$	$[M+FA-H]^-$	13.22	800.677	845.675	0	2.00E-03	-48.2%
Sphingolipids	SM	d41:2	$C_{46}H_{91}N_2O_6P$	$[M+FA-H]^-$	12.16	798.662	843.661	1	4.42E-03	-55.3%
Sphingolipids	SM	d42:1	$C_{47}H_{95}N_2O_6P$	$[M+FA-H]^-$	13.49	814.693	859.689	2	3.92E-04	-52.5%

* Sodium ion for PC (32:0) was also discriminatory (omitted from Table to only show one peak per feature).

** Results based on non-predominant isotope (main isotope was observed but not significantly discriminatory).

Key: C, carbon atom; Cer, ceramide; CE, cholesterol ester; d, dihydroxy; FA, formate; H, hydrogen atom; LysoPC, lysophosphatidylcholine; LysoPE, lysophosphatidylethanolamine; M, molecular ion; m/z, mass/charge ratio; N, Nitrogen atom; O, Oxygen atom; O-, O-alkyl-bond; P, Phosphorus atom; P-, O-alk-1-enyl-bond; PC, phosphatidylcholine; PE, phosphatidylehtanolamine; ppm, parts per million; PS, phosphatidylserine; RT, retention time; SM, sphingomyelin; TG, triglyeride. *Modified with permission from Lamour, et al. (2015)[275] Copyright 2015 Lamour, et al.*

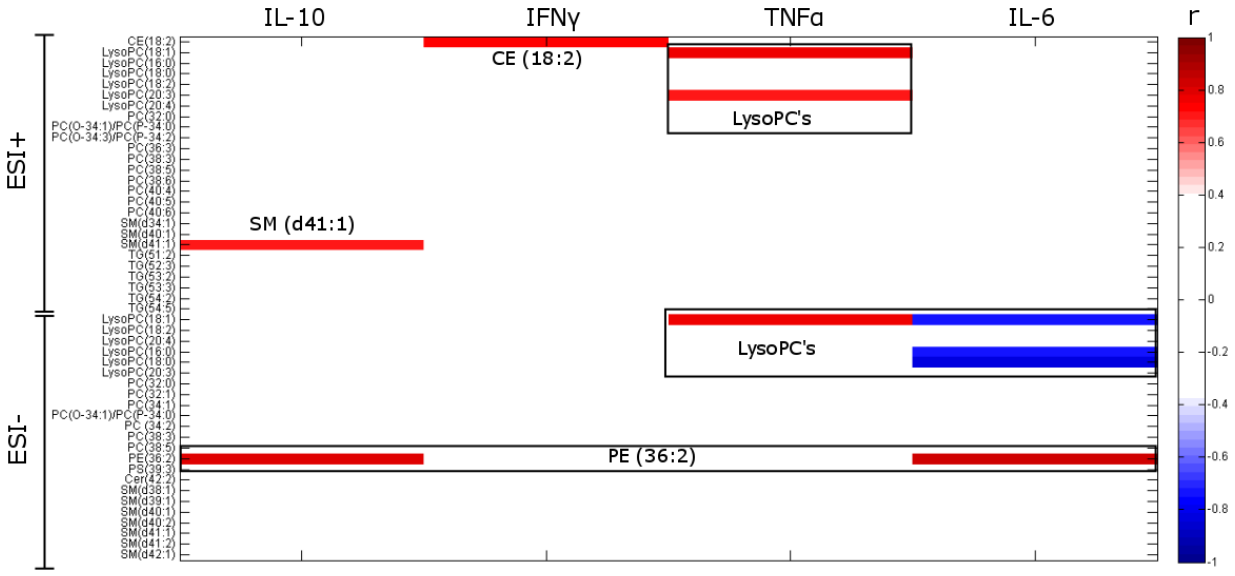


Figure A.8: Correlations between Plasma Cytokines and UPLC-MS Markers prior to FDR Correction

2D Pearson-based correlations performed between integrals of lipid UPLC-MS features and cytokine concentrations, measured in plasma of HAT patients (n=8). Positive correlations are shown in red and negative ones in blue. For all displayed associations, p<0.05 prior to Benjamini-Hochberg FDR but were no longer significant after FDR correction.

Table A.9: Accuracy of the Collection of CSF NMR Discriminatory Markers for HAT

	Using All 5 Markers (R^2Y=0.341, Q^2Y=0.308)	
	Actual HAT Patients	Actual Controls
Predicted HAT Patients	31	0
Predicted Controls	20	9
	Sensitivity	Specificity
	60.78%	100%
Standard error	6.84%	0.00%
Upper 95% CI	74.18%	100%
Lower 95% CI	43.78%	100%

Confusion matrix showing calculations for sensitivity and specificity, based on CSF sample HAT/control assignment from O-PLS-DA models, using discriminatory ^{1}H NMR metabolite integrals features as loading variables. Key and calculations as described for Table 5.3.

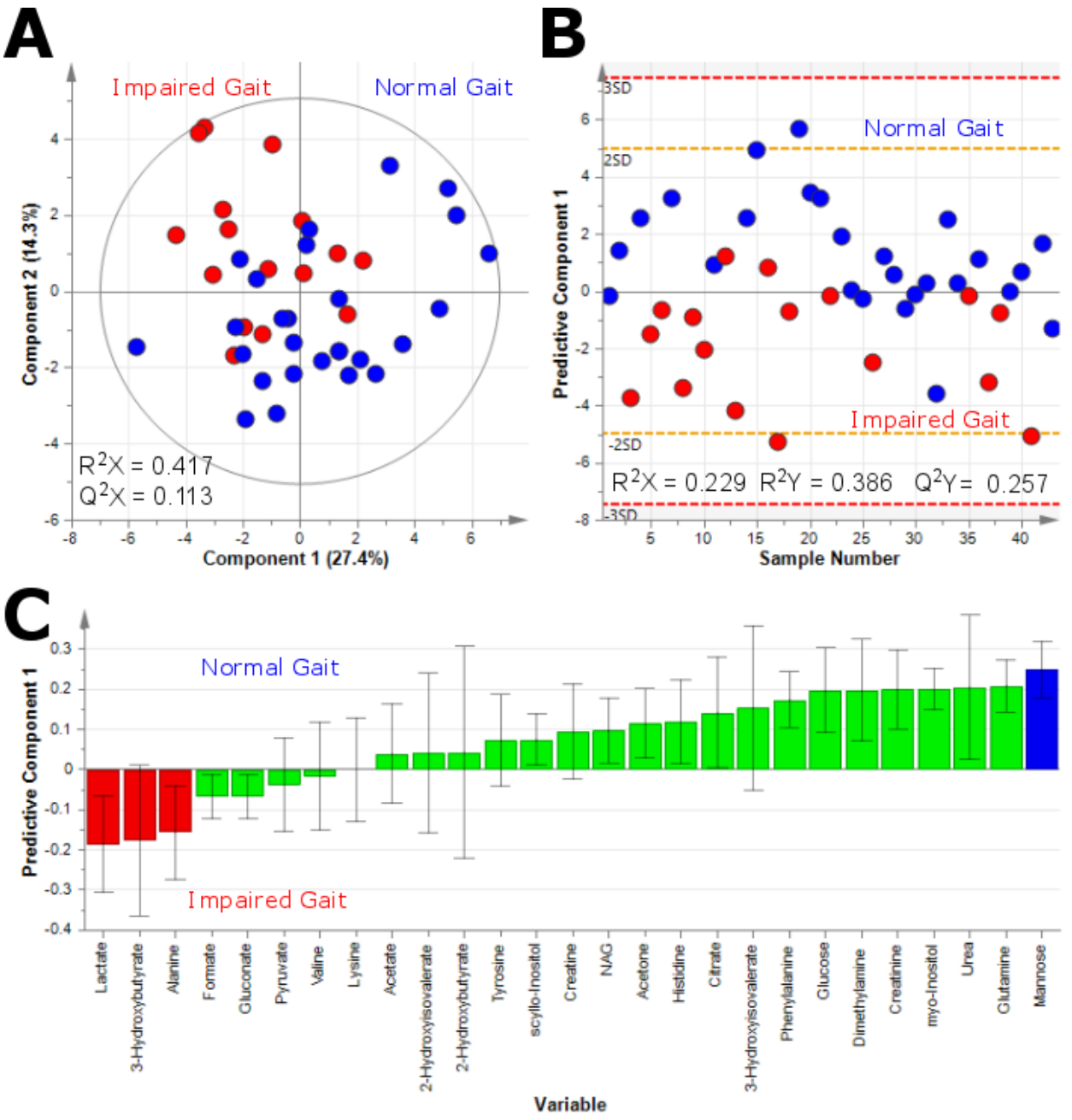


Figure A.9: Metabolic Changes in CSF Associated with Impaired Gait

(A) PCA scores plot, comparing HAT patients with impaired gait (in red, n=17) vs. those with normal gait (in blue, n=26). (B) O-PLS-DA scores plot and corresponding loadings plot (C), showing differences between the two groups. Metabolites that also tested as significantly different *via* Welch T-test and FDR correction (p<0.05) are shown in red (higher in symptomatic patients) and blue (higher in patoients with normal gait).

Publications

Journal Articles

1. Lamour SD, Gomez-Romero M, Vorkas PA, Alibu VP, Saric J, Holmes E, Sternberg JM. (2015) **Discovery of Infection Associated Metabolic Markers in Human African Trypanosomiasis.** PLOS Neglecteed Tropical Diseases, Vol: 9 (10): e0004200

    - Begins on page 246
    - Open Access article reprinted under the terms of the Creative Commons Attribution License. Copyright 2015 Lamour *et al.*

2. Lamour SD, Veselkov KA, Posma JM, Giraud E, Rogers ME, Croft S, Marchesi JR, Holmes E, Seifert K, Saric J. (2015) **Metabolic, Immune, and Gut Microbial Signals Mount a Systems Response to Leishmania major Infection.** Journal of Proteome Research, Vol: 14, Pages: 318-329

    - Begins on page 263
    - Reprinted with permission from Lamour, *et al.* (2014) Copyright 2014 American Chemical Society.

3. Lamour SD, Straschil U, Saric J, Delves MJ. (2014) **Changes in Metabolic Phenotypes of Plasmodium falciparum in vitro Cultures during Gametocyte Development.** Malaria Journal, Vol: 13

    - Includes journal article and supplementary files. Begins on page 275
    - Open Access article reprinted under the terms of the Creative Commons Attribution License 4.0. Copyright 2014 Lamour *et al.*

4. Lamour SD, Choi B-S, Keun HC, Mueller I, Saric J. (2012) **Metabolic Characterization of Leishmania major Infection in Activated and Nonactivated Macrophages.** Journal of Proteome Research, Vol: 11, Pages: 4211-4222

    - Begins on page 290
    - Reprinted under the Standard ACS AuthorChoice/Editors' Choice usage agreement. Copyright 2012 American Chemical Society.

Reports

- Lamour SD (2013) **Laboratory Clinical Diagnostics in Côte d'Ivoire. Case Studies: Methodist Hospital of Dabou & Taabo General Hospital**
- Begins on page 302

PLOS | NEGLECTED TROPICAL DISEASES

RESEARCH ARTICLE

Discovery of Infection Associated Metabolic Markers in Human African Trypanosomiasis

Sabrina D. Lamour[1], Maria Gomez-Romero[2], Panagiotis A. Vorkas[1], Vincent P. Alibu[2], Jasmina Saric[1], Elaine Holmes[1‡], Jeremy M. Sternberg[3‡*]

1 Section of Biomolecular Medicine, Division of Computational and Systems Medicine, Department of Surgery and Cancer, Imperial College London, London, United Kingdom, **2** Section of Hepatology and Gastroenterology, Department of Medicine, Imperial College London, London, United Kingdom, **3** Institute of Biological and Environmental Sciences, University Of Aberdeen, Aberdeen, United Kingdom

‡ These authors are joint senior authors on this work.
* j.sternberg@abdn.ac.uk

CrossMark
click for updates

OPEN ACCESS

Citation: Lamour SD, Gomez-Romero M, Vorkas PA, Alibu VP, Saric J, Holmes E, et al. (2015) Discovery of Infection Associated Metabolic Markers in Human African Trypanosomiasis. PLoS Negl Trop Dis 9(10): e0004200. doi:10.1371/journal.pntd.0004200

Editor: Philippe Büscher, Institute of Tropical Medicine, BELGIUM

Received: June 22, 2015

Accepted: October 7, 2015

Published: October 27, 2015

Copyright: © 2015 Lamour et al. This is an open access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability Statement: All relevant data are within the paper and its Supporting Information files.

Funding: JMS and VPA were supported by a Wellcome Trust (www.wellcome.ac.uk) project grant 082786. SDL held a MRC (www.marc.ac.uk) Doctoral Training Award G1000390. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: The authors have declared that no competing interests exist.

Abstract

Human African trypanosomiasis (HAT) remains a major neglected tropical disease in Sub-Saharan Africa. As clinical symptoms are usually non-specific, new diagnostic and prognostic markers are urgently needed to enhance the number of identified cases and optimise treatment. This is particularly important for disease caused by *Trypanosoma brucei rhodesiense*, where indirect immunodiagnostic approaches have to date been unsuccessful. We have conducted global metabolic profiling of plasma from *T.b.rhodesiense* HAT patients and endemic controls, using ^{1}H nuclear magnetic resonance (NMR) spectroscopy and ultra-performance liquid chromatography, coupled with mass spectrometry (UPLC-MS) and identified differences in the lipid, amino acid and metabolite profiles. Altogether 16 significantly disease discriminatory metabolite markers were found using NMR, and a further 37 lipid markers via UPLC-MS. These included significantly higher levels of phenylalanine, formate, creatinine, *N*-acetylated glycoprotein and triglycerides in patients relative to controls. HAT patients also displayed lower concentrations of histidine, sphingomyelins, lysophosphatidylcholines, and several polyunsaturated phosphatidylcholines. While the disease metabolite profile was partially consistent with previous data published in experimental rodent infection, we also found unique lipid and amino acid profile markers highlighting subtle but important differences between the host response to trypanosome infections between animal models and natural human infections. Our results demonstrate the potential of metabolic profiling in the identification of novel diagnostic biomarkers and the elucidation of pathogenetic mechanisms in this disease.

Author Summary

Metabolic profiling of biofluids and tissues in disease and healthy individuals is a powerful approach to discover new markers for diagnosis. We have applied these techniques to the protozoan infection human African trypanosomiasis (HAT), otherwise known as sleeping sickness. The form of HAT endemic in East Africa, caused by *Trypanosoma brucei*

rhodesiense, requires technically demanding direct microscopic diagnosis, as there are no indirect rapid diagnostic tests available. We studied the metabolite profiles in plasma from HAT patients and controls. Clear biochemical differences were discovered between control individuals and patients, including changes in the overall lipid composition and concentration of certain amino acids. These may have been caused by the inflammatory immune response to infection and the uptake of particular molecules by the parasites, although further research will be required for confirmation. We demonstrate that plasma metabolic profiles are characteristic for *T. b. rhodesiense* infection. While some of these changes are consistent with those observed in an experimental mouse infection model of HAT, many are unique to this clinical study and indicate the necessity of validating experimental animal study data in clinical disease studies. Our results also reveal biochemical changes in patients that will help us understand the development of disease.

Introduction

Human African Trypanosomiasis (HAT) is caused by infection with either of two subspecies of *Trypanosoma brucei. Trypanosoma brucei (T. b.) gambiense* causes chronic disease (that can last months or years) and is endemic in Western Africa while *T. b. rhodiesiense* causes a more acute illness in Eastern and Southern Africa, which is typically fatal within less than a year of infection if untreated [1,2]. Diagnosis is a critical challenge for treatment and control of this disease as patients display an array of non-specific inflammatory symptoms, often indistinguishable from other endemic illnesses such as malaria or enteric fever [1,3]. Infection with *T. b. gambiense* is routinely screened using the Card Agglutination Test for Trypanosomiasis (CATT) and also a new generation of lateral-flow rapid diagnostic tests are being deployed based on the host-response to commonly expressed variant surface glycoproteins (VSG) [4]. However no such serological approach has been successful for *T. b. rhodesiense* probably due to the greater diversity of VSG expression. New diagnostic techniques are urgently needed for this disease in order to meet the World Health Organization (WHO) target of elimination by 2020 [5,6].

In search for novel markers of disease, we have conducted global untargeted metabolic profiling of plasma from *rhodesiense* trypanosomiasis patients from Uganda and matched controls. Metabolic profiling (also termed metabonomics/metabolomics) can be used to characterise biochemical patterns associated with specific physiological and/or pathological states [7]. In the field of parasitic infections, this approach has shown capacity for characterising infection-induced metabolic changes in the host, within infected tissues and systemically (as measured by plasma or urine profiles). As yet, parasitic metabolic profiling studies have largely focused on *in vitro* assays [8–10] and experimental rodent models [11–15]. There have been very few examples of identifying the metabolic signature of a specific parasitic infection in humans [16], largely because human profiles are confounded by strong genetic and environmental variation, and are often superimposed upon a background of other concurrent endemic infections. In mice, infection with *T. b. brucei* (a subspecies of *T. brucei* that is not infective to humans) resulted in augmented plasma levels of lactate, acetylglycoproteins and creatine, and reductions in phosphatidylcholine and lipoproteins, as well as elevated pro-inflammatory cytokines [11,17]. Increases in plasma pro-inflammatory cytokines have also been measured in clinical samples [18,19], although metabolic similarities between the mouse model and humans remain to be investigated. In this study we present the first characterisation the metabolic effects of *T. brucei rhodesiense* infection in humans, and highlight both similarities and differences to results from published mouse model infections with *T.brucei brucei*.

Plasma samples were analysed by proton nuclear magnetic resonance (^{1}H NMR) spectroscopy and reversed-phase ultra-performance liquid chromatography, coupled to mass spectrometry (RP-UPLC-MS). ^{1}H NMR is a highly reproducible method optimal for analysing complex biological mixtures, such as plasma, with minimal sample preparation [20]. Evidence from a range of experimental models have highlighted marked changes in plasma lipids following *T. brucei ssp* infection, such as hypertriglyceridaemia [11,21–23]. Consequently, lipid profiles were separately characterised via UPLC-MS, to provide complementary information on the different lipid species altered in disease. We hypothesised that the metabolic phenotype of individuals with HAT *rhodesiense* infection would be distinct from that of uninfected individuals, yielding a range of discriminatory markers that may be of diagnostic value.

Methods

Ethics Statement

The human plasma samples in this study were collected under protocols approved by ethics committees in Uganda (UNCST) and UK (North of Scotland Research Ethics Committee), conforming to the principles of the Declaration of Helsinki. Ethical consent forms and information sheets were designed in English and translated into local languages. Informed consent was given as a signature or a thumb-print (as approved by the UNCST) after verbal explanation. For those aged under 18, affirmative consent as well as the consent of their legal guardian was obtained.

Patient Information and Sample Collection

A total of 46 HAT patients and 21 controls were recruited at Lwala Hospital, Kaberamaido District and Serere Health Centre, Serere District in Eastern Uganda, between November 2008 and March 2010. Controls were healthy individuals who were confirmed as non-infected with either trypanosomes or *Plasmodium sp*. having undergone parasitological assessment. Study sites, recruitment protocols, treatment regimens, disease progression characteristics and clinical examination methods have been published elsewhere [24]. Patients with intercurrent infections of malaria, filariasis or schistosomiasis were excluded. Plasma and cerebral spinal fluid (CSF) samples were collected from all patients prior to treatment as part of normal diagnostic and staging procedures. Staging was carried out in accordance with WHO criteria [25] defining late stage by the presence of parasites in the lumbar CSF and/or a CSF white blood cell count (WBC) $>$ 5/μl. Aliquots of 1–2 ml of plasma were immediately frozen and then maintained in liquid nitrogen until transfer to the UK. After air-freight (24 hours on dry ice) samples were maintained at -80°C until analysis. A summary of the parasitological and demographic characteristics of the patients is presented in supplementary S1 Table.

^{1}H NMR Acquisition and Processing

All sample preparation was performed in one batch in a randomised order. Prior to acquisition, plasma samples were diluted 1:1 with plasma buffer (0.142 M $NaHPO_4$, 2 mM NaN_3, 0.08% (volume/volume i.e. v/v) 3-(Trimethylsilyl)propionic-2,2,3,3-d_4 acid sodium salt (TSP) solution, 20% (v/v) D_2O; all SIGMA-Aldrich, Germany), and transferred into 5 mm NMR tubes. ^{1}H NMR data were acquired (over 3 days) on a Bruker Avance 600 MHz (14.1 T) NMR Spectrometer with BB probe head and refrigerated SampleJet autosampler (Bruker, Germany), using Topsin (3.1) software (Bruker BioSpin, Germany). Samples were acquired at 300 K (~27°C) using a standard one-dimensional pulse with water suppression program for general metabolite screening [26]; sequence as follows: 2 second (s) relaxation delay (RD)– 90° pulse–

4 μs delay—90° pulse– 100 μs mixing time– 90° pulse– 2.73 s acquisition (ACQ) of free induction decay (FID) signal. Number of dummy scans (ds) was 4, and number of acquired FID's (ns) was 64. To enhance the visualisation of low molecular weight molecules, a Carr-Purcell Meiboom-Gill (CPMG) spin-echo pulse spectrum was also acquired for each sample [26], with the following pulse sequence: 2 s RD– 90° pulse–$(\tau$– 180°–$\tau)_n$-ACQ, where spin echo time τ = 300 μs, number of loops n = 128, total spin echo time $2n\tau$ = 76.8 ms, ds = 4 and ns = 64. All FID's were multiplied by an exponential function with a 0.3 Hz line broadening prior to Fourier transformation into spectral data (spectral width of 18 kHz collected over 131,072 data points). Spectra were processed using automatic phasing, baseline correction and calibrated to the TSP peak via an in-house algorithm in MATLAB (version R2013a, Mathworks Inc.) and Topspin. The residual water signal was removed prior to automatic spectral alignment [27], and probabilistic quotient normalisation [28], using an in-house MATLAB script. Additionally, artefactual peaks arising from EDTA buffer [29] present in blood collection tubes were removed from all plasma spectra for final analysis. The regions removed were 2.53–2.62, 2.7–2.735, 3.08–3.256, 3.60–3.66 ppm which may have resulted in the loss of some endogenous metabolic information. Spectral assignments of metabolite peaks were identified using Chenomx NMR suite profiler 8.1 software and known assignments from in-house NMR databases and literature [30,31].

Plasma Lipid Profiling UPLC-MS Acquisition and Processing

Due to sample limitations, additional aliquots of only a subset of plasma samples were used for lipid profiling (n = 30, see supplementary S1 Table). As for NMR, sample preparation for UPLC-MS was performed in one day in one batch, in a randomised manner. Samples underwent protein precipitation using isopropanol, as described by Sarafian et al, 2014 [32]. Briefly, 50 μl plasma was mixed 1:3 (v/v) with ice-cold isopropanol (SIGMA-Aldrich), incubated for 10 mins at room temperature, and then overnight at -20°C. Samples were then centrifuged at 20,817 x g (14,000 rpm) at 4°C for 20 mins and 100 μl of the supernatant (subsequently referred to as test samples) were transferred to MS vials (Waters Corp., UK) for analysis and placed in the auto-sampler (kept at 8°C). Additional aliquots from each plasma-isopropanol sample were pooled to generate a quality control (QC) master sample. The QC was injected before starting the run to optimise the injection volume and to condition the column, and then injected every 8 samples for assurance of a stable run.

Lipid profiling was performed on an Acquity UPLC system (Waters Ltd. Elstree, UK) coupled to a Q-TOF Premier mass spectrometer (Waters Ltd., Manchester, UK). Mass spectra for all samples were acquired continuously, firstly in positive and then negative ion electrospray (ESI+ and ESI-) modes, over the course of 3 days. The separation conditions for reversed-phase UPLC have been previously established [33,34]. Gradient elution was performed using a CSH C_{18} (1.7 μm, 2.1 x 100 mm) column (Waters Corporation, Milford, USA) kept at 55°C. The mobile phases consisted of 0.1% (v/v) formic acid and 10 mM ammonium formate in 60:40 (v/v) acetonitrile (ACN)/HPLC-grade water (mobile phase A) and 0.1% formic acid (v/v) and 10 mM ammonium formate in 90:10 (v/v) IPA/ACN (mobile phase B) at a flow rate of 0.4 ml/min (all solvents purchased from SIGMA-Aldrich except water which was from Fisher, Germany). The sample injection volume was 5 μl and 10 μl for ESI+ and ESI-, respectively.

The MS parameters were set as follows: capillary voltage, 3 kV (ESI+) and 2.5 kV (ESI-); sample cone voltage, 30 V (ESI+) and 25 V (ESI-); source temperature, 120°C; desolvation temperature, 400°C; desolvation gas flow, 800 L/h and cone gas flow, 25 L/h. For mass accuracy, a 0.2 ng/μl leucine encephalin (mass/charge ratio (m/z) 556.2771 in ESI+ and m/z 554.2615 in ESI–) solution at 20 μl/min was used as the lock mass. Data were collected in centroid mode

with a scan range of 50–1200 m/z, with lock mass scans collected every 30 seconds and averaged over 3 scans to perform mass correction. Additionally, data-dependent acquisition (DDA) of the QC sample was performed for structural elucidation for each ionisation mode.

Chromatograms and spectra were displayed using MassLynx (version 4.1; Waters Corp.). The MassLynx package 'Databridge' was used to convert the data files into NetCDF format. The package XCMS (R software) and an in-house developed script was used for the pre-processing of the data (peak-picking, alignment, grouping, zero filling, and median-fold change normalisation) [35]. Integrals of metabolic features with a coefficient of variation (CV) greater than 30% in the QC samples ($n = 6$) were removed in Microsoft Excel (2013, Microsoft, USA) prior to analysis.

Statistical Analysis

Metabolic data from plasma ^{1}H NMR spectra and lipid profiling UPLC-MS data were initially analysed in SIMCA software (V. 13.0, Umetrics) via Principal Component Analysis (PCA) [36], a non-supervised descriptive model used to investigate inherent similarities/differences across the data set and identify outliers. Analyses resulted in the removal of one NMR sample with an acquisition artefact thus bringing the final total to $n = 66$: 45 HAT and 21 controls (see S1 Table). PCA was additionally used to identify and remove contaminant/artefactual features, as described by Vorkas et al, 2015 [34], and to evaluate experimental reproducibility by assessing QC sample clustering for both ESI+ and ESI- runs. This was followed by Partial Least Squares Discriminant Analysis (PLS-DA), and Orthogonal PLS-DA (O-PLS-DA) [37]. Both are multivariate regression models used to assess the spectral differences that contribute towards classification of patients vs. controls, and to determine how well these differences can predict class separation. Contribution of within-class variation that is uncorrelated to the separation between classes is minimised in O-PLS-DA models, which were used for final extraction of discriminating metabolites. See supplementary protocols in supporting information (S1 File) for further details on data processing, analyses, and model validation.

For NMR data, metabolites integrals were then compared between HAT patients and controls via the two-tailed Welch's (unequal variance) T-test in Microsoft Excel, followed by Benjamini-Hochberg multiple test correction [38]. Those where $p<0.05$ was observed were labelled as statistically significant discriminatory markers. For UPLS-MS data, threshold criteria for initial selection of markers were variables with a correlation value p(corr)[1]>0.5 and covariance p[1]>0.05, as visualized on the OPLS-DA S-plot, to extract the most discriminatory and robust features (see supporting protocols). This was followed by conducting a two-tailed Welch's T-test with Benjamini-Hochberg correction, whereby the final statistical discriminatory criteria was also $p<0.05$ after multiple testing correction. Lipid species assignment of discriminatory features was based on the accurate mass-to-charge (m/z) ratio, searched against online MS database METLIN (https://metlin.scripps.edu) and previous published results generated in-house [34,39], as well as on the MS/MS information obtained in the DDA experiments of QC samples.

Integrals of these discriminatory variables were re-imported into SIMCA as new loadings for O-PLS-DA models, to evaluate the sensitivity and specificity of the combination of these metabolites for biomarker potential. O-PLS-DA models was performed twice, separately for NMR and UPLC-MS data. Firstly, models were generated using integrals of all significantly discriminatory markers, and secondly using only integrals of the top 5 discriminating markers, defined as those which displayed the greatest % difference between patients and controls. Calculations are further explained in supplementary protocols (see Supporting Information S1 File).

To explore links between potentially complementary metabolic datasets, NMR metabolite integrals were correlated with corresponding integral of discriminatory lipid profiling features from matching patient samples ($n = 16$), via Pearson correlation in MATLAB, followed by Benjamini-Hochberg correction. Similarly Pearson -based correlation analyses (with multiple test correction) was performed between metabolic (NMR and/or MS) data sets. Relationships between patient demographic data and metabolic datasets were assessed using Spearman correlation analysis, Mann-Whitney tests, and mixed linear models as appropriate, in JMP (v.10.0, SAS, USA).

Results

^{1}H NMR Spectroscopy Shows Marked Differences in Plasma between Human African Trypanosomiasis Patients and Controls

Inherent differences between *T. b. rhodesiense* infected HAT patients and uninfected controls dominated overall dataset variability, demonstrated by the partial separation of these two populations along the first component in PCA (Fig 1A), indicative that the HAT infection produced a strong metabolic response. No significant differences between metabolic profiles based on demographic factors (gender, age, or diagnostic stage) were found. Metabolic disparities between patients and controls were further examined using O-PLS-DA (a supervised modelling method that maximises differences between classes whilst minimising within-class variation), which showed complete segregation between patients and controls, also along the first component (Fig 1B). The model had a Q^2Y (model predictive capability parameter) value of 0.708, indicating that the model was robust. Further validation through permutation testing found that the likelihood of achieving the same discriminatory result by chance was zero ($p = 0$ for both R^2Y and Q^2Y). It is noteworthy that although the first two components of the PCA model only cumulatively account for 18.1% of the total variance (R^2) in the data set, indicating variability in sample composition, the O-PLS-DA model indicates that the metabolic signature is robust.

Examination of the corresponding O-PLS-DA loadings revealed that the lipid peaks accounted for the majority of the discrimination. In total, 34 endogenous metabolites were identified from NMR spectra though only 31 of these displayed distinct peaks that were used for integral calculation (signals from isoleucine, leucine and lysine were superimposed by overlapping lipid peaks), and compared between patients and controls (shown in S2 Table). Those metabolites that were significantly higher in disease are shown in Fig 2A, whilst metabolites significantly lower in disease are shown in Fig 2B. The majority of lipid moieties detected in plasma were significantly reduced in patients, including those containing unsaturated bonds (CH = CH and C = CCH_2C = C), lysyl group in plasma albumin, and cholesterol (Fig 2B). In contrast, lipid glyceryl and ketene (CH_2CO) groups were significantly higher in patients than in controls (Fig 2A). Several plasma amino acids were also significantly discriminatory between HAT and controls, including histidine and alanine, which were found in lower concentrations in patients, and phenylalanine, which was augmented. Other significant changes included an increase of creatinine, *N*-acetyl glycoprotein, formate and *myo*-inositol in HAT patients whilst citrate was higher in controls. No associations between discriminatory metabolites and patient demographic data (age, gender or diagnostic stage) were found though the sample size in this study limited the power of this analysis (this was particularly the case for staging as our cohort had relatively few early cases—see S1 Table).

Additional O-PLS-DA models were built, firstly using integrals of all 16 discriminatory metabolites, followed by models using only the top 5 metabolites. The latter were defined as those which displayed the biggest measured percentage difference between HAT and controls,

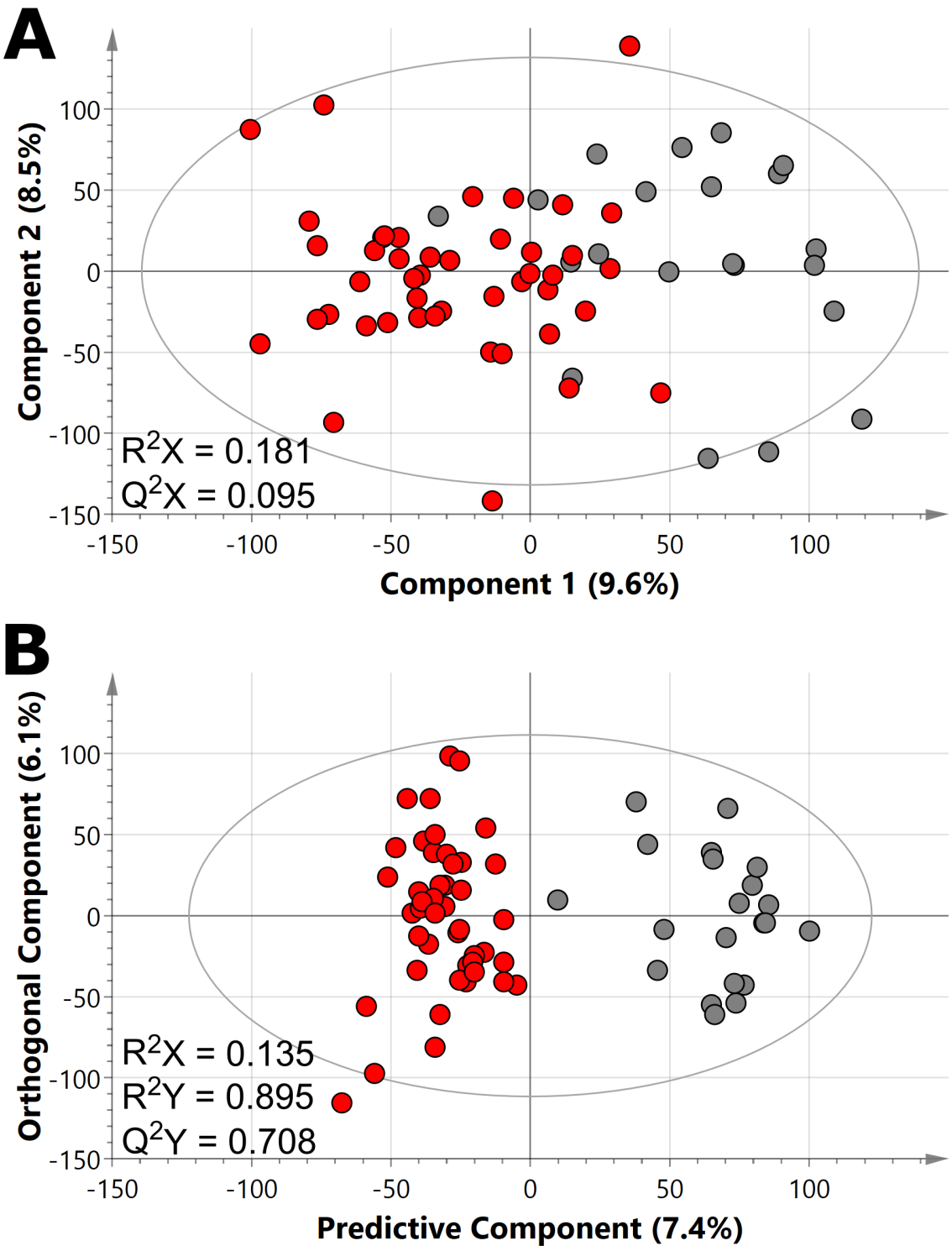


Fig 1. Plasma in HAT patients display different ^{1}H NMR metabolic profiles compared with controls. PCA model (A) and O-PLS-DA model (B) score plots of plasma ^{1}H NMR spectra across HAT patients and controls. Each circle represents a spectra from one sample, whereby patients are presented in red (n = 45) and controls in dark grey (n = 21). Abbreviations: R^2X, model fit parameter for variation in spectral data; R^2Y, model fit parameter for variation in classifier data (for O-PLS-DA); Q^2, model predictive parameter for spectral data in PCA (Q^2X) and for classifier data in O-PLS-DA (Q^2Y). Individual component contribution of R^2X are shown on the axes as percentage.

doi:10.1371/journal.pntd.0004200.g001

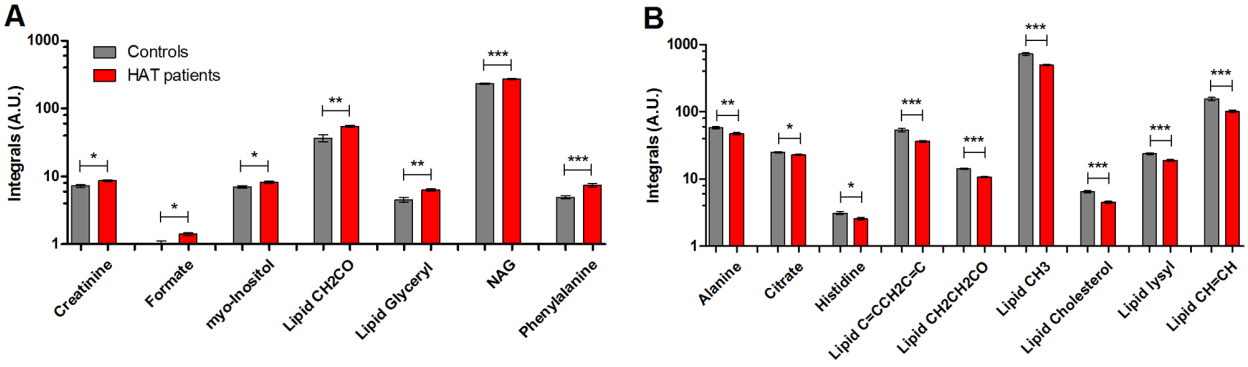


Fig 2. Differences in plasma metabolites between HAT patients and controls detected by NMR. Bar-charts show relative levels of plasma metabolites that were significantly altered between patients (shown in red, $n = 45$) and controls (shown in grey, $n = 21$), as measured by ^{1}H NMR spectroscopy. (A) Metabolites higher in patients than controls. (B). Metabolites lower in patients than controls. Bars represent group mean average with standard error of the mean as error bars. Significance measured via Welch T-test with multiple test correction, shown as asterisks; * $p<0.05$, ** $p<0.01$, *** $p<0.001$. A.U., arbitrary units; NAG, *N*-acetyl glycoproteins.

doi:10.1371/journal.pntd.0004200.g002

representing metabolites of most practical diagnostic potential. These consisted of phenylalanine, formate and lipid moieties CH_2CO, CH = CH, and glyceryl. In both O-PLS-DA models, results provided identical specificity scores for classifying patients and controls samples, namely 95.24% (with 95% confidence intervals [CI] of 86.1–100%). Model sensitivity was 88.89% (79.7–98.1% CI) using all NMR discriminatory markers, and 84.44% (73.9–95.0% CI) using just the top 5 metabolites (see S3 Table for confusion matrices).

Lipid Profiling UPLC-MS Reveals Compositional Differences in Plasma Lipid Species between HAT and Controls

Lipid plasma profiles were further examined using reversed-phase UPLC-MS. In total, 8771 metabolic features were detected in positive ionisation mode (ESI+), whilst 2543 features were observed in negative mode (ESI-). Following the exclusion of peaks with CV > 30% across the QC's, as well as the removal of contaminant peaks present in either QC's or blank controls, 6952 features in ESI+ and 1714 features in ESI- were retained for final analyses. PCA models showed tight clustering of QC samples compared with the remaining test samples (supplementary information S1 Fig), assuring the robustness of the MS experimental runs.

As noted with NMR, lipid profiling results showed an inherent separation between the two classes already in the first two components of PCA, observed in both ionisation modes (see S1 Fig). These differences were confirmed in the corresponding O-PLS-DA models (S1 Fig). Permutation testing confirmed model robustness (ESI+, $p = 0.004$ for R^2Y and $p = 0$ for Q^2Y and for ESI-, $p = 0$ for both R^2Y and Q^2Y). Metabolic features responsible for class discrimination (i.e. HAT patients vs. controls) were visualised in S-plots in SIMCA, with the cut-off values of p(corr)[1]>0.5 and covariance p[1]>0.05 were used as selection criteria (see Methods), shown in Fig 3A and 3B. Integrals of selected discriminatory features highlighted in the S-plots are shown in Fig 3C and 3D, demonstrating significant decreases in lysophosphatidylcholines 16:0 and 18:0, as well as increases in phosphatidylcholine 32:0 and 34:1.

Overall, lipid profiling analysis identified 37 unique features that were statistically different between HAT and controls, 12 of which were found to be significantly different in both of these modes (Table 1). Features included six lysophosphatidylcholines (LysoPCs 16:0, 18:0, 18:1, 18:2, 20:3 and 20:4), four phosphatidylcholines (PCs 32:0, O-34:1, 38:3 and 38:5) and two sphingomyelin species (SMs d40:1 and d41:1). Further information on UPLC-MS assignments

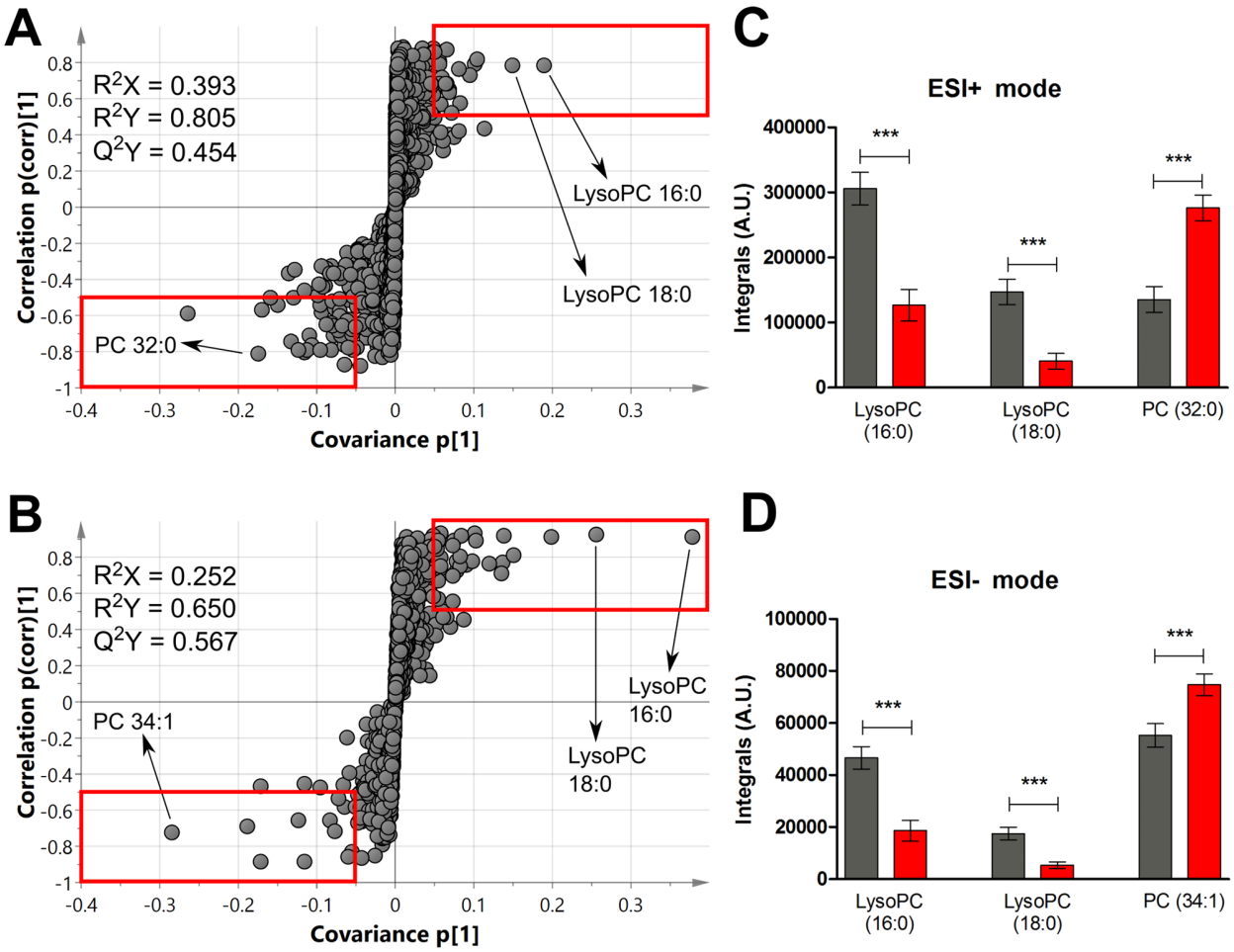


Fig 3. Plasma in HAT patients display different lipid profiles compared with controls. S-plots of O-PLS-DA model for plasma lipid profiling UPLC-MS features detected in ESI+ mode (A) and ESI- mode (B), whereby each circle represents one feature with a unique combination of m/z and retention time values. Discriminatory features selected surpassed p[1] and p(corr)[1] threshold criteria, highlighted in red boxes (see Methods). See S1 Fig for corresponding scores plots. Integrals of features highlighted in the S-plots are shown as bar-chart showing mean averages ± standard error of the mean, for both ESI+ (C) and ESI- (D). Patients levels are shown in red (n = 16) and controls in dark grey (n = 14). Significant differences, as measured via Welch T-test with multiple test correction, are labelled with asterisks, where *** p<0.001. Abbreviations: A.U., arbitrary units; LysoPC, lysophosphatidylcholine; PC, phosphatidylcholine.

doi:10.1371/journal.pntd.0004200.g003

of features can be found in S4 Table. The majority of lipid classes studied were glycerophospholipids (LysoPCs and PCs), followed by sphingolipids (mainly sphingomyelins) and triglycerides (see Table 1 and S4 Table). Free fatty acids were detected using our MS methodology but did not significantly differ between patients and controls.

In total, six significantly altered lysoPCs measured in this study were found to be lower in patients than controls. Significantly lower intensities were also observed for eight out of nine sphingolipids, including seven sphingomyelins (SM) as well as nervonic ceramide (Cer d42:2). Only SM d34:1, the smallest of the analysed SM species, was found to be higher in patients than controls. Additionally, patients had augmented levels of six triglycerides (TG) species and depleted levels of cholesterol ester (CE) 18:2 in patients, in line with NMR findings where we detected higher glyceryl moieties and lower cholesterol in HAT.

Table 1. Discriminatory lipid species observed in plasma via UPLC-MS.

	Lipid Species	ESI+		ESI-	
		P-value	% Difference	P-value	% Difference
Increased in HAT	Cer (d42:2)	-	-	2.51E-03	54.30%
	PC (32:0)	2.49E-05	104.20%	2.58E-03	111.70%
	PC (32:1)	-	-	1.29E-04	62.50%
	PC (34:1)	-	-	8.21E-03	35.00%
	PC (34:2)	-	-	1.29E-04	40.20%
	PC(O-34:1)/ PC(P-34:0)	3.93E-04	86.20%	2.00E-03	84.30%
	PE (36:2)	-	-	2.51E-03	54.40%
	PS (39:3)	-	-	4.42E-03	37.50%
	SM (d34:1)	1.75E-02	96.10%	-	-
	TG (51:2)	6.52E-05	98.00%	-	-
	TG (52:3)	4.40E-03	45.40%	-	-
	TG (53:2)	6.73E-04	86.10%	-	-
	TG (53:3)	4.02E-04	77.50%	-	-
	TG (54:2)	3.74E-04	70.30%	-	-
	TG (54:5)	2.20E-03	68.20%	-	-
Decreased in HAT	CE (18:2)	1.12E-04	-53.10%	-	-
	LysoPC (16:0)	1.93E-05	-58.60%	5.42E-04	-60.10%
	LysoPC (18:0)	1.11E-04	-72.60%	4.42E-03	-69.50%
	LysoPC (18:1)	1.98E-05	-66.50%	1.00E-04	-62.50%
	LysoPC (18:2)	3.28E-04	-80.50%	3.18E-04	-78.20%
	LysoPC (20:3)	5.97E-05	-75.00%	7.00E-04	-77.90%
	LysoPC (20:4)	4.36E-06	-75.60%	2.62E-04	-73.10%
	PC (40:5)	1.13E-02	-39.20%	-	-
	PC (36:3)	6.56E-03	-33.90%	-	-
	PC (38:3)	1.93E-03	-43.00%	2.51E-03	-30.70%
	PC (38:5)	1.71E-02	-40.90%	1.29E-04	-28.80%
	PC (38:6)	1.63E-02	-33.30%	-	-
	PC (40:4)	3.25E-03	-47.90%	-	-
	PC (40:6)	2.94E-02	-32.70%	-	-
	PC(O-34:3)/PC(P-34:2)	8.43E-04	-52.80%	-	-
	SM (d38:1)	-	-	3.46E-03	-37.60%
	SM (d39:1)	-	-	4.42E-03	-52.10%
	SM (d40:1)	1.01E-03	-43.90%	3.46E-03	-48.60%
	SM (d40:2)	-	-	2.58E-03	-60.20%
	SM (d41:1)	1.06E-03	-58.60%	2.00E-03	-48.20%
	SM (d41:2)	-	-	4.42E-03	-55.30%
	SM (d42:1)	-	-	3.92E-04	-52.50%

Molecular species that not detected in a particular electro-spray ionisation (ESI) mode were omitted from the table. For more description of detected lipid species (including retention times and mass/charge ratios) see supplementary S2 Table.

Abbreviations: -, no discriminatory marker detected that ionisation mode; Cer, ceramide; CE, cholesterol ester; LysoPC, lysophosphatidylcholine; LysoPE, lysophosphatidylethanolamine; NS, not statistically significant; PC, phosphatidylcholine; PE, phosphatidylehtanolamine; PS, phosphatidylserine; RT, retention time; SM, sphingomyelin.

doi:10.1371/journal.pntd.0004200.t001

In general, differences in lipid species concentrations between patients and controls were similar across the lipid subclasses, with the exception of the phosphatidylcholines. In contrast, our UPLC-MS results indicate that PCs that contained a high number of saturated bonds were significantly decreased in patients (e.g. PCs 38:5, 38:6, 40:5), whilst those with two or fewer unsaturated bonds were increased (e.g. PCs 32:0, 32:1, and 34:1, see Table 1). These results mirror the NMR findings, where lipid moieties with unsaturated bonds (CH = CH and C = CCH2C = C) were significantly lower in patients than controls (Fig 2).

Similar to the NMR data, additional O-PLS-DA models were generated using the integrals all 37 markers from both ionisation modes, followed by further models using only the top 5 features with the greatest % difference (as shown in Table 1). As most of the largest changes were observed in the positive mode, we selected the feature from this ionisation mode only. These included PC(34:3), PC(32:0), TG(51:2), PC(O-34:1) and SM(d34:1). Both O-PLS-DA models with either all markers or top 5 gave identifical sensitivity and specificity scores, namely 93.75% (81.9–100% CI) and 85.71% respectively (67.4–100% CI; see S5 Table for confusion matrices).

Potential links between UPLC-MS and NMR data was explored by correlating lipid metabolic features with NMR metabolite levels measured via NMR. Whilst multiple trends were observed, including links between lipid MS species and several NMR lipid moieties, 3-hydroxybutyrate (a ketone body generated from fatty acid break-down), and energy associated metabolites, e.g. creatinine and lactate), none of these associations remained statistically significant after Benjamini-Hochberg multiple test correction was performed, due to lack of statistical power. No significant links were observed between MS data and age and/or gender data (analyses could not be performed for diagnostic staging due to limited numbers of early stage within this subset of data—see S2 Table).

Discussion

The HAT infection was characterised by strong metabolic differences in plasma profiles, based on both NMR and UPLC-MS, which shows the potential of this approach in the development of novel diagnostic marker sets. The discriminatory profile was largely driven by variation in lipids and amino acids. Phenylalanine was the single most discriminatory metabolite as detected by NMR, with the largest percentage difference between HAT and controls in concentration. Raised levels of phenylalanine have been frequently linked with infection and inflammation [40]. Phenylalaninemia has previously been reported in patients with enteric fever [41], and with malaria [42] Evidence suggests that this phenomenon may be related to neurological effects seen in cerebral malaria [42], since high levels of phenylalanine are known to be neurotoxic [43]. Other differences in plasma amino acid concentrations included significant decreases in histidine and alanine in HAT. Histidine concentrations have been shown to be negatively correlated with inflammatory disease markers in humans [44–46] and this would be consistent with clinical evidence of systemic inflammatory responses in HAT [3].

Creatinine concentration was also found to be significantly higher in HAT, potentially a sign of renal dysfunction which is one of the clinical features observed in sleeping sickness [47,48]. Additionally, concentrations of *N*-acetyl groups of plasma glycoproteins were also higher in patients. Most of these have been shown to belong to α1 acid glycoprotein [49,50], another inflammatory marker known to be associated with the acute-phase response. This result is similar to what has been found with experimental murine infection of *T. b. brucei* where plasma acetylated glycoproteins (dominated by *O*-acetyls) were found to be augmented [11].

Some of the most pronounced changes observed in the NMR spectra were the clear disparities in the plasma lipid composition between the two groups. In addition to their roles as energy reserves and major constituents of membranes, lipids also represent highly biologically

active metabolites, with involvement (or function) in signalling and a range of inflammatory processes, e.g. prostaglandins and leukotrienes synthesis [51]. Our results showed that the majority of the concentrations of discriminatory lipids were significantly reduced in patients. These primarily included lipids containing unsaturated bonds (CH = CH and $C = CCH_2C = C$), CH_3 and CH_2CH_2CO moieties, and albumin lysyl groups, and high-density lipoprotein (HDL) cholesterol. Dyslipidaemia has been observed in a range of infections, measured primarily using lipid-specific biochemical serum assays, as opposed to comprehensive metabolic profiling, whereby significantly reduced HDL and total cholesterol have been reported in malaria [52], enteric fever [53] and others [54,55]. Decreased lipoprotein CH_3 moieties were also observed in experimental murine *T. b. brucei* infection [11], confirming the need for deeper assessment into the variety of lipid species, to help dissect their different roles during infection.

Further investigation of the lipid profile with UPLC-MS indicated that HAT patients displayed significantly lower levels of polyunsaturated phosphatidylcholines (e.g. PC 40:6, 40:5, 38:5 and 38:3), although PC molecules with 2 or fewer unsaturated bonds were in contrast increased. This profile presents a more complex picture than was obtained in experimental mouse infections with *T. b. brucei*, where an overall decrease in plasma PC concentrations was measured by NMR [11].

A range of lysophophatidylcholines (e.g. LysoPC 16:0, 18:0, 18:2) intensities was also significantly lower in patients. Despite the fact that *Trypanosoma brucei ssp* have the capability to produce many of their lipids *de novo*, they are known to scavenge several lipid precursor molecules from the host and have been shown to require lipid uptake for optimal growth [56,57]. More specifically, these include fatty acids and lyso-phospholipids, which are assembled to make phospholipids, the most dominant lipid component of *T. brucei ssp* [58]. As fatty acyl chain composition in *T. b. rhodesiense* bloodstream forms have a high proportion of long-chain polyunsaturated fatty acids [59,60], it could be hypothesised that the lower plasma concentrations of polyunsaturated lipids in HAT patients may be due to increased uptake by the parasites.

Numerous sphingomyelin species were also found in lower concentrations in infection (e.g. SM d40:1, d41:1 and d42:1), accompanied by a significant increase of ceramide. The sphingolipid turnover is regulated by the balance between enzymatic degradation and synthesis, with ceramide being a major break-down product. Oxidative stress, inflammatory cytokines and infection can trigger sphingomyelinase enzymes to increase ceramide generation. Higher levels of ceramide, in turn, have been linked to increased cell autophagy and apoptosis (the latter particularly in endothelial cells leading to increased vascular permeability), additional pro-inflammatory cytokine and chemokine synthesis, and other metabolic disorders (e.g. insulin resistance and obesity) [61]. Thus, the changes in sphingolipids observed in HAT further highlights the widespread inflammatory effects on the host, associated with *T. b. rhodesiense* infection.

Another characteristic of HAT patients was hypertriglyceridaemia. Not only is this result is in agreement with a range of rabbit and non-human primate models of *T. brucei spp* infections [21,22], but is also consistent with marked increases in serum triglycerides in *T. b. gambiense* HAT [62]. However, the disease discriminatory potential of hypertriglyceridaemia (often combined with increases in circulating free fatty acids) will require further investigation as it has been shown to be a common feature across numerous infections in humans [52–54].

In summary, the presented work is the first to apply a comprehensive metabolic profiling approach to the investigation of *Trypanosoma brucei* infection in humans and highlights the potential of metabonomic technology for developing a diagnostic platform for HAT. The models generated based on both NMR and UPLC-MS were robust, thus the utility of this

technology in an exploratory capacity is evident. Similarities but importantly significant differences in discriminatory NMR metabolites were identified to previously published metabolic profiles in experimental rodent *T. brucei spp*. infections. We identified a range of discriminatory plasma metabolites in clinical HAT that collectively have diagnostic disease marker potential (up to 93.75% sensitivity and 95.24% specificity). While some of these markers have been associated with other infections or inflammatory processes, the effects of specific infection on a combined panel such as in this study has not previously been reported and may prove diagnostically effective in differentiating HAT from other endemics disease. However, it is obvious that neither NMR spectroscopy nor UPLC-MS are platforms that can be used at point-of-care in developing countries. Therefore the functionality of these platforms lies in characterisation of the metabolic consequences of a parasitic infection, with subsequent development of simpler clinical assays such as a biochemical dipstick based on the strongest differentiating biomarkers. Of particular interest from our results were differences in lipids, observed in both NMR and UPLC-MS, that dominated the variation in metabolic profiles between HAT and controls. Further investigation of inflammatory lipid mediators such prostaglandins and leukotrienes would also be of particular interest, to better understand pathology and provide key insights into links between lipid metabolism and immune responses in the host. Detailed lipid sub-species characterisation with the use of lipid reference standards and MS/MS are required to further isolate novel diagnostic candidates to ultimately aid in the detection and management of *T. b. rhodiesiense* HAT, moving closer towards the WHO goal of disease eradication.

Supporting Information

S1 Fig. Differences observed in plasma lipid-MS profiles between HAT patients and controls. Multivariate model scores plots based on lipid MS features of plasma samples. PCA models based on first two component of all plasma test samples (as light grey circles, $n = 30$) and QC's (as blue circles, $n = 6$) in positive ionisation (ESI+) mode (A) and negative (ESI-) mode (B). PCA models of the first two components of plasma test samples only, comparing HAT patients (red circles, $n = 16$) vs. controls (dark grey circles, $n = 14$), in ESI+ (C) and ESI- (D). O-PLS-DA models of test samples comparing HAT vs. controls (labelling and group sizes as described for PCA), in ESI+ (E) and ESI- (F). O-PLS-DA based on both predictive and orthogonal components for ESI+ but only on the first predictive component for ESI- (the orthogonal component for this model reduced the predictive ability of model and was thus omitted). Abbreviations: R^2X, model fit parameter for variation in spectral data; R^2Y, model fit parameter for variation in classifier data (for O-PLS-DA models); Q^2, model predictive parameter (for spectral data Q^2X in PCA/for classifier data Q^2Y in O-PLS-DA), SD, standard deviation.
(TIF)

S1 Table. Patient and control cohort demographic information.
(PDF)

S2 Table. Plasma metabolite levels measured via ^{1}H NMR spectroscopy.
(PDF)

S3 Table. Confusion matrices of HAT vs. control classification ability of NMR discriminatory markers.
(PDF)

S4 Table. List of Discriminatory Lipid Species detected via UPLC-MS in both ionisation modes.
(PDF)

S5 Table. Confusion matrices of HAT vs. control classification ability of NMR discriminatory markers.
(PDF)

S1 File. Supplementary protocols for statistical analyses.
(DOCX)

Author Contributions

Conceived and designed the experiments: JMS SDL MGR JS EH. Performed the experiments: SDL MGR. Analyzed the data: SDL PAV MGR JMS. Contributed reagents/materials/analysis tools: JMS VPA. Wrote the paper: SDL JMS EH MGR JS PAV. Contributed to data processing: SDL PAV.

References

1. Chappuis F, Loutan L, Simarro P, Lejon V, Buscher P. (2005) Options for field diagnosis of human african trypanosomiasis. Clin Microbiol Rev 18: 133–146. PMID: 15653823
2. Fèvre E,M., Wissmann Bv, Welburn SC, Lutumba P. (2008) The burden of human african trypanosomiasis. PLoS Neglected Tropical Diseases 2: e333. doi: 10.1371/journal.pntd.0000333 PMID: 19104653
3. Sternberg JM. (2004) Human african trypanosomiasis: Clinical presentation and immune response. Parasite Immunol 26: 469–476. PMID: 15771682
4. Sternberg JM, Gierliński M, Biéler S, Ferguson MAJ, Ndung'u JM. (2014) Evaluation of the diagnostic accuracy of prototype rapid tests for human african trypanosomiasis. PLoS Negl Trop Dis 8: e3373. doi: 10.1371/journal.pntd.0003373 PMID: 25521120
5. World Health Organization. (2012) Accelerating work to overcome the global impact of neglected tropical diseases: A roadmap for implementation.
6. Simarro PP, Franco JR, Diarra A, Ruiz Postigo JA, Jannin J. (2013) Diversity of human african trypanosomiasis epidemiological settings requires fine-tuning control strategies to facilitate disease elimination. Res Rep Trop Med 4: 1–6.
7. Lindon JC, Nicholson JK, Holmes E, Antti H, Bollard ME, et al. (2003) Contemporary issues in toxicology the role of metabonomics in toxicology and its evaluation by the COMET project. Toxicol Appl Pharmacol 187: 137–146. PMID: 12662897
8. Teng R, Junankar PR, Bubb WA, Rae C, Mercier P, et al. (2009) Metabolite profiling of the intraerythrocytic malaria parasite plasmodium falciparum by 1H NMR spectroscopy. NMR Biomed 22: 292–302. doi: 10.1002/nbm.1323 PMID: 19021153
9. Lamour SD, Choi B, Keun H, Mueller I, Saric J. (2012) Metabolic characterization of *Leishmania major* infection in activated and non-activated macrophages. J Proteome Res.
10. Vincent IM, Creek DJ, Burgess K, Woods DJ, Burchmore RJ, et al. (2012) Untargeted metabolomics reveals a lack of synergy between nifurtimox and eflornithine against trypanosoma brucei. PLoS Negl Trop Dis 6: e1618. doi: 10.1371/journal.pntd.0001618 PMID: 22563508
11. Wang Y, Utzinger J, Saric J, Li JV, Burckhardt J, et al. (2008) Global metabolic responses of mice to trypanosoma brucei brucei infection. Proc Natl Acad Sci U S A 105: 6127–6132. doi: 10.1073/pnas.0801777105 PMID: 18413599
12. Ghosh S, Sengupta A, Sharma S, Sonawat HM. (2013) Metabolic perturbations of kidney and spleen in murine cerebral malaria: ^{1}H NMR-based metabolomic study. PLoS ONE 8: e73113. doi: 10.1371/journal.pone.0073113 PMID: 24039868
13. Girones N, Carbajosa S, Guerrero NA, Poveda C, Chillon-Marinas C, et al. (2014) Global metabolomic profiling of acute myocarditis caused by trypanosoma cruzi infection. PLoS Negl Trop Dis 8: e3337. doi: 10.1371/journal.pntd.0003337 PMID: 25412247
14. Tritten L, Keiser J, Karwa T, Utzinger J, Holmes E, et al. (2014) Comparing systemic metabolic responses in mice to single or dual infection with plasmodium berghei and heligmosomoides bakeri. Mol BioSyst 10: 2358–2367. doi: 10.1039/c4mb00097h PMID: 24960299
15. Lamour SD, Veselkov KA, Posma JM, Giraud E, Rogers ME, et al. (2015) Metabolic, immune, and gut microbial signals mount a systems response to leishmania major infection. J Proteome Res 14: 318–329. doi: 10.1021/pr5008202 PMID: 25369177

16. Balog CIA, Meissner A, Göraler S, Bladergroen MR, Vennervald BJ, et al. (2011) Metabonomic investigation of human schistosoma mansoni infection. Mol BioSyst: 1742–206X.

17. Saric J, Li JV, Swann JR, Utzinger J, Calvert G, et al. (2010) Integrated cytokine and metabolic analysis of pathological responses to parasite exposure in rodents. J Proteome Res 9: 2255–2264. doi: 10.1021/pr901019z PMID: 20092362

18. Lejon V, Lardon J, Kenis G, Pinoges L, Legros D, et al. (2002) Interleukin (IL)-6, IL-8 and IL-10 in serum and CSF of trypanosoma brucei gambiense sleeping sickness patients before and after treatment. Transactions of The Royal Society of Tropical Medicine and Hygiene 96: 329–333. PMID: 12174791

19. MacLean L, Chisi JE, Odiit M, Gibson WC, Ferris V, et al. (2004) Severity of human african trypanosomiasis in east africa is associated with geographic location, parasite genotype, and host inflammatory cytokine response profile. Infection and Immunity 72: 7040–7044. PMID: 15557627

20. Dumas M, Maibaum EC, Teague C, Ueshima H, Zhou B, et al. (2006) Assessment of analytical reproducibility of 1H NMR spectroscopy based metabonomics for large-scale epidemiological research: The INTERMAP study. Anal Chem 78: 2199–2208. PMID: 16579598

21. Rouzer CA, Cerami A. (1980) Hypertriglyceridemia associated with trypanosoma brucei brucei infection in rabbits: Role of defective triglyceride removal. Mol Biochem Parasitol 2: 31–38. PMID: 7464858

22. Nakamura Y. (1998) Alterations of serum lipid, lipoprotein and inflammatory cytokine profiles of rabbits infected with trypanosoma brucei brucei. Vet Parasitol 80: 117–125. PMID: 9870364

23. Gaithuma AK, Karanja SM, Ngotho M, Maathai RG, Kagira JM, et al. (2012) Lipid metabolism and other metabolic changes in vervet monkeys experimentally infected with trypanosoma brucei rhodesiense. J Med Primatol 41: 75–81. doi: 10.1111/j.1600-0684.2011.00523.x PMID: 22070162

24. MacLean LM, Odiit M, Chisi JE, Kennedy PGE, Sternberg JM. (2010) Focus-specific clinical profiles in human african trypanosomiasis caused by *Trypanosoma brucei rhodesiense*. PLoS Negl Trop Dis 4: e906. doi: 10.1371/journal.pntd.0000906 PMID: 21151878

25. World Health Organization. (1998) Control and surveillance of african trypanosomiasis: Report of a WHO expert committee. WHO Technical Report Series.

26. Beckonert O, Keun HC, Ebbels TM, Bundy J, Holmes E, et al. (2007) Metabolic profiling, metabolomic and metabonomic procedures for NMR spectroscopy of urine, plasma, serum and tissue extracts. Nat Protoc 2: 2692–2703. PMID: 18007604

27. Veselkov KA, Lindon JC, Ebbels TM, Crockford D, Volynkin VV, et al. (2009) Recursive segment-wise peak alignment of biological (1)h NMR spectra for improved metabolic biomarker recovery. Anal Chem 81: 56–66. doi: 10.1021/ac8011544 PMID: 19049366

28. Dieterle F, Ross A, Schlotterbeck G, Senn H. (2006) Probabilistic quotient normalization as robust method to account for dilution of complex biological mixtures. application in 1H NMR metabonomics. Anal Chem 78: 4281–4290. PMID: 16808434

29. Barton RH, Waterman D, Bonner FW, Holmes E, Clarke R, et al. (2010) The influence of EDTA and citrate anticoagulant addition to human plasma on information recovery from NMR-based metabolic profiling studies. Mol Biosyst 6: 215–224. doi: 10.1039/b907021d PMID: 20024083

30. Nicholson JK, Foxall PJD, Spraul M, Farrant RD, Lindon JC. (1995) 750 MHz 1H and 1H-13C NMR spectroscopy of human blood plasma. Anal Chem 67: 793–811. PMID: 7762816

31. Liu M, Tang H, Nicholson JK, Lindon JC. (2002) Use of 1H NMR-determined diffusion coefficients to characterize lipoprotein fractions in human blood plasma. Magn Reson Chem 40: S83–S88.

32. Sarafian MH, Gaudin M, Lewis MR, Martin F, Holmes E, et al. (2014) Objective set of criteria for optimization of sample preparation procedures for ultra-high throughput untargeted blood plasma lipid profiling by ultra performance liquid Chromatography–Mass spectrometry. Anal Chem 86: 5766–5774. doi: 10.1021/ac500317c PMID: 24820162

33. Isaac G, McDonald S, Astarita G. (2011) Lipid separation using UPLC with charged surface hybrid technology. Waters App note.

34. Vorkas PA, Isaac G, Anwar MA, Davies AH, Want EJ, et al. (2015) Untargeted UPLC-MS profiling pipeline to expand tissue metabolome coverage: Application to cardiovascular disease. Anal Chem 87: 4184–4193. doi: 10.1021/ac503775m PMID: 25664760

35. Smith CA, Want EJ, O'Maille G, Abagyan R, Siuzdak G. (2006) XCMS:Processing mass spectrometry data for metabolite profiling using nonlinear peak alignment, matching, and identification. Anal Chem 78: 779–787. PMID: 16448051

36. Hotelling H. (1933) Analysis of a complex of statistical variables into principal components. J Educ Psychol 24: 417.

37. Bylesjö M, Rantalainen M, Cloarec O, Nicholson JK, Holmes E, et al. (2006) OPLS discriminant analysis: Combining the strengths of PLS-DA and SIMCA classification. J Chemometrics 20: 341–351.

38. Benjamini Y, Hochberg Y. (1995) Controlling the false discovery rate: A practical and powerful approach to multiple testing. Journal of the Royal Statistical Society.Series B (Methodological) 57: 289–300.

39. Vorkas PA, Shalhoub J, Isaac G, Want EJ, Nicholson JK, et al. (2015) Metabolic phenotyping of atherosclerotic plaques reveals latent associations between free cholesterol and ceramide metabolism in atherogenesis. J Proteome Res 14: 1389–1399. doi: 10.1021/pr5009898 PMID: 25565173

40. Wannemacher RW, Klainer AS, Dinterman RE, Beisel WR. (1976) The significance and mechanism of an increased serum phenylalanine-tyrosine ratio during infection. The American Journal of Clinical Nutrition 29: 997–1006. PMID: 822705

41. Nasstrom E, Vu Thieu NT, Dongol S, Karkey A, Voong Vinh P, et al. (2014) Salmonella typhi and salmonella paratyphi A elaborate distinct systemic metabolite signatures during enteric fever. Elife 3: doi: 10.7554/eLife.03100

42. Lopansri BK, Anstey NM, Stoddard GJ, Mwaikambo ED, Boutlis CS, et al. (2006) Elevated plasma phenylalanine in severe malaria and implications for pathophysiology of neurological complications. Infect Immun 74: 3355–3359. PMID: 16714564

43. Kaufman S. (1989) An evaluation of the possible neurotoxicity of metabolites of phenylalanine. J Pediatr 114: 895–900. PMID: 2654351

44. Enwonwu CO, Afolabi BM, Salako LO, Idigbe EO, Bashirelahi N. (2000) Increased plasma levels of histidine and histamine in falciparum malaria: Relevance to severity of infection. J Neural Transm 107: 1273–1287. PMID: 11145003

45. Watanabe M, Suliman ME, Qureshi AR, Garcia-Lopez E, Bárány P, et al. (2008) Consequences of low plasma histidine in chronic kidney disease patients: Associations with inflammation, oxidative stress, and mortality. The American Journal of Clinical Nutrition 87: 1860–1866. PMID: 18541578

46. Niu YC, Feng RN, Hou Y, Li K, Kang Z, et al. (2012) Histidine and arginine are associated with inflammation and oxidative stress in obese women. Br J Nutr 108: 57–61. doi: 10.1017/S0007114511005289 PMID: 21996294

47. Brun R, Blum J, Chappuis F, Burri C. (2010) Human african trypanosomiasis. The Lancet 375: 148–159.

48. Omeje JN, Anene BM. (2012) Comparative serum biochemical changes induced by experimental infection of T. brucei and T. congolense in pigs. Vet Parasitol 190: 368–374. doi: 10.1016/j.vetpar.2012.07.008 PMID: 22858639

49. Bell JD, Brown JCC, Nicholson JK, Sadler PJ. (1987) Assignment of resonances for 'acute-phase' glycoproteins in high resolution proton NMR spectra of human blood plasma. FEBS Lett 215: 311–315. PMID: 2438159

50. Fischer K, Kettunen J, WÃ¼rtz P, Haller T, Havulinna AS, et al. (2014) Biomarker profiling by nuclear magnetic resonance spectroscopy for the prediction of all-cause mortality: An observational study of 17,345 persons. PLoS Med 11: e1001606. doi: 10.1371/journal.pmed.1001606 PMID: 24586121

51. Calder PC. (2001) Polyunsaturated fatty acids, inflammation, and immunity. Lipids 36: 1007–1024. PMID: 11724453

52. Visser BJ, Wieten RW, Nagel IM, Grobusch MP. (2013) Serum lipids and lipoproteins in malaria-a systematic review and meta-analysis. Malaria J 12: 1–16.

53. Khosla SN, Goyle N, Seth RK. (1991) Lipid profile in enteric fever. J Assoc Physicians India 39: 260–262. PMID: 1880095

54. Alvarez C, Ramos A. (1986) Lipids, lipoproteins, and apoproteins in serum during infection. Clin Chem 32: 142–145. PMID: 3940695

55. Soares NM, Leal TF, Fiuza MC, Reis EA, Souza MA, et al. (2010) Plasma lipoproteins in visceral leishmaniasis and their effect on leishmania-infected macrophages. Parasite Immunol 32: 259–266. doi: 10.1111/j.1365-3024.2009.01187.x PMID: 20398226

56. Black S, Vandeweerd V. (1989) Serum lipoproteins are required for multiplication of trypanosoma brucei brucei under axenic culture conditions. Mol Biochem Parasitol 37: 65–72. PMID: 2615788

57. Coppens I, Levade T, Courtoy PJ. (1995) Host plasma low density lipoprotein particles as an essential source of lipids for the bloodstream forms of trypanosoma brucei. J Biol Chem 270: 5736–5741. PMID: 7890701

58. Richmond GS, Gibellini F, Young SA, Major L, Denton H, et al. (2010) Lipidomic analysis of bloodstream and procyclic form trypanosoma brucei. Parasitology 137: 1357–1392. doi: 10.1017/S0031182010000715 PMID: 20602846

59. Dixon H, Williamson J. (1970) The lipid composition of blood and culture forms of trypanosoma lewisi and trypanosoma rhodesiense compared with that of their environment. Comp Biochem Physiol 33: 111–128. PMID: 5440923

60. Smith TK, Bütikofer P. (2010) Lipid metabolism in trypanosoma brucei. Mol Biochem Parasitol 172: 66–79. doi: 10.1016/j.molbiopara.2010.04.001 PMID: 20382188

61. Maceyka M, Spiegel S. (2014) Sphingolipid metabolites in inflammatory disease. Nature 510: 58–67. doi: 10.1038/nature13475 PMID: 24899305

62. Huet G, Lemesre JL, Grard G, Noireau F, Boutignon F, et al. (1990) Serum lipid and lipoprotein abnormalities in human african trypanosomiasis. Trans R Soc Trop Med Hyg 84: 792–794. PMID: 2096509

Journal of proteome research

Article

pubs.acs.org/jpr

Metabolic, Immune, and Gut Microbial Signals Mount a Systems Response to *Leishmania major* Infection

Sabrina D. Lamour,[†] Kirill A. Veselkov,[†] Joram M. Posma,[†] Emilie Giraud,[‡] Matthew E. Rogers,[§] Simon Croft,[‡] Julian R. Marchesi,[∥,⊥] Elaine Holmes,[†] Karin Seifert,[‡] and Jasmina Saric*[,†]

[†]Division of Computational and Systems Medicine, Department of Surgery and Cancer, Imperial College London, Exhibition Road, London SW7 2AZ, United Kingdom

[‡]Department of Immunology and Infection, London School of Hygiene & Tropical Medicine, Keppel Street, London, WC1E 7HT, United Kingdom

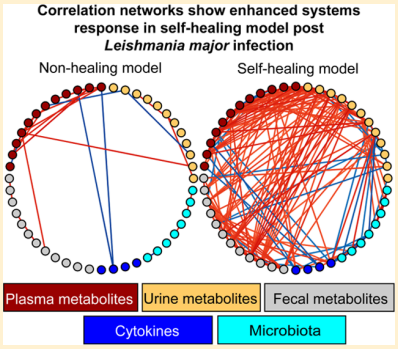
[§]Department of Disease Control, London School of Hygiene & Tropical Medicine, Keppel Street, London, WC1E 7HT, United Kingdom

[∥]Cardiff School of Biosciences, Division of Microbiology, Cardiff University, Museum Avenue, Cardiff, CF10 3AT, United Kingdom

[⊥]Centre for Digestive and Gut Health, Imperial College London, Exhibition Road, London, SW7 2AZ, United Kingdom

Supporting Information

ABSTRACT: Parasitic infections such as leishmaniasis induce a cascade of host physiological responses, including metabolic and immunological changes. Infection with *Leishmania major* parasites causes cutaneous leishmaniasis in humans, a neglected tropical disease that is difficult to manage. To understand the determinants of pathology, we studied *L. major* infection in two mouse models: the self-healing C57BL/6 strain and the nonhealing BALB/c strain. Metabolic profiling of urine, plasma, and feces via proton NMR spectroscopy was performed to discover parasite-specific imprints on global host metabolism. Plasma cytokine status and fecal microbiome were also characterized as additional metrics of the host response to infection. Results demonstrated differences in glucose and lipid metabolism, distinctive immunological phenotypes, and shifts in microbial composition between the two models. We present a novel approach to integrate such metrics using correlation network analyses, whereby self-healing mice demonstrated an orchestrated interaction between the biological measures shortly after infection. In contrast, the response observed in nonhealing mice was delayed and fragmented. Our study suggests that trans-system communication across host metabolism, the innate immune system, and gut microbiome is key for a successful host response to *L. major* and provides a new concept, potentially translatable to other diseases.



KEYWORDS: *host, response, infection, metabolic, profiling, microbiota, cytokine, multivariate, correlation, Leishmania*

INTRODUCTION

The leishmaniases describe a spectrum of diseases that are caused by protozoan parasites of the genus *Leishmania*. The sum of diseases results in significant health burden globally, estimated at approximately 2.4 million DALYs.[1,2] There are severe limitations to the current management of these, including variable clinical manifestations and severity of disease (including the cutaneous and visceral forms), suboptimal diagnosis, limited treatments, and increases in drug resistance.[3] A more fundamental understanding of the pathogenesis is thus key to progressing novel avenues for more effective identification and treatment.

Experimental research on murine models for cutaneous leishmaniasis (CL) has shown that disease outcome depends to a large extent on the immune-phenotype of the host. A dominant Th1 response, prevalent in C57BL/6 mice, with marked increased IFNγ and IL-12 cytokine production, has been associated with disease resolution.[4,5] In contrast, a dominant Th2 response with elevated IL-4 levels, as observed in BALB/c mice, has been linked with disease progression.[6,7]

Recent evidence has shown that a range of different leukocytes, including different Th subsets, and signaling molecules significantly contribute to the host response to *Leishmania* infection.[8−10] Furthermore, host defense to *Leishmania* is not solely built on immune components but also relies on regulatory metabolic messengers, such as the arginine pathway intermediates.[11,12] This special link between immune and metabolic system has only found attention in the past few years and hence still represents a largely unexplored source of potential new information on infection mechanisms.[13−15]

Special Issue: Environmental Impact on Health

Received: August 5, 2014
Published: November 4, 2014

ACS Publications © 2014 American Chemical Society

Metabolic profiling has shown great promise in discovering parasite-specific imprints on global host metabolism but has not yet been investigated with *Leishmania* infection in vivo. Such studies on other parasitic infections have conventionally relied on ^{1}H nuclear magnetic resonance (NMR) spectroscopy-based screening of urine and plasma samples, taken from infected rodent hosts for diagnostic biomarker research.[16,17] NMR-based analyses, combined with advances in multivariate statistical tools, have shown to yield highly reproducible data, with coefficients of variances (CVs) of major metabolite markers typically <10%.[18,19] This approach in infection studies has expanded successively to include tissues and, more recently, to combine metrics from both the metabolic and immune system. Such a combined approach enables a more systemic view on host response to a pathogen and characterizes the immune−metabolic interface during infection.[20]

A recurrent theme across a range of these parasite−rodent models characterized by metabolic profiling is the association between parasitic infection and an altered host microbiome, as indicated by a set of microbiota-associated metabolites detected in the urine and plasma.[16,17,21] Additionally, *L. major* infection models have demonstrated that germ-free mice did not resolve lesions as well as their conventional counterparts, despite generating an appropriate Th1 immune response.[22] Results strongly implicate that the microbiota are key for mounting a successful host response to the parasite.

The current study was performed to characterize the infection-specific response in a self-healing C57BL/6 and a nonhealing BALB/c experimental model for CL (Figure 1). We adopted a holistic approach to investigate the early phase of *L. major* infection by combining metrics of the immune, metabolic, and gut microbial response in the host, including metabolic profiles from urine, plasma, and feces, selected peripheral cytokines, and fecal bacterial composition. By using this combined set of phenotypic metrics, we aim to gain a better understanding of the interactions that lead to the complex pathology of leishmaniasis. Moreover, by taking the successfully resistant mouse phenotype as a reference system, we intend to define the biological parameters of the host that are associated with disease resolution.

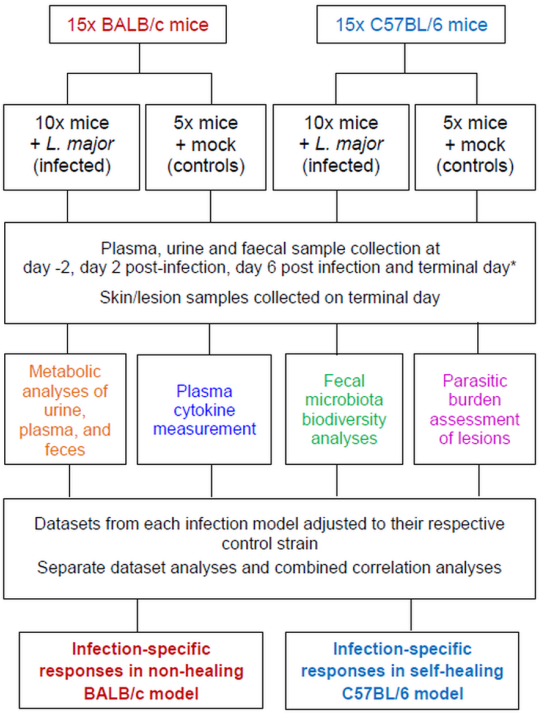


Figure 1. Schematic overview of study. Study compares experimental cutaneous leishmaniasis in nonhealing BALB/c mice versus self-healing C57BL/6 mice during early phases of *L. major* infection. (See the Materials and Methods.) *Terminal day defined for each infected mouse on the day shortly before lesion ulceration (days 7−13) or on day 12 for control mice.

MATERIALS AND METHODS

Ethics Statement

All animal experiments were performed at Imperial College London, adhering to local and national animal handling guidelines, under project license PPL70/6997. The protocols were approved by the animal ethics committees at London School of Hygiene & Tropical Medicine, Imperial College London, and the U.K. Home Office.

Experimental Model and Sample Collection

L. major infection was established in 7 to 8 week old female C57BL/6J mice and BALB/cJ mice (n = 15 per strain, 10 infected, and 5 controls), purchased from Charles River, U.K. Mice were fed on standard rodent diet (Rat and Mouse No. 1 Maintenance diet, Special Diets Service, U.K.). Following 1 week acclimatization, mice were randomly allocated into infection and control groups. Mice in the infection group received subcutaneous (s.c.) injections with 2×10^{6} stationary phase *L. major* promastigotes (strain MHOM/SA/85/JISH118, passage 2; culture methods detailed in the Supporting Information) in 100 μL of Schneider's Insect Medium (Sigma-Aldrich) into the shaven rump, whereas the control group received medium only. Plasma, urine, and fecal samples were taken from each mouse 2 days prior to infection and postinfection at days 2 and 6 and at the terminal day of the study, defined as the day the lesion had developed before ulceration occurred (7−13 days postinfection, depending on each infected animal) or 12 days postinjection for control animals. Samples were snap-frozen in liquid nitrogen and stored at −40 °C for metabolic analysis. Additional fecal pellets for microbial assessment were taken at all four time points as well as extra sections of terminal day lesion (or skin for controls) for measurement of parasitic burden, which were all fixed in 10% formalin and subsequently stored at −20 °C.

^{1}H NMR Acquisition and Processing

Plasma, urine, and fecal samples were prepared for ^{1}H NMR analyses based on protocols as described by Beckonert et al., 2007,[23] where methods for tissue metabolite extraction were adjusted for preparation of feces (Supporting Information). ^{1}H NMR data were acquired on a Bruker Avance 600 NMR spectrometer with TXI probe head (Bruker), using XWIN-NMR software (Bruker BioSpin). Acquisition settings are outlined in the Supporting Information. Spectra were preprocessed using automatic phasing, baseline correction, and reference peak calibration via an in-house algorithm in MATLAB (version R2012b, Mathworks) and Topspin 3.1 software (Bruker BioSpin). Water (and urea regions for urine) were removed prior to automatic spectral alignment and probabilistic quotient normalization,[24] using an in-house MATLAB script.

dx.doi.org/10.1021/pr5008202 | J. Proteome Res. 2015, 14, 318−329

Plasma Cytokine Assessment

Following NMR acquisition, plasma samples were tested in duplicate via the ultrasensitive plasma Th1/Th2 Cytokine 9-plex assay kit (Meso Scale Discovery/MSD) for levels of IFNγ, IL-1β, IL-2, IL-4, IL-5, CXCL1, IL-10, IL-12, and TNFα. Protocols were based on manufacturer's instructions, with sample incubation time increased to 3 h for increased detection. Assay plates were read on Sector Imager 2400, and concentrations were calculated using accompanying Discovery Workbench MSD software.

Fecal Microbiota Diversity Analysis with Next-Generation Sequencing

Fecal DNA was extracted using QIAmp Stool Mini Kit (QIAGEN) based on manufacturer's instructions with minor protocol amendments (Supporting Information). The presence of microbial genes was confirmed via 35 cycles of PCR [30 s at 95 °C, 40 s at 55 °C, 60 s at 72 °C] using primers against the V1−V3 region of bacterial 16S rRNA genes, as described by Lewis et al., 2013.[25] DNA samples were sequenced via next-generation pyrosequencing using a 454 FLX Genome Sequencer (Roche) by Research and Testing Laboratory (Texas, USA), and data were processed and analyzed according to company guidelines. To be able to discriminate between functionally distinctive groups of bacteria, as opposed to purely genotypic differences, when assessing changes at lower taxonomic levels, we analyzed the presented data at the class level, comparing relative percentage abundance of bacteria between samples. The top eight bacterial classes with highest abundance were used for final analyses, representing 99.1% of all measured classes across the data set.

Parasitic Burden Assessment via Real-Time PCR

DNA was extracted from lesion samples using DNeasy Blood & Tissue Kit (QIAGEN), following manufacturer's instructions, with minor modifications (PBS wash steps before extraction and doubling proteinase K for lesion samples). DNA concentration and purity were measured using Nanodrop (Thermo Scientific). Real-time PCR was performed with triplicate samples using SYBR-Green kit (Applied BioSystems)[26] and primers against mouse arginase 1 to quantify genomic DNA and *Leishmania*-specific RV1/RV2[27] to detect parasites. Parasitic burden for each sample was assigned as average fluorescence intensity of parasitic DNA as a multiple of the lowest positive reading, normalized to arginase levels and calculated with Qbase Plus 2.4 software (Biogazelle).

Statistical Analysis and Data Integration

Metabolic spectra of infected and control groups from both strains were initially analyzed by principal component analysis (PCA).[28] Graphical representations of the first few components were used to explore overall similarities/differences in metabolic phenotypes using SIMCA P+ software (V. 12.0, Umetrics). Mouse strain differences at the preinfection time-point as well as infection-associated differences between infected and time-matched control mice for each strain were then explored using orthogonal partial lease squares discriminatory analysis (O-PLS-DA)[29] which includes a seven-fold cross-validation for each biological matrix at each time-point. Final analyses focused on the comparison between the infection-related responses of the two mouse models, whereby to account for the strain differences, metabolic spectra, cytokine, and microbiome data sets from infected mice, were expressed in relation to their respective control strain by subtracting the uninfected control group median spectra level from infected data. The strain-adjusted metabolic data were further analyzed via O-PLS-DA. Assigned discriminatory metabolites from O-PLS-DA models were selected when the associated Pearson-product moment correlation coefficient (r) of a given metabolite peak surpassed the critical values for $p < 0.05$ and its ROC sensitivity/specificity score was at least 0.8. Spectral assignments of metabolite peaks were performed based on Chenomx NMR suite profiler 7.0 software and using known assignments from in-house NMR databases and literature.[21]

The differences in individual cytokines and class microbiota levels were analyzed using nonparametric univariate tests: Kruskal−Wallis ANOVA and pairwise Mann−Whitney U-test for time-matched strain comparisons and Friedman tests and Wilcoxon matched-pair tests for comparative analyses between time-points for each mouse model. Bonferroni correction was applied to account for multiple testing. Baseline plasma samples from the control groups were lacking, thus no control-strain adjustment could be performed for plasma metabolic and cytokine data at this time point, required for spectral O-PLS-DA and cytokine univariate analyses, respectively.

Two-dimensional-matrix, Pearson-based, correlation analyses were performed between time-matched spectral data and cytokine levels, spectral data and microbiota class percentages, and cytokine levels and microbiota using in-house MATLAB code (Dr. Judith Fonville). The script included a 10 000-fold random permutation of sample order. A p value was calculated as the number of times the absolute value of correlation coefficient of the resampled data exceeded the original correlation coefficient value. Only correlative pairs with $p < 0.05$ were displayed. Twelve metabolites were selected each from urine, plasma, and feces based on their level of correlation with host cytokines or microbiome and on their discriminatory role in differentiating between the two CL mouse models. The integrals were calculated and coanalyzed alongside time-matched microbiota class percentage abundance and cytokine levels via PCA and PLS-DA to determine how the different data sets (48 variables in total) drive the separation between the two infection models.

These same variables were correlated via Pearson correlation with relative levels of parasitic burden for each mouse model to determine direct associations between infection and host measures. Furthermore, Pearson correlation network graphs were drawn for each time point in both strains individually to determine specific host interactions between variables, in response to infection, using in-house MATLAB code based on MetaboNetworks software.[30] A generous cutoff for the q values of 0.3 using the Storey−Tibshirani false discovery rate[31] was applied for multiple testing correction in the parasite−host correlation as well as correlation network analyses for maximal exploration of potential biologically relevant host interactions in this study. Although the sample size of this study was small, and future studies are needed to perform targeted analyses to test these association, the majority of associations are still expected to be true discoveries.

RESULTS

BALB/c and C57BL/6 Mice Have Different Metabolic, Immune, and Gut Microbial Phenotypes in Response to *L. major* Infection

Among the physiological and pathological measures that were monitored for the two strains within the first 2 weeks of

dx.doi.org/10.1021/pr5008202 | J. Proteome Res. 2015, 14, 318−329

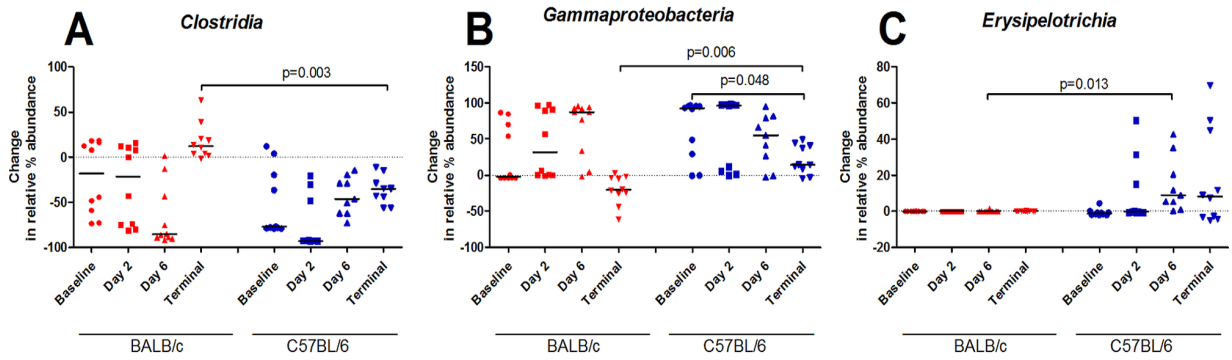


Figure 2. Changes in fecal microbial composition between the two mouse CL models. Infection-related changes in composition of the bacterial classes *Clostridia* (A), *Gammaproteobacteria* (B), and *Erysipelotrichia* (C) in feces of infected mice, adjusted to their respective control strain. Horizontal bars represent group median averages, and each symbol represents a fecal sample (n = 10 for all groups except C57BL/6 Day 6 where n = 9).

infection, only the hematocrit differed significantly (Table S1 in the Supporting Information). No significant changes were observed in weight, day of lesion onset, or lesion size (although trends showed that the BALB/c had larger overall lesions). Infection with *L. major* was confirmed via real-time PCR of *L. major* DNA in lesions taken at the terminal time point (days 7−13 post infection, the most acute phase of the infection in our study), with no significant differences observed between strains.

In contrast with the minor pathological changes, the strains displayed clear strain and infection-related differences in their metabolism across the three assessed biofluids: urine, plasma, and fecal extracts. Preliminary multivariate analyses using PCA and O-PLS-DA revealed that strain differences prior to infection included significantly higher levels of glucose and lactate with lower lipid moieties in plasma of C57BL/6 mice compared with BALB/c. BALB/c also displayed relatively higher fecal lipid levels, along with lower amounts of urinary citrate, succinate, and 2-oxoadipate. Infection resulted in significant decreases in plasma glucose and urinary hippurate across both strains as well as decreases in plasma lipids in nonhealing BALB/c mice. To investigate the differences in metabolic infection-related responses between the two models, we based finalized results (Table S2 in the Supporting Information) for O-PLS-DA models on spectra that took into account strain differences unrelated to infection by adjusting each infected group to their respective uninfected control strain, a method that was also applied to subsequent cytokine and microbiome data (Materials and Methods). Metabolites with significantly higher levels in the nonhealing strain for at least one postinfection time point included alanine (in plasma and feces), succinate (feces), 3-hydroxybutyrate (3-HB), and glucose (in plasma) as well as the gut-microbiota-associated metabolites phenylacetylglcyine (PAG) and trimethylamine (TMA) in the urine. In contrast, metabolites found in higher concentrations in the self-healing strain after infection included urinary acetate and several fatty acid moieties from lipoproteins in plasma.

Out of the nine cytokines initially tested in plasma, four displayed detectable levels, namely, IL-1β, IL-10, IL-12, and CXCL1. IL-1β was the only cytokine to be significantly different between the two strains prior to infection (higher in BABL/c mice, p = 0.006). Following strain adjustment, infection-related differences between the models (Figure S1 in the Supporting Information) showed that nonhealing mice generally had higher levels of the proinflammatory cytokines CXCL1 and IL-1β than the self-healing model over all time-points, although statistical significance was reached only for IL-1β, 6 days postinfection (d 6 p.i., p = 0.036). They also had higher levels of cytokine IL-12 than self-healing mice, 2 days post infection (d 2 p.i., p = 0.013).

To compare fecal microbial composition between animals, we assessed changes in percentage abundance of bacterial classes (further outlined in the Materials and Methods). *Clostridia* (phylum *Firmicutes*) and *Gammaproteobacteria* (phylum *Proteobacteria*) were the most dominant classes of fecal bacteria observed in this study across both strains, accounting for over 85% relative abundance of the total data. Initial trends showed that *Clostridia* were present in higher concentrations in BALB/c mice while *Gammaproteobacteria* were higher in C57BL/6 mice before infection, yet this was not significant. Following strain-adjustment, infection-related microbial levels for *Clostridia* remained consistently higher in the BALB/c strain, reaching significance at the terminal time point (p = 0.003, Figure 2A). A considerable decrease in relative *Gammaproteobacteria* levels was observed in both strains by the end of the study (significant in C57BL/6 mice, p = 0.048), although levels in C57BL/6 mice remained significantly higher than those in BALB/c mice (p = 0.006, Figure 2B).

Erysipelotrichia represented the third most abundant bacterial class, which was largely absent in BALB/c mice and, where present, was significantly lower than in self-healing C57BL/6 mice, d 6 p.i. (p = 0.013; Figure 2C, Table S3 in the Supporting Information). The remaining bacterial classes were only detected at low levels (typically <5% relative abundance), although significant changes relative to baseline were detected in self-healing mice, for example, increases in *Alphaproteobacteria* (d 2 p.i., p = 0.048) and *Bacteroidia* (d 2 p.i., p = 0.012 and terminal, p = 0.048) and decreases in *Actinobacteria* (d 6 p.i., p = 0.027) and *Bacilli* (terminal day, p = 0.048; Table S4 in the Supporting Information). No significant changes in fecal bacterial composition were evident over the study period in nonhealing mice, indicating a more static microbial population than that observed in the self-healing model (Tables S3 and S4 in the Supporting Information).

Peripheral Cytokines Correlate with the Systemic Host Metabolism

Correlation analysis was performed between cytokine and NMR spectral data to identify associations between the host

dx.doi.org/10.1021/pr5008202 | J. Proteome Res. 2015, 14, 318−329

Table 1. Immune−Metabolic Correlates in Urine, Plasma, and Feces of the Self-Healing and Non-Healing CL Infection Models[a]

		infected BALB/c				infected C57BL/6			
biofluid	metabolite correlate	IL-10	IL-12	IL-1β	CXCL1	IL-10	IL-12	IL-1β	CXCL1
urine	2-oxoadipate*					−		−	
	2-oxoglutarate							−	
	2-oxoisocaporate	+	−	−	−		+		
	acetate				−	−			
	cis-aconitate			−					
	citrate		−				+		+
	creatine				−	±			+
	creatinine			−					
	fumarate				-				
	guanidoacetate		−		+				
	hippurate	−	−			−			
	lactate	+	−	−		−		−	
	PAG		+		±		+	+	+
	succinate				−			−	
	taurine	−	−		±	±	−	+	+
	TMA							−	−
	TMAO					+	+		±
	ureidopropanoate		−	−	−				
plasma	3-HB								+
	choline		+						
	glucose	+	+		−	±		+	±
	glycine					−			
	GPC		−						
	lactate	±	−	−				+	
	lipid CH_2		−			+			+
	lipid CH_3		−			+			+
	lipid* 3.22(m)					+			+
	scyllo-inositol*			−					
feces	butyrate		−		−	+			
	glutamate						−		
	lactate						+		±
	lipid CH_3						+		
	propionate		+				+		+
	succinate							−	+
	taurine						+		

[a]Key: +, positive correlations; −, negative correlations; ±, direction of correlation alternated between different time-points; *, tentative assignment; 3-HB, 3-hydroxybutyrate; GPC, glycerophoshocholine; (m), unassigned multiplet (accompanied by chemical shift region); PAG, phenylacetylglycine; TMA, trimethylamine; TMAO; trimethylamine-*N*-oxide.

immune and metabolic system in response to infection for each of the murine phenotypes. Results showed that multiple interactions existed between the cytokine and the metabolite data for all three biofluids (Table 1), with a higher number of correlations determined in the self-healing strain. Certain immune-metabolic correlates were common to both the self-healing and nonhealing models, for example, a negative association between urinary hippurate and IL-10. Other correlations were specific to a particular CL model: urinary ureidopropanoate, for instance, was found to be negatively correlated with IL-12, IL-1β, and CXCL1 in the nonhealing strain but did not correlate with cytokines in the self-healing model. In several instances, particularly in nonhealing mice, metabolites that were correlated with IL-12, IL-1β, or CXCL1 produced opposing or no correlative associations with regulatory cytokine IL-10. These disparate correlation patterns included 2-oxoisocaproate, PAG, and lactate in the urine.

Clostridia and *Gammaproteobacteria* Display Opposing Associations with Host Metabolites

Extending further from the immune-metabolic correlations, associations between the host fecal microbiome and metabolism were similarly investigated using percentage abundance of the different bacterial classes and spectral NMR data as input parameters. Results for plasma and feces have been summarized in Table 2 (Table S5 in the Supporting Information for urine). The majority of the metabolites that were negatively linked with the *Clostridia* class were positively linked to *Gammaproteobacteria* and vice versa in both mouse models. Such correlative metabolites included choline, glycerophosphocholine (GPC), plasma lipids, lactate, the gut-microbiota-associated molecules hippurate, PAG and TMA, and metabolites related to the tricarboxylic acid cycle (e.g., acetate and fumarate) in the urine. *Erysipelotrichia* displayed similar correlation patterns with host urinary and fecal metabolites as *Clostridia* in both CL models as well as with plasma metabolites in the self-healing C57BL/6 strain.

dx.doi.org/10.1021/pr5008202 | *J. Proteome Res.* 2015, 14, 318−329

Table 2. Microbial−Metabolic Correlates in Plasma and Feces in the Self-Healing and Non-Healing Infection Models[a]

biofluid	metabolic correlate	clos	gamma	erysi	alpha	beta	bacil	actino	bacter
BALB/c: plasma	3-HB			+	+			+	
	choline	+	−	+	+			+	+
	glucose	−		+	− −	−	+	−	+
	GPC	−	+	+	−	−			−
	lactate				±		−	−	
	lipid* 3.22(m)			+	−			+	
	lipid CH_2		+	+	±	−		+	
	lipid CH_3			+	±	−		+	
	s-inositol*				−	±		+	−
BALB/c: feces	acetate			−					
	butyrate	−	+	−		+	+		−
	lactate						−		
	propionate		+	−			−		
	succinate						−		
C57BL/6: plasma	3-HB	−	+	−	−				
	choline	−	+						
	glucose			−	−			− −	++
	GPC	−	+						
	lactate	−						−	
	lipid* 3.22(m)	−	+	−					
	lipid CH_2	−	+	−					
	lipid CH_3	−		−					
C57BL/6: feces	acetate					+	+	−	
	butyrate							−	
	DMA			−					
	glutamate			−			−		
	lactate	−	+	−	−				+
	propionate	−	+					−	
	succinate	−	+	−	−				

[a]Key: +, positive correlations at one time-point; ++, positive correlations for at least two time-points; −, negative correlations at one time-point; − −; negative correlation for at least two time-points; ±, direction of correlation alternating between different time-points; *, tentative assignment; 3.22(m), multiplet spectral peak at δ 3.22; 3-HB, 3-hydroxybutyrate; actino, *Actinobacteria*; alpha, *Alphaproteobacteria*; bacil, *Bacilli*; bacter, *Bacteroidia*; beta, *Betaproteobacteria*; clos, *Clostridia*; DMA; dimethylamine; erysi, *Erysipelotrichia*; gamma, *Gammaproteobacteria*; GPC, glycerophoshocholine; (m), unassigned multiplet (accompanied by chemical shift region); PAG, phenylacetylglycine; *s*-, *scyllo*.

Table 3. Immune−Microbial Correlates Across the Self-Healing and Non-Healing Infection Models[a]

	infected BALB/c				infected C57BL/6			
microbial class	IL-10	IL-12	IL-1β	CXCL1	IL-10	IL-12	IL-1β	CXCL1
Clostridia	+	++				+	++	
Gammaproteobacteria			±	+	+		+	+
Erysipelotrichia		+		++	+	+	+	
Alphaproteobacteria	++	+				+	++	
Betaproteobacteria	±					+	++	
Bacilli		+	+		+	+	+	
Actinobacteria	++	+			++		+	
Bacteroidia	++	+			+	+	+	

[a]Key: +, positive correlations at one time-point; ++ positive correlations for at least two time-points; ±, direction of correlation alternating between different time-points.

Multiple Links between Cytokines and Fecal Microbial Classes in Both Infection Models

Details of the correlation analyses between cytokine concentrations and microbial class abundances are summarized in Table 3. IL-1β displayed the highest number of correlations with the microbiota in the self-healing phenotype. IL-1β was found to be positively correlated with all microbial classes for at least one time point in self-healing mice yet was only correlated with *Bacilli* and *Gammaproteobacteria* in nonhealing mice. IL-12 and IL-10 displayed the most immune-microbial correlations out of all the cytokines in the nonhealing mice. *Actinobacteria* and *Bacteroidia* classes were strongly positively correlated with IL-10 levels in both strains, suggesting these commensal bacteria may exert IL-10-dependent anti-inflammatory effects on the host.

Differences in Metabolic and Microbial Responses to *L. major* Infection Dominate Separation between the Self-Healing and the Nonhealing Phenotype

To ascertain the relative contribution of the host metabolic, immunological, and microbial backgrounds toward differ-

dx.doi.org/10.1021/pr5008202 | *J. Proteome Res.* 2015, 14, 318−329

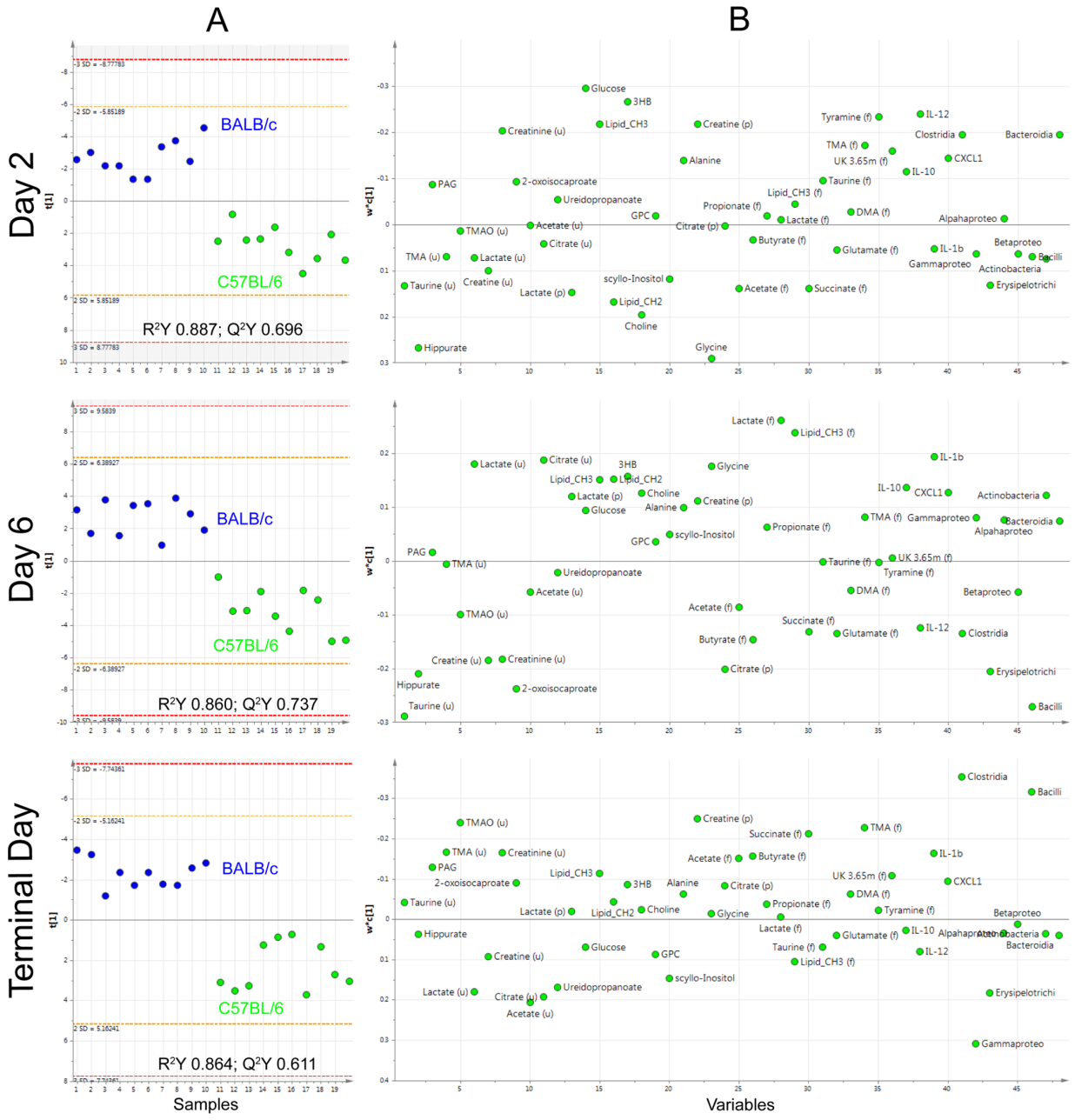


Figure 3. Discriminatory variables between two infection models for each postinfection time-point assessed via PLS-DA. Scores plots (A) and corresponding loadings plots (B) of partial least-squares discriminatory analyses (PLS-DA) between the two infected strains are presented based on the first component vector for 2 days and 6 days postinfection and at the terminal study day. (A) Each circle represents one sample. BALB/c (blue) and C57BL/6 (green) samples are separated into the two infection models, vertically along the first PLS vector $t[1]$. Horizontal hashed lines represented two and three standard deviations (SD) from the mean. (B) Relative contribution of each of the discriminatory variables (circles) toward the vertical mouse model separation in the associated scores plot, composed of 48 markers in the following order: 12 urinary (u), 12 plasma (p), and 12 fecal metabolites (f), 4 cytokines, and 8 percentage abundance of eight microbial classes. Origin of metabolites present in more than biofluid additionally labeled with (u), (p), or (f). R^2Y and Q^2Y describe fit and predictive value of each model, respectively. Key: 3HB, 3-hydroxybutyrate; GPC, glycerophosphocholine; PAG, pheynylacetylglycine; TMA, trimethylamine; TMAO, trimethylamine-*N*-oxide; UK 3.65m, unassigned multiplet at δ 3.65.

entiating the two mouse strains' responses to *L. major* infection, we coanalyzed metabolite integrals, cytokine concentrations, and bacterial percentage abundances by partial least-squares discriminatory analyses (PLS-DA), shown in Figure 3. Average coefficient of variances (CVs) for the selected metabolites across all postinfection time points were 27.7 (urine), 39.2 (plasma), and 16.4% for BALB/C mice and 32.8 (urine), 35.3 (plasma), and 20.5% for C57BL/6 mice. Analyses showed that metabolites in the urine and plasma represented the major drivers of separation between host responses of the two strains to infection, particularly at the beginning of infection; for example, urinary hippurate and plasma glycine were found in

dx.doi.org/10.1021/pr5008202 | *J. Proteome Res.* 2015, 14, 318−329

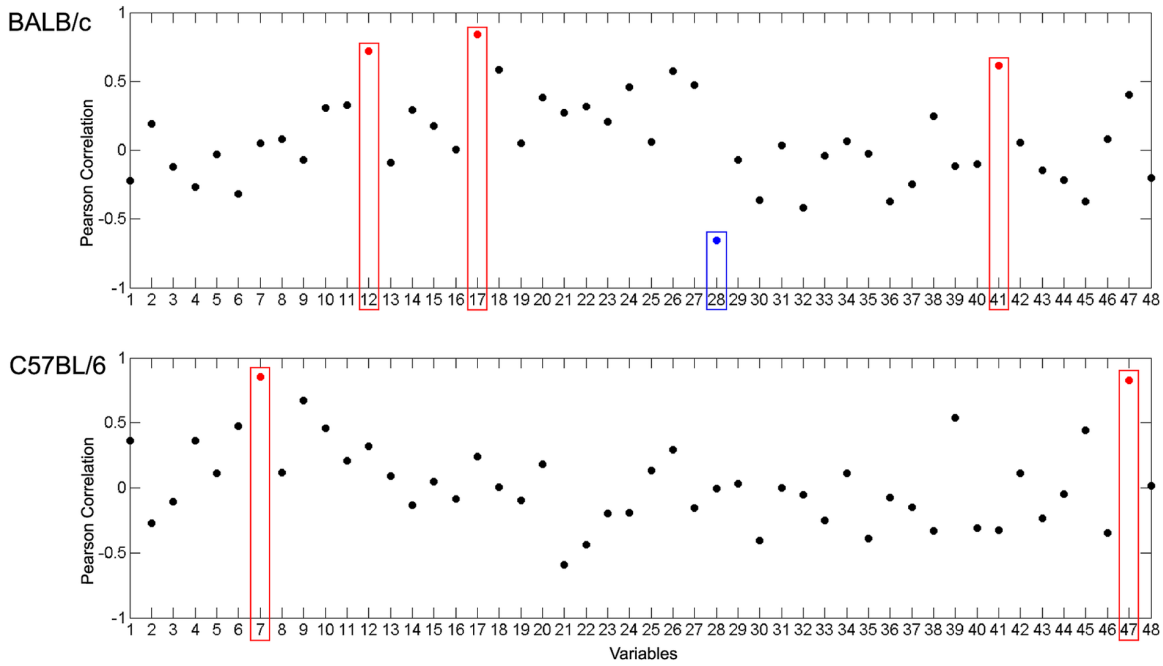


Figure 4. Correlation between parasitic burden and host response markers in nonhealing and self-healing CL models. Plot shows Pearson correlation between relative amount of *L. major* DNA in lesions (Table S1 in the Supporting Information) from infected BALB/c and C57BL/6 mice and levels of infection-related metabolic, immune, and fecal microbial variables (terminal time-point). Each variable is represented by a small circle, where significantly positively correlated variables are boxed in red and negative in blue ($q < 0.3$). Key: fecal metabolites: 1, unassigned multiplet peak at δ 3.65; 2, tyramine; 3, trimethylamine (TMA); 4, dimethylamine; 5, glutamate; 6, taurine; 7, succinate; 8, lipid (CH3); 9, lactate; 10, propionate; 11, butyrate; 12, acetate. Plasma metabolites: 13, citrate; 14, glycine; 15, creatine; 16, alanine; 17, scyllo-inositol; 18, glycerophophocholine; 19, choline; 20, 3-hydroxybutyrate; 21, lipid (CH2); 22, lipid (CH3); 23, glucose; 24, lactate. Urine metabolites: 25, ureidopropanoate; 26, citrate; 27, acetate; 28, 2-oxoisocaproate; 29, creatinine; 30, creatine; 31, lactate; 32, trimethylamine-*N*-oxide; 33, TMA; 34, phenylacetylglycine; 35, hippurate; 36, taurine. Bacterial classes: 37, *Clostridia*; 38, *Gammaproteobacteria*; 39, *Erysipelotrichia*; 40, *Alphaproteobacteria*; 41, *Betaproteobacteria*; 42, *Bacilli*; 43, *Actinobacteria*; 44, *Bacteroidia*. Cytokines: 45, CXCL1; 46, IL-1ß; 47, IL-12; 48, IL-10.

higher concentrations in the self-healing phenotype (Figure 3). Hippurate remained consistently higher in this strain throughout the study in addition to urinary creatine, fecal glutamate, and levels of the *Erysipelotrichia* and *Betaproteobacteria*. In contrast, CXCL1, urinary PAG, and plasma levels of 3-HB and lipid methyl groups were consistently higher in nonhealing mice throughout infection. With infection progression, the fecal microbiota became the most discriminatory variables between the two strains. By the terminal time point, members of the *Gammaproteobacteria* class were strongly linked to self-healing mice, while *Clostridia* and *Bacilli* classes were associated with the nonhealing phenotype (Figure 3).

Parasitic Load Associated with a Select Set of Metabolic, Immunological, and Microbial Markers

To assess whether parasitic burden was linked to host metabolism, immune status, or fecal microbiota, we correlated relative fluorescence intensities of lesion-derived *L. major DNA* with the same variables as used for the previous PLS-DA models (Figure 4). Fecal succinate and IL-12 were found to be positively linked to parasitic load in the self-healing mouse strain. In contrast, plasma *scyllo*-inositol, fecal acetate, and *Betaproteobacteria* were positively associated with parasite counts in nonhealing mice, whereas urinary 2-oxoisocaproate was negatively correlated.

L. major Results in a CL Model-Specific Systems Response between Host Metabolism, Immune System, and Microbiota

Pair-wise Pearson correlation analysis was performed between selected variables to assess potential associations between the different host responses within each infection model. A separate model was built for each time-point and plotted as correlation network wheels, shown in Figure 5. The level of interconnectivity between the variables differed between the two strains, even prior to infection, where the nonhealing strain displayed hardly any correlations between variables (Figure S2 in the Supporting Information). Infection resulted in a substantial increase in the number of correlative associations in the self-healing strain at d 2 p.i., the highest observed in this study, whereby each variable was directly correlated with at least one other variable. At d 6 p.i., nonhealing mice displayed more associations between variables than earlier time points, while correlations between variables in self-healing mice became more confined to their individual biological matrices. Multiple correlations were still observed in both models at the terminal time-point, whereby the two strains displayed many similarities in their correlation patterns.

While the constrained format of the correlation wheels allows for direct comparisons of variables between strains or time points, results were additionally depicted as network maps (Figure 6 and Figure S3 in the Supporting Information). The topology of the network maps is more conducive to interpretation than the correlation wheels, favoring the

dx.doi.org/10.1021/pr5008202 | *J. Proteome Res.* 2015, 14, 318−329

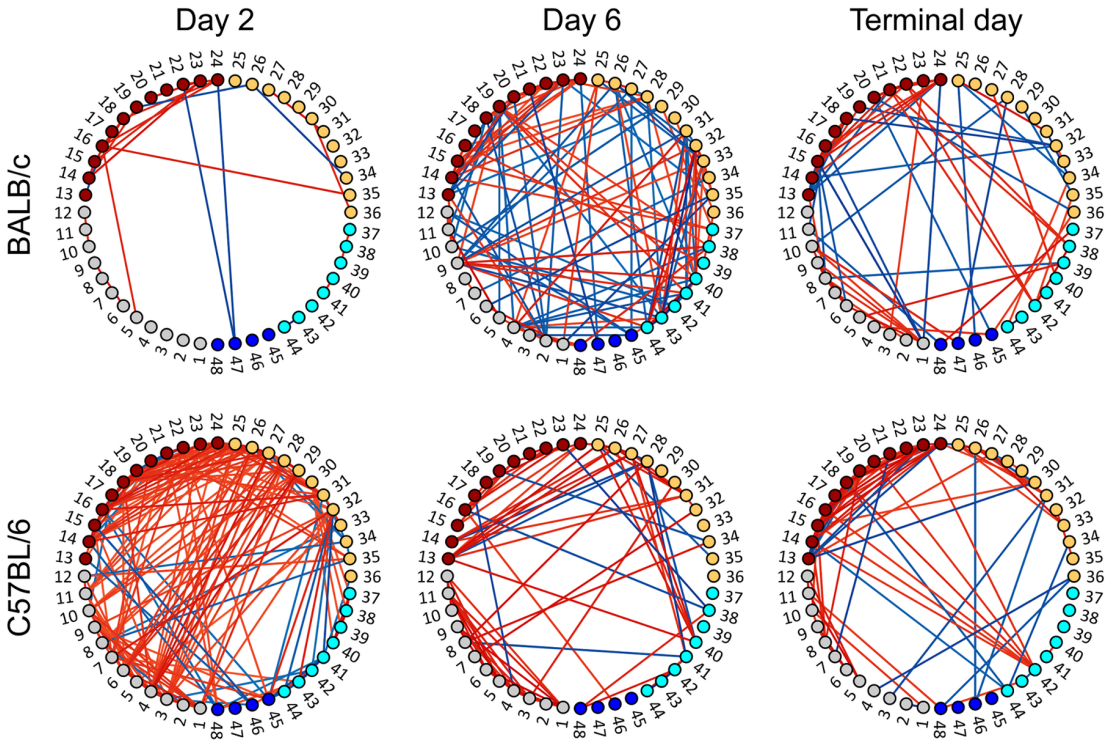


Figure 5. Correlation wheels of metabolic, immune, and microbial markers during acute infection in two CL models. Correlation networks for each measured time-point postinfection based on pairwise Pearson correlation between selected variables for *L.-major*-infected BALB/c and C57BL/6 mice (see Figure S3 in the Supporting Information for preinfection results). Red and blue connecting lines between variables represent significant positive and negative correlations, respectively ($q < 0.3$). The structure of the wheels is constrained such that variables have a fixed position and have color-coded nodes along the circumference. Key: Gray, fecal metabolites; 1, unassigned multiplet peak at 3.65 ppm; 2, tyramine; 3, trimethylamine (TMA); 4, dimethylamine; 5, glutamate; 6, taurine; 7, succinate; 8, lipid (CH3); 9, lactate; 10, propionate; 11, butyrate; 12, acetate. Red, plasma metabolites; 13, citrate; 14, glycine; 15, creatine; 16, alanine; 17, scyllo-inositol; 18, glycerophophocholine; 19, choline; 20, 3-hydroxybutyrate; 21, lipid (CH2); 22, lipid (CH3); 23, glucose; 24, lactate. Yellow, urine metabolites; 25, ureidopropanoate; 26, citrate; 27, acetate; 28, 2-oxoisocaproate; 29, creatinine; 30, creatine; 31, lactate; 32, trimethylamine-*N*-oxide; 33, TMA; 34, phenylacetylglycine; 35, hippurate; 36, taurine. Cyan, bacterial classes; 37, *Clostridia*; 38, *Gammaproteobacteria*; 39, *Erysipelotrichia*; 40, *Alphaproteobacteria*; 41, *Betaproteobacteria*; 42, *Bacilli*; 43, *Actinobacteria*; 44, *Bacteroidia*. Dark blue, cytokines; 45, CXCL1; 46, IL-1ß; 47, IL-12; 48, IL-10.

visualization of major hubs and chains of correlation within the networks, and thus more aligned with mapping to biological pathways. The maps demonstrated that metabolites in plasma and urine represented the core frameworks within the networks during the course of infection in both models, displaying the highest number of inter- and intramatrix correlations. Acetate, citrate, and alanine represented major network hubs in the self-healing mouse model, with some of the highest numbers of direct associations across all three postinfection time-points. Unlike C57BL/6, network hubs in the nonhealing strain were less stable through time and alternated throughout the course of the study, with only plasma lactate remaining highly correlated with other variables at the three postinfection time points.

With the exception of d 2 p.i. in the BALB/c mice, IL-10 was consistently connected with multiple host components within the core network of both phenotypes, highlighting its systemic regulatory role within the host. The pro-inflammatory cytokine CXCL1 was directly correlated with numerous bacterial classes and metabolites across different biological matrices in both infected strains. IL-12 produced negligible correlations in the self-healing strain in the current study. Within the microbiota data, *Bacilli* and *Bacteroidia* classes were consistently disconnected from the core network hubs and were connected largely with themselves, if correlated at all. *Clostridia* and *Gammaproteobacteria* were negatively correlated with each other in both models across multiple time-points.

DISCUSSION

The current study presents a systematic characterization of the global metabolic profile, peripheral cytokine status, and the fecal microbiome in both susceptible BALB/c and self-healing C57BL/6 mouse models of CL. The specific objective of the work was to reveal the immune−metabolic−microbial phenotype that underlies a successful response to *L. major*, using an integrated network approach. Within the immunological context of this study, the first assessed postinfection time point (d 2 p.i.) represents an immune response that is entirely driven by innate factors. The two later time points assessed (d 6 p.i. and terminal) are generally thought to be driven by the host's adaptive immunity.[32] While our study focused only on the early stages of disease, trends in increased lesion size and parasitic burden were already apparent in the nonhealing model.

Key differences in strain-related metabolic responses toward infection included changes in glucose and lipid metabolism. The nonhealing BALB/c mice appeared to rely predominantly on fatty acids rather than glucose as source of energy during infection, reflected by the relative increases in plasma glucose and 3-HB: a ketone body generated from fatty acid break-

dx.doi.org/10.1021/pr5008202 | *J. Proteome Res.* 2015, 14, 318−329

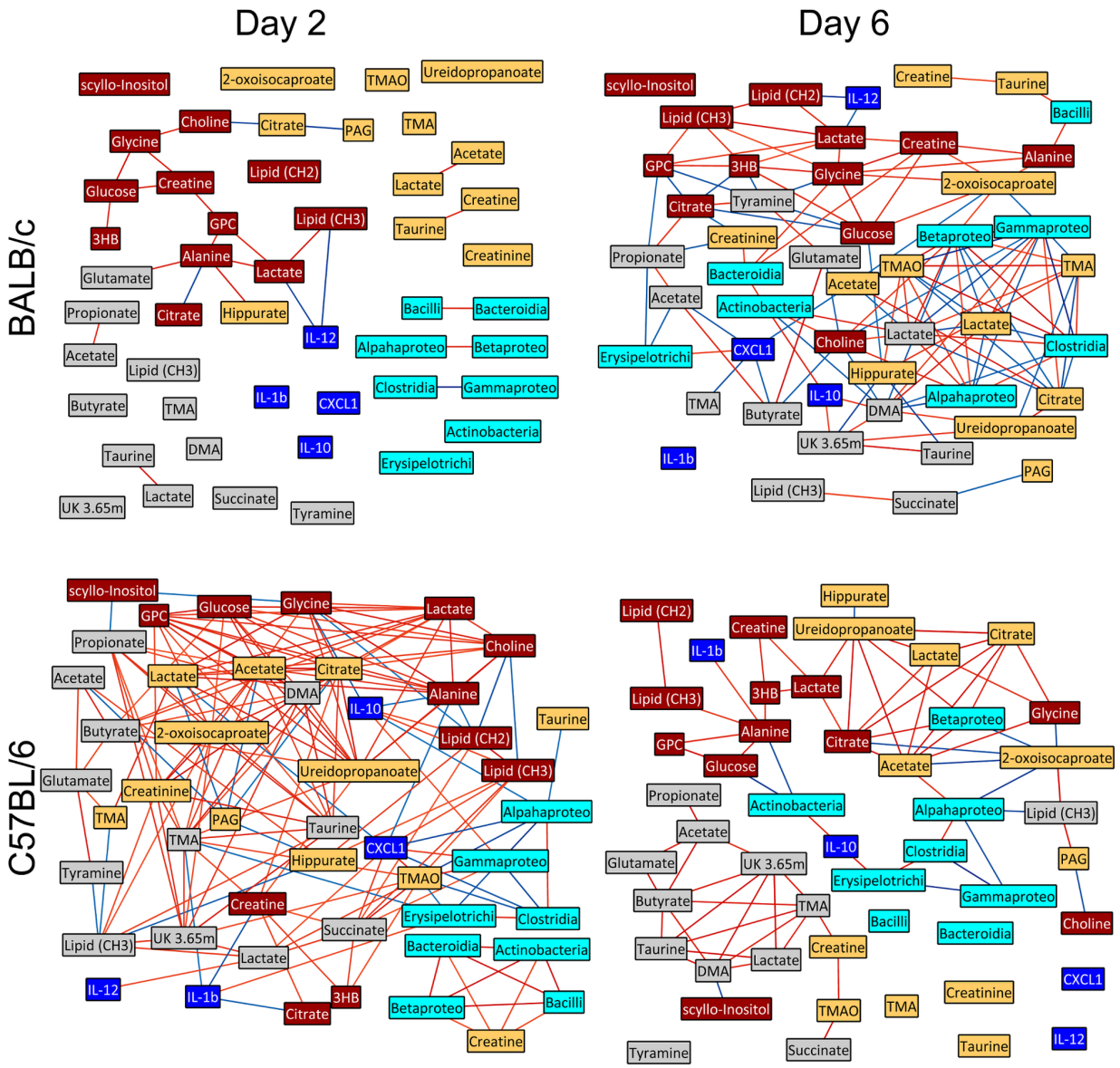


Figure 6. Correlation network map of metabolic, immunologic and microbial markers in response to early infection in two mouse CL models. Correlation networks based on pairwise Pearson correlation between all variables for BALB/c and C57BL/6 mice infected with *L. major* at 2 and 6 days postinfection (see Figure S3 in the Supporting Information for terminal). Red and blue connecting lines between variables represent significant positive or negative correlations, respectively (q < 0.3). Variables are grouped according to correlation clusters and are color-coded in boxes according to their compartments: urine in yellow, plasma in red, cytokines in dark blue, and bacterial classes in cyan. Key: 3HB, 3-hydroxybutyrate; DMA, dimethylamine; GPC, glycerophosphocholine; PAG, phenylacetylglycine; TMA, trimethylamine; TMAO, trimethylamine-*N*-oxide; UK 3.65m, unassigned multiplet peak at δ 3.65.

down.[33] Biochemical blood results from the supplier indicate that BALB/c mice typically have nearly twice as much circulating triglycerides and lower levels of glucose in their blood than C57BL/6 mice.[34,35] This finding was in line with our results, where baseline plasma lipid moieties were also higher in BALB/c mice than in the C57BL/6 strain. However, *L. major* infection induced a pronounced decrease in lipids BALB/c mice only. It may thus be speculated that because of their high bioavailability metabolism of fatty acids represents a major alternative for energy production in BALB/c mice not only in general but also more importantly in response to infection.

Alanine and succinate concentrations were found to be consistently higher across multiple biofluids (in feces and for alanine, also in plasma) in the nonhealing phenotype throughout infection. Both metabolites represent major waste products of *Leishmania* species,[36] and our results showed that succinate was also directly correlated with parasitic burden. Earlier in vitro studies have shown concentrations of alanine and succinate to be significantly higher in cell media from *L.-major*-infected macrophages when compared with uninfected macrophages,[37] highlighting their potential role as biomarkers of *Leishmania* infection. It is thus plausible that these metabolites may in fact be directly derived from the parasite. Urinary hippurate levels decreased upon infection in both

dx.doi.org/10.1021/pr5008202 | *J. Proteome Res.* 2015, 14, 318−329

strains and, upon control-strain adjustment, were consistently found to be lower in the nonhealing infected BALB/C mice when compared with the self-healing strain. Interestingly, a decrease in hippurate has also been observed in previous experimental rodent parasitic models (e.g., in *Trypanosoma brucei brucei* and *Schistosoma mansoni* infections), indicating that this may be a more general marker of parasitic infection.[16,17]

Immunological investigation of the two infected strains confirmed the importance of the innate immune system driving the differential response during early stages of infection, whereby the nonhealing model displayed an inherent proinflammatory phenotype compared with the self-healing strain. In contrast with published evidence,[38] defining the role of Th1 cytokine IL-12 in the current study proved difficult, not only owing to the variable concentrations measured over the course of infection in both strains but also because of its observed positive association with parasitic load.

Following infection, marked changes in fecal bacterial composition were observed, most notably, the expression levels of members of the *Clostridia* class, which were higher in the nonhealing strain, and *Gammaproteobacteria* classes, which were higher in self-healing mice. Our results indicate that the abundances of *Clostridia* and *Gammaproteobacteria* were inversely correlated, a finding that was also reported following weight-loss through bariatric surgery in rats by Li and colleagues.[39] In our study, the two bacterial classes are often linked with the same metabolite but demonstrate contrasting directionality, a phenomenon that was apparent in both strains of mice.

Parasitic load within the lesion of infected mice produced significant correlations with select biomarkers across the different responses assessed in this study. Twice as many correlates were observed in the nonhealing mice. Plasma *scyllo*-inositol, an organic osmolyte that has recently been found to be associated brain disorders,[40] represented the strongest associated variable with parasite counts in this strain.

Correlation network analyses between the host metabolic, cytokine, and fecal microbial responses during infection uniquely demonstrated direct statistical interactions between host components. Remarkable differences in the level of inter- and intramatrix correlations were observed between the two strains and sampling time-points. The self-healing C57BL/6 strain displayed the maximal number of correlations between cytokines, microbiota, and metabolites 2 days after infection. Acetate, citrate, and alanine represented major network hubs in the self-healing strain, implying that these metabolites may prove significant in orchestrating a systemic response toward *Leishmania*.

Acetate represents a key metabolite central to the metabolism of both carbohydrates and fatty acids, used predominantly for the formation of acetyl coenzyme A. Additionally, it is the main short-chain fatty acid readily reabsorbed from the colon back into systemic circulation and thus also acts as a marker of colonic fermentation of dietary carbohydrate.[41] Both citrate (a tricarboxylic acid intermediate) and alanine (a gluconeogenic amino acid via the Cahill pathway[42]) represent markers linked with host energy metabolism. Thus, results from our study emphasize the focal role of carbohydrate metabolism within the self-healing model during infection, presumably to account for the high-energy requirements to mount a successful system response toward the pathogen. In the nonhealing mice, the highest number of associations between variables was observed 6 days post-infection, with few correlations observed at earlier time points. The associations in the nonhealing phenotype appeared to be connected at random (disordered) rather than as functionally structured hubs, as in self-healing mice. It could thus be hypothesized that these highly connected interactions between multiple compartments within the host are vital for mounting a rapid and effective innate immune response toward *L. major*, as seen in the self-healing state.

CONCLUSIONS

In summary, our systemic approach to the characterization of two experimental models of a *L. major* infection adds a novel aspect to the assessment of host defense, which can be applied to other disease models. Tangible differences between the infection models included differences in energy metabolism, inflammatory status, and gut microbial composition. Furthermore, our study revealed parasitic-linked metabolites that were higher in the nonhealing model and provide novel candidates to be further investigated as potential biomarkers of *Leishmania* infection. Globally we have shown that already during the innate immune response an efficiently orchestrated network of metabolic, immunological, and gut microbial features may be linked to a successful pathogen defense. Infection in the self-healing model displayed a comprehensive correlation network between multiple biological responses toward the parasite that was not present in the nonhealing model, which may explain the superiority of this strain in controlling disease. We have also identified a suite of metabolic candidates that represented major network hubs linked to a successful immune response to *L. major*, which can be further explored to ascertain potential indices for immune-modulatory or leishmanicidal activity.

ASSOCIATED CONTENT

Supporting Information

Figure S1. *L. major* infection-related changes in cytokine levels in plasma from two CL models. Figure S2. Statistical interconnectivity of metabolic, immune, and microbial markers in mouse models prior to infection. Figure S3. Correlation network map of metabolic, immune and microbial markers of CL models (terminal day). Table S1. Physiological and pathological measures of CL mouse models. Table S2. Infection-related discriminatory metabolic markers between the two CL models. Table S3. Infection-related differences in fecal microbiotal composition between the two CL models. Table S4. Infection-related differences in fecal microbiotal composition over time within each CL model. Table S5. Microbial-metabolic correlates in urine of CL models. Supporting Protocols. This material is available free of charge via the Internet at http://pubs.acs.org.

AUTHOR INFORMATION

Corresponding Author

*Tel: +44 20 7594-3899. E-mail: jasmina.saric@imperial.ac.uk.

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

We thank the Wellcome Trust for financial support of the research presented in the current manuscript (Sir Henry Wellcome Fellowship JS, Award number 089002/B/09/Z) and

the MRC and Imperial College for the MRC Doctoral Training Award for S.D.L.'s studentship funds. K.A.V. acknowledges the junior research fellowship scheme by Imperial College London. We thank The Royal Society for funding the compute cluster HIVE at Cardiff University and the Biotechnology and Biological Sciences Research Council for support for MER and EG (David Phillips Fellowship MER, award number BB/H022406/1). The funding bodies had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

■ ABBREVIATIONS

CL, cutaneous leishmaniaisis; CXCL, chemokine (C-X-C motif) ligand; DALY, disability-adjusted life year; IL, interleukin; PCR, polymerase chain reaction; Th, T helper cell

■ REFERENCES

(1) Desjeux, P. *Nat. Rev. Microbiol.* **2004**, *9*, 692−693.
(2) Alvar, J.; Velez, I. D.; Bern, C.; Herrero, M.; Desjeux, P.; Cano, J.; Jannin, J.; den Boer, M. WHO Leishmaniasis Control Team. *PLoS One* **2012**, *5*, e35671.
(3) Croft, S. L.; Olliaro, P. *Clin. Microbiol. Infect.* **2011**, *10*, 1478−1483.
(4) Scott, P. *J. Immunol.* **1991**, *9*, 3149−3155.
(5) Sypek, J. P.; Chung, C. L.; Mayor, S. E.; Subramanyam, J. M.; Goldman, S. J.; Sieburth, D. S.; Wolf, S. F.; Schaub, R. G. *J. Exp. Med.* **1993**, *6*, 1797−1802.
(6) Sadick, M. D.; Heinzel, F. P.; Holaday, B. J.; Pu, R. T.; Dawkins, R. S.; Locksley, R. M. *J. Exp. Med.* **1990**, *1*, 115−127.
(7) Kopf, M.; Brombacher, F.; Köhler, G.; Kienzle, G.; Widmann, K. H.; Lefrang, K.; Humborg, C.; Ledermann, B.; Solbach, W. *J. Exp. Med.* **1996**, *3*, 1127−1136.
(8) Belkaid, Y.; Piccirillo, C. A.; Mendez, S.; Shevach, E. M.; Sacks, D. L. *Nature* **2002**, *6915*, 502−507.
(9) Woelbing, F.; Kostka, S. L.; Moelle, K.; Belkaid, Y.; Sunderkoetter, C.; Verbeek, S.; Waisman, A.; Nigg, A. P.; Knop, J.; Udey, M. C.; von Stebut, E. *J. Exp. Med.* **2006**, *1*, 177−188.
(10) Peters, N. C.; Egen, J. G.; Secundino, N.; Debrabant, A.; Kimblin, N.; Kamhawi, S.; Lawyer, P.; Fay, M. P.; Germain, R. N.; Sacks, D. *Science* **2008**, *5891*, 970−974.
(11) Iniesta, V.; Carcelen, J.; Molano, I.; Peixoto, P. M. V.; Redondo, E.; Parra, P.; Mangas, M.; Monroy, I.; Campo, M. L.; Gomez Nieto, C.; Corraliza, I. *Infect. Immun.* **2005**, *9*, 6085−6090.
(12) Kropf, P.; Fuentes, J. M.; Fahnrich, E.; Arpa, L.; Herath, S.; Weber, V.; Soler, G.; Celada, A.; Modolell, M.; Muller, I. *FASEB J.* **2005**, *8*, 1000−1002.
(13) Li, P.; Yin, Y.; Li, D.; Woo Kim, S.; Wu, G. *Br. J. Nutr.* **2007**, *02*, 237−252.
(14) Saric, J. *Parasitology* **2010**, No. Special Issue 09, 1451.
(15) Ferrante, A. W., Jr. *J. Clin. Invest.* **2013**, *12*, 4992−4993.
(16) Wang, Y.; Holmes, E.; Nicholson, J. K.; Cloarec, O.; Chollet, J.; Tanner, M.; Singer, B. H.; Utzinger, J. *Proc. Natl. Acad. Sci. U. S. A.* **2004**, *34*, 12676−12681.
(17) Wang, Y.; Utzinger, J.; Saric, J.; Li, J. V.; Burckhardt, J.; Dirnhofer, S.; Nicholson, J. K.; Singer, B. H.; Brun, R.; Holmes, E. *Proc. Natl. Acad. Sci. U. S. A.* **2008**, *16*, 6127−6132.
(18) Dumas, M.; Maibaum, E. C.; Teague, C.; Ueshima, H.; Zhou, B.; Lindon, J. C.; Nicholson, J. K.; Stamler, J.; Elliott, P.; Chan, Q.; Holmes, E. *Anal. Chem.* **2006**, *7*, 2199−2208.
(19) Schreier, C.; Kremer, W.; Huber, F.; Neumann, S.; Pagel, P.; Lienemann, K.; Pestel, S. *Biomed. Res. Int.* **2013**, 878374.
(20) Saric, J.; Li, J. V.; Swann, J. R.; Utzinger, J.; Calvert, G.; Nicholson, J. K.; Dirnhofer, S.; Dallman, M. J.; Bictash, M.; Holmes, E. *J. Proteome Res.* **2010**, *5*, 2255−2264.
(21) Saric, J.; Li, J. V.; Wang, Y.; Keiser, J.; Bundy, J. G.; Holmes, E.; Utzinger, J. *PLoS Neglected Trop. Dis.* **2008**, *7*, e254.
(22) Oliveira, M. R.; Tafuri, W. L.; Afonso, L. C.; Oliveira, M. A.; Nicoli, J. R.; Vieira, E. C.; Scott, P.; Melo, M. N.; Vieira, L. Q. *Parasitology* **2005**, *4*, 477−488.
(23) Beckonert, O.; Keun, H. C.; Ebbels, T. M.; Bundy, J.; Holmes, E.; Lindon, J. C.; Nicholson, J. K. *Nat. Protoc.* **2007**, *11*, 2692−2703.
(24) Dieterle, F.; Ross, A.; Schlotterbeck, G.; Senn, H. *Anal. Chem.* **2006**, *13*, 4281−4290.
(25) Lewis, D. A.; Brown, R.; Williams, J.; White, P.; Jacobson, S. K.; Marchesi, J.; Drake, M. J. *Front. Cell. Infect. Microbiol.* **2013**, *41*, 10.3389/fcimb.2013.00041.
(26) de La Llave, E.; Lecoeur, H.; Besse, A.; Milon, G.; Prina, E.; Lang, T. *Cell. Microbiol.* **2011**, *1*, 81−91.
(27) Solcà, M. d. S.; Guedes, C. E. S.; Nascimento, E. G.; Oliveira, G. G. d. S.; Santos, W. L. C. d.; Fraga, D. B. M.; Veras, P. S. T. *Vet. Parasitol.* **2012**, *2−4*, 133−140.
(28) Hotelling, H. *J. Educ. Psychol.* **1933**, *6*, 417.
(29) Trygg, J.; Holmes, E.; Lundstedt, T. *J. Proteome Res.* **2007**, *2*, 469−479.
(30) Posma, J. M.; Robinette, S. L.; Holmes, E.; Nicholson, J. K. *Bioinformatics* **2014**, *30* (6), 893−895.
(31) Storey, J. D.; Tibshirani, R. *Proc. Natl. Acad. Sci. U.S.A.* **2003**, *16*, 9440−9445.
(32) Janeway, C. A. J.; Travers, P.; Walport, M.; Shlomchik, M. J. Principles of Innate and Adaptive Immunity. In *Immunobiology: The Immune System in Health and Disease*; Garland Science: New York, 2001.
(33) Morris, A. A. M. *J. Inherited Metab. Dis.* **2005**, *2*, 109−121.
(34) Charles River Laboratories Clinical Pathology Data for BALB/c Mouse Colonies in North America for January 2008 - December 2012. http://www.criver.com/files/pdfs/rms/balbc/rm_rm_r_balb-c_mouse_clinical_pathology_data.aspx (accessed August 5, 2014).
(35) Charles River Laboratories Biochemistry and Hematology for C57BL/6NCrl Mouse Colonies in North American for January 2008 − December 2012. http://www.criver.com/files/pdfs/rms/c57bl6/rm_rm_r_c57bl6_mouse_clinical_pathology_data.aspx (accessed August 5, 2014).
(36) McConville, M. J.; de Souza, D.; Saunders, E.; Likic, V. A.; Naderer, T. *Trends Parasitol.* **2007**, *8*, 368−375.
(37) Lamour, S. D.; Choi, B.; Keun, H. C.; Müller, I.; Saric, J. *J. Proteome Res.* **2012**, *8*, 4211−4222.
(38) Sacks, D.; Noben-Trauth, N. *Nat. Rev. Immunol.* **2002**, *11*, 845.
(39) Li, J. V.; Ashrafian, H.; Bueter, M.; Kinross, J.; Sands, C.; le Roux, C. W.; Bloom, S. R.; Darzi, A.; Athanasiou, T.; Marchesi, J. R.; Nicholson, J. K.; Holmes, E. *Gut* **2011**, *9*, 1214−1223.
(40) Griffith, H. R.; den Hollander, J. A.; Stewart, C. C.; Evanochko, W. T.; Buchthal, S. D.; Harrell, L. E.; Zamrini, E. Y.; Brockington, J. C.; Marson, D. C. *NMR Biomed.* **2007**, *8*, 709−716.
(41) Hijova, E.; Chmelarova, A. *Bratisl. Lek. Listy* **2007**, *8*, 354−358.
(42) Felig, P.; Pozefsk, T.; Marlis, E.; Cahill, G. F. *Science* **1970**, *3920*, 1003−1004.

Lamour et al. Malaria Journal 2014, 13:468
http://www.malariajournal.com/content/13/1/468

MALARIA JOURNAL

RESEARCH Open Access

Changes in metabolic phenotypes of *Plasmodium falciparum in vitro* cultures during gametocyte development

Sabrina D Lamour[1], Ursula Straschil[2], Jasmina Saric[1] and Michael J Delves[2*]

Abstract

Background: Gametocytes are the *Plasmodium* life stage that is solely responsible for malaria transmission. Despite their important role in perpetuating malaria, gametocyte differentiation and development is poorly understood.

Methods: To shed light on the biochemical changes that occur during asexual and gametocyte development, metabolic characterization of media from *in vitro* intra-erythrocytic *Plasmodium falciparum* cultures was performed throughout gametocyte development by applying ^{1}H nuclear magnetic spectroscopy, and using sham erythrocyte cultures as controls. Spectral differences between parasite and sham cultures were assessed via principal component analyses and partial-least squares analyses, and univariate statistical methods.

Results: Clear parasite-associated changes in metabolism were observed throughout the culture period, revealing differences between asexual parasites and gametocyte stages. With culture progression and development of gametocytes, parasitic release of the glycolytic end products lactate, pyruvate, alanine, and glycerol, were found to be dramatically reduced whilst acetate release was greatly increased. Also, uptake of lipid moieties CH_2, CH_3, and $CH=CH\text{-}CH_2\text{-}CH_2$ increased throughout gametocyte development, peaking with maturity.

Conclusions: This study uniquely presents an initial characterization of the metabolic exchange between parasite and culture medium during *in vitro P. falciparum* gametocyte culture. Results suggest that energy metabolism and lipid utilization between the asexual stages and gametocytes is different. This study provides new insights for gametocyte-specific nutritional requirements to aid future optimization and standardization of *in vitro* gametocyte cultivation, and highlights areas of novel gametocyte cell biology that deserve to be studied in greater detail and may yield new targets for transmission-blocking drugs.

Keywords: *Plasmodium*, Gametocyte, Gametocytogenesis, Culture, Metabolic, Asexual, Sexual, Maturation, Differentiation, Nutritional

Background

Plasmodium, the causative agent of malaria was reported to be responsible for 627,000 deaths in 2012 [1]. Disease symptoms are largely caused by progressive cycles of parasite invasion of erythrocytes in the bloodstream through what is known as the asexual cycle. With every 48-hour asexual cycle, ~0.1-0.2% of *Plasmodium falciparum* asexual parasites become committed to an alternative developmental path, forming the sexual stages of the parasite called gametocytes [2]. The trigger for this transformation is not well understood, however it is thought to be upregulated in response to stress signals induced by high parasitaemia in the human host [3] and is regulated by the Api2G transcription factor [4,5]. Gametocytes are the *Plasmodium* life stage solely responsible for onward transmission to the mosquito and do not participate in disease symptoms [6].

In *P. falciparum*, gametocytes take about ten days to develop to maturity [2]. During this time they undergo distinct morphological changes defined as Stages I-V. Stages I-III are thought to share similar metabolic activity to asexual parasites, e.g., utilizing erythrocyte haemoglobin as a major amino acid source. Furthermore, they

* Correspondence: michael.delves@imperial.ac.uk
[2]Department of Life Sciences, Imperial College London, London SW7 2AZ, UK
Full list of author information is available at the end of the article

BioMed Central

© 2014 Lamour et al.; licensee BioMed Central Ltd. This is an Open Access article distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by/4.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly credited. The Creative Commons Public Domain Dedication waiver (http://creativecommons.org/publicdomain/zero/1.0/) applies to the data made available in this article, unless otherwise stated.

also share similar sensitivity to chloroquine treatment which inhibits haem detoxification [7]. As gametocytes mature, they dramatically reduce their activity and become quiescent, waiting for uptake into the mosquito during a blood feed before continuing development with explosive colonization of the mosquito midgut [8]. Likely due to their arrested state, mature Stage V gametocytes are insensitive to most currently used anti-malarials [9-11]. This leads to a situation where an infected individual can be cured of their disease symptoms but still be infectious to mosquitoes, thus able to perpetuate the cycle of infection. The complex and changing cell biology of the gametocyte highlights the utmost importance of better understanding gametocyte development *in vitro,* to best enable the development and implementation of transmission-blocking interventions.

The laboratory study of *P. falciparum* asexual parasites has been greatly facilitated by their ability to be continuously cultured *in vitro* if supplied regularly with fresh erythrocytes to invade, and supported by appropriate medium [12]. Although simplified protocols for gametocyte cultivation exist and external factors have been somewhat optimized, the infectivity of these cultured gametocytes varies wildly both between experiments [13] and externally between research groups, despite using similar protocols [14]. Considering the differences in length of development, cell biology and extracellular environment that asexual parasites and gametocytes experience *in vivo*, it is likely that the metabolic activity and nutritional requirements of asexual parasites and developing gametocytes is different.

Metabolic profiling studies, using spectroscopic methods such as nuclear magnetic resonance (NMR), mass spectrometry or capillary electrophoresis-based methods, provide powerful tools in capturing the relative levels of whole variety of small metabolites simultaneously. Such work has initially been applied to a variety of *in vivo* models, resulting in the discovery of complex metabolic host-parasite interactions and candidate biomarkers associated with infection and/or disease, including several malaria rodent models [15,16]. More recently, this approach has been applied to *in vitro* infection models to address more targeted metabolic questions on specific infection-induced changes in host cells [17,18] or responses towards drug treatment [19], and has been increasingly implemented to characterize blood-stage *P. falciparum* cultures [20-23].

Here a standard ^{1}H NMR metabolic profiling approach is used to characterize changes in medium over the entire course of *P. falciparum* gametocyte development *in vitro*, from the generation of gametocytes from an initial asexual population until their final maturation stage. The aim of this explorative study is to identify changes in metabolic activity throughout culture and identify key metabolic markers associated with growth and/or differentiation of the various gametocyte stages, which may shed light on optimizing and standardizing *in vitro P. falciparum* gametocyte culture.

Methods

Study design

The study was based on two separate experiments. In the first study, eight replicate parasite cultures were grown for 14 days until functionally mature Stage V gametocytes (as evidenced by their ability to undergo gamete formation) predominated. Cultures were induced by seeding *P. falciparum* 3D7 strain asexual parasites at 1% parasitaemia and 4% haematocrit in 10 ml total volume in 25 sq cm flasks under 3% O_2, 5% CO_2, 92% N_2 gas. Culture medium consisted of RPMI medium supplemented with 25 mM HEPES (Life Technologies), 50 mg/l hypoxanthine (Sigma), 2 g/l sodium bicarbonate (Sigma), 10% A+ pooled human serum (Interstate Bloodbank). Total medium replacement was performed every 24 hours whilst maintaining a culture temperature of 37°C. Aliquots of fresh culture medium not exposed to parasites were used as a baseline control ($n = 8$). Each day, during routine daily medium replacement, a 1 ml sample of spent medium from each culture was collected. Samples were then centrifuged at 12,000 × *g* for 5 minutes to pellet any contaminating cell debris and supernatants were transferred to fresh tubes and stored at -80°C until data acquisition.

To control for any potential changes in the baseline metabolism of the uninfected erythrocyte over the 14-day culture period which could confound the analysis of the parasite metabolism, a second sham experiment was performed in which uninfected erythrocytes were cultured identically as for the first experiment. Culture media samples were collected on days 1, 4, 7, 8, and 14 of the study ($n = 8$ per time-point, 40 in total), along with aliquots of the fresh medium, taken at the same time-point as for the control erythrocytes ($n = 40$ in total). Samples were processed identically as described for the first study. In parallel, three gametocyte cultures were initiated, as before, and thin blood smears taken on days 1, 4, 7, 8, and 14. Smears were stained with Giemsa and parasite stages were counted and recorded microscopically. The mean of five fields at x500 magnification was calculated for each slide and parasitaemia expressed as a percentage of total erythrocyte number.

^{1}H NMR acquisition and processing

Media samples were prepared for NMR analyses based on protocols as described by Beckonert and colleagues [24]. Briefly, 300 μl media samples were mixed 1:1 with NMR phosphate buffer (50% v/v D20 (GOSS Scientific, UK, 0.01% v/v sodium 3-(trimethylsilyl) propionic acid 2,2,3,3-d4 ([TSP)], pH 7.4) and centrifuged at 12,470 × *g*

for 5 minutes. Some 550 μl of sample was transferred into 5-mm NMR tubes (Bruker, Germany) and all samples from both studies were run on a Bruker Avance 600 NMR Spectrometer with TXI probe head (Bruker), using XWIN-NMR software (Bruker Biospin, Germany). ^{1}H NMR data were acquired by applying a standard one-dimensional (1D) pulse programme for 128 scans (after eight dummy scans), that included water irradiation during the recycle delay, set at 2 seconds (s). The pulse sequence was set to: recycle delay-90°-t-90°-tm-90°-ACQ, whereby 90° pulse length was set to between 16.5 μs, t (short delay) = 2 s, tm (mixing time) = 100 ms and ACQ (acquisition period) = at 2.73 s per scan. Spectral data underwent baseline correction, internal reference (TSP) peak calibration and phasing, using an in-house MATLAB algorithm (version R2012b, Mathworks Inc, USA) and Topspin 3.1 software (Bruker BioSpin, Germany). Water regions and HEPES buffer peaks were removed followed by to automatic spectral alignment and probabilistic quotient normalization [25] in MATLAB.

Statistical data analysis

Metabolic spectra of raw data for each study, as well as data of the two combined studies, were analysed via principal component analysis (PCA), an unsupervised pattern-recognition model used to show an unbiased overview of overall spectral variability of the dataset [26]. Using PCA, the total spectral information of each individual sample becomes down-projected as a single coordinate on the PCA scores plot. Metabolic information on those components responsible for driving the groupings/separations in the scores plot can be derived from the corresponding PCA loadings plot. Data were then subjected to partial-least squares discriminatory analyses (PLS-DA), a supervised model that incorporates class assignment, in order to distinguish any subtle metabolic differences between groups [26]. Both PCA and PLS-DA were performed with SIMCA P+ software (Version 13.0, Umetrics, Sweden).

Specific spectral differences between media from parasite samples and fresh media were further analysed using orthogonal PLS-DA [26], to assess metabolites that were consumed/released into the media during parasite culture. Significantly differential metabolites were selected based on their peak having a Pearson-product moment correlation coefficient (R) that surpassed 0.5, corresponding to the approximate minimum critical values for $p < 0.05$, and ROC sensitivity/specificity score of at least 0.8. Metabolites were assigned using NMR suite profiler 7.0 software (Chenomx, USA) and referring to inhouse NMR databases and RPMI media formulation sheets from supplier.

Levels of differential metabolites were calculated by measuring metabolite peak integrals across parasite cultures, uninfected erythrocytes and fresh media samples. To account for media batch differences between the two studies, mean average integrals of fresh culture media from each study were subtracted from each metabolite integral of their respective cultures, for each time-point. Media-adjusted metabolite integrals from parasite and erythrocyte cultures were compared for each metabolite, at each matching time-point, by applying non-parametric Mann-Whitney U-test, followed by Benjamini-Hochberg false discovery rate correction [27], to determine significant changes in metabolism linked with *P. falciparum* infection. Parasite-associated changes in metabolites (i.e., media-adjusted levels from gametocytes minus media-adjusted levels from time-matched control erythrocytes) were subsequently divided by total parasitaemia percentages for each corresponding time-point.

Results

Parasite developmental changes are reflected by alterations in metabolism

Media from parasite cultures were collected every 24 hours throughout the 14-day culture period until gametocyte maturation, and metabolically phenotyped using ^{1}H NMR spectroscopy. PCA analysis showed a clear separation between the cultures and fresh medium (Figure 1A), which according to the loadings was driven primarily by higher glucose levels in the fresh medium (Additional file 1). Furthermore, striking changes were also observed between time-points within the cultures. From Day 1 the model showed a strong shift in metabolic phenotype along the first component, relative to fresh medium, until Day 4 of culture. Parasitaemia counts revealed that Day 4 corresponded to the peak in asexual population (Figure 1B and C) as well as initial appearance of 'early' gametocytes (Stage I-II). As the asexual parasite populations reduced from 8.1% (± 1.3% standard error, SEM) at Day 4 to 3.6% (± 0.8% SEM) by Day 7, early gametocyte numbers continued to increase. During this period, the overall culture metabolism reverted back, similar to the initial phenotype (Figure 1A).

Interestingly, a different metabolic shift was observed from Day 8 onwards, correlating with continued gametocyte generation and the emergence of intermediate stage (III-IV) gametocytes (Figure 1B and C). Metabolic profiles from Days 12 to 14 largely clustered together, with little difference to fresh medium along the first component, thus confirming that the final stages of gametocyte maturation share a similar metabolic activity. By Day 14, asexual parasites were completely absent from the culture and most gametocytes were mature (Stage V; see Figure 1C), resulting in a final mature gametocytaemia of 2.8% (± 0.8% SEM). Corresponding PCA loadings revealed that release of lactate and pyruvate was closely linked with the early shift during Days 1-4, whilst acetate

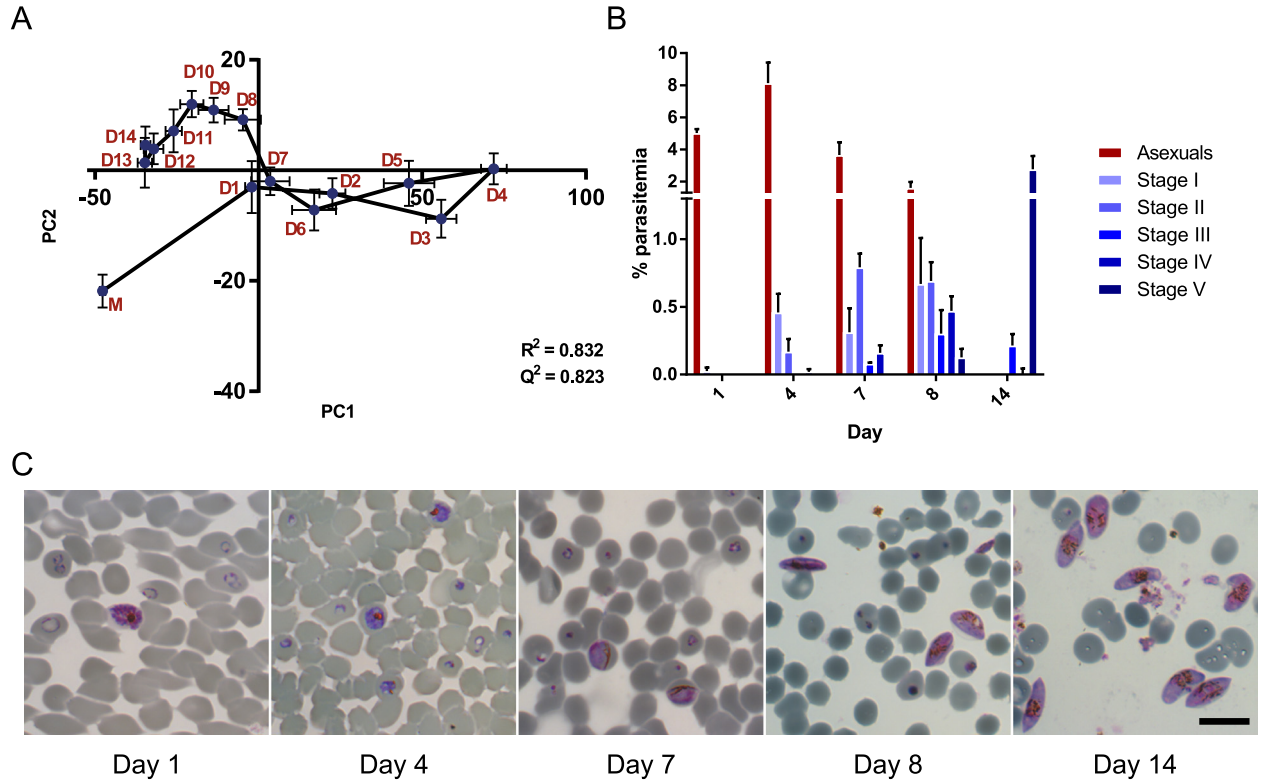


Figure 1 The relationship between metabolic and developmental changes during *in vitro* parasite culture. (A) Metabolic time-trajectory showing the first two components (PC1 and PC2) of a PCA scores plot on fresh culture medium (M) and parasite-conditioned medium from intra-erythrocytic *P. falciparum* cultures, throughout the 14-day culture period (D1-D14). $n = 8$ samples for each time-point. Coordinates and error bars represent mean averages for each time point and standard error of the mean (SEM), respectively. R^2 and Q^2 represent the values of the degrees of model fit and predictive values of the PCA model for components 1 and 2, respectively. **(B)** Parasite populations at culture days representing major shifts in trajectory (D1, D4, D7, D8, D14) were assessed by manual counting of Giemsa-stained culture smears (n = 3 cultures, showing mean ± SEM). Days 1-4 were typified by an increasing asexual population; Days 4-7 showed a decline in asexuals and emergence of early gametocytes; Days 7-14 showed gametocyte maturation; by Day 14, the culture consisted mainly of mature Stage V gametocytes. **(C)** Representative microscope images showing parasite culture composition at the culture days assessed. Scale bar = 10 μm.

production was associated with the later stages of development to maturity (Additional file 1).

Plasmodium falciparum infection associated with increased intra-erthrocytic metabolic activity

To adjust for the potential basal metabolic changes of uninfected erythrocytes, which comprise the large bulk of the parasite cultures, sham cultures of uninfected erythrocytes were generated using identical conditions. Having established that major metabolic changes occur at Days 1, 4, 7, 8, and 14 of parasite culture (Figure 1), medium from these time-points was selected for further analysis. Spectra of parasite cultures, uninfected erythrocytes and fresh culture medium from all assessed time-points were analysed using PCA. Gametocyte cultures clearly separated from the other two groups, whilst control erythrocytes and fresh medium broadly co-localized (Figure 2A). This suggests that a relatively low background metabolic activity was exerted by uninfected erythrocytes, with only a minor contribution to the total observed metabolic activity in the parasite cultures. Further investigation with PLS-DA showed an enhanced separation of parasite cultures from the other samples but also revealed partial separation between the uninfected cultures and fresh medium (Figure 2B). Additional analyses on fresh medium samples alone via PCA revealed subtle batch differences, particularly between the two studies (see Additional file 2) and was thus accounted for in subsequent comparative analyses between the gametocyte and control erythrocyte cultures (see Methods).

Differential metabolic markers are associated with asexual growth and gametocyte maturation

Multivariate discriminatory analysis (using orthogonal PLS-DA) was performed between fresh medium and spent medium of parasite cultures, to identify metabolites that were consumed and/or released during culture. In order to separate parasite-associated changes in metabolism from that of uninfected erythrocytes, integrals of these were calculated and differences between time-

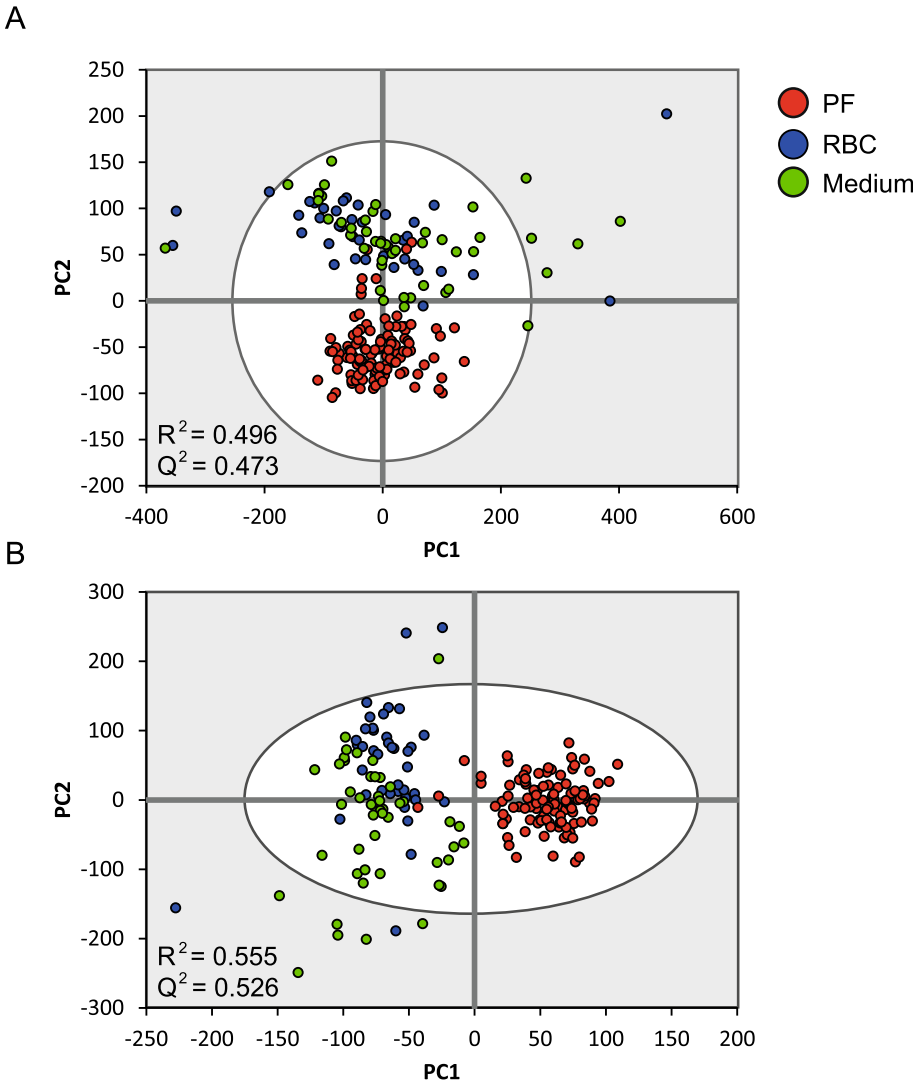


Figure 2 Variance of dataset dominated by infection status. PCA model **(A)** and PLS-DA model **(B)** score plots of ^{1}H NMR spectra across all time-points from both studies. *Plasmodium falciparum* parasite cultures (PF) in red ($n = 112$), uninfected erythrocyte cultures (RBC) in blue ($n = 40$) and fresh media samples are coloured green ($n = 48$). Cumulative values of model fit (R^2) and model prediction (Q^2) for the first two components are denoted on both plots.

point-matched parasite and control erythrocyte cultures were investigated (Figure 3 and Table 1).

One of the most notable changes appears to be linked to glucose utilization. Glucose consumption was significantly higher in parasite cultures from Days 4-14, peaking at Day 4 (Figure 3B), consistent with maximal parasitaemia (Figure 3A). Following normalization to parasite counts, relative stable glucose consumption from Days 4-14 was observed (Additional file 3), suggesting that asexuals and gametocytes of all stages of development have similar glucose requirements. Known metabolites of glucose utilization in *Plasmodium* asexual parasites – alanine, glycerol, lactate, and pyruvate [28] were all maximally released by the parasites at Day 4, before declining to the basal erythrocyte level (or below) by Day 14 (Figure 3B). When total culture parasitaemia was accounted for, all still showed a maximal peak of release at Day 4, with little to no release by Day 14 (Additional file 3). In contrast to other glycolytic end-products, acetate, which has been linked to glucose utilization in gametocytes [22] was released by the parasites only at very low levels early in culture, was maximal later in culture, before slightly declining again at Day 14 (Figure 3B). This trajectory was mirrored even after normalizing for parasitaemia (Additional file 3).

The amino acids histidine, phenylalanine, tyrosine, valine, and leucine were all maximally released into the medium at Day 4 before sharply declining to basal erythrocyte levels by

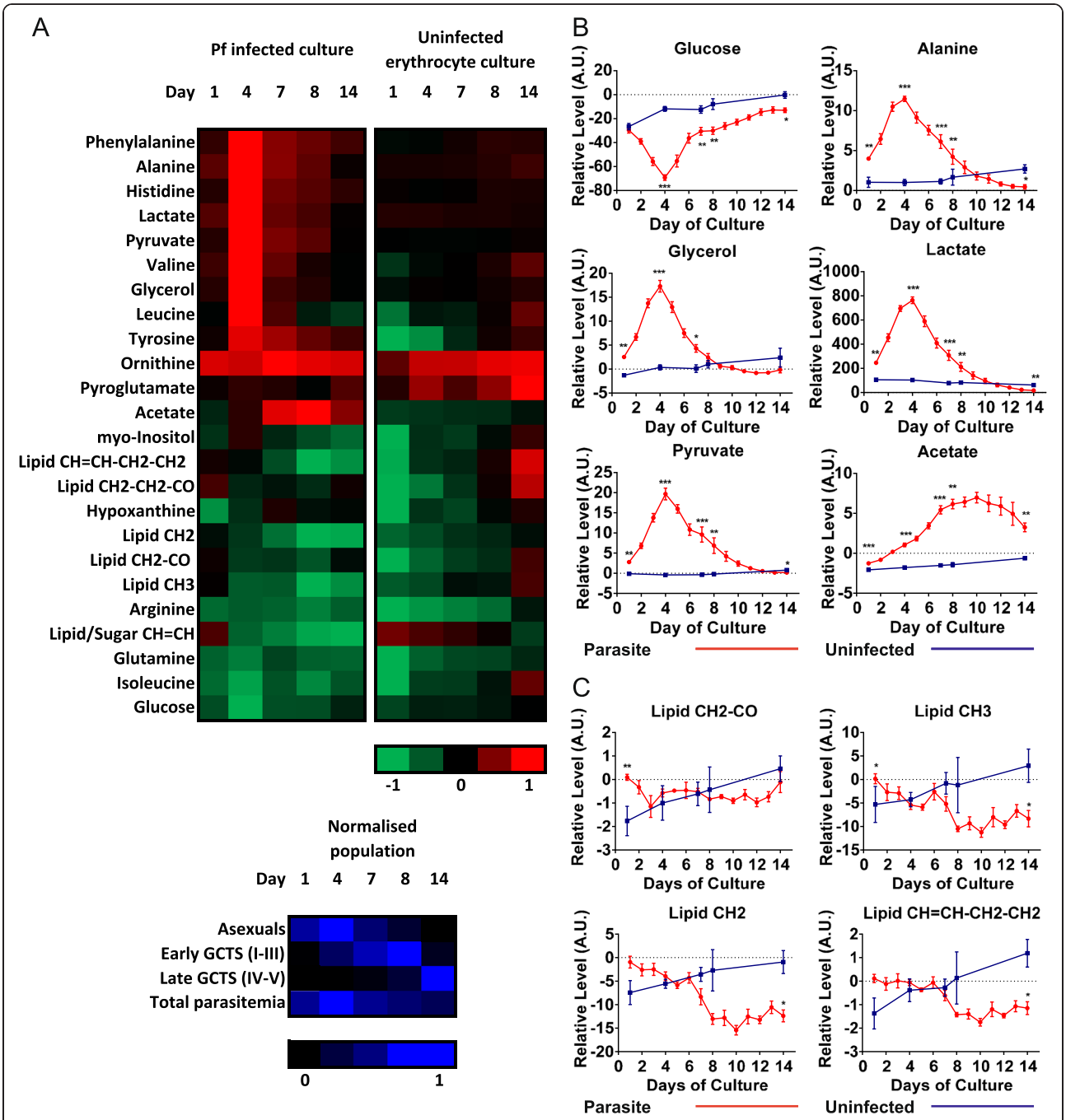


Figure 3 Metabolites associated with different developmental stages of *Plasmodium falciparum* within *in vitro* gametocyte cultures. **(A)** Medium metabolite levels for *P. falciparum* (Pf) cultures and uninfected erythrocyte cultures sampled during culture progression. Red indicates metabolite consumed from medium, green indicates released into medium. For comparison, culture parasite stage composition for the different time points is indicated in blue; GCT, gametocytes. **(B)** Culture metabolite trajectories for glucose and reported *Plasmodium* glycolytic endpoints. **(C)** Culture metabolite trajectories for selected lipid moieties. Graphs show means ± SEM as error bars. *P < 0.05; **P < 0.01; ***P < 0.001 Mann-Whitney U test.

Day 14 (Figure 3A, Additional file 3). Accounting for total parasitaemia, peak histidine and phenylalanine release occurred at Day 4 and gradually declined. Leucine, tyrosine and valine release was relatively stable at Days 1-7 before sharply dropping to basal erythrocyte levels or below at Days 8-14 (Additional file 3). In contrast, isoleucine, an amino acid known to be essential for parasite growth [29], was consumed at all timepoints, with statistical significance observed on Days 4 and 14. When adjusted for culture parasitaemia, isoleucine consumption was maximal at Day 14 (Figure 3A, Additional file 3).

Table 1 Metabolites significantly altered by *Plasmodium falciparum* infection throughout gametocyte culture period

	Metabolite	*p*-values				
		Day 1	Day 4	Day 7	Day 8	Day 14
Metabolites released into medium	Acetate	0.0046	0.0005	0.0008	0.0024	0.0024
	Alanine	0.0012	0.0005	0.0008	NS	0.0188 (-)
	Glycerol	0.0012	0.0005	0.0153	NS	NS
	Histidine	0.0129	0.0005	0.0008	0.0036	NS
	Lactate	0.0012	0.0005	0.0008	0.0024	0.0024 (-)
	Ornithine	0.0129	NS	NS	NS	NS
	Phenylalanine	0.0014	0.0005	0.0008	0.0188	NS
	Pyroglutamate	NS	NS	NS	0.0444 (-)	0.0210 (-)
	Pyruvate	0.0012	0.0005	0.0008	0.0024	0.0250 (-)
	Tyrosine	0.0046	0.0007	0.0090	NS	NS
	Valine	0.0033	0.0005	0.0153	NS	NS
Metabolites consumed from medium	Arginine	0.0166 (-)	NS	NS	NS	NS
	Glucose	NS	0.0005	0.0038	0.0091	0.0152
	Glutamine	0.0033 (-)	NS	NS	NS	NS
	Hypoxanthine	NS	NS	0.0125 (-)	NS	NS
	myo-Inositol	NS	NS	NS	NS	0.0210 (-)
	Isoleucine	NS	0.0012	NS	NS	0.0250
	Lipid/Sugar CH = CH	NS	NS	NS	NS	NS
	Lipid CH = CH-CH2-CH2	NS	NS	NS	NS	0.0414
	Lipid CH2	NS	NS	NS	NS	0.0414
	Lipid CH2-CH2-CO	0.0222 (-)	NS	NS	NS	NS
	Lipid CH2-CO	0.0033 (-)	NS	NS	NS	NS
	Lipid CH3	0.0397 (-)	NS	NS	NS	0.0180
Variable consumption/release	Leucine	0.0065 (-)	0.0005	0.0296	NS	0.0188 (-)

Statistical differences between metabolite levels of medium from parasite cultures and uninfected erythrocytes, assessed for each time-point as measured via Mann-Whitney U-test. *P*-values indicate significantly higher metabolite release/consumption by parasite cultures, unless marked with (-), where release/consumption was lower than erythrocytes. NS, no significant difference.

Generally, identified lipid moieties fell into two categories: those that showed little or no difference from basal erythrocyte consumption levels, such as CH_2-CO, CH_2-CH_2-CO and CH = CH; and those that increased in consumption with gametocyte maturation from Day 8 onwards (Figure 3A and C, Additional file 3). The latter category included lipid methyl groups (CH_3), methylene groups (CH_2,) and CH = CH-CH_2-CH_2 fractions (Figure 3C), whereby the increase in consumption was significant by Day 14 of culture (Table 1).

Discussion

The generation of *P. falciparum* gametocytes *in vitro* involves the initial rapid growth of intra-erythrocytic asexual parasites, sexual commitment, which is then followed by gradual maturation [3]. This study surveys the extracellular metabolic changes throughout gametocyte culture development through analysis of spent culture medium *via* ^{1}H NMR spectroscopy. Using this approach has identified clear differences in metabolic activity between asexual and gametocyte growth which is broadly consistent with other studies [22,28].

The most distinct difference observed appeared to be associated with parasite glucose utilization. Given the abundance of glucose in serum, the asexual parasite relies entirely on glycolysis for energy production [30]. Four previously confirmed glycolytic fermentation end-products (lactate, pyruvate, alanine, and glycerol [28]) were identified as released into the culture medium coinciding with the peak asexual population. Whilst culture glucose consumption continued throughout gametocyte maturation, as the asexual population declined the release of these metabolites decreased and was replaced by the production of acetate. Acetate is a common glycolytic end-product of other eukaryotic parasites [31] and has been recently linked to Stage III gametocyte glucose

metabolism [22]. Whilst not proven, this shift, linked to *Plasmodium* sexual development, is hypothesized to be a result of increased acetyl-CoA synthesis in the gametocyte mitochondrion or apicoplast [22]. Alternatively, recent work has provided evidence that acetyl-CoA synthetase and (indirectly) histone deacetylation may also contribute to the intracellular acetate pool in asexual parasites [32], and perhaps this is upregulated during gametocyte development. Additionally, acetate may be released during parasite GPI anchor biosynthesis [33].

Indicative of an abundance of amino acids available from haemoglobin digestion, asexual growth appears to be generally linked to the release of amino acids into the culture medium rather than their consumption. In support of this, a previous metabolomic study that focused on asexual intra-erythrocytic development using mass spectroscopy found similarly that most amino acids are released into the culture medium during the asexual cycle [21]. Presented here, the release of the majority of amino acids was at Day 4 when peak asexual parasitaemia is reached and gametocytes first arise (Figure 3A). Given that asexual parasitaemia subsequently begins to sharply decline, it cannot be ruled out that the bulk metabolite release at this time-point may simply be caused by cell lysis through apoptosis/necrosis rather than productive metabolism. In either case, elevated levels of these metabolites could signal to induce an upregulation of commitment to gametocyte development or alternatively could be toxic and detrimental to *in vitro* gametocyte culture (as has recently shown for lactate which retards asexual growth [34]).

Medium depletion of several lipid moieties was found to be associated with gametocyte maturation (Figure 3C). *Plasmodium* has the ability to generate some fatty acids *de novo* in the apicoplast [35] but most are scavenged from host serum and modified or utilized directly [36]. Transcriptomic analysis indicates that five of the six members of the type II fatty acid pathway are upregulated during gametocyte development [37]. Also, it has recently been shown that the lipid transporter gABC2 plays a role in the accumulation of neutral lipids within the parasite and is expressed predominantly in maturing female gametocytes [38]. Both of these support the assertion that lipid metabolism is important for gametocyte development and is different from asexual parasites.

One of the most remarkable morphological changes that occurs during gametocyte development is extensive mitochondrial elongation, branching and the formation of multiple mitochondrial cristate structures [39,40]. Supported by the shift towards glycolytic acetate production, which suggests fundamental changes in the way parasites generate energy during gametocyte development, it is tempting to speculate that this mitochondrial transformation burdens the developing gametocyte with increased requirements for membrane production and concomitantly increases their demand for exogenous lipids. As such, parasite energy production or lipid metabolism may be vulnerable targets in the generally drug-insensitive stage V mature gametocyte and could suggests avenues for drug discovery.

Conclusions

Very little is known about gametocyte development at the biochemical level. This preliminary study identifies and links the developmental flux of key metabolites to parasites in *in vitro* gametocyte cultures. It is clear that the metabolism of gametocytes is divergent from asexual parasites. Despite these differences, the composition of conventional parasite culture medium remains constant throughout the culture period. Clearly viable gametocytes develop under these conditions, as evidenced by their onward infectivity to mosquitoes, however cultured gametocytes are less infective than those that have developed in the human host [41] suggesting suboptimal growth conditions. With this in mind, it is hoped that data presented will help focus attention towards the development of gametocyte-specific medium formulations.

To this effect, supporting evidence is provided that gametocytes utilize acetate as an end-product for glucose metabolism and that gametocyte lipid requirements are divergent from asexual parasites. Consequently, it is suggested that further investigations of acetate and fatty acid metabolism in particular would significantly benefit current understanding on gametocyte developmental cell biology. Additional ^{13}C-labelling metabolic studies and/or proteomic investigations may also help elucidate role of acetate metabolism, whilst other targeted metabolic profiling techniques, such as reverse phase liquid chromatography-mass spectrometry (LC-MS), could provide a more comprehensive characterization and identification of the different lipid species in question. Together, such studies would generate a better understanding of gametocyte development and will greatly aid both optimization of expensive and variable gametocyte culture for future laboratory investigations, helping to shed light on potential new targets for transmission-blocking drugs.

Additional files

Additional file 1: Key metabolites responsible for phenotypic shifts observed within PCA time-trajectory of parasite cultures. PCA loadings plot showing spectral peaks responsible for driving the variability of parasite conditioned media in the first two components (p[1] and p[2]), observed in the corresponding scores plot shown in Figure 1A. Metabolites of the peaks the most strongly associated with the observed shifts in metabolic phenotypes have been labelled on the plot. Abbreviations: BCAA, branched chain amino acid peaks (valine, leucine, isoleucine).

Additional file 2: Variability in media composition between the two studies. PCA scores plot showing components 1 *vs* 2 **(A)** and 3 vs 4 **(B)** of fresh media samples from the two studies. Partial separation of Study 1 Day 0 media samples from the remaining samples apparent from components 3 and 4. R2 and Q2 represent the values of the degrees of model fit and predictive values of the PCA model, respectively, for each of the four components.

Additional file 3: Trajectories of highlighted metabolites adjusted for parasitaemia. Graphs showing culture metabolite levels over time for media from parasite-infected cultures (red) and control uninfected erythrocyte cultures (blue) as determined by multivariate discriminatory analysis (orthogonal PLAS-DA). *n*=8, with standard error of the mean as error bars. Control culture metabolite levels were subtracted from parasite-infected cultures and then normalised to the total parasitemia counts derived from Figure 1B (green).

Competing interests
The authors declare that they have no competing interests.

Authors' contributions
Study design was by MJD, SDL and JS. US, assisted by MJD, prepared all parasite samples, followed by NMR acquisition and analysis by SDL, assisted by JS. Manuscript was written by MJD, SDL and JS including feedback from US. All authors read and approved the final manuscript.

Acknowledgements
We would like to thank the Wellcome Trust (Sir Henry Wellcome Fellowship JS, Award number 089002/B/09/Z), the MRC and Imperial College for financial support of the presented work, as well as providing funds for SDL (MRC Doctoral Training Award). MJD is supported by the Bill and Melinda Gates Foundation (OPP1043501) and US is supported by the Medicines for Malaria Venture (MMV 08/2800/01).

Author details
[1]Division of Computational and Systems Medicine, Department of Surgery and Cancer, Imperial College London, London SW7 2AZ, UK. [2]Department of Life Sciences, Imperial College London, London SW7 2AZ, UK.

Received: 9 October 2014 Accepted: 24 November 2014
Published: 1 December 2014

References
1. WHO: *World Malaria Report 2013*. Geneva: World Health Organization; 2013. http://www.who.int/malaria/publications/world_malaria_report_2013/en/.
2. Sinden RE: **Sexual development of malarial parasites.** *Adv Parasitol* 1983, **22:**153–216.
3. Baker DA: **Malaria gametocytogenesis.** *Mol Biochem Parasitol* 2010, **172:**57–65.
4. Kafsack BFC, Rovira-Graells N, Clark TG, Bancells C, Crowley VM, Campino SG, Williams AE, Drought LG, Kwiatkowski DP, Baker DA, Cortés A, Llinás M: **A transcriptional switch underlies commitment to sexual development in malaria parasites.** *Nature* 2014, **507:**248–252.
5. Sinha A, Hughes KR, Modrzynska KK, Otto TD, Pfander C, Dickens NJ, Religa AA, Bushell E, Graham AL, Cameron R, Kafsack BFC, Williams AE, Llinás M, Berriman M, Billker O, Waters AP: **A cascade of DNA-binding proteins for sexual commitment and development in Plasmodium.** *Nature* 2014, **507:**253–257.
6. Dixon MWA, Thompson J, Gardiner DL, Trenholme KR: **Sex in Plasmodium: a sign of commitment.** *Trends Parasitol* 2008, **24:**168–175.
7. Sinden RE: **Gametocytogenesis of *Plasmodium falciparum* in vitro: ultrastructural observations on the lethal action of chloroquine.** *Ann Trop Med Parasitol* 1982, **76:**15–23.
8. Sinden RE, Smalley ME: **Gametocytogenesis of *Plasmodium falciparum* in vitro: the cell-cycle.** *Parasitology* 1979, **79:**277–296.
9. Smalley ME: ***Plasmodium falciparum* gametocytes: The effect of chloroquine on their development.** *Trans R Soc Trop Med Hyg* 1977, **71:**526–529.
10. Chutmongkonkul M, Maier WA, Seitz HM: **A new model for testing gametocytocidal effects of some antimalarial drugs on *Plasmodium falciparum* in vitro.** *Ann Trop Med Parasitol* 1992, **86:**207–215.
11. Kumar N, Zheng H: **Stage-specific gametocytocidal effect in vitro of the antimalaria drug qinghaosu on *Plasmodium falciparum*.** *Parasitol Res* 1990, **76:**214–218.
12. Trager W, Jensen JB: **Human malaria parasites in continuous culture.** *Science* 1976, **193:**673–675.
13. Ponnudurai T, Lensen AH, Van Gemert GJ, Bensink MP, Bolmer M, Meuwissen JH: **Infectivity of cultured *Plasmodium falciparum* gametocytes to mosquitoes.** *Parasitology* 1989, **98**(Pt 2):165–173.
14. Churcher TS, Blagborough AM, Delves M, Ramakrishnan C, Kapulu MC, Williams AR, Biswas S, Da DF, Cohuet A, Sinden RE: **Measuring the blockade of malaria transmission - an analysis of the standard membrane feeding assay.** *Int J Parasitol* 2012, **42:**1037–1044.
15. Li JV, Wang Y, Saric J, Nicholson JK, Dirnhofer S, Singer BH, Tanner M, Wittlin S, Holmes E, Utzinger J: **Global metabolic responses of NMRI mice to an experimental *Plasmodium berghei* infection.** *J Proteome Res* 2008, **7:**3948–3956.
16. Ghosh S, Sengupta A, Sharma S, Sonawat HM: **Multivariate modelling with (1) H NMR of pleural effusion in murine cerebral malaria.** *Malar J* 2011, **10:**330.
17. Lamour SD, Choi B-S, Keun HC, Müller I, Saric J: **Metabolic characterization of *Leishmania major* infection in activated and nonactivated macrophages.** *J Proteome Res* 2012, **11:**4211–4222.
18. Birungi G, Chen SM, Loy BP, Ng ML, Li SFY: **Metabolomics approach for investigation of effects of dengue virus infection using the EA.hy926 cell line.** *J Proteome Res* 2010, **9:**6523–6534.
19. Vincent IM, Creek DJ, Burgess K, Woods DJ, Burchmore RJS, Barrett MP: **Untargeted metabolomics reveals a lack of synergy between nifurtimox and eflornithine against *Trypanosoma brucei*.** *PLoS Negl Trop Dis* 2012, **6:**e1618.
20. Teng R, Junankar PR, Bubb WA, Rae C, Mercier P, Kirk K: **Metabolite profiling of the intraerythrocytic malaria parasite *Plasmodium falciparum* by (1)H NMR spectroscopy.** *NMR Biomed* 2009, **22:**292–302.
21. Olszewski KL, Morrisey JM, Wilinski D, Burns JM, Vaidya AB, Rabinowitz JD, Llinás M: **Host-parasite interactions revealed by *Plasmodium falciparum* metabolomics.** *Cell Host Microbe* 2009, **5:**191–199.
22. MacRae JI, Dixon MW, Dearnley MK, Chua HH, Chambers JM, Kenny S, Bottova I, Tilley L, McConville MJ: **Mitochondrial metabolism of sexual and asexual blood stages of the malaria parasite *Plasmodium falciparum*.** *BMC Biol* 2013, **11:**67.
23. Sana TR, Gordon DB, Fischer SM, Tichy SE, Kitagawa N, Lai C, Gosnell WL, Chang SP: **Global mass spectrometry based metabolomics profiling of erythrocytes infected with *Plasmodium falciparum*.** *cbvPLoS One* 2013, **8:**e60840.
24. Beckonert O, Keun HC, Ebbels TMD, Bundy J, Holmes E, Lindon JC, Nicholson JK: **Metabolic profiling, metabolomic and metabonomic procedures for NMR spectroscopy of urine, plasma, serum and tissue extracts.** *Nat Protoc* 2007, **2:**2692–2703.
25. Dieterle F, Ross A, Schlotterbeck G, Senn H: **Probabilistic quotient normalization as robust method to account for dilution of complex biological mixtures. Application in 1H NMR metabonomics.** *Anal Chem* 2006, **78:**4281–4290.
26. Trygg J, Holmes E, Lundstedt T: **Chemometrics in metabonomics.** *J Proteome Res* 2007, **6:**469–479.
27. Benjamini Y, Hochberg Y: **Controlling the false discovery rate: a practical and powerful approach to multiple testing.** *J R Stat Soc Ser B Methodol* 1995, **57:**289–300.
28. Lian L-Y, Al-Helal M, Roslaini AM, Fisher N, Bray PG, Ward SA, Biagini GA: **Glycerol: An unexpected major metabolite of energy metabolism by the human malaria parasite.** *Malar J* 2009, **8:**38.
29. Istvan ES, Dharia NV, Bopp SE, Gluzman I, Winzeler EA, Goldberg DE: **Validation of isoleucine utilization targets in *Plasmodium falciparum*.** *Proc Natl Acad Sci U S A* 2011, **108:**1627–1632.
30. van Schalkwyk DA, Priebe W, Saliba KJ: **The inhibitory effect of 2-halo derivatives of d-glucose on glycolysis and on the proliferation of the human malaria parasite *Plasmodium falciparum*.** *J Pharmacol Exp Ther* 2008, **327:**511–517.
31. Tielens AGM, van Grinsven KWA, Henze K, van Hellemond JJ, Martin W: **Acetate formation in the energy metabolism of parasitic helminths and protists.** *Int J Parasitol* 2010, **40:**387–397.
32. Cobbold SA, Vaughan AM, Lewis IA, Painter HJ, Camargo N, Perlman DH, Fishbaugher M, Healer J, Cowman AF, Kappe SHI, Llinás M: **Kinetic flux profiling elucidates two independent acetyl-CoA biosynthetic pathways in *Plasmodium falciparum*.** *J Biol Chem* 2013, **288:**36338–36350.

33. *PF3D7_0624700: (N-acetylglucosaminylphosphatidylinositol deacetylase)*; http://plasmodb.org/plasmo/showRecord.do?name=GeneRecordClasses.GeneRecordClass&source_id=PF3D7_0624700&project_id=PlasmoDB.
34. Hikosaka K, Hirai M, Komatsuya K, Ono Y, Kita K: **Lactate retards the development of erythrocytic stages of the human malaria parasite *Plasmodium falciparum*.** *Parasitol Int* 2014, Epub ahead of print.
35. Lim L, McFadden GI: **The evolution, metabolism and functions of the apicoplast.** *Philos Trans R Soc Lond B Biol Sci* 2010, **365:**749–763.
36. Mi-Ichi F, Kita K, Mitamura T: **Intraerythrocytic *Plasmodium falciparum* utilize a broad range of serum-derived fatty acids with limited modification for their growth.** *Parasitology* 2006, **133**(Pt 4):399–410.
37. Young JA, Fivelman QL, Blair PL, de la Vega P, Le Roch KG, Zhou Y, Carucci DJ, Baker DA, Winzeler EA: **The *Plasmodium falciparum* sexual development transcriptome: A microarray analysis using ontology-based pattern identification.** *Mol Biochem Parasitol* 2005, **143:**67–79.
38. Tran PN, Brown SHJ, Mitchell TW, Matuschewski K, McMillan PJ, Kirk K, Dixon MWA, Maier AG: **A female gametocyte-specific ABC transporter plays a role in lipid metabolism in the malaria parasite.** *Nat Commun* 2014, **5:**4773.
39. Krungkrai J: **The multiple roles of the mitochondrion of the malarial parasite.** *Parasitology* 2004, **129**(Pt 5):511–524.
40. Okamoto N, Spurck TP, Goodman CD, McFadden GI: **Apicoplast and mitochondrion in gametocytogenesis of *Plasmodium falciparum*.** *Eukaryot Cell* 2009, **8:**128–132.
41. Schneider P, Bousema JT, Gouagna LC, Otieno S, van de Vegte-Bolmer M, Omar SA, Sauerwein RW: **Submicroscopic *Plasmodium falciparum* gametocyte densities frequently result in mosquito infection.** *Am J Trop Med Hyg* 2007, **76:**470–474.

doi:10.1186/1475-2875-13-468
Cite this article as: Lamour *et al.*: Changes in metabolic phenotypes of *Plasmodium falciparum in vitro* cultures during gametocyte development. *Malaria Journal* 2014 **13**:468.

Submit your next manuscript to BioMed Central and take full advantage of:

- **Convenient online submission**
- **Thorough peer review**
- **No space constraints or color figure charges**
- **Immediate publication on acceptance**
- **Inclusion in PubMed, CAS, Scopus and Google Scholar**
- **Research which is freely available for redistribution**

Submit your manuscript at
www.biomedcentral.com/submit

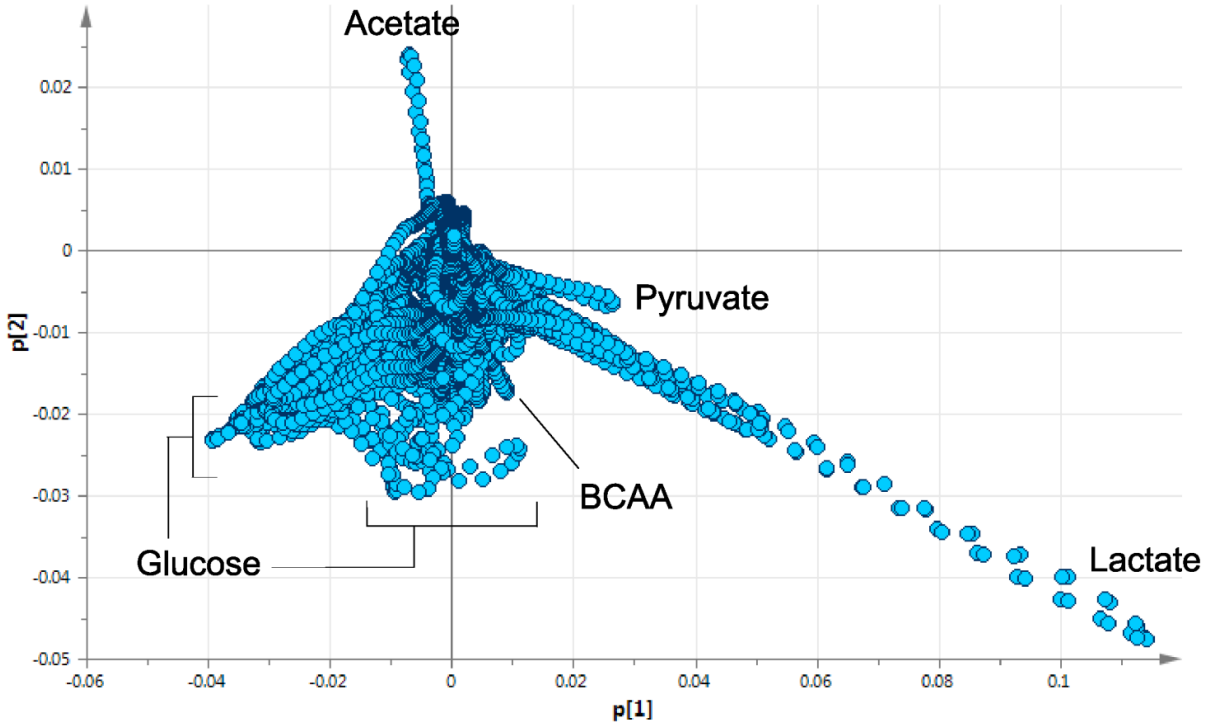
BioMed Central

Changes in metabolic phenotypes of *Plasmodium falciparum in vitro* cultures during gametocyte development

Additional files

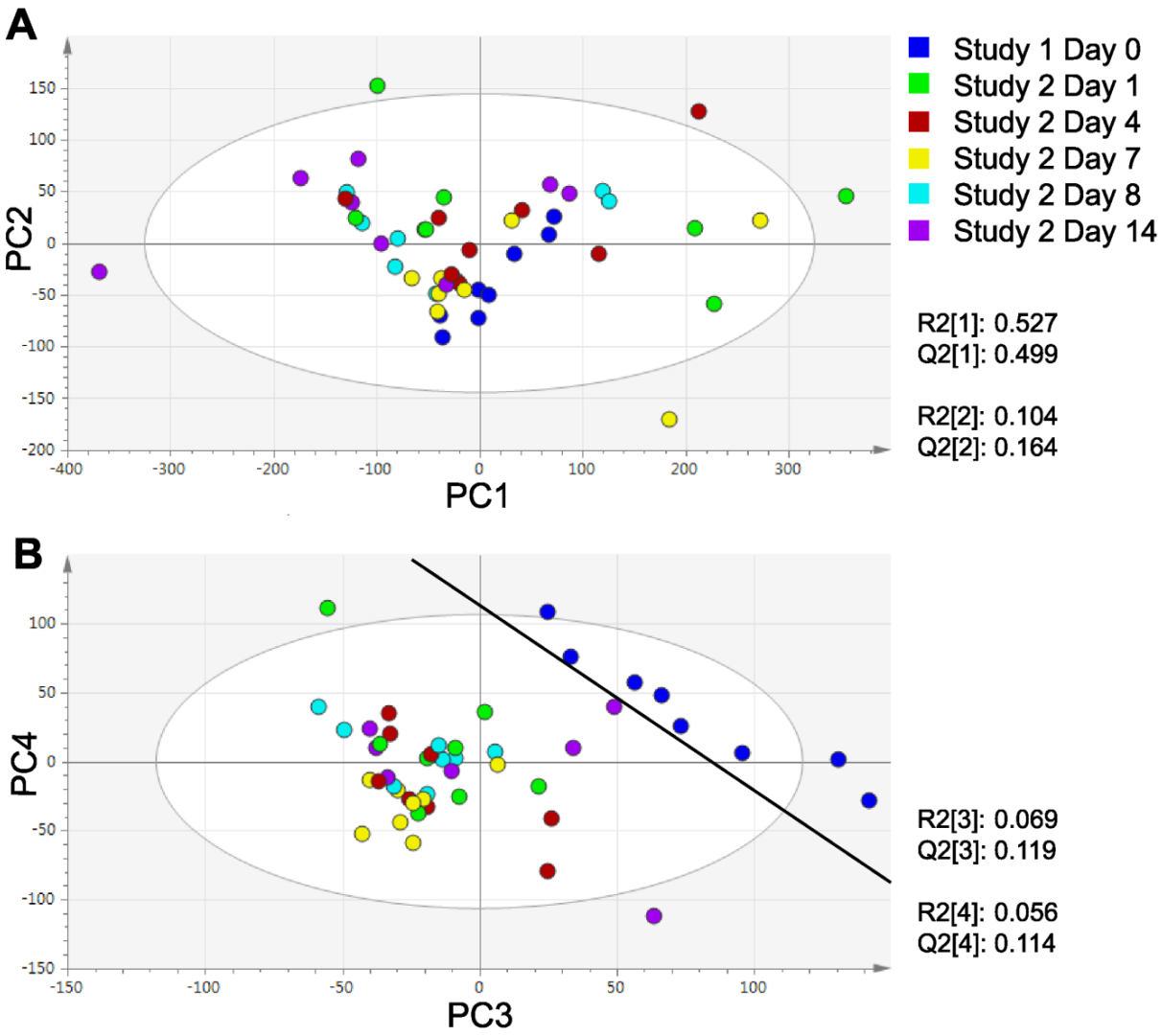
Additional file 1 – Key metabolites responsible for phenotypic shifts observed within PCA time-trajectory of parasite cultures

PCA loadings plot showing spectral peaks responsible for driving the variability of parasite-conditioned media in the first two components (p[1] and p[2]), observed in the corresponding scores plot shown in Figure 1A. Metabolites of the peaks the most strongly associated with the observed shifts in metabolic phenotypes have been labelled on the plot. Abbreviations: BCAA, branched-chain amino acid peaks (valine, leucine, isoleucine).



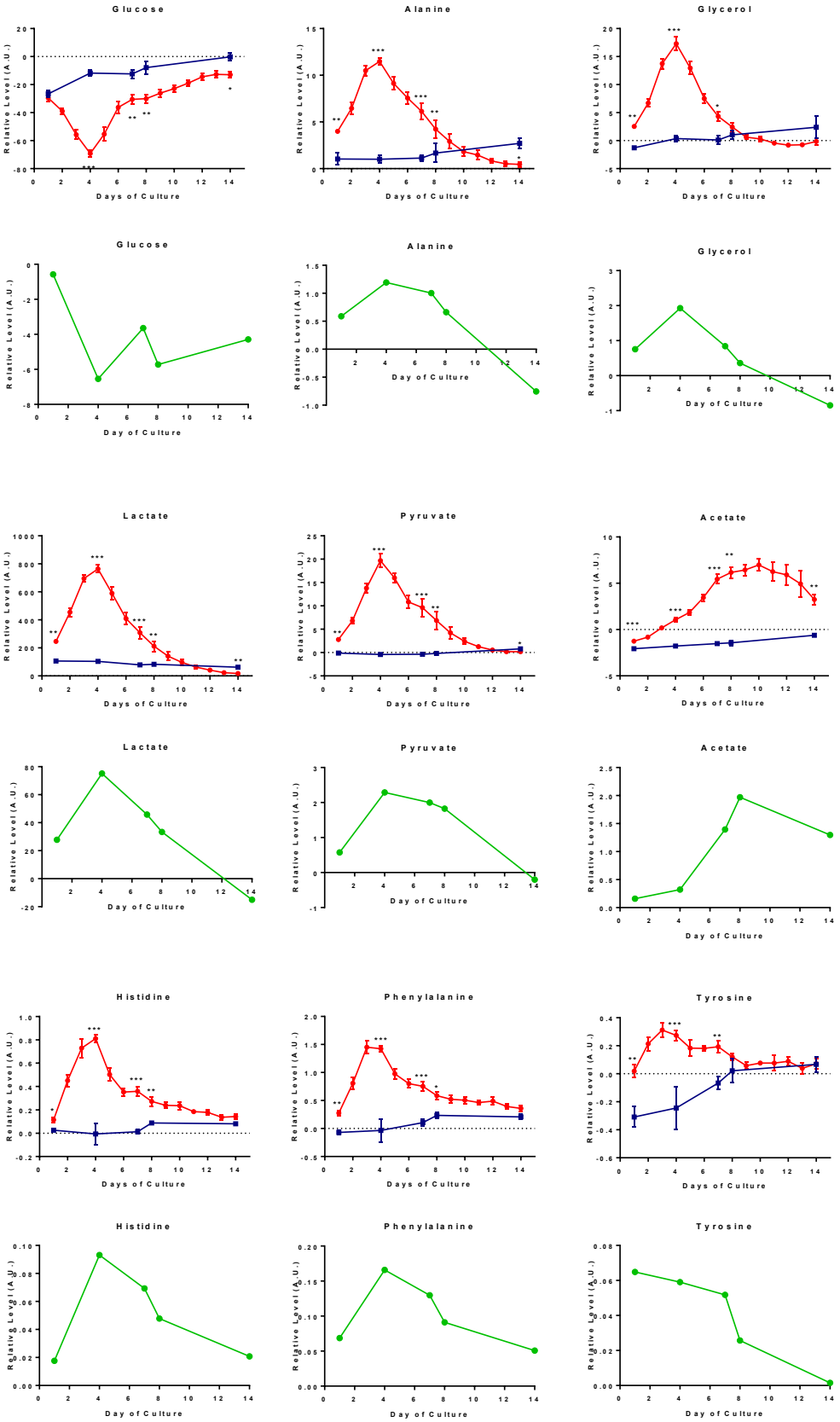
Additional file 2 – Variability in media composition between the two studies

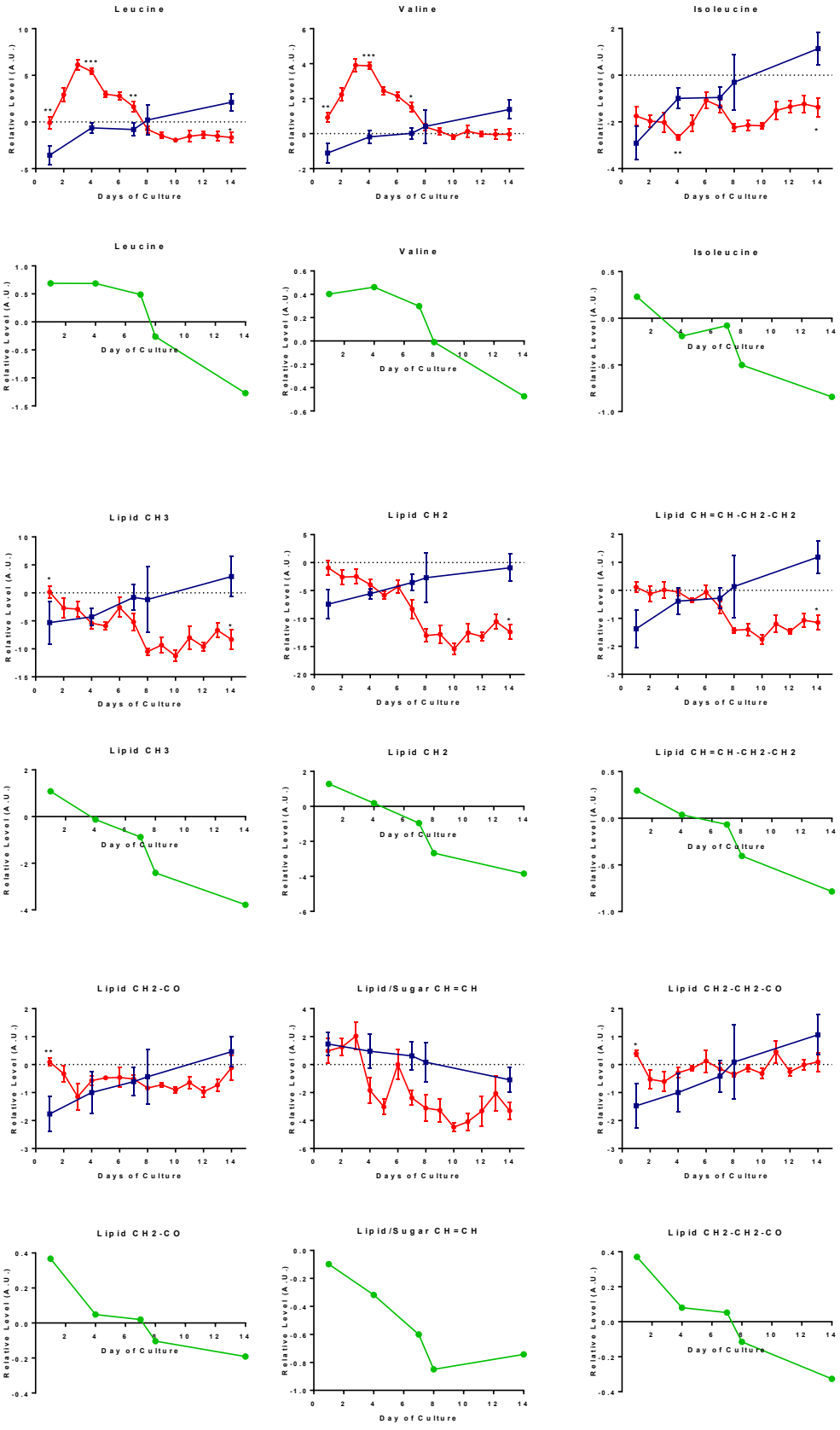
PCA scores plot showing components 1 *vs* 2 (A) and 3 vs 4 (B) of fresh media samples from the two studies. Partial separation of Study 1 Day 0 media samples from the remaining samples apparent from components 3 and 4. R^2 and Q^2 represent the values of the degrees of model fit and predictive values of the PCA model, respectively, for each of the four components.



Additional file 3 – Trajectories of highlighted metabolites adjusted for parasitaemia

Graphs showing culture metabolite levels over time for media from parasite-infected cultures (red) and control uninfected erythrocyte cultures (blue) as determined by multivariate discriminatory analysis (orthogonal PLAS-DA). *n*=8, with standard error of the mean as error bars. Control culture metabolite levels were subtracted from parasite-infected cultures and then normalised to the total parasitemia counts derived from Figure 1B (green).


Glucose
Alanine
Glycerol
Glucose
Alanine
Glycerol
Lactate
Pyruvate
Acetate
Lactate
Pyruvate
Acetate
Histidine
Phenylalanine
Tyrosine
Histidine
Phenylalanine
Tyrosine
Relative Level (A.U.)
Days of Culture
Day of Culture


Leucine
Valine
Isoleucine
Relative Level (A.U.)
Days of Culture
Day of Culture
Lipid CH3
Lipid CH2
Lipid CH=CH-CH2-CH2
Lipid CH2-CO
Lipid/Sugar CH=CH

Journal of proteome research

Article

pubs.acs.org/jpr

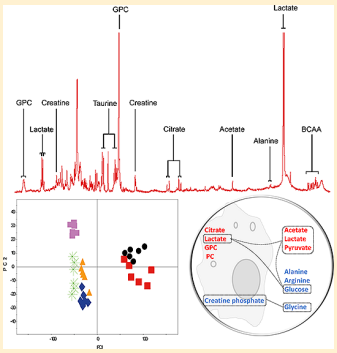
Metabolic Characterization of *Leishmania major* Infection in Activated and Nonactivated Macrophages.

Sabrina D. Lamour,[†] Beak-San Choi,[‡] Hector C. Keun,[†] Ingrid Müller,*[,‡] and Jasmina Saric*[,†]

[†]Biomolecular Medicine, Department of Surgery and Cancer, Faculty of Medicine, Imperial College London, Sir Alexander Fleming Building, South Kensington, London SW7 2AZ, United Kingdom

[‡]Department of Medicine, Section of Immunology, Faculty of Medicine, Imperial College London, London W2 1PG, United Kingdom

ABSTRACT: Infection with *Leishmania spp.* can lead to a range of symptoms in the affected individual, depending on underlying immune-metabolic processes. The macrophage activation state hereby plays a key role. Whereas the L-arginine pathway has been described in detail as the main biochemical process responsible for either nitric oxide mediated parasite killing (classical activation) or amplification of parasite replication (alternative activation), we were interested in a wider characterization of metabolic events *in vitro*. We therefore assessed cell growth medium, parasite extract, and intra- and extracellular metabolome of activated and nonactivated macrophages, in presence and absence of *Leishmania major*. A metabolic profiling approach was applied combining ^{1}H NMR spectroscopy with multi- and univariate data treatment. Metabolic changes were observed along both conditional axes, that is, infection state and macrophage activation, whereby significantly higher levels of potential parasite end products were found in parasite exposed samples including succinate, acetate, and alanine, compared to uninfected macrophages. The different macrophage activation states were mainly discriminated by varying glucose consumption. The presented profiling approach allowed us to obtain a metabolic snapshot of the individual biological compartments in the assessed macrophage culture experiments and represents a valuable read out system for further multiple compartment *in vitro* studies.



KEYWORDS: *metabolic profiling, metabolomics, Leishmania major, macrophage activation, ^{1}H NMR spectroscopy, multivariate data analysis*

INTRODUCTION

The genus *Leishmania spp.* describes a cluster of protozoan parasites that cause a variety of related disease patterns with main relevance in low income countries of the tropics and subtropics and around the Mediterranean basin. More recent figures indicate 50000 deaths and 2 million new infections each year[1] mainly due to suboptimal disease management including diagnosis, prognosis, and treatment.

Infective *Leishmania* promastigotes are injected by the sandfly vector (*Lutzomya spp.* and *Phlebotomus spp.*) into the skin of the mammalian host and are taken up by mononuclear phagocytic cells, especially macrophages and neutrophils. The main place of residence and proliferation of *Leishmania* (*L.*) *spp.* is the macrophage phagolysosome[2] where the physico-chemical conditions, in particular the high temperature and low pH initiate transformation from flagellated promastigote to immobile amastigote stage which proliferates via binary fission. Completion of the life cycle is granted when parasite-containing macrophages are taken up by the vector during a blood meal and the amastigotes are transformed back into promastigotes in the midgut. The promastigotes differentiate into infective metacyclics and migrate to the proboscis to be injected into the mammalian host.[3]

Disease development is strongly linked to parasite species and immune status of the host. While *L. aethiopica, L. major* and *L. tropica* are dermotropic species for which proliferation is usually restricted to the primary site of inoculation, the dermo-mucotropic (*L. braziliensis, L. panamanensis*) and viscerotropic species (*L. donovani, L. infantum*) can spread to mucosa and internal organs, respectively.[4,5]

Although the tropism of the various *Leishmania* strains is a major contributor to disease outcome, the state of the host's immune system is also influential and can, especially in the immune-suppressed individual, induce a spread of parasites to unusual locations such as the gastro-intestinal or respiratory tract during the visceral form of disease.[6]

Early observations on murine *L. major* infection models showed a varying disease outcome in different mouse strains, whereby the self-healing course in C57BL/6 and the progressive infection in BALB/c mice have probably been most characterized and scientifically exploited.[7−10] Due to the clear distinctive course of pathology depending on the rodent model chosen, *L. major* has been an important tool in defining the mechanisms of an effective immune response to intra-

Received: April 5, 2012
Published: June 22, 2012

ACS Publications © 2012 American Chemical Society

cellular pathogens. The system has substantially contributed to basic immunological understanding in infection by confirming the general mechanistic trends of T helper cell subpopulations[11] and their specific role in micro- and macro-parasite defense. The original classification into two dichotomous T helper (Th) cell populations had to be reshaped in the past decade due to other emerging Th subsets including Th17 and T regulatory cells[12] and the substantiation of the emerging concept of T cell plasticity.[13,14] However, the original simplification can still offer a gross mechanistic guidance on the cellular as well as on the metabolic level.

The L-arginine pathway is a crucial immune-metabolic cascade responsible for either control or explosion of intracellular parasite populations, which is tightly controlled by the Th cell response and the corresponding cytokine environment. Classically activated macrophages (caMΦ), stimulated with TNF-α, IL-12, and IFN-γ, typically referred to as Th1 cytokines, will generate reactive nitrogen species including nitric oxide (NO) and its derivatives NO_2, HNO_2 and $ONOO^-$, *via* the action of inducible nitric oxide synthase (iNOS). By contrast, macrophage exposure to IL-4, IL-13 and IL-21, classically representing Th2 cytokines, stimulates transcription of arginase and the subsequent generation of polyamines via L-ornithine, which support parasite growth and replication. These macrophages are considered as alternatively activated macrophages (aaMΦ).[15,16]

The major underlying immune-metabolic processes of macrophage mediated defense to *L. major* have been described in depth, particularly with regard to the L-arginine pathway and its major metabolic regulators including picolinic acid, an intermediate of the tryptophan degradation (L-kynurenine pathway),[17] phenylalanine and taurine.[18,19] Substantial progress has been made indeed to further dissect the role of selected pathway components, particularly of the role of L-arginine, arginase and iNOS.[16,20−22] However, an approach for simultaneous global metabolic assessment of all participating metabolic entities, that is, parasite, macrophage and cell environment has not been described in depth so far.

Metabolic profiling based on ^{1}H nuclear magnetic resonance (NMR) spectroscopy combined with multivariate modeling strategies and correlation analysis has shown the capacity to shed light on parasite-host interaction *in vivo* in an integrative and nonassumptive way. Despite being applied only in recent years within parasitology, this approach has opened a new avenue to assess parasite-host interplay from a global perspective and allowed unexpected insight into remote parasite effects and systemic metabolic events in the host.[23,24] By coanalyzing and correlating relative cytokine levels with the ^{1}H NMR derived metabolic information, new comprehension on immune-metabolic processes in the host can potentially be obtained.[25,26]

Although *in vivo* research has been the main focus of interest in the field of metabolic profiling, the importance and simplicity of *in vitro* systems has been acknowledged and has offered a simple tool to address more targeted questions. Such research has been introduced more recently as a complementary component in a variety of disease scenarios, including infectious disease and cancer.[27−29]

Here we assess a well-defined cytokine-regulated macrophage *in vitro* system and combine targeted enzymatic and biochemical assays with a ^{1}H NMR-based screening of the single compartments involved, including parasite, external and internal macrophage metabolome, and cell growth supernatant, in order to comprehend the metabolic communication between the metabolically active entities. Three different macrophage activation states were characterized and compared, specifically caMΦ, aaMΦ and nonactivated MΦ (naMΦ) in an uninfected state and during *L. major* exposure. Our studies define the wider metabolic impact of the cytokine induced activation on parasite, macrophage and cell environment.

EXPERIMENTAL PROCEDURES

Culture of Bone Marrow Derived Macrophages

Bone marrow derived macrophages were obtained from naïve 6−8 week old female BALB/c mice (Charles River, U.K.). Bone marrow was harvested by flushing femurs and tibias and 5×10^6 bone marrow precursor cells were differentiated in hydrophobic Teflon bags in 50 mL of macrophage differentiation medium into mature macrophages. The macrophage differentiation medium used was Dulbecco's modified Eagle medium (DMEM, Gibco) supplemented with 10% heat-inactivated fetal calf serum (FCS), 5% horse serum, 50 IU/ml penicillin, 50 μg/mL streptomycin, 2 mM glutamine, (Gibco), 0.05 mM β-mercaptoethanol (SIGMA) and (10% v/v) L929 fibroblast culture supernatant, the latter contains macrophage colony stimulating factor (M-CSF) necessary for the differentiation of the precursor cells into macrophages. The single cell suspensions were differentiated for 8 days in Teflon bags, harvested, washed in DMEM, counted and adjusted to the required cell concentration.

Macrophage Stimulation and Infection

Mature bone marrow derived macrophages (MΦ) were activated in the presence or absence of *L. major* LV39 (MRHO/SU/59/P-strain) promastigotes at a multiplicity of infection of 10 parasites per MΦ. Macrophage cultures were prepared in 6 replicates in 6-well tissue culture plates (Costar) and seeded at 2.5×10^6 cells/well in 5 mL DMEM, supplemented with heat-inactivated 10% FCS, 50 IU/mL penicillin, 50 μg/mL streptomycin, 2 mM glutamine, (Gibco), and 0.05 mM β-mercaptoethanol (SIGMA). Six different groups were assessed and compared, namely alternatively activated macrophages (aaMΦ), classically activated macrophages (caMΦ) and nonactivated macrophages (naMΦ). Each activation state was set up in absence and presence of *L. major* and all cultures were incubated for 48 h at 37 °C, 5% CO_2. The 48 h culture period was selected to ensure optimal viability of macrophages used in the six different groups. Activation stimuli included IFN-γ and TNF-α for caMΦ (5×10^5/mL mature MΦ with 100 U/ml recombinant murine IFN-γ and 500 U/ml recombinant mouse TNF-α; PeproTech) whereas no stimuli were used for naMΦ. IL-4 was supplemented for generation of aaMΦ (5×10^5/mL mature MΦ with 20 U/mL recombinant mouse IL-4 PeproTech).

Parasite Maintenance and Culture

L. major parasites LV39 (MRHO/SU/59/P-strain) were maintained in a virulent state by monthly passage in BALB/c mice. Parasites were isolated monthly from lesions of *L. major* infected BALB/c mice, and used for an average of 6−8 *in vitro* passages. The freshly isolated parasites were cultured and maintained at 26 °C, 5% CO_2 in solid phase blood agar, overlaid with 5 mL DMEM containing 50 IU/ml penicillin and 50 μg/mL streptomycin (P/S). Late log phase/stationary phase *L. major* promastigotes were collected from culture, washed three times in DMEM and an aliquot of the parasite suspension

dx.doi.org/10.1021/pr3003358 | J. Proteome Res. 2012, 11, 4211−4222

was fixed in 2% paraformaldehyde and used for counting parasites using a hemacytometer and adjusted to the required density.

Parasite Transformation

After 48 h of incubation at 37 °C, 10% CO_2 infected macrophages were washed and incubated with freshly prepared sterile lysis medium (DMEM supplemented with 0.008% w/v SDS) for 7−20 min at 37 °C, 10% CO_2. During this time, macrophage disintegration was monitored in regular intervals and when the host cells were lysed, the reaction was stopped by adding DMEM supplemented with 20% heat-inactivated FCS, 20 mM HEPES, 4 mM sodium bicarbonate, 50IU/ml penicillin and 50 μg/mL streptomycin. The lysates containing the liberated *L. major* amastigotes were transferred into blood agar cultures and incubated at 26 °C, 5% CO_2 to allow the amastigotes to transform back into promastigotes and to proliferate. *Leishmania* promastigotes were counted in a separate experiment based on four replicates per activation group, using a hemacytometer after 3−5 days of culture.

For metabolic profile assessment, *L. major* parasites were counted, adjusted to 4.4×10^8 promastigotes and washed three times in ice-cold sterile PBS. Supernatant was removed and the residual parasite pellet was snap frozen on dry ice and subsequently stored at −40 °C prior to NMR analysis.

Preparation of Cell Culture Supernatant and Cell Extracts

A volume of 5 mL of the cell culture medium (= supernatant) was transferred from tissue culture plates into sterile 15 mL Falcon tubes and centrifuged at 2500 rpm at room temperature (RT) to pellet any cell debris. Supernatants were transferred again into a new set of tubes and stored at −40 °C prior to ^{1}H NMR acquisition.

After cell culture medium was removed, macrophage monolayers were quickly washed once with 1 mL/well PBS (kept at RT) which was discarded. This was followed by a cell quenching step using 960 μL of ice cold HPLC grade methanol (Sigma-Aldrich) per well and subsequently placing the plates at 4 °C for 2−5 min to permit cell lysis. Lysed cells and debris were harvested from the bottom of the wells using a cell scraper and the suspension was transferred into 2 mL Eppendorf tubes, whereby the recovered amount was ∼600−650 μL from each well due to the rapid evaporation of the methanol. A further 460 μL methanol were used to rinse each well and subsequently transfer to the cell lysate suspension kept on ice to obtain maximum metabolic yield. Lysates were dried overnight in the SpeedVac at 45 °C, followed by 80% methanol extraction, performed twice, as follows: dried lysates were resuspended in 500 μL cold HPLC grade methanol and vortexed for ∼30 s; 130 μL of distilled water was added into each tube and samples were vortexed again for ∼30 s. Tubes were kept on ice for 10−20 min, spun down at 13000 rpm for 10 min and lysate supernatants transferred into new tubes on ice. Extraction with 80% methanol was performed a second time on the remaining pellets, as described. Tubes were spun down again at 13000 rpm for 10 min. Supernatants were pooled with cell lysate samples from the first methanol extraction and dried overnight in the SpeedVac at 45 °C. Dried cell extracts were stored at −40 °C prior to preparation for ^{1}H NMR analysis.

^{1}H NMR Spectroscopic Assessment

Equal amounts of cell supernatant and phosphate buffer (0.2 M Na_2HPO_4, 0.043 M NaH_2PO_4, D_2O:H_2O = 7:3, v/v, 0.01% of sodium 3-(trimethylsilyl) propionic acid 2,2,3,3-d4 ([TSP)], pH = 7.4), were mixed to a total amount of 600 μL in Eppendorf tubes by vortexing. Samples were spun down at 13000 rpm for 5 min and 550 μL was transferred into 5 mm NMR tubes.

Cell extracts were resuspended in 600 μL phosphate buffer via vortexing. Samples were spun down at 13000 rpm for 5 min and 550 μL of the supernatant was transferred into 5 mm NMR tubes.

A frozen pellet of 4.4×10^8 parasites was transferred immediately into a ZrO_2 rotor and ∼25 μL D_2O were added. After inserting a spacer to adjust for the small amount of biomass, the rotor was sealed with a Kel-F rotor cap.

All cell culture supernatants were assessed on a Bruker Avance 600 NMR spectrometer with TXI probe head (Bruker; Rheinstetten, Germany) operating at 600.13 MHz for proton frequency, whereby 128 scans per sample were sufficient for obtaining optimal signal output.

The cell extracts were acquired using a cryo probe head (CO TCI) in order to minimize electric noise and augment the signal-to-noise ratio, whereby the number of scans was adjusted to 384 for each sample. A standard 1-dimensional (1D) pulse program was applied (recycle delay (RD)-90°-t1-90°-tm-90°-acquisition time (AQ)) and the water impact on the baseline was minimized by irradiating the water frequency during the RD (2 s). The 90° pulse length was set to 14.75 s for all cell supernatants and 14.62 s for the cell extracts. The spectra were recorded in 32768 data points within a spectral width of 20 ppm and a 0.3 Hz line-broadening factor was applied to the free induction decays (FIDs) prior to Fourier transformation.

In contrast to supernatants and cell extracts, the *L. major* pellet was acquired via high resolution magic angle spinning (MAS)-NMR to maximize metabolic information from such a small amount of biomass.[30] The rotor was therefore placed in a MAS probe at an angle of 54.7° and spun at 5000 kHz in order to compensate for the line broadening effects introduced by the restricted motion of the analyte and the sample heterogeneity. Additionally to 1D spectral acquisition, a Carr−Purcell−Meibom−Gill (CPMG)[31] pulse sequence was applied to the parasite pellet in order to account for the smaller metabolic weight components. MAS spectra were acquired in 256 scans and recorded in 65536 data points, whereby a 90° pulse length of 10.43 s was applied. All other parameters remained the same as described above.

Spectral assignments were made by consulting the literature and in-house databases confirmed via statistical total correlation spectroscopy (STOCSY)[32] and using NMR suite profiler software (Chenomx, USA).

Data Processing and Statistical Treatment

Spectral data were processed using nmrproc, an in-house developed MATLAB algorithm (Dr. Tim Ebbels and Dr. Hector Keun, MathWorks) which performs optimized phasing, baseline alignment and automatic calibrating to the TSP reference peak at δ 0.00. To exclude the reference peak and residual resonances of the water peak from the data modeling process, only the spectral region from δ 0.5−9 was imported into MATLAB and the region affected by the water was cut out, that is, δ 4.54−4.98 for extracts and δ 4.7−5 for the supernatants. An in house adapted script for probabilistic quotient normalization was applied prior to statistical analysis (Dr. Kirill Veselkov).[33,34]

Multivariate analysis including principal component analysis (PCA) and partial least-squares (PLS)-related methods were

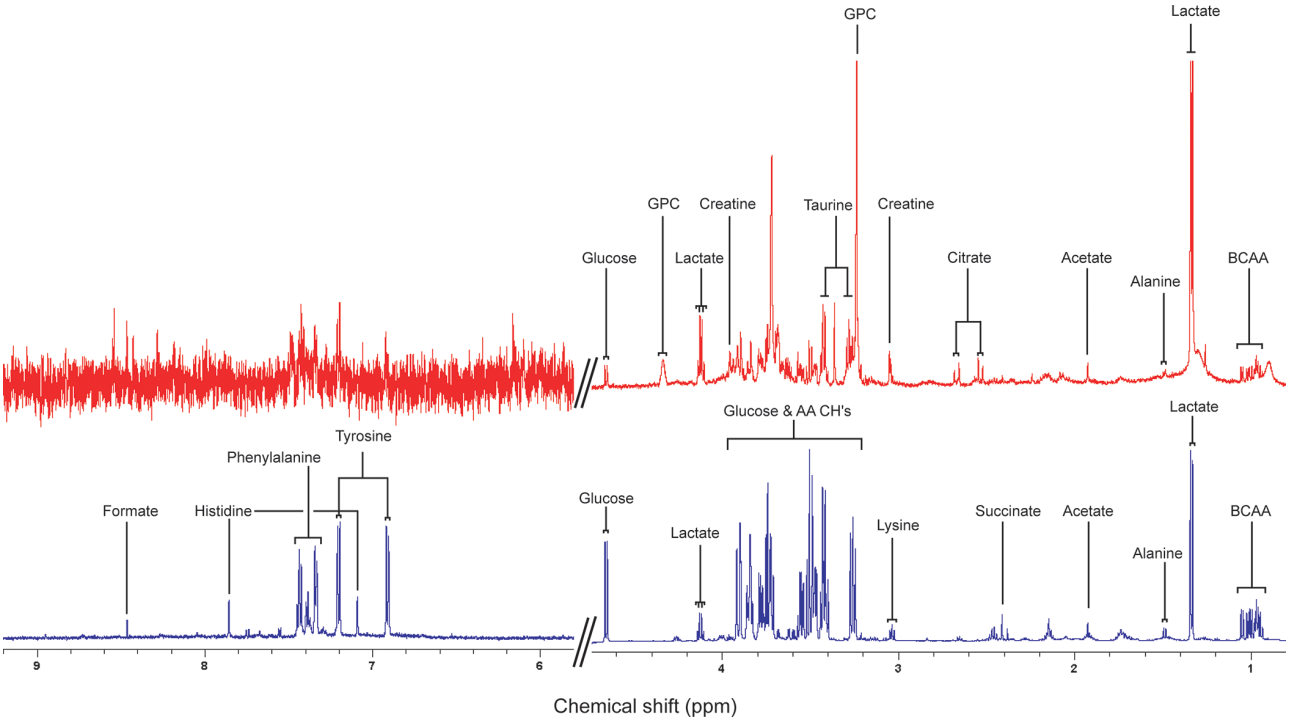


Figure 1. [1]H NMR spectrum of cell extract acquired by using a standard pulse sequence on cell extracts (red) and cell culture supernatants (blue). The metabolite concentration differes substantially between the two matrixes assessed which leads to a generally lower signal-to-noise ration in the spectra acquired from the cell extracts. Key: AA's, amino acids; BCAA, branched chain amino acids; GPC, glycerophosphocholine.

applied to the data using a SIMCA (Umetrics, Sweden) and MATLAB interface. As unsupervised method, PCA offers an unbiased overview on the general data distribution of a data matrix (X) by compressing the total amount of information into one single data point onto a plane.[32,35] The resulting scores plot allows visualization of groupings, outliers and time trajectories hence PCA is the ideal tool for defining further analytical strategy. PLS on the other hand relates the data in X to a second matrix (Y) containing a separate set of information which can be quantitative, such as cytokine levels, parasitaemia, inflammatory markers, etc., or qualitative (e.g., affiliation to class) in which case a PLS discriminant analysis (PLS-DA) would be applied.[35,36] O-PLS-DA is a further optimization and is currently being applied for biomarker identification, comparing two classes of samples.[37,38] Whereas the PLS algorithm maximizes the class differences and facilitates identification of the loadings (metabolites) responsible, the inbuilt orthogonal filter removes systemic variation unrelated to class-affiliation.

Additional univariate approaches have been applied to further evaluate the candidate biomarkers. Integrals from selected nonoverlapped regions from each potential biomarker were therefore calculated in MATLAB and imported into GraphPad PRISM (GraphPad, USA) and SPSS (IBM Corporation, USA) for graphical representations and univariate statistical tests which included two tailed Mann−Whitney U-test with Bonferroni correction for pair wise comparison and Kruskal−Wallis for multiple group comparisons.

Arginase Activity

In addition to the 6 well cultures for metabolic profiling, macrophage cultures (1 mL/well) were set up in four replicates in 24 well tissue culture plates and stimulated and infected in parallel. Bone-marrow derived macrophages (5×10^5/mL) were assessed in unstimulated, classically, and alternatively activated state (as above) in the absence or presence of *L. major* promastigotes at a multiplicity of infection (MOI) of 10:1. After 48 h of culture at 37 °C, 10% CO_2, macrophages were washed and lysed.

Arginase activity was measured in macrophage lysates by measuring the conversion of L-arginine to urea, as described previously.[16,39] Briefly, cells were lysed with 0.1% Triton X-100 (Sigma) and 50 mM Tris-HCl (pH 7.5, Sigma), which contains the arginase cofactor manganese chloride (Sigma) and the enzyme activated at 56 °C for 7 min. The hydrolysis was conducted by the addition of 0.5 M L-arginine (pH 9.7) and incubated at 37 °C for 15−120 min. The reaction was stopped using an acid mix consisting of phosphoric acid (85%, Fluka), sulphuric acid (98%, Fluka) and distilled water at a v/v/v ratio of 1/3/7. The urea production via the hydrolysis of L-arginine by arginase was measured by the addition of 6% α-isonitrosopropiophenone and incubated at 100 °C for 45 min. Urea production by L-arginine hydrolysis in samples was measured against a urea standard and the OD of the reactions was measured at 550 nm. Arginase activity was calculated in mU per sample and expressed per 1×10^6 macrophages, whereby the lowest detectable level of arginase was 2.1 mU/1×10^6 macrophages. One unit of enzyme activity defined as the amount of enzyme that catalyzes the formation of 1 μM urea per minute.

Nitrite Production

After two days at 37 °C, 10% CO_2, culture supernatants were harvested from the different macrophage cultures in 24 well plates. Nitrite accumulation was used as an indicator of nitric oxide (NO) production in culture supernatants and measured

dx.doi.org/10.1021/pr3003358 | J. Proteome Res. 2012, 11, 4211−4222

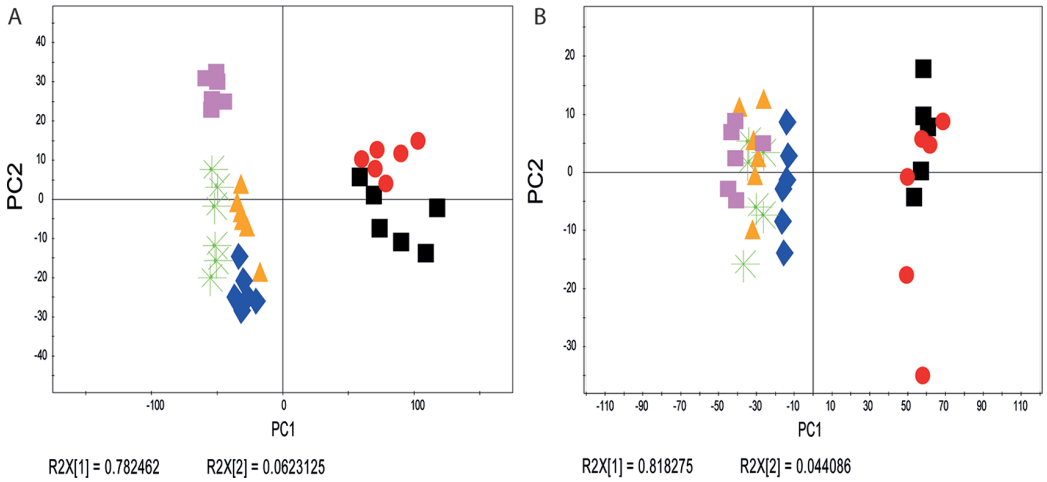


Figure 2. PCA analysis of (A) cell extracts and (B) cell supernatants. The groups in the scores plots represent different infection and activation states of the bone marrow derived macrophages. R2X explaines the variance in the model for each component. Key: blue, aaMΦ infected; green, aaMΦ uninfected; yellow, naMΦ infected; pink, naMΦ uninfected; red, caMΦ infected; black, caMΦ uninfected.

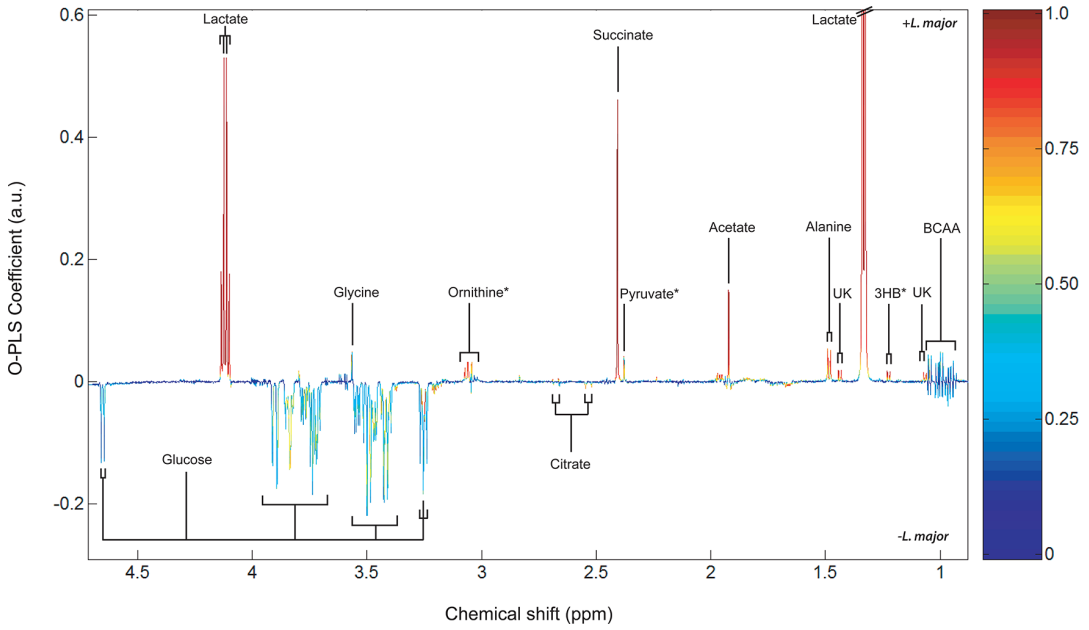


Figure 3. O-PLS-DA analysis comparing spectra of cell supernatants of infected and uninfected aaMΦ. Metabolites which are different between the two states are color coded according to significance, whereby red denotes highest and blue no discriminatory power. Upward pointing peaks are positively correlated with the infection, whereas downward pointing peaks are anti correlated with *L. major* infection. Key: a.u., arbitrary units; BCAA, branched chain amino acids; 3HB, 3-hydroxybutyrate; UK, unknown metabolite; *, tentatively assigned.

by using the Griess reagent. Equal volumes of macrophage culture supernatants and Griess reagent (1% sulphanilamide/0.1% *N*-(1-naphtyl) ethlyenediamine dihydrochloride/2.5% H_3PO_4) were incubated at room temperature for 10 min. Absorbance was measured in a microplate reader at 550 nm. Nitrite concentration was determined using $NaNO_2$ as standard. The limit of detection was set at 5 μM nitrite.

RESULTS

Global Data Distribution and Biomarker Identification

Spectral data from macrophage extracts and cell supernatants (Figure 1) were assessed separately for differences between infection and activation state. PCA of cell extract spectra revealed clear groupings between the different macrophage activation states as well as between *L. major* infected and uninfected cell cultures (Figure 2A). The most striking separation was observed between caMΦ and all other *in vitro* conditions. Analysis of the cell culture supernatants also indicated a metabolic separation between the activation states along the first principal component (PC1), again showing a distinction between caMΦ from all other data points (Figure 2B). Infection induced a shift of the aaMΦ data points along PC1 but no clear infection-related group separation was found within caMΦ and naMΦ spectra.

To identify the metabolites responsible for group separation O-PLS-DA was applied in a first step (Figure 3). To further validate the statistical significance of the identified metabolic

Table 1. Discriminating Metabolites Associated with *L. major* Infection[a]

		caMΦ's		aaMΦ's		naMΦ's	
	Metabolite	Fold change relative to noninfected	*p*	Fold change relative to noninfected	*p*	Fold change relative to noninfected	*p*
Cell culture supernatants	Acetate	0.32	**	0.32	**	0.32	NS
	Alanine	0.07	*	0.07	**	0.07	NS
	Citrate	0.02	NS	0.02	*	0.02	NS
	Fumarate	0.65	*	0.65	*	0.65	NS
	Glucose	−0.06	NS	−0.06	**	−0.06	NS
	Lactate	−0.01	NS	−0.01	**	−0.01	NS
	Ornithine	−0.04	NS	−0.04	**	−0.04	NS
	Pyruvate	0.28	**	0.28	**	0.28	NS
	Succinate	0.84	**	0.84	**	0.84	*
Cell Extracts	Acetate	−0.10	NS	−0.10	**	−0.10	**
	β-Alanine	0.05	NS	0.05	*	0.05	**
	Citrate	−0.10	NS	−0.10	*	−0.10	**
	Creatine	0.53	**	0.53	**	0.53	**
	Creatine Phosphate	0.14	NS	0.14	**	0.14	**
	Glycerophosphocholine	0.03	NS	0.03	**	0.03	**
	Glycine	0.01	NS	0.01	NS	0.01	**
	Lactate	−0.18	NS	−0.18	**	−0.18	**
	Phosphocholine	−0.15	NS	−0.15	**	−0.15	**
	Taurine	0.04	NS	0.04	NS	0.04	**

[a]Metabolites were selected based on the O-PLS-DA output and the difference in between the median values of metabolite concentrations (in mM, estimated from integral calculations, relative to TSP) between infected group and non-infected group was calculated for each. Changes in median concentration and fold changes were calculated relative to non-infected states, whereby increases are shown as positive values and decreases are shown as negative values. Statistical significance (p) measured *via* two-tailed Mann-Whitney test (with Bonferroni correction for individual MΦ subsets) whereby *, $p < 0.05$; **, $p < 0.01$; ***, $p < 0.001$; ****, $p < 0.0001$; and NS, not statistically significant.

Table 2. Discriminating Metabolites in the Cell Culture Supernatant Associated with Macrophage Activation[a]

	noninfected MΦ's				*L. major*-infected MΦ's			
	Median concentration and range (mM)				Median concentration and range (mM)			
Metabolite	CA	AA	NA	*p*	CA	AA	NA	*p*
Acetate	0.152 (0.148−0.159)	0.136 (0.129−0.139)	0.168 (0.157−0.212)	***	0.200 (0.195−0.220)	0.169 (0.166−0.169)	0.209 (0.174−0.210)	**
Alanine	0.887 (0.853−0.924)	1.002 (0.956−1.022)	1.016 (0.946−1.054)	**	0.945 (0.932−1.029)	1.079 (1.065−1.086)	1.072 (1.045−1.098)	**
Citrate	0.126 (0.122−0.154)	0.161 (0.152−0.169)	0.128 (0.124−0.141)	**	0.128 (0.124−0.141)	0.153 (0.150−0.157)	0.137 (0.126−0.141)	**
Glucose	10.261 (9.276−10.717)	15.701 (14.437−15.179)	15.609 (14.701−16.663)	**	9.670 (9.070−10.523)	13.848 (13.524−13.977)	15.067 (14.825−16.540)	**
Glycine	0.289 (0.273−0.296)	0.308 (0.305−0.320)	0.329 (0.313−0.343)	**	0.276 (0.269−0.302)	0.314 (0.311−0.325)	0.329 (0.328−0.336)	***
Lactate	8.149 (7.734−8.791)	2.018 (1.830−2.119)	1.551 (1.503−2.244)	***	8.048 (7.684−8.864)	3.228 (3.069−3.350)	2.231 (1.447−2.278)	***
Ornithine	0.216 (0.210−0.239)	0.361 (0.330−0.376)	0.226 (0.211−0.237)	**	0.209 (0.203−0.234)	0.395 (0.384−0.399)	0.226 (0.218−0.233)	***
Pyruvate[b]	0.095 (0.087−0.103)	0.073 (0.070−0.076)	0.097 (0.095−0.105)	**	0.121 (0.111−0.132)	0.081 (0.078−0.086)	0.097 (0.92−0.98)	***
Succinate	0.091 (0.088−0.095)	0.090 (0.085−0.092)	0.094 (0.088−0.192)	NS	0.167 (0.162−0.181)	0.186 (0.182−0.192)	0.193 (0.097−0.196)	*

[a]Metabolite levels are displayed as median concentration and range per condition (in mM), estimated from integral calculations, relative to TSP. Statistical significance (p) measured *via* Kruskal-Wallis test across three activation classes, whereby *, $p < 0.05$; **, $p < 0.01$; ***, $p < 0.001$; and NS, not statistically significant. [b]Tentatively assigned.

markers, the integrals were calculated for each, and subjected to univariate nonparametric tests (see methods). In Table 1, those metabolites are represented that were found to be statistically different between infected and noninfected samples, whereas Tables 2 and 3 show the metabolic differences between activation states, that is, aaMΦ, caMΦ and naMΦ. Only metabolites that reached statistical significance as validated via uni- and multivariate methods are documented.

The infection state showed particular discrimination in cell culture supernatants of aaMΦ and the cell extracts of naMΦ (Table 1), whereas the naMΦ supernatants did not appear to differ significantly between the infected and uninfected state with the exception of increased succinate after *L. major*

dx.doi.org/10.1021/pr3003358 | *J. Proteome Res.* 2012, 11, 4211−4222

Table 3. Discriminating Metabolites in Cell Extracts Associated with Macrophage Activation[a]

	noninfected MΦ's				*L. major*-infected MΦ's			
	Median concentration and range (mM)				Median concentration and range (mM)			
Metabolite	CA	AA	NA	*p*	CA	AA	NA	*p*
Acetate	0.012 (0.011–0.014)	0.010 (0.009–0.012)	0.010 (0.007–0.010)	**	0.011 (0.009–0.012)	0.014 (0.013–0.015)	0.013 (0.011–0.014)	**
β-Alanine	0.028 (0.021–0.029)	0.027 (0.023–0.039)	0.012 (0.011–0.014)	**	0.029 (0.020–0.034)	0.052 (0.031–0.056)	0.035 (0.023–0.036)	**
Citrate	0.045 (0.035–0.048)	0.024 (0.019–0.032)	0.015 (0.010–0.016)	***	0.040 (0.035–0.050)	0.032 (0.030–0.037)	0.028 (0.026–0.057)	*
Creatine	0.021 (0.017–0.023)	0.022 (0.014–0.025)	0.006 (0.004–0.006)	**	0.032 (0.025–0.043)	0.055 (0.046–0.058)	0.026 (0.023–0.030)	**
Creatine Phosphate	0.012 (0.010–0.015)	0.026 (0.015–0.028)	0.005 (0.004–0.006)	***	0.013 (0.011–0.015)	0.055 (0.050–0.057)	0.026 (0.020–0.029)	***
Glucose	0.080 (0.068–0.146)	0.123 (0.094–0.162)	0.093 (0.114–0.217)	NS	0.057 (0.042–0.088)	0.099 (0.068–0.134)	0.115 (0.165–0.261)	**
Glycerophosphocholine	0.074 (0.061–0.082)	0.010 (0.007–0.012)	0.007 (0.006–0.008)	***	0.077 (0.056–0.091)	0.027 (0.022–0.028)	0.026 (0.024–0.029)	**
Glycine	0.010 (0.008–0.012)	0.008 (0.006–0.021)	0.007 (0.004–0.007)	*	0.010 (0.005–0.028)	0.010 (0.009–0.013)	0.010 (0.009–0.011)	NS
Lactate	0.153 (0.140–0.247)	0.055 (0.038–0.071)	0.031 (0.018–0.034)	***	0.126 (0.096–0.142)	0.090 (0.075–0.104)	0.063 (0.049–0.083)	**
Phosphocholine	0.014 (0.007–0.016)	0.007 (0.004–0.007)	0.003 (0.003–0.003)	***	0.012 (0.010–0.019)	0.012 (0.011–0.013)	0.010 (0.009–0.012)	NS
Taurine	0.127 (0.098–0.163)	0.206 (0.129–0.224)	0.108 (0.087–0.132)	**	0.132 (0.100–0.145)	0.211 (0.182–0.226)	0.207 (0.174–0.232)	**

[a]Metabolite levels are displayed as median concentration and range per condition (in mM), estimated from integral calculations, relative to TSP. Statistical significance (p) measured *via* Kruskal-Wallis test across three activation classes, whereby *, $p < 0.05$; **, $p < 0.01$; ***, $p < 0.001$; and NS, not statistically significant.

exposure. The caMΦ cell extracts did not seem to undergo major metabolic changes either upon infection and showed solely an increase of creatine. The cell supernatants on the other hand revealed elevated levels of acetate, alanine, fumarate, pyruvate and succinate in those samples which have been incubated with *L. major*.

The activation state was reflected strongly in the metabolic background, that is, cell extracts and culture supernatant in both *L. major* infected and uninfected macrophage cultures (Tables 2 and 3). A depletion of glucose, alanine, and glycine (cell culture supernatant) and a subsequent increase of lactate (cell culture supernatant or extract) was the most general change comparing caMΦ with either naMΦ or aaMΦ in both states, infected and uninfected (Tables 2 and 3, Figures 3 and 4). More specific differences included relative lower acetate levels in aaMΦ cell supernatants compared to naMΦ supernatants during infection, or a general higher intracellular level of taurine in both aaMΦ and naMΦ compared to caMΦ during infection.

A few biomarkers that were identified in the cell culture supernatants were also found in the cell extracts, namely acetate, citrate and lactate, whereas others were specific to extracts (e.g., creatine, creatine phosphate and choline derivatives) and supernatants (e.g., succinate, alanine) (Tables 1–3), which is in line with the general metabolite recovery (Table 4).

Metabolic Characterization of *In Vitro* Compartments

The metabolites detected by ^{1}H NMR spectroscopy of macrophage extracts, macrophage culture supernatants, fresh cell culture growth media and *L. major* showed a high overlap. A list of metabolites identified in each compartment is presented in Table 4. Lactate and the branched chain amino acids (BCAA) leucine, isoleucine and valine, were identified in all four compartments, and a large amount of metabolites was observed in three compartments, including acetate, glucose, glycine, phenylalanine and tyrosine. Uracil on the other hand was only present in the *L. major* metabolome and formate and fumarate were specific to the external macrophage metabolome (Figures 1 and 5 and Table 4). Some residual lipid resonances were observed only in MAS NMR spectra of *L. major* (Figure 5) as a result of analyzing intact parasites and not further discussed in comparisons to solution-state analyses of cell extracts which focused only on polar metabolites. Alanine, acetate, citrate, succinate, 3-hydroxybutyrate, formate and fumarate were all found in the supernatant after incubation with cellular material.

L-Arginine Metabolism and Parasite Replication

Arginase activity, NO production and parasite counts were assessed in two separate experiments including four replicates in order to confirm that the different stimulation assays resulted in the expected macrophage activation state and the expected effect to infection intensity. The nitric oxide assay did not show detectable amounts of NO in naMΦ and aaMΦ (limit of detection 5 μM) but revealed higher levels in *L. major* infected cells (76.1 ± 4.38 μM) compared to the uninfected control samples (58.1 ± 1.99 μM) as validated by the nonparametric Mann–Whitney U test ($p < 0.05$).

Pronounced arginase activity was detected in aaMΦ (*L. major* (−): 580.8 ± 16.40; *L. major* (+): 666.0 ± 2.06 mU/10^6 cells), whereas naMΦ (*L. major* (−): 17.6 ± 0.29; *L. major* (+): 21.1 ± 0.17 mU/10^6 cells) and caMΦ (*L. major* (−): 6.7 ± 1.59; *L. major* (+): 47.8 ± 0.78 mU/10^6 cells) only expressed background levels of arginase activity. Significance of the effect was confirmed *via* Kruskal–Wallis test ($p < 0.01$). Additional

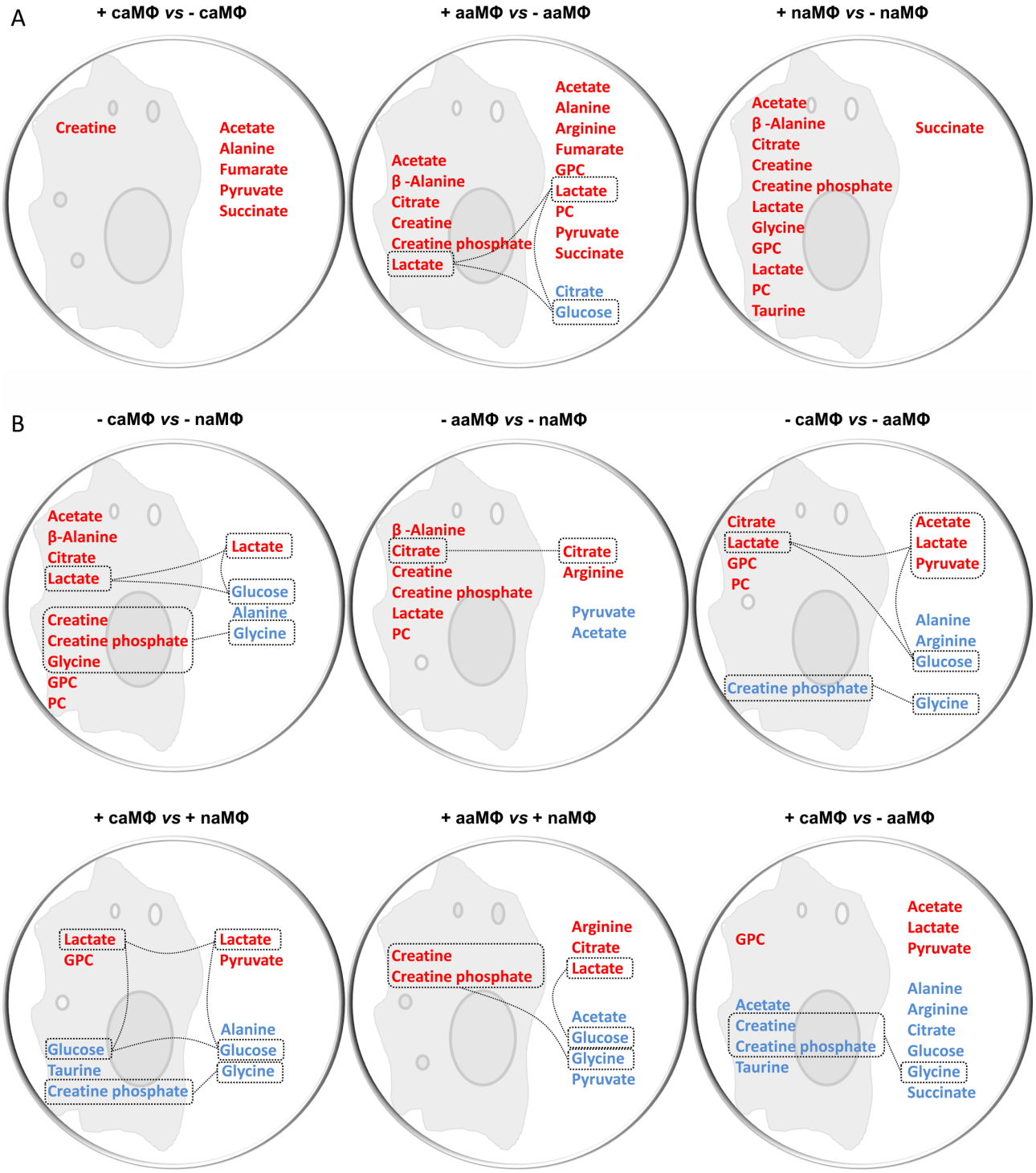
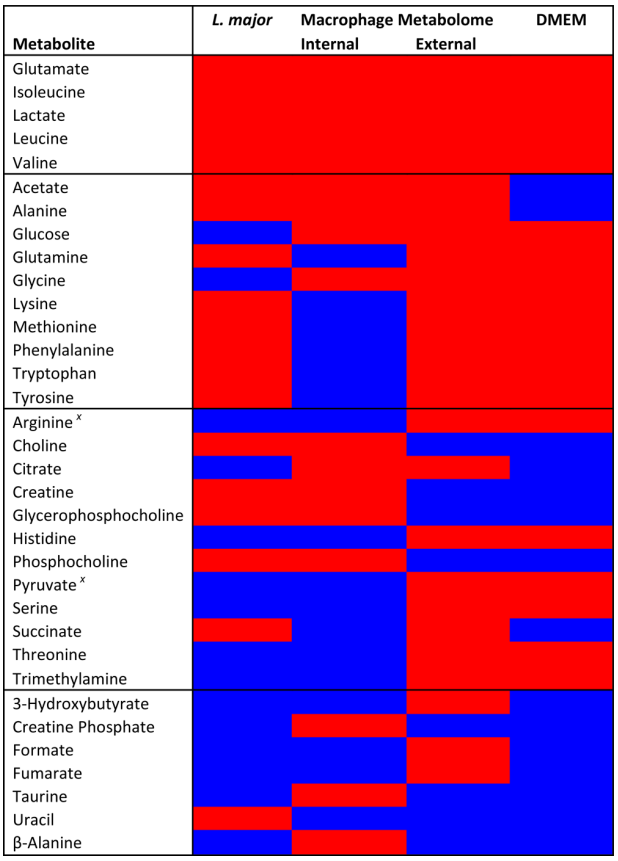


Figure 4. Biomarkers represented in a schematic *in vitro* context arranged in pairwise comparisons: Those metabolites within the macrophage were identified as biomarkers derived from the analysis on the cell extracts, whereas the metabolites outside the cell denote changes characterized in the corresponding cell supernatants. A) Effect of infection within each activation state, whereby metabolites in red indicate a relative increase of the metabolite during infection whereas blue metabolites were found to be relatively lower in macrophages infected with *L. major*. B) Effect of activation state within uninfected cells (top row) and infected macrophages (bottom row). The color code (red, increased; blue, decreased) referes to the first mentioned activation state in each pairwise comparison, for example, in caMΦ vs naMΦ, the red metabolites would refer to caMΦ, indicating that in caMΦ red is relatively increased compared to naMΦ, etc. Key: +, infected with *L. major*; −, uninfected control cells.

Mann−Whitney tests between infection and control groups within each activation class revealed that NO and arginase activity were significantly different between infected and uninfected MΦ throughout all activation groups in which detectable levels were obtained. Whereas NO was found to be relatively higher in *L. major* (+) caMΦ than in *L. major* (−)

dx.doi.org/10.1021/pr3003358 | *J. Proteome Res.* 2012, 11, 4211−4222

Table 4. Metabolites Identified via ^{1}H NMR Spectroscopy in *L. major*, Macrophage Extracts (Internal Metabolome), Macrophage Culture Supernatant (External Metabolome) and Pure DMEM, to Compare Overlap of Metabolic Information between the Compartments[a]



[a]The color code indicates presence (red) and no confirmed detection of a metabolite (blue). GPC, Glycerophosphocholine; *x*, tentatively assigned.

caMΦ, arginase levels were significantly higher in *L. major* (+) aaMΦ and naMΦ compared to each corresponding uninfected set of cells but were found to be relatively lower in *L. major* (+) caMΦ compared to *L. major* (−) caMΦ.

Parasite transformation and determination of promastigote survival after 48 h intracellular exposure to the different activation states confirmed the well established association of NO with parasite killing and arginase with promotion of parasite growth. About 70% fewer promastigotes were recovered from caMΦ ($5 \times 10^6 \pm 0.66 \times 10^6$) whereas higher promastigote numbers were recovered from aaMΦ ($28.3 \times 10^6 \pm 3.12 \times 10^6$) as compared to baseline levels in naMΦ ($22.8 \times 10^6 \pm 1.32 \times 10^6$). The parasite decrease in caMΦ was confirmed as significant by the nonparametric Kruskal–Wallis test between all three activation groups.

DISCUSSION

Macrophage activation plays a critical role in determining disease progress during the course of *L. major* infection. Whereas classical activation leads to clearance of the parasite via generation of reactive nitrogen species, in particular NO, alternative activation will induce increased arginase-mediated conversion of L-arginine to ornithine and polyamines, which are used as substrates for parasite growth and replication.[15,16,20] The measured NO production and arginase activity in the here presented study confirmed that the different stimulation conditions used, resulted in alternatively or classically activated macrophages. The parasite counts furthermore confirmed the leishmanicidal effect of the classical activation pathway products. The metabolic screening of cell extracts via ^{1}H NMR has revealed clear metabolic differences between aaMΦ's, caMΦ's and naMΦ's in *L. major*-infected and uninfected state, whereby the classical activation state separated most clearly from all other conditions. The corresponding cell culture supernatant confirmed the striking separation of the caMΦ's from all other states in the global PCA (Figure 2), indicating that classical activation of the macrophages has a more profound effect on the global cellular metabolism than the coexistence of an intracellular pathogen. The more obvious group separation of *L. major*-exposed cells in the extracts, compared to the cell supernatants, might be due to the exposure of any secreted components to the relatively large volume of external growth medium. A potential dilution of cell derived metabolites might render the detection of more subtle changes to intracellular metabolic pathways more difficult in the extracellular milieu. However, subsequent in depth analysis using O-PLS discriminant analysis was able to detect those subtle differences and reveal the biomarkers responsible for discrimination of infection and activation state. Moreover the biomarker yield in each compartment was dependent on activation state but overall quantitatively comparable (Figure 4).

O-PLS-DA assessment of infection markers within each of the three activation states separately, showed that *L. major* infection induced a similar metabolic fingerprint in all macrophage populations, which included different combinations of relatively elevated levels of actetate, alanine, pyruvate, succinate and lactate in internal and external macrophage metabolome and may, as main end products of *Leishmania*, reflect a direct contribution of the parasite (Figures 3 and 4, Table 1).[2,40]

The strongest biomarkers that were recovered from the analysis of the macrophage extracts included glycerophosphocholine, phosphocholine, creatine phosphate and creatine which were measured at relative higher concentrations in infected samples from aa and naMΦ, compared to the noninfected corresponding samples (Figure 4; Table 1). The increase of choline species might indicate a certain degree of membrane decomposition and subsequent lipid degradation during invasion of *L. major*,[41] which did not seem to have been the case after classical activation. The higher cellular uptake of creatine (aa, ca. and naMΦ's) and creatine phosphate (aa and naMΦ's), on the other hand might reflect the higher energy demand of the macrophages in order to cope with the infection, since creatine phosphate represents a major cellular short-term energy reserve for ATP generation.[42] The caMΦ showed, however, overall the least metabolic differences between *L. major*-infected and uninfected state, which may be explained by the confirmed effective intracellular killing of the parasite and hence a relatively minor accumulation of metabolic end products and intracellular pathway interruptions in contrast to the uncontrolled growth and replication of *L. major* in naMΦ and aaMΦ. Parallel evaluations of arginase and NO indeed confirmed relatively higher leishmanicidal activity and significantly lower parasite counts in caMΦ.

dx.doi.org/10.1021/pr3003358 | *J. Proteome Res.* 2012, 11, 4211–4222

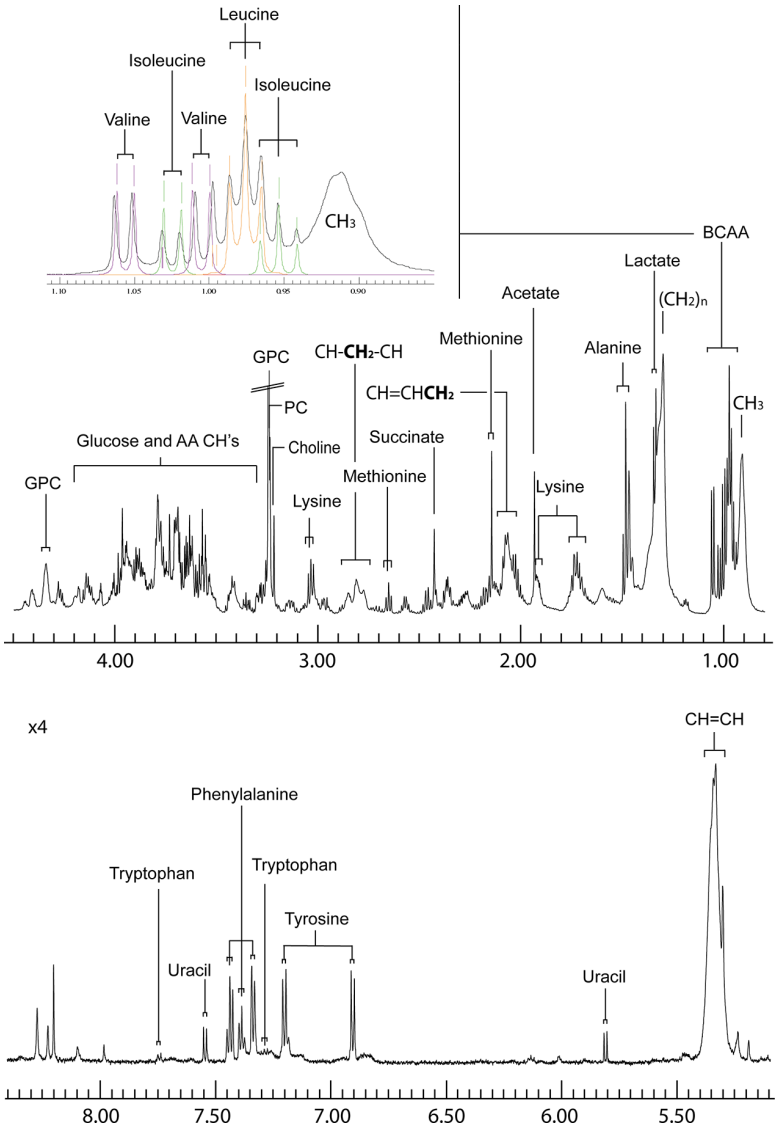


Figure 5. High resolution ^{1}H MAS NMR acquisition of *L. major* pellet using a cpmg pulse sequence to obtain overview on the small metabolic weight components of the parasite. The top figure indicates the alphatic region of a ^{1}H NMR spectrum and the bottom represents CH groups in aromatic systems. The standards for the BCAA assignements were used and overlapped in Amix, Version 3.9.3. Key: AA, Amino acid; BCAA, Branched chain amino acid; GPC, Glycerophosphocholine; PC, Phosphocholine.

The most striking effect related to macrophage activation state was the clear separation of the caMΦ's from both aa and naMΦ's (Figure 2, Tables 2 and 3). Biomarker assessment within each infection state (*L. major* present and absent) revealed that a depletion of glucose in the cell culture supernatants and a subsequent increase of lactate in both macrophage extract and culture supernatants were the main metabolic drivers for the separation and indicate that caMΦ has a significantly higher energy demand compared to the other two groups, to successfully defend infection. An increased amount of ATP is necessary, for instance, for enzyme activity such as the NADPH oxidase which is crucial for the production of reactive oxygen and nitrogen species.[43] It has been shown that the excessive NO production upon classical activation can lead to complete arrest of ATP production via the respiratory chain. Anaerobic glycolysis therefore undergoes an increase to meet the ATP demand in the macrophage.[44] The increased glycolytic rate and the subsequent accumulation of lactate, in turn, might be directly responsible for the observed pH decrease in the cell supernatants of infected caMΦ (caMΦ 6–7; aa and naMΦ's 8–9). The potential contribution of parasitic D-lactate[40] seems herby unlikely, since no lactate increase was observed in infected caMΦ compared to uninfected samples (Figure 4, Table 1).

The majority of metabolites which were found prevalent in all or most compartments assessed were amino acids and glucose to serve as source for energy and anabolic pathways. Choline and its derivatives phosphocholine and glycerophosphocholine (GPC), on the other hand, were found in the intracellular metabolome of parasite and macrophage but not in the external fluids, which might reflect accumulation of the intermediates for the intracellular recycling process of remodelling the cell membrane, since choline is substrate for the phospatidylcholine anabolism.[41,45]

The tricarboxylic acid cycle intermediates citrate, fumarate, pyruvate and succinate were unequally distributed throughout the compartments whereby presence in the external macrophage metabolome was the only commonality. Only succinate was detected in the protozoa via ^{1}H NMR, which might reflect back to the fact that the compound is one of the main metabolic end products of *L. major* with relatively higher accumulation in the parasite when compared with fumarate and citrate for instance.[2,40]

The culture growth medium and the external macrophage metabolome showed the highest degree of intercompartment overlap (i.e., 79.4%). Qualitative differences were detected in only 7 out of 34 metabolites and include acetate, alanine, citrate, succinate, 3-hydroxybutyrate, formate and fumarate which were all present in the macrophage footprint but absent in pure DMEM. Whereas acetate, alanine and succinate are major end products of the *Leishmania* metabolism and might be directly derived from the parasite, the remaining components are likely to reflect metabolic waste products of both, macrophage and parasite metabolism.

CONCLUSION

Work in previous parasite-rodent models has shown the capacity of metabolic profiling in discovery of infection-specific candidate diagnostic biomarkers and in characterizing immune-metabolic codevelopment during infection. Whereas *in vivo*, immune measures such as cytokines are being coanalyzed with the metabolic background in a systemic, untargeted manner, the present *in vitro* approach is exploring metabolic change in macrophages upon specific cytokine stimulation, which represents a more targeted and complementary approach. We are describing a highly efficient and reproducible read out system for characterizing the gross metabolic background of all metabolically relevant compartments involved (i.e., macrophage, parasite and growth media) via ^{1}H NMR spectroscopy. The well-defined compartmentalization and the known baseline metabolic compositions allow much clearer conclusion toward biomarker origin and metabolic translocation across the compartments, compared to complex mammalian systems. The approach therefore holds promise to offer substantial support in identifying parasite-derived compounds as candidate diagnostic biomarkers. Moreover, the efficiency of the presented testing system may be further exploited for evaluating potential parasiticidal metabolic supplements, novel drug candidates, and varying cytokine stimuli and can be easily adaptable to other metabolic profiling platforms if need be. We believe that integration of this more focused *in vitro* strategy into the landscape of untargeted metabolic profiling will help us to access the wealth of information waiting to be explored within the immune-metabolic interface.

AUTHOR INFORMATION

Corresponding Author

*Tel.: +44 20 7594-3899. E-mail: jasmina.saric@imperial.ac.uk. Tel.: +44 20 7594-3732. E-mail: i.muller@imperial.ac.uk.

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

We thank the Wellcome Trust for financial support of the research presented in the current manuscript (Sir Henry Wellcome Fellowship JS, Award number P23665) and the MRC and The Faculty of Medicine at Imperial College for the MRC studentship which provided additional funds for SL. We would also like to acknowledge Dr. Kirill Veselkov for offering updated data processing scripts and general statistical advice.

REFERENCES

(1) WHO *The world health report 2004*; WHO, 2004.
(2) McConville, M. J.; de Souza, D.; Saunders, E.; Likic, V. A.; Naderer, T. Living in a phagolysosome; metabolism of *Leishmania* amastigotes. *Trends Parasitol.* **2007**, *23* (8), 368.
(3) Oliveira, F.; Jochim, R. C.; Valenzuela, J. G.; Kamhawi, S. Sand flies, *Leishmania*, and transcriptome-borne solutions. *Parasitol. Int.* **2009**, *58* (1), 1.
(4) Reithinger, R.; Dujardin, J. C.; Louzir, H.; Pirmez, C.; Alexander, B.; Brooker, S. Cutaneous leishmaniasis. *Lancet Infect. Dis.* **2007**, *7* (9), 581.
(5) Dedet, J. P. P., F. In *Manson's tropical diseases*; Cook, G. C. Z., A. I., Ed. 2009.
(6) Rosenthal, E.; Marty, P.; del Giudice, P.; Pradier, C.; Ceppi, C.; Gastaut, J. A.; Le Fichoux, Y.; Cassuto, J. P. HIV and *Leishmania* coinfection: a review of 91 cases with focus on atypical locations of Leishmania. *Clin. Infect. Dis.* **2000**, *31* (4), 1093.
(7) Beil, W. J.; Meinardus-Hager, G.; Neugebauer, D. C.; Sorg, C. Differences in the onset of the inflammatory response to cutaneous leishmaniasis in resistant and susceptible mice. *J. Leukoc. Biol.* **1992**, *52* (2), 135.
(8) Louis, J.; Moedder, E.; Behin, R.; Engers, H. Recognition of protozoan parasite antigens by murine T lymphocytes. I. Induction of specific T lymphocyte-dependent proliferative response to *Leishmania tropica*. *Eur. J. Immunol.* **1979**, *9* (11), 841.
(9) Mitchell, G. F.; Curtis, J. M.; Scollay, R. G.; Handman, E. Resistance and abrogation of resistance to cutaneous leishmaniasis in reconstituted BALB/c nude mice. *Aust. J. Exp. Biol. Med. Sci.* **1981**, *59* (Pt 5), 539.
(10) Howard, J. G.; Hale, C.; Chan-Liew, W. L. Immunological regulation of experimental cutaneous leishmaniasis. 1. Immunogenetic aspects of susceptibility to *Leishmania tropica* in mice. *Parasite Immunol.* **1980**, *2* (4), 303.
(11) Mosmann, T. R.; Coffman, R. L. TH1 and TH2 cells: different patterns of lymphokine secretion lead to different functional properties. *Annu. Rev. Immunol.* **1989**, 7145.
(12) Tato, C. M.; Laurence, A.; O'Shea, J. J. Helper T cell differentiation enters a new era: le roi est mort; vive le roi! *J. Exp. Med.* **2006**, *203* (4), 809.
(13) O'Shea, J. J.; Paul, W. E. Mechanisms underlying lineage commitment and plasticity of helper CD4+ T cells. *Science* **2010**, *327* (5969), 1098.
(14) Murphy, K. M.; Stockinger, B. Effector T cell plasticity: flexibility in the face of changing circumstances. *Nat. Immunol.* **2010**, *11* (8), 674.
(15) Noel, W.; Raes, G.; Hassanzadeh Ghassabeh, G.; De Baetselier, P.; Beschin, A. Alternatively activated macrophages during parasite infections. *Trends Parasitol.* **2004**, *20* (3), 126.
(16) Kropf, P.; Fuentes, J. M.; Fahnrich, E.; Arpa, L.; Herath, S.; Weber, V.; Soler, G.; Celada, A.; Modolell, M.; Muller, I. Arginase and polyamine synthesis are key factors in the regulation of experimental leishmaniasis *in vivo*. *FASEB J.* **2005**, *19* (8), 1000.
(17) Alberati-Giani, D.; Malherbe, P.; Ricciardi-Castagnoli, P.; Kohler, C.; Denis-Donini, S.; Cesura, A. M. Differential regulation of indoleamine 2,3-dioxygenase expression by nitric oxide and inflammatory mediators in IFN-gamma-activated murine macrophages and microglial cells. *J. Immunol.* **1997**, *159* (1), 419.
(18) Li, P.; Yin, Y. L.; Li, D.; Kim, S. W.; Wu, G. Amino acids and immune function. *Br. J. Nutr.* **2007**, *98* (2), 237.
(19) Wu, G.; Meininger, C. J. Regulation of nitric oxide synthesis by dietary factors. *Annu. Rev. Nutr.* **2002**, 2261.
(20) Choi, B. S.; Martinez-Falero, I. C.; Corset, C.; Munder, M.; Modolell, M.; Muller, I.; Kropf, P. Differential impact of L-arginine

deprivation on the activation and effector functions of T cells and macrophages. *J. Leukoc. Biol.* **2009**, *85* (2), 268.
(21) Kropf, P.; Baud, D.; Marshall, S. E.; Munder, M.; Mosley, A.; Fuentes, J. M.; Bangham, C. R.; Taylor, G. P.; Herath, S.; Choi, B. S.; Soler, G.; Teoh, T.; Modolell, M.; Muller, I. Arginase activity mediates reversible T cell hyporesponsiveness in human pregnancy. *Eur. J. Immunol.* **2007**, *37* (4), 935.
(22) Modolell, M.; Choi, B. S.; Ryan, R. O.; Hancock, M.; Titus, R. G.; Abebe, T.; Hailu, A.; Muller, I.; Rogers, M. E.; Bangham, C. R.; Munder, M.; Kropf, P. Local suppression of T cell responses by arginase-induced L-arginine depletion in nonhealing leishmaniasis. *PLoS Negl. Trop. Dis.* **2009**, *3* (7), e480.
(23) Saric, J.; Li, J. V.; Wang, Y.; Keiser, J.; Veselkov, K.; Dirnhofer, S.; Yap, I. K.; Nicholson, J. K.; Holmes, E.; Utzinger, J. Panorganismal metabolic response modeling of an experimental *Echinostoma caproni* infection in the mouse. *J. Proteome Res.* **2009**, *8* (8), 3899.
(24) Wang, Y.; Utzinger, J.; Saric, J.; Li, J. V.; Burckhardt, J.; Dirnhofer, S.; Nicholson, J. K.; Singer, B. H.; Brun, R.; Holmes, E. Global metabolic responses of mice to *Trypanosoma brucei brucei* infection. *Proc. Natl. Acad. Sci. U.S.A.* **2008**, *105* (16), 6127.
(25) Saric, J.; Li, J. V.; Swann, J. R.; Utzinger, J.; Calvert, G.; Nicholson, J. K.; Dirnhofer, S.; Dallman, M. J.; Bictash, M.; Holmes, E. Integrated cytokine and metabolic analysis of pathological responses to parasite exposure in rodents. *J. Proteome Res.* **2010**, *9* (5), 2255.
(26) Saric, J.; Li, J. V.; Utzinger, J.; Wang, Y.; Keiser, J.; Dirnhofer, S.; Beckonert, O.; Sharabiani, M. T.; Fonville, J. M.; Nicholson, J. K.; Holmes, E. Systems parasitology: effects of *Fasciola hepatica* on the neurochemical profile in the rat brain. *Mol. Syst. Biol.* **2010**, 6396.
(27) Birungi, G.; Chen, S. M.; Loy, B. P.; Ng, M. L.; Li, S. F. Metabolomics approach for investigation of effects of dengue virus infection using the EA.hy926 cell line. *J. Proteome Res.* **2010**, *9* (12), 6523.
(28) Ellis, J. K.; Chan, P. H.; Doktorova, T.; Athersuch, T. J.; Cavill, R.; Vanhaecke, T.; Rogiers, V.; Vinken, M.; Nicholson, J. K.; T, M. D. E.; Keun, H. C. Effect of the histone deacetylase inhibitor trichostatin a on the metabolome of cultured primary hepatocytes. *J. Proteome Res.* **2010**, *9* (1), 413.
(29) Ellis, J. K.; Athersuch, T. J.; Cavill, R.; Radford, R.; Slattery, C.; Jennings, P.; McMorrow, T.; Ryan, M. P.; Ebbels, T. M.; Keun, H. C. Metabolic response to low-level toxicant exposure in a novel renal tubule epithelial cell system. *Mol. Biosyst.* **2011**, *7* (1), 247.
(30) Beckonert, O.; Coen, M.; Keun, H. C.; Wang, Y.; Ebbels, T. M.; Holmes, E.; Lindon, J. C.; Nicholson, J. K. High-resolution magic-angle-spinning NMR spectroscopy for metabolic profiling of intact tissues. *Nat. Protoc.* **2010**, *5* (6), 1019.
(31) Meiboom, S. G. D. Modified spin-echo method for measuring nuclear relaxation times. *Rev. Sci. Instrum.* **1958**, 29688.
(32) Cloarec, O.; Dumas, M. E.; Craig, A.; Barton, R. H.; Trygg, J.; Hudson, J.; Blancher, C.; Gauguier, D.; Lindon, J. C.; Holmes, E.; Nicholson, J. Statistical total correlation spectroscopy: an exploratory approach for latent biomarker identification from metabolic ^{1}H NMR data sets. *Anal. Chem.* **2005**, *77* (5), 1282.
(33) Veselkov, K. A.; Vingara, L. K.; Masson, P.; Robinette, S. L.; Want, E.; Li, J. V.; Barton, R. H.; Boursier-Neyret, C.; Walther, B.; Ebbels, T. M.; Pelczer, I.; Holmes, E.; Lindon, J. C.; Nicholson, J. K. Optimized preprocessing of ultra-performance liquid chromatography/mass spectrometry urinary metabolic profiles for improved information recovery. *Anal. Chem.* **2011**, *83* (15), 5864.
(34) Dieterle, F.; Ross, A.; Schlotterbeck, G.; Senn, H. Probabilistic quotient normalization as robust method to account for dilution of complex biological mixtures. Application in ^{1}H NMR metabonomics. *Anal. Chem.* **2006**, *78* (13), 4281.
(35) Eriksson, L.; Antti, H.; Gottfries, J.; Holmes, E.; Johansson, E.; Lindgren, F.; Long, I.; Lundstedt, T.; Trygg, J.; Wold, S. Using chemometrics for navigating in the large data sets of genomics, proteomics, and metabonomics (gpm). *Anal. Bioanal. Chem.* **2004**, *380* (3), 419.
(36) Trygg, J.; Holmes, E.; Lundstedt, T. Chemometrics in metabonomics. *J. Proteome Res.* **2007**, *6* (2), 469.
(37) Cloarec, O.; Dumas, M. E.; Trygg, J.; Craig, A.; Barton, R. H.; Lindon, J. C.; Nicholson, J. K.; Holmes, E. Evaluation of the orthogonal projection on latent structure model limitations caused by chemical shift variability and improved visualization of biomarker changes in ^{1}H NMR spectroscopic metabonomic studies. *Anal. Chem.* **2005**, *77* (2), 517.
(38) Trygg, J. W. S. Orthogonal projections to latent structures (O-PLS). *J. Chemom.* **2002**, *16* (3), 119.
(39) Munder, M.; Eichmann, K.; Modolell, M. Alternative metabolic states in murine macrophages reflected by the nitric oxide synthase/arginase balance: competitive regulation by CD4+ T cells correlates with Th1/Th2 phenotype. *J. Immunol.* **1998**, *160* (11), 5347.
(40) Opperdoes, F. R.; Coombs, G. H. Metabolism of Leishmania: proven and predicted. *Trends Parasitol.* **2007**, *23* (4), 149.
(41) Li, Z.; Vance, D. E. Phosphatidylcholine and choline homeostasis. *J. Lipid Res.* **2008**, *49* (6), 1187.
(42) Wyss, M.; Kaddurah-Daouk, R. Creatine and creatinine metabolism. *Physiol. Rev.* **2000**, *80* (3), 1107.
(43) Naderer, T.; McConville, M. J. The *Leishmania*-macrophage interaction: a metabolic perspective. *Cell Microbiol.* **2008**, *10* (2), 301.
(44) Garedew, A.; Henderson, S. O.; Moncada, S. Activated macrophages utilize glycolytic ATP to maintain mitochondrial membrane potential and prevent apoptotic cell death. *Cell Death Differ.* **2010**, *17* (10), 1540.
(45) Zeisel, S. H.; Mar, M. H.; Howe, J. C.; Holden, J. M. Concentrations of choline-containing compounds and betaine in common foods. *J. Nutr.* **2003**, *133* (5), 1302.

Laboratory Clinical Diagnostics in Côte d'Ivoire:

Case Studies: Methodist Hospital of Dabou & Taabo General Hospital

By Sabrina D. Lamour*,
* Department Surgery & Cancer, Imperial College London, London SW7 2AZ
Contact email: sabrina.lamour10@imperial.ac.uk

1. Overview Methodist Hospital of Dabou



Fig 1. Map of Côte d'Ivoire (modified from ClimateZone.com, 2004)

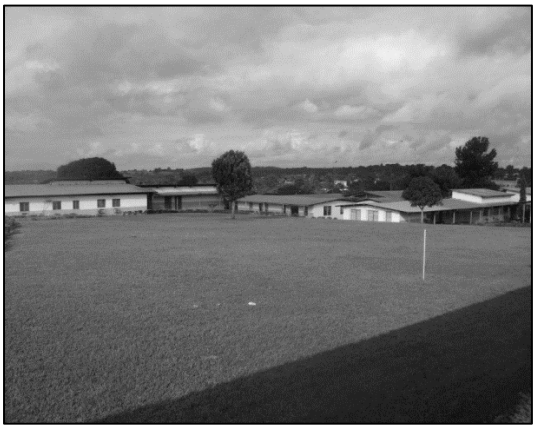


Fig 2. Photo of HDM grounds (S. Lamour, taken Oct 2013)

Established in 1968, the Methodist Hospital of Dabou (HDM) is a private hospital located approximately 50 km west from Côte d'Ivoire's former capital city, Abidjan (Fig. 1). The hospital is spread over several buildings (Fig. 2) accompanied by large grounds which also include housing and schools for many of the staff and their families. In total, HDM employs approximately 140 staff members – including six doctors – and offers over 100 beds (UMCOR, 2011). Due to the wide-range of services (with prices comparable to those of the state-owned and subsidized General Dabou Hospital), HDM's patient base stretches far beyond the town of Dabou itself (+80,000 inhabitants; Wikipedia), receiving patients from all over the southern region of Côte d'Ivoire. These services include:

- General Medicine
- General Surgery
- Paediatrics
- Gynaecology and Obstetrics
- Family planning
- Dental Care
- Cardiology
- Radiology
- Clinical Laboratory
- Pharmacy
- Outpatient care
- Counselling, Campaigns and teaching (e.g. HIV/AIDS, sexual health, etc.)

Despite the high patient turn-over at this hospital, it is important to note that financial resources of HDM still remain limited, with many patients struggling to pay for the costs of tests, treatments and/or operations. Much of the infrastructure, equipment and drugs at the hospital have been made available or improved over the years through collaborations and donations from international sources. Despite regular servicing, the majority of their machines and facilities could use some updating.

2. Clinical Laboratory Services available at HDM

The clinical laboratory staff at HDM consists of approximately ten staff members, of which over half are permanent staff (primarily laboratory technicians and a few doctors) and the rest are trainees. Each staff member is placed on a weekly rota responsible for running and recording the various types of clinical tests. The laboratory itself is spread over four rooms (next to the operating block). The first point of entry is a small room designated for blood sample collection, followed by the main open-plan lab shown in Fig. 3A,

used for sample deposition, patient/sample registration and records, serum preparation (centrifugation) and some basic blood tests. The third (and largest) room (Fig. 3B) is used for slide staining and microscopic evaluation (bacterial and parasitic tests), immunological and biochemical analyses, electrophoreses and also contains a separate office (for the lab manager) and a curtained area with a bed, for vaginal sample swabbing. Lastly, the fourth room (Fig. 3C) is where full blood counts are performed and it is the designated area for HIV-associated tests (including CD4 T cell counts and use of HIV test strips). The full range of clinical tests at the laboratory have been summarised in Table 1.

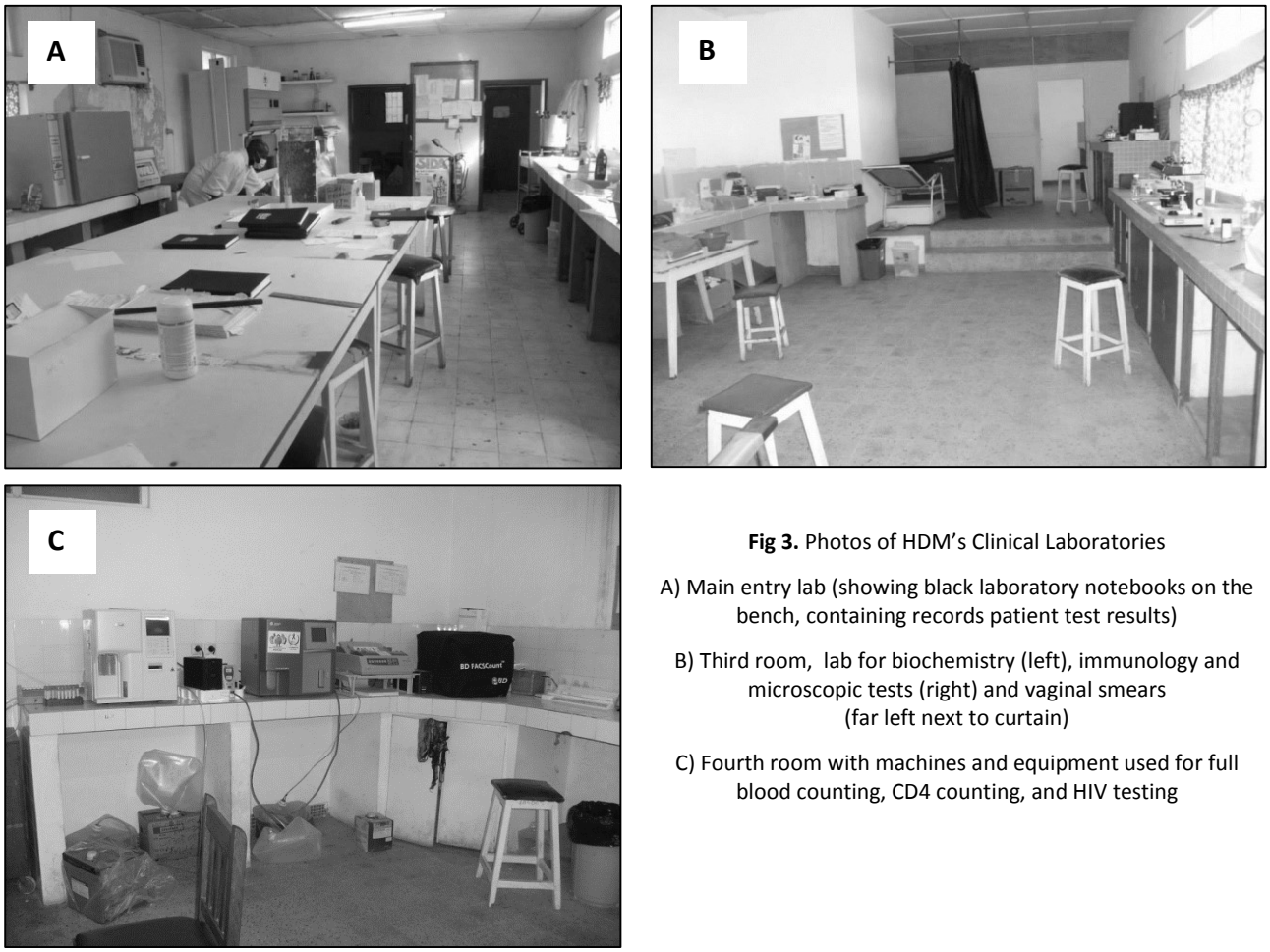


Fig 3. Photos of HDM's Clinical Laboratories

A) Main entry lab (showing black laboratory notebooks on the bench, containing records patient test results)

B) Third room, lab for biochemistry (left), immunology and microscopic tests (right) and vaginal smears (far left next to curtain)

C) Fourth room with machines and equipment used for full blood counting, CD4 counting, and HIV testing

Typically, the patient arrives to the lab with his/her form obtained from the doctor (either at HDM or elsewhere), accompanied by the receipt to prove that they have paid for the tests, though in many cases nurses can come directly to the lab bringing these forms for multiple patients. Patients wait are then called to give blood or other samples, which are tested the same day. The vast majority of results from these tests are reported back (on the same forms) to the relevant hospital staff and/or patient a few hours later, yet in some circumstances – notably the bacteriology tests – results may take several days later as they may require long follow-up tests (e.g. bacterial cultures by the microbiology department or are done out-of-house) or in some instances, due to shortage of reagents and/or broken-down equipment.

Abbreviations for Table 1 (in English and French): Ab= Antibody, Ag= Antigen, ASO= Anti-Streptolysine-O, BK= "Bacille de Koch" test, BW= Bordet-Wassermann test, CRP= C-reactive Protein, FS= "Frottis Sanguin", GE= "Goutte Epaisse", Hb= Haemoglobin, KOP=" Kyerte Ouef "Parasite, NFS= "Numeration Formule Sanguine", RBC= Red Blood Cell, TP-TCK= time for prothrombin and thromboplastin activation, WBC= White Blood Cell, TB= Tuberculosis, TPHA= Treponema pallidum Haemagglutination test, VDRL= Venereal Disease Research Laboratory, VS= "Vitesse de sedimentation"

Table 1. Clinical Tests performed at HDM Laboratory

Clinical Area	Available Tests	Purpose	Sample Type
Bacteriology (*via* microscopic examination)	Vaginal smear	Test for presence of pathogenic bacteria (e.g. due to sexual transmitted diseases), including, *Gonococci*, *Trichonomas vagina*, *Monilia* and Gram positive/negative staining of bacteria	Vaginal/urethral fluid
	BK test	Identification of *Mycobacterium* (TB)	Sputum
Biochemistry (*via* spectrophotometric tests)	Albumin	Test for markers of general well-being, including kidney and liver function, fasting glucose (diabetes), and various other metabolites	Serum and/or urine
	Bilirubin (total & conjugated)		
	Calcium		
	Chlorides		
	Cholesterol (total, HDL, VDL)		
	Creatinine		
	Glucose (fasting)		
	Magnesium		
	Phosphorus		
	Phosphatases		
	Potassium		
	Protein (total)		
	Sodium		
	Transaminases		
	Triglycerides		
	Urea		
	Uric Acid		
Other Urine Tests	Acetone "Acetest"	Detection of acetonuria	Urine
	Glucose "Clinitest"	Detecting of glucose (diabetes)	
	Bilirubin "Ictotest"	Detection of bilirubin and by-products	
	Urobilinogen test		
	Microscopic examination (cytology-bacteriology)	Detection of presence of anomalies including leukocytes, blood, epithelial cells, uric crystals, parasites, etc.)	Urine (whole or on supernatant and/or sediment following centrifugation)
Haematology	Full blood count (NFS) using automatic blood counter (Results include WBC count, RBC count, platelet count, Hb and Haematocrit levels, and counts/% abundance of the different types of WBC's)	General markers of immune system, anaemia and platelet counts	Whole Blood
	Sedimentation rate (VS)	General marker of blood thickness (indicative for blood anomalies)	
	Hb Electrophoresis	Determination of haemoglobinopathies (sickle-cell disease, thalassemia, etc.)	
Immunology (*via* Ab-Ag agglutination tests)	TP-TCK + Fibrenemia	Determination of speed/ability of blood (thrombin, fibrin) coagulation (indicative for blood anomalies)	Serum or Whole Blood
	ASO	Inflammatory marker	
	CRP	Inflammatory marker	
	VDRL + THPA (BW test)	Detection of Abs against Syphilis (*T. pallidum* bacteria)	
	Widal & Felix test	Detection of Abs against *Salmonella typhi* Ags (TO & TH) and/or *Paratyphi* Ags (AO, BO, CO, AH, BH, CH), markers of typhoid fever	
	Toxoplasma test	Detection of Abs against *Toxoplasma* infection	
	Rubella test	Detection of Abs against Rubella infection	
	Blood Group test	Determination of Blood Type (A, B, AB, O) and Rhesus state (positive/negative)	Whole Blood

Immunology (*via* other non-agglutination tests)	HIV test (via test strip)	Determination of HIV infection	Whole Blood or Serum
	HIV Type test (via test strip)	Determination of HIV-1 and/or HIV-2 infection	
	CD4 T cell count (via test kit & BD FACS counter)	Determination of CD4 T cell count (marker of HIV/AIDS pathology)	HIV positive Whole Blood
	Hepatitis test	Determination of Hepatitis Infection	Whole Blood or Serum
	Hepatitis Type test	Determination of Hepatitis A, B or C infection	Hepatitis positive Whole Blood
	Pregnancy test (via test strip)	Pregnancy	Urine
Parasitology (*via* Microscopic Examination)	Thick Blood Smear test (GE)	Malarial parasite detection (burden)	Whole Blood
	Thin Blood Smear test (FS)	Malarial parasite detection (species and stages)	Whole Blood
	Stool Ova and Parasite (KOP) test	Detection of gastro-intestinal parasitic infections	Faeces
Miscellaneous	Spermogram	Assessing male fertility	Sperm Fluid

Sample processing and waste management

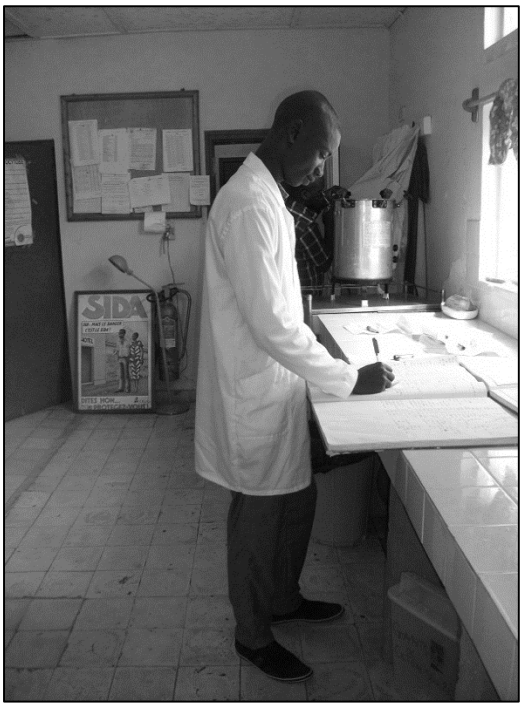


Fig 4. Patient registration performed by one of the staff members (next to bench where blood samples have been deposited). The autoclave machine can be seen in the far right corner, used sterilisation of sample waste

Based on their records, 600-900 patient samples are typically processed by the laboratory each month. All patient samples (blood, faeces, sputum, urine, etc.) are deposited in the main open-plan lab (Fig. 4), whatever the health status of the patient (including HIV+ and TB+ samples). Blood samples are collected using disposable needles into standard BD Vacutainer® blood collection tubes, whilst other biological fluids are placed in plastic sealable containers. Samples are labelled with requested tests, patient names and sample numbers and recorded into the registration book (Fig. 4). Copies of all test results are recorded manually into laboratory notebooks (Fig. 3A). All lab users are required to wash and disinfect their hands using soap and disinfectant. Lab coats and gloves are (generally) worn by lab workers throughout sample handling, whereby all samples except blood are disposed (into waste bins) after laboratory tests have been completed. Blood and serum sample tubes are stored at room temperature for up 1 week on the bench (in case a subsequent repeat test measurement is required) and then disposed by autoclaving tubes (Fig. 4), and pouring waste down the drain.

Test strips (e.g. for HIV, pregnancy, etc.) are disinfected overnight in diluted bleach (10% of "L'eau de Javel", a sodium hypochlorite bleaching/disinfecting agent) and then discarded. Due to limited resources, pipette tips, test tubes, glass slides, etc. are re-used, following a procedure of overnight disinfection in diluted bleach, rinsing and air drying. All other solid waste placed into waste bins which are regularly collected by waste management team and are incinerated (though due to lack of proper incineration equipment, these are currently burned *via* more crude measures and then covered over with earth, located in designated areas of the hospital grounds). It is well worth noting that despite high patient turn-over, the

3. Overview of Taabo General Hospital



Fig 5. Photo of Taabo General Hospital

Taabo Cité is a small town (~8000 inhabitants) is situated approximately 85 km south-east of the current capital Yamassoukro (Fig. 1). It is most known for its hydro-electric dam around Lake Taabo, constructed in 1979, which is currently responsible for up to 35% of the nation's electric power (Wikipedia). In this largely rural setting, the small Taabo General Hospital (Fig. 5) is the only hospital in the area, providing basic primary and secondary health care services to the local population of Taabo Cité and neighbouring villages. These services include:

- General Medicine
- General Surgery
- Clinical Laboratory
- Pharmacy
- Buruli Ulcer Clinic
- Vaccination programs

Due to a number of reasons – including a high economic burden for a relatively poor community, a high proportion of illiteracy and other cultural factors – the hospital services are rarely used to their full capacity. The busiest day for the hospital coincides with Taabo's market day (every Wednesday), when people from all neighbouring villages come to sell their local produce and arrange their hospital visits in the morning.

4. Clinical Laboratory Services available at Taabo Hospital

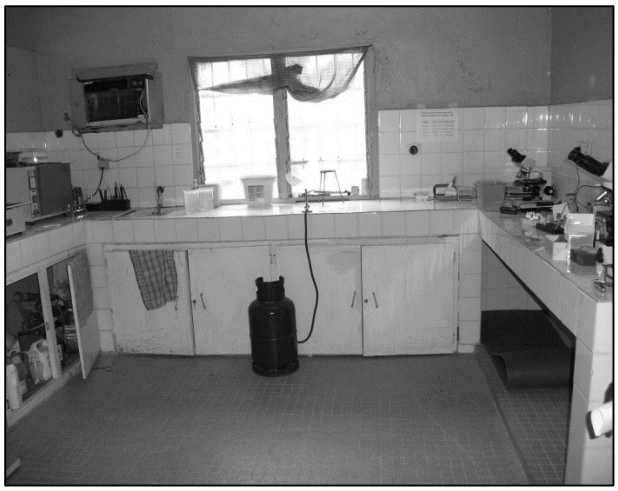


Fig 6. Photo of Taabo Hospital's Clinical Laboratory (Main room)

In comparison to the HDM, the laboratory at Taabo General hospital is much more resource-limited, reflected in relatively older and more basic equipment and infrastructure, as well as more limited reagents (Fig. 6). There is also a shortage of some of the personal protective equipment, including limited number of clean lab coats and gloves. The laboratory is manned by two people – a senior and a junior laboratory technician – and is spread over two rooms (and an office for the senior technician). The first main room (Fig. 6) is used for sample collection (primarily blood samples), microscopic examinations (e.g. of blood smears, etc.), blood type and immunological tests. The second room contains the automatic equipment required for biochemical analyses, full blood counts and a centrifuge. The list of clinical tests available at Taabo Laboratory are shown below in Table 2:

Table2. Clinical Tests performed at Taabo Hospital Laboratory

Clinical Area	Available Tests	Purpose	Sample Type
Cytology (*via* microscopic examination and/or culture methods)	Cytology- Bacteriology Exam (ECBU)	Detection of abnormal presence/abundance of bacteria, leukocytes and other cell types	Urine
	Cytology- Haematology Exam		Whole Blood
Biochemistry	Albumin	Test for markers of general well-being, etc.	Serum and/or urine
	Creatinine		
	Glucose (fasting)		
	Transaminases		
	Urea		
	Uric Acid		Urine
	Acetonuria		
	Proteinuria		
Haematology	Full blood count (NFS) using automatic blood counter	General markers of immune system, anaemia and platelet counts	Whole Blood
	Sedimentation rate (VS)	General marker of blood thickness (indicative for blood anomalies)	
	Hb Electrophoresis	Determination of haemoglobinopathies (sickle-cell disease, thalassemia, etc.)	
	Hb- Haematocrit	General marker of blood thickness (indicative for blood anomalies)	
	Indirect Coomb's Test	Test for autoimmune haemolytic anaemia	Serum
Immunology	CRP	Inflammatory marker	Serum or Whole Blood
	VDRL + THPA (BW test)	Detection of Abs against Syphilis (*T. pallidum* bacteria)	
	Widal & Felix test	Detection of Abs against *Salmonella typhi* Ags (TO & TH) and/or *Paratyphi* Ags (AO, BO, CO, AH, BH, CH), markers of typhoid fever	
	Toxoplasma test	Detection of Abs against *Toxoplasma* infection	
	Rubella test	Detection of Abs against Rubella infection	
	Blood Group test	Determination of Blood Type (A, B, AB, O) and Rhesus state (positive/negative)	Whole Blood
	HIV test (via test strip)	Determination of HIV infection	Whole Blood
	HIV Type test (via test strip)	Determination of HIV-1 and/or HIV-2 infection	
Parasitology (*via* Microscopic Examination)	Thick Blood Smear test (GE)	Malarial parasite detection (burden)	Whole Blood
	Thin Blood Smear test (FS)	Malarial parasite detection (species and stages)	Whole Blood
	Stool Ova and Parasite (KOP) test	Detection of gastro-intestinal parasitic infections (cysts, eggs, worms)	Faeces
Miscellaneous	Protein Electrophoresis	(Unspecified)	(Unspecified)
	Throat sampling	(Unspecified)	(Unspecified)

Abbreviations for Table 2 as for Table 1, ECBU= Examen cytobactériologique des urines

Sample processing and waste management

Similar to HDM, patient test results are recorded into one laboratory book, though the process is much more ad hoc, due to the low number of patient turn over (typically only up to five patients a day come to visit the clinical laboratory). The vast majority of samples tested are blood samples, collected using disposable needles and BD Vacutainer® tubes. Tests are run shortly after, though unlike at HDM, glass slides and tubes are not recycled (new tubes/slides are used each time). All waste into a large waste bin that is incinerated by waste management (using a hospital-owned incinerator).

5. Summary

The presented report gives a brief outline of the clinical laboratory services in an urban and rural setting in southern Côte d'Ivoire. In both cases, financial resources for health care services are limited (e.g. in contrast to services available in Western Europe, e.g. Audit Scotland, 2008; an example of clinical services available by the NHS), putting several restraints on the availability and/or feasibility of clinical tests that can be run. This is especially apparent in the rural setting of Taabo, whereby the incapacity to pay for the health care severely hinders the local population in approaching the hospital for help. Despite these shortcomings, the staff have consistently shown considerable amount of resilience and pride in their work, and continue to strive to improve their work and ultimately help patients, *via* their motto:

"*Doing the best we can with what we have."*

6. References

Audit Scotland, 2008: http://www.audit-scotland.gov.uk/docs/health/2008/nr_081120_diagnostic_services.pdf

ClimateZone.com, 2004. URL: http://www.climate-zone.com/climate/ivory-coast/

ONUCI 2013. URL http://www.onuci.org/spip.php?article9745

UMCOR, 2011. URL: http://new.gbgm-umc.org/umcor/work/health/international/hospitalsandclinics/dabou/

Wikipedia. URL 1: http://fr.wikipedia.org/wiki/Dabou
URL 2: http://fr.wikipedia.org/wiki/Taabo

Please note that unless referenced, all photographs were taken personally by Sabrina Lamour October 2013, following consent from the pictured individuals and/or hospital staff.

7. Acknowledgements

The author would like to acknowledge and thank the *Medical Research Council* and the *British Society for Parasitology* for their support in funding the author's visit to Côte d'Ivoire through their international fieldwork awards. The author would also like to thank the *Centre Suisse de Recherches Scientifiques en Côte d'Ivoire* and *Swiss Tropical and Public Health Institute* for their support and their guidance, particularly in the organisation of the trip.

Permissions for Third Party Copyright Works

Table A.10: Summary of Permission for Third Party Copyright Works

Page Number	Type of work: text, figure, map, etc.	Source work	Copyright holder & year	Permission to reproduce
Page 25	Figure	Kaye, P and Scott, P (2011) Nature Reviews Microbiology , vol 9 (8), pp. 604-615, Figure 1	© 2011 Nature Publishing Group	✓
Page 25	Text	Kaye, P and Scott, P (2011) Nature Reviews Microbiology , vol 9 (8), pp. 604-615, Figure 1 legend	© 2011 Nature Publishing Group	✓
Page 28	Figure & text*	Saunders, E et al. (2010) Parasitology, vol 137 (09), pp. 1303-1313, Figure 2 and legend	© 2010 Cambridge University Press	✓
Page 32	Figure	Brun, R et al. (2010) The Lancet, vol 375 (9709), pp. 148-159, Figure 1	© 2010 Elsevier	✓
Page 34	Table	Brun, R et al. (2010) The Lancet, vol 375 (9709), pp. 148-159, Table 1	© 2010 Elsevier	✓
Page 37	Figure	Langousis, G and Hill, K L (2014) Nature Reviews Microbiology, 12(7), pp. 505-518, Box 1 Figure	© 2014 Nature Publishing Group	✓

*Permission license for Figure legend attached separately. Disclaimer for permission of re-use of the Figure is as follows:
Cambridge University Press grants permission freely for the reproduction in another work of a short prose extract (less than 400 words), a single figure or a single table in which it holds rights (see the important caveat in the Notes below). In such cases a request for permission need not be submitted, but the reproduced material must be accompanied by a full citation of the original source.
Note also that Cambridge University Press may not be able to grant permission for the use of all the material it publishes. In particular, illustrations, tables, poetry or other extensive quotations included in Cambridge publications may have been reproduced by Cambridge with the permission of another rights holder. Always check the caption or citation to ensure that Cambridge is the rights holder before proceeding.
`http://www.cambridge.org/about-us/rights-permissions/permissions/permissions-requests`

Please see attached licenses for all other the items listed above, in the order as shown in the table.

NATURE PUBLISHING GROUP LICENSE
TERMS AND CONDITIONS

Jan 14, 2016

This is a License Agreement between Sabrina D Lamour ("You") and Nature Publishing Group ("Nature Publishing Group") provided by Copyright Clearance Center ("CCC"). The license consists of your order details, the terms and conditions provided by Nature Publishing Group, and the payment terms and conditions.

All payments must be made in full to CCC. For payment instructions, please see information listed at the bottom of this form.

License Number	3787670113808
License date	Jan 14, 2016
Licensed content publisher	Nature Publishing Group
Licensed content publication	Nature Reviews Microbiology
Licensed content title	Leishmaniasis: complexity at the host–pathogen interface
Licensed content author	Paul Kaye and Phillip Scott
Licensed content date	Jul 11, 2011
Volume number	9
Issue number	8
Type of Use	reuse in a dissertation / thesis
Requestor type	academic/educational
Format	print and electronic
Portion	figures/tables/illustrations
Number of figures/tables/illustrations	1
High-res required	no
Figures	FIGURE 1. The life cycle of Leishmania parasites
Author of this NPG article	no
Your reference number	fig1_lifecycle
Title of your thesis / dissertation	Investigation of Host Metabolic Interactions in Trypanosomatid Infections
Expected completion date	Jan 2016
Estimated size (number of pages)	320
Total	0.00 GBP
Terms and Conditions	

Terms and Conditions for Permissions

Nature Publishing Group hereby grants you a non-exclusive license to reproduce this material for this purpose, and for no other use,subject to the conditions below:

1. NPG warrants that it has, to the best of its knowledge, the rights to license reuse of this material. However, you should ensure that the material you are requesting is original to Nature Publishing Group and does not carry the copyright of another entity (as credited in the published version). If the credit line on any part of the material you have requested

indicates that it was reprinted or adapted by NPG with permission from another source, then you should also seek permission from that source to reuse the material.

2. Permission granted free of charge for material in print is also usually granted for any electronic version of that work, provided that the material is incidental to the work as a whole and that the electronic version is essentially equivalent to, or substitutes for, the print version.Where print permission has been granted for a fee, separate permission must be obtained for any additional, electronic re-use (unless, as in the case of a full paper, this has already been accounted for during your initial request in the calculation of a print run).NB: In all cases, web-based use of full-text articles must be authorized separately through the 'Use on a Web Site' option when requesting permission.

3. Permission granted for a first edition does not apply to second and subsequent editions and for editions in other languages (except for signatories to the STM Permissions Guidelines, or where the first edition permission was granted for free).

4. Nature Publishing Group's permission must be acknowledged next to the figure, table or abstract in print. In electronic form, this acknowledgement must be visible at the same time as the figure/table/abstract, and must be hyperlinked to the journal's homepage.

5. The credit line should read:
   Reprinted by permission from Macmillan Publishers Ltd: [JOURNAL NAME] (reference citation), copyright (year of publication)
   For AOP papers, the credit line should read:
   Reprinted by permission from Macmillan Publishers Ltd: [JOURNAL NAME], advance online publication, day month year (doi: 10.1038/sj.[JOURNAL ACRONYM].XXXXX)

   Note: For republication from the *British Journal of Cancer,* the following credit lines apply.
   Reprinted by permission from Macmillan Publishers Ltd on behalf of Cancer Research UK: [JOURNAL NAME] (reference citation), copyright (year of publication)For AOP papers, the credit line should read:
   Reprinted by permission from Macmillan Publishers Ltd on behalf of Cancer Research UK: [JOURNAL NAME], advance online publication, day month year (doi: 10.1038/sj.[JOURNAL ACRONYM].XXXXX)

6. Adaptations of single figures do not require NPG approval. However, the adaptation should be credited as follows:

   Adapted by permission from Macmillan Publishers Ltd: [JOURNAL NAME] (reference citation), copyright (year of publication)

   Note: For adaptation from the *British Journal of Cancer,* the following credit line applies.
   Adapted by permission from Macmillan Publishers Ltd on behalf of Cancer Research UK: [JOURNAL NAME] (reference citation), copyright (year of publication)

7. Translations of 401 words up to a whole article require NPG approval. Please visit http://www.macmillanmedicalcommunications.com for more information.Translations of up to a 400 words do not require NPG approval. The translation should be credited as follows:

   Translated by permission from Macmillan Publishers Ltd: [JOURNAL NAME] (reference citation), copyright (year of publication).

   Note: For translation from the *British Journal of Cancer,* the following credit line applies.
   Translated by permission from Macmillan Publishers Ltd on behalf of Cancer Research UK: [JOURNAL NAME] (reference citation), copyright (year of publication)

We are certain that all parties will benefit from this agreement and wish you the best in the use of this material. Thank you.
Special Terms:

NATURE PUBLISHING GROUP LICENSE
TERMS AND CONDITIONS

Jan 14, 2016

This is a License Agreement between Sabrina D Lamour ("You") and Nature Publishing Group ("Nature Publishing Group") provided by Copyright Clearance Center ("CCC"). The license consists of your order details, the terms and conditions provided by Nature Publishing Group, and the payment terms and conditions.

All payments must be made in full to CCC. For payment instructions, please see information listed at the bottom of this form.

License Number	3787670594306
License date	Jan 14, 2016
Licensed content publisher	Nature Publishing Group
Licensed content publication	Nature Reviews Microbiology
Licensed content title	Leishmaniasis: complexity at the host–pathogen interface
Licensed content author	Paul Kaye and Phillip Scott
Licensed content date	Jul 11, 2011
Volume number	9
Issue number	8
Type of Use	reuse in a dissertation / thesis
Requestor type	academic/educational
Format	print and electronic
Portion	Text extract
Extract size	101 to 250 words
Circulation/distribution	>50,000
Author of this NPG article	no
Your reference number	fig1_legend
Title of your thesis / dissertation	Investigation of Host Metabolic Interactions in Trypanosomatid Infections
Expected completion date	Jan 2016
Estimated size (number of pages)	320
Total	0.00 GBP
Terms and Conditions	

Terms and Conditions for Permissions

Nature Publishing Group hereby grants you a non-exclusive license to reproduce this material for this purpose, and for no other use,subject to the conditions below:

1. NPG warrants that it has, to the best of its knowledge, the rights to license reuse of this material. However, you should ensure that the material you are requesting is original to Nature Publishing Group and does not carry the copyright of another entity (as credited in the published version). If the credit line on any part of the material you have requested indicates that it was reprinted or adapted by NPG with permission from another source, then you should also seek permission from that source to reuse the material.

2. Permission granted free of charge for material in print is also usually granted for any electronic version of that work, provided that the material is incidental to the work as a whole and that the electronic version is essentially equivalent to, or substitutes for, the print version.Where print permission has been granted for a fee, separate permission must be obtained for any additional, electronic re-use (unless, as in the case of a full paper, this has already been accounted for during your initial request in the calculation of a print run).NB: In all cases, web-based use of full-text articles must be authorized separately through the 'Use on a Web Site' option when requesting permission.

3. Permission granted for a first edition does not apply to second and subsequent editions and for editions in other languages (except for signatories to the STM Permissions Guidelines, or where the first edition permission was granted for free).

4. Nature Publishing Group's permission must be acknowledged next to the figure, table or abstract in print. In electronic form, this acknowledgement must be visible at the same time as the figure/table/abstract, and must be hyperlinked to the journal's homepage.

5. The credit line should read:
   Reprinted by permission from Macmillan Publishers Ltd: [JOURNAL NAME] (reference citation), copyright (year of publication)
   For AOP papers, the credit line should read:
   Reprinted by permission from Macmillan Publishers Ltd: [JOURNAL NAME], advance online publication, day month year (doi: 10.1038/sj.[JOURNAL ACRONYM].XXXXX)

   Note: For republication from the *British Journal of Cancer,* the following credit lines apply.
   Reprinted by permission from Macmillan Publishers Ltd on behalf of Cancer Research UK: [JOURNAL NAME] (reference citation), copyright (year of publication)For AOP papers, the credit line should read:
   Reprinted by permission from Macmillan Publishers Ltd on behalf of Cancer Research UK: [JOURNAL NAME], advance online publication, day month year (doi: 10.1038/sj.[JOURNAL ACRONYM].XXXXX)

6. Adaptations of single figures do not require NPG approval. However, the adaptation should be credited as follows:

   Adapted by permission from Macmillan Publishers Ltd: [JOURNAL NAME] (reference citation), copyright (year of publication)

   Note: For adaptation from the *British Journal of Cancer,* the following credit line applies.
   Adapted by permission from Macmillan Publishers Ltd on behalf of Cancer Research UK: [JOURNAL NAME] (reference citation), copyright (year of publication)

7. Translations of 401 words up to a whole article require NPG approval. Please visit http://www.macmillanmedicalcommunications.com for more information.Translations of up to a 400 words do not require NPG approval. The translation should be credited as follows:

   Translated by permission from Macmillan Publishers Ltd: [JOURNAL NAME] (reference citation), copyright (year of publication).

   Note: For translation from the *British Journal of Cancer,* the following credit line applies.
   Translated by permission from Macmillan Publishers Ltd on behalf of Cancer Research UK: [JOURNAL NAME] (reference citation), copyright (year of publication)

We are certain that all parties will benefit from this agreement and wish you the best in the use of this material. Thank you.
Special Terms:
v1.1

Questions? customercare@copyright.com or +1-855-239-3415 (toll free in the US) or +1-978-646-2777.

CAMBRIDGE UNIVERSITY PRESS LICENSE TERMS AND CONDITIONS

Jan 14, 2016

This is a License Agreement between Sabrina D Lamour ("You") and Cambridge University Press ("Cambridge University Press") provided by Copyright Clearance Center ("CCC"). The license consists of your order details, the terms and conditions provided by Cambridge University Press, and the payment terms and conditions.

All payments must be made in full to CCC. For payment instructions, please see information listed at the bottom of this form.

License Number	3787671280542
License date	Jan 14, 2016
Licensed content publisher	Cambridge University Press
Licensed content publication	Parasitology
Licensed content title	Central carbon metabolism of *Leishmania* parasites
Licensed content author	ELEANOR C. SAUNDERS, DAVID P. DE SOUZA, THOMAS NADERER, MARIJKE F. SERNEE, JULIE E. RALTON, MARIA A. DOYLE, JAMES I. MACRAE, JENNY L. CHAMBERS, JOANNE HENG, AMSHA NAHID, VLADIMIR A. LIKIC and MALCOLM J. MCCONVILLE
Licensed content date	Feb 17, 2010
Volume number	137
Issue number	09
Start page	1303
End page	1313
Type of Use	Dissertation/Thesis
Requestor type	Not-for-profit
Portion	Text extract
Number of pages requested	1
Order reference number	Fig2_legend_text
Territory for reuse	World
Title of your thesis / dissertation	Investigation of Host Metabolic Interactions in Trypanosomatid Infections
Expected completion date	Jan 2016
Estimated size(pages)	320
Billing Type	Invoice
Billing Address	Sabrina D Lamour Computational and Systems Medicine Sir Alexander Fleming Building Exhibition Road London, United Kingdom SW7 2AZ Attn: Sabrina D Lamour
TAX (0.00%)	0.00 GBP
Total	0.00 GBP

Terms and Conditions

TERMS & CONDITIONS

Cambridge University Press grants the Licensee permission on a non-exclusive non-transferable basis to reproduce, make available or otherwise use the Licensed content 'Content' in the named territory 'Territory' for the purpose listed 'the Use' on Page 1 of this Agreement subject to the following terms and conditions.

1. The License is limited to the permission granted and the Content detailed herein and does not extend to any other permission or content.
2. Cambridge gives no warranty or indemnity in respect of any third-party copyright material included in the Content, for which the Licensee should seek separate permission clearance.
3. The integrity of the Content must be ensured.
4. The License does extend to any edition published specifically for the use of handicapped or reading-impaired individuals.
5. The Licensee shall provide a prominent acknowledgement in the following format: author/s, title of article, name of journal, volume number, issue number, page references, , reproduced with permission.

Other terms and conditions:
v1.0

Questions? customercare@copyright.com or +1-855-239-3415 (toll free in the US) or +1-978-646-2777.

ELSEVIER LICENSE
TERMS AND CONDITIONS

Jan 14, 2016

This is a License Agreement between Sabrina D Lamour ("You") and Elsevier ("Elsevier") provided by Copyright Clearance Center ("CCC"). The license consists of your order details, the terms and conditions provided by Elsevier, and the payment terms and conditions.

All payments must be made in full to CCC. For payment instructions, please see information listed at the bottom of this form.

Supplier	Elsevier Limited The Boulevard,Langford Lane Kidlington,Oxford,OX5 1GB,UK
Registered Company Number	1982084
Customer name	Sabrina D Lamour
Customer address	Computational and Systems Medicine London, SW7 2AZ
License number	3787670803117
License date	Jan 14, 2016
Licensed content publisher	Elsevier
Licensed content publication	The Lancet
Licensed content title	Human African trypanosomiasis
Licensed content author	Reto Brun,Johannes Blum,Francois Chappuis,Christian Burri
Licensed content date	9–15 January 2010
Licensed content volume number	375
Licensed content issue number	9709
Number of pages	12
Start Page	148
End Page	159
Type of Use	reuse in a thesis/dissertation
Intended publisher of new work	other
Portion	figures/tables/illustrations
Number of figures/tables/illustrations	2
Format	both print and electronic
Are you the author of this Elsevier article?	No
Will you be translating?	No
Order reference number	fig1_map_tab
Original figure numbers	Figure 1 and Table 1
Title of your	Investigation of Host Metabolic Interactions in Trypanosomatid

thesis/dissertation	Infections
Expected completion date	Jan 2016
Estimated size (number of pages)	320
Elsevier VAT number	GB 494 6272 12
Permissions price	0.00 GBP
VAT/Local Sales Tax	0.00 GBP / 0.00 GBP
Total	0.00 GBP

Terms and Conditions

INTRODUCTION

1. The publisher for this copyrighted material is Elsevier. By clicking "accept" in connection with completing this licensing transaction, you agree that the following terms and conditions apply to this transaction (along with the Billing and Payment terms and conditions established by Copyright Clearance Center, Inc. ("CCC"), at the time that you opened your Rightslink account and that are available at any time at http://myaccount.copyright.com).

GENERAL TERMS

2. Elsevier hereby grants you permission to reproduce the aforementioned material subject to the terms and conditions indicated.
3. Acknowledgement: If any part of the material to be used (for example, figures) has appeared in our publication with credit or acknowledgement to another source, permission must also be sought from that source. If such permission is not obtained then that material may not be included in your publication/copies. Suitable acknowledgement to the source must be made, either as a footnote or in a reference list at the end of your publication, as follows:
"Reprinted from Publication title, Vol /edition number, Author(s), Title of article / title of chapter, Pages No., Copyright (Year), with permission from Elsevier [OR APPLICABLE SOCIETY COPYRIGHT OWNER]." Also Lancet special credit - "Reprinted from The Lancet, Vol. number, Author(s), Title of article, Pages No., Copyright (Year), with permission from Elsevier."
4. Reproduction of this material is confined to the purpose and/or media for which permission is hereby given.
5. Altering/Modifying Material: Not Permitted. However figures and illustrations may be altered/adapted minimally to serve your work. Any other abbreviations, additions, deletions and/or any other alterations shall be made only with prior written authorization of Elsevier Ltd. (Please contact Elsevier at permissions@elsevier.com)
6. If the permission fee for the requested use of our material is waived in this instance, please be advised that your future requests for Elsevier materials may attract a fee.
7. Reservation of Rights: Publisher reserves all rights not specifically granted in the combination of (i) the license details provided by you and accepted in the course of this licensing transaction, (ii) these terms and conditions and (iii) CCC's Billing and Payment terms and conditions.
8. License Contingent Upon Payment: While you may exercise the rights licensed immediately upon issuance of the license at the end of the licensing process for the transaction, provided that you have disclosed complete and accurate details of your proposed use, no license is finally effective unless and until full payment is received from you (either by publisher or by CCC) as provided in CCC's Billing and Payment terms and conditions. If full payment is not received on a timely basis, then any license preliminarily granted shall be deemed automatically revoked and shall be void as if never granted. Further, in the event that you breach any of these terms and conditions or any of CCC's Billing and Payment terms and conditions, the license is automatically revoked and shall be void as if never

granted. Use of materials as described in a revoked license, as well as any use of the materials beyond the scope of an unrevoked license, may constitute copyright infringement and publisher reserves the right to take any and all action to protect its copyright in the materials.

9. Warranties: Publisher makes no representations or warranties with respect to the licensed material.

10. Indemnity: You hereby indemnify and agree to hold harmless publisher and CCC, and their respective officers, directors, employees and agents, from and against any and all claims arising out of your use of the licensed material other than as specifically authorized pursuant to this license.

11. No Transfer of License: This license is personal to you and may not be sublicensed, assigned, or transferred by you to any other person without publisher's written permission.

12. No Amendment Except in Writing: This license may not be amended except in a writing signed by both parties (or, in the case of publisher, by CCC on publisher's behalf).

13. Objection to Contrary Terms: Publisher hereby objects to any terms contained in any purchase order, acknowledgment, check endorsement or other writing prepared by you, which terms are inconsistent with these terms and conditions or CCC's Billing and Payment terms and conditions. These terms and conditions, together with CCC's Billing and Payment terms and conditions (which are incorporated herein), comprise the entire agreement between you and publisher (and CCC) concerning this licensing transaction. In the event of any conflict between your obligations established by these terms and conditions and those established by CCC's Billing and Payment terms and conditions, these terms and conditions shall control.

14. Revocation: Elsevier or Copyright Clearance Center may deny the permissions described in this License at their sole discretion, for any reason or no reason, with a full refund payable to you. Notice of such denial will be made using the contact information provided by you. Failure to receive such notice will not alter or invalidate the denial. In no event will Elsevier or Copyright Clearance Center be responsible or liable for any costs, expenses or damage incurred by you as a result of a denial of your permission request, other than a refund of the amount(s) paid by you to Elsevier and/or Copyright Clearance Center for denied permissions.

LIMITED LICENSE

The following terms and conditions apply only to specific license types:

15. **Translation**: This permission is granted for non-exclusive world **English** rights only unless your license was granted for translation rights. If you licensed translation rights you may only translate this content into the languages you requested. A professional translator must perform all translations and reproduce the content word for word preserving the integrity of the article.

16. **Posting licensed content on any Website**: The following terms and conditions apply as follows: Licensing material from an Elsevier journal: All content posted to the web site must maintain the copyright information line on the bottom of each image; A hyper-text must be included to the Homepage of the journal from which you are licensing at http://www.sciencedirect.com/science/journal/xxxxx or the Elsevier homepage for books at http://www.elsevier.com; Central Storage: This license does not include permission for a scanned version of the material to be stored in a central repository such as that provided by Heron/XanEdu.

Licensing material from an Elsevier book: A hyper-text link must be included to the Elsevier homepage at http://www.elsevier.com . All content posted to the web site must maintain the copyright information line on the bottom of each image.

Posting licensed content on Electronic reserve: In addition to the above the following clauses are applicable: The web site must be password-protected and made available only to bona fide students registered on a relevant course. This permission is granted for 1 year only. You may obtain a new license for future website posting.

17. **For journal authors:** the following clauses are applicable in addition to the above:

Preprints:

A preprint is an author's own write-up of research results and analysis, it has not been peer-reviewed, nor has it had any other value added to it by a publisher (such as formatting, copyright, technical enhancement etc.).

Authors can share their preprints anywhere at any time. Preprints should not be added to or enhanced in any way in order to appear more like, or to substitute for, the final versions of articles however authors can update their preprints on arXiv or RePEc with their Accepted Author Manuscript (see below).

If accepted for publication, we encourage authors to link from the preprint to their formal publication via its DOI. Millions of researchers have access to the formal publications on ScienceDirect, and so links will help users to find, access, cite and use the best available version. Please note that Cell Press, The Lancet and some society-owned have different preprint policies. Information on these policies is available on the journal homepage.

Accepted Author Manuscripts: An accepted author manuscript is the manuscript of an article that has been accepted for publication and which typically includes author-incorporated changes suggested during submission, peer review and editor-author communications.

Authors can share their accepted author manuscript:

- immediately
  - via their non-commercial person homepage or blog
  - by updating a preprint in arXiv or RePEc with the accepted manuscript
  - via their research institute or institutional repository for internal institutional uses or as part of an invitation-only research collaboration work-group
  - directly by providing copies to their students or to research collaborators for their personal use
  - for private scholarly sharing as part of an invitation-only work group on commercial sites with which Elsevier has an agreement
- after the embargo period
  - via non-commercial hosting platforms such as their institutional repository
  - via commercial sites with which Elsevier has an agreement

In all cases accepted manuscripts should:

- link to the formal publication via its DOI
- bear a CC-BY-NC-ND license - this is easy to do
- if aggregated with other manuscripts, for example in a repository or other site, be shared in alignment with our hosting policy not be added to or enhanced in any way to appear more like, or to substitute for, the published journal article.

Published journal article (JPA): A published journal article (PJA) is the definitive final record of published research that appears or will appear in the journal and embodies all value-adding publishing activities including peer review co-ordination, copy-editing, formatting, (if relevant) pagination and online enrichment.

Policies for sharing publishing journal articles differ for subscription and gold open access articles:

Subscription Articles: If you are an author, please share a link to your article rather than the full-text. Millions of researchers have access to the formal publications on ScienceDirect, and so links will help your users to find, access, cite, and use the best available version.

Theses and dissertations which contain embedded PJAs as part of the formal submission can be posted publicly by the awarding institution with DOI links back to the formal publications on ScienceDirect.

If you are affiliated with a library that subscribes to ScienceDirect you have additional

private sharing rights for others' research accessed under that agreement. This includes use for classroom teaching and internal training at the institution (including use in course packs and courseware programs), and inclusion of the article for grant funding purposes.
Gold Open Access Articles: May be shared according to the author-selected end-user license and should contain a CrossMark logo, the end user license, and a DOI link to the formal publication on ScienceDirect.
Please refer to Elsevier's posting policy for further information.
18. **For book authors** the following clauses are applicable in addition to the above: Authors are permitted to place a brief summary of their work online only. You are not allowed to download and post the published electronic version of your chapter, nor may you scan the printed edition to create an electronic version. **Posting to a repository:** Authors are permitted to post a summary of their chapter only in their institution's repository.
19. **Thesis/Dissertation**: If your license is for use in a thesis/dissertation your thesis may be submitted to your institution in either print or electronic form. Should your thesis be published commercially, please reapply for permission. These requirements include permission for the Library and Archives of Canada to supply single copies, on demand, of the complete thesis and include permission for Proquest/UMI to supply single copies, on demand, of the complete thesis. Should your thesis be published commercially, please reapply for permission. Theses and dissertations which contain embedded PJAs as part of the formal submission can be posted publicly by the awarding institution with DOI links back to the formal publications on ScienceDirect.

Elsevier Open Access Terms and Conditions
You can publish open access with Elsevier in hundreds of open access journals or in nearly 2000 established subscription journals that support open access publishing. Permitted third party re-use of these open access articles is defined by the author's choice of Creative Commons user license. See our open access license policy for more information.
Terms & Conditions applicable to all Open Access articles published with Elsevier:
Any reuse of the article must not represent the author as endorsing the adaptation of the article nor should the article be modified in such a way as to damage the author's honour or reputation. If any changes have been made, such changes must be clearly indicated.
The author(s) must be appropriately credited and we ask that you include the end user license and a DOI link to the formal publication on ScienceDirect.
If any part of the material to be used (for example, figures) has appeared in our publication with credit or acknowledgement to another source it is the responsibility of the user to ensure their reuse complies with the terms and conditions determined by the rights holder.
Additional Terms & Conditions applicable to each Creative Commons user license:
CC BY: The CC-BY license allows users to copy, to create extracts, abstracts and new works from the Article, to alter and revise the Article and to make commercial use of the Article (including reuse and/or resale of the Article by commercial entities), provided the user gives appropriate credit (with a link to the formal publication through the relevant DOI), provides a link to the license, indicates if changes were made and the licensor is not represented as endorsing the use made of the work. The full details of the license are available at http://creativecommons.org/licenses/by/4.0.
CC BY NC SA: The CC BY-NC-SA license allows users to copy, to create extracts, abstracts and new works from the Article, to alter and revise the Article, provided this is not done for commercial purposes, and that the user gives appropriate credit (with a link to the formal publication through the relevant DOI), provides a link to the license, indicates if changes were made and the licensor is not represented as endorsing the use made of the work. Further, any new works must be made available on the same conditions. The full details of the license are available at http://creativecommons.org/licenses/by-nc-sa/4.0.
CC BY NC ND: The CC BY-NC-ND license allows users to copy and distribute the Article, provided this is not done for commercial purposes and further does not permit distribution of the Article if it is changed or edited in any way, and provided the user gives appropriate

credit (with a link to the formal publication through the relevant DOI), provides a link to the license, and that the licensor is not represented as endorsing the use made of the work. The full details of the license are available at http://creativecommons.org/licenses/by-nc-nd/4.0. Any commercial reuse of Open Access articles published with a CC BY NC SA or CC BY NC ND license requires permission from Elsevier and will be subject to a fee.
Commercial reuse includes:

- Associating advertising with the full text of the Article
- Charging fees for document delivery or access
- Article aggregation
- Systematic distribution via e-mail lists or share buttons

Posting or linking by commercial companies for use by customers of those companies.

20. **Other Conditions**:

v1.8

Questions? customercare@copyright.com or +1-855-239-3415 (toll free in the US) or +1-978-646-2777.

NATURE PUBLISHING GROUP LICENSE
TERMS AND CONDITIONS

Jan 14, 2016

This is a License Agreement between Sabrina D Lamour ("You") and Nature Publishing Group ("Nature Publishing Group") provided by Copyright Clearance Center ("CCC"). The license consists of your order details, the terms and conditions provided by Nature Publishing Group, and the payment terms and conditions.

All payments must be made in full to CCC. For payment instructions, please see information listed at the bottom of this form.

License Number	3787671061113
License date	Jan 14, 2016
Licensed content publisher	Nature Publishing Group
Licensed content publication	Nature Reviews Microbiology
Licensed content title	Motility and more: the flagellum of Trypanosoma brucei
Licensed content author	Gerasimos Langousis, Kent L. Hill
Licensed content date	Jun 16, 2014
Volume number	12
Issue number	7
Type of Use	reuse in a dissertation / thesis
Requestor type	academic/educational
Format	print and electronic
Portion	figures/tables/illustrations
Number of figures/tables/illustrations	1
High-res required	no
Figures	Box 1 Trypanosoma brucei life cycle
Author of this NPG article	no
Your reference number	box1_tryps_lifecycle
Title of your thesis / dissertation	Investigation of Host Metabolic Interactions in Trypanosomatid Infections
Expected completion date	Jan 2016
Estimated size (number of pages)	320
Total	0.00 GBP
Terms and Conditions	

Terms and Conditions for Permissions

Nature Publishing Group hereby grants you a non-exclusive license to reproduce this material for this purpose, and for no other use,subject to the conditions below:

1. NPG warrants that it has, to the best of its knowledge, the rights to license reuse of this material. However, you should ensure that the material you are requesting is original to Nature Publishing Group and does not carry the copyright of another entity (as credited in the published version). If the credit line on any part of the material you have requested

indicates that it was reprinted or adapted by NPG with permission from another source, then you should also seek permission from that source to reuse the material.

2. Permission granted free of charge for material in print is also usually granted for any electronic version of that work, provided that the material is incidental to the work as a whole and that the electronic version is essentially equivalent to, or substitutes for, the print version.Where print permission has been granted for a fee, separate permission must be obtained for any additional, electronic re-use (unless, as in the case of a full paper, this has already been accounted for during your initial request in the calculation of a print run).NB: In all cases, web-based use of full-text articles must be authorized separately through the 'Use on a Web Site' option when requesting permission.

3. Permission granted for a first edition does not apply to second and subsequent editions and for editions in other languages (except for signatories to the STM Permissions Guidelines, or where the first edition permission was granted for free).

4. Nature Publishing Group's permission must be acknowledged next to the figure, table or abstract in print. In electronic form, this acknowledgement must be visible at the same time as the figure/table/abstract, and must be hyperlinked to the journal's homepage.

5. The credit line should read:
   Reprinted by permission from Macmillan Publishers Ltd: [JOURNAL NAME] (reference citation), copyright (year of publication)
   For AOP papers, the credit line should read:
   Reprinted by permission from Macmillan Publishers Ltd: [JOURNAL NAME], advance online publication, day month year (doi: 10.1038/sj.[JOURNAL ACRONYM].XXXXX)

   Note: For republication from the *British Journal of Cancer,* the following credit lines apply.
   Reprinted by permission from Macmillan Publishers Ltd on behalf of Cancer Research UK: [JOURNAL NAME] (reference citation), copyright (year of publication)For AOP papers, the credit line should read:
   Reprinted by permission from Macmillan Publishers Ltd on behalf of Cancer Research UK: [JOURNAL NAME], advance online publication, day month year (doi: 10.1038/sj.[JOURNAL ACRONYM].XXXXX)

6. Adaptations of single figures do not require NPG approval. However, the adaptation should be credited as follows:

   Adapted by permission from Macmillan Publishers Ltd: [JOURNAL NAME] (reference citation), copyright (year of publication)

   Note: For adaptation from the *British Journal of Cancer,* the following credit line applies.
   Adapted by permission from Macmillan Publishers Ltd on behalf of Cancer Research UK: [JOURNAL NAME] (reference citation), copyright (year of publication)

7. Translations of 401 words up to a whole article require NPG approval. Please visit http://www.macmillanmedicalcommunications.com for more information.Translations of up to a 400 words do not require NPG approval. The translation should be credited as follows:

   Translated by permission from Macmillan Publishers Ltd: [JOURNAL NAME] (reference citation), copyright (year of publication).

   Note: For translation from the *British Journal of Cancer,* the following credit line applies.
   Translated by permission from Macmillan Publishers Ltd on behalf of Cancer Research UK: [JOURNAL NAME] (reference citation), copyright (year of publication)

We are certain that all parties will benefit from this agreement and wish you the best in the use of this material. Thank you.
Special Terms:

v1.1

Questions? customercare@copyright.com or +1-855-239-3415 (toll free in the US) or +1-978-646-2777.