Investigation of circulatory and tissue-specific metabolomic biomarkers in valvular heart disease using mass spectrometry

Student: Daniel Wakiri Mutithu

Student Number: MTTDAN008

Submitted in fulfilment of the requirements for the degree of

Doctor of Philosophy

Department of Medicine
Faculty of Health Sciences
UNIVERSITY OF CAPE TOWN

Date of submission (17.05.2025)

Supervisor [s]:

Professor Ntobeko Ntusi; Department of Medicine, Faculty of Health Sciences, University of Cape Town, and Groote Schuur Hospital.

Professor Richard Naidoo; Department of Pathology, Faculty of Health Sciences, University of Cape Town, and National Health Laboratory Services.

Professor Sebastian Skatulla; Department of Civil Engineering, Faculty of Engineering and the Built Environment, University of Cape Town.

Declaration of originality

I, DANIEL WAKIRI MUTITHU, hereby declare that the work on which this thesis is based is my original work (except where acknowledgement indicate otherwise) and that neither the whole work nor any part of it has been, is being, or is to be submitted for another degree in this or any other university.

I empower the university to reproduce for the purpose of research either the whole or any portion of the contents in any manner whatsoever.

Signature: ……

Date: ……

Declaration of publications inclusion

I, DANIEL WAKIRI MUTITHU, confirm that I have been granted permission by the University of Cape Town's Doctoral Degrees Board to include the following publication(s) in my thesis, and where co-authorships are involved, my co-authors have agreed that I may include the publication(s):

1. **Mutithu, D. W.**, Aremu, O. O., Mokaila, D., Bana, T., Familusi, M., Taylor, L.,... & Ntusi, N. A. (2024). A study protocol to characterise pathophysiological and molecular markers of rheumatic heart disease and degenerative aortic stenosis using multiparametric cardiovascular imaging and multi-omics techniques. PLOS ONE, 19(5), e0303496. https://doi.org/10.1371/journal.pone.0303496
2. **Mutithu, D. W.**, Kirwan, J. A., Adeola, H. A., Aremu, O. O., Lumngwena, E. N., Wiesner, L., ... & Ntusi, N. A. (2023). High-Throughput Metabolomics Applications in Pathogenesis and Diagnosis of Valvular Heart Disease. Reviews in Cardiovascular Medicine, 24(6), 169. https://doi.org/10.31083/j.rcm2406169
3. **Mutithu D. W.**, Roberts R., Manganyi R., Ntusi N. A. B. (2022). Chronic rheumatic heart disease with recrudescence of acute rheumatic fever on histology: a case report. Eur Hear J - Case Reports, 6(7), ytac278. https://doi.org/10.1093/ehjcr/ytac278
4. **Mutithu, D. W.**, Kirwan, J. A., Adeola, H. A., Aremu, O. O., Lumngwena, E. N., Familusi, M., ... & Ntusi, N. A. (2024). Serum metabolic profiling in rheumatic heart disease and degenerative aortic stenosis [Manuscript submitted for peer review]. JBTS, JBTS022324-0135.
5. **Mutithu, D. W.**, Kirwan, J. A., Adeola, H. A., Aremu, O. O., Lumngwena, E. N., Familusi, M., ... & Ntusi, N. A. (2024). Major energetic metabolites associated with echocardiographic and histological features of aortic and mitral valve disease. [Manuscript submitted for peer review].

Signature: ……

Date: ……

Declaration of ethical approval

I, DANIEL WAKIRI MUTITHU, hereby declare that the research in this thesis has the approval of the Human Research Ethics Committee of the University of Cape Town, South Africa. Ethics clearance reference number: HREC REF 574/2018 linked to 554/2017.

Signature: ……

Date: ……

Dedication

I dedicate this work to all whom who had to go out of their ways to hold the door for me to achieve my dreams. Through your actions I got a sit at the table I never knew existed. You are many but I remember each one of you. Thank you!

I dedicate it to my wonderful parents Mr. and Mrs Mutithu who at all odds fed me with hope when even hoping seemed unachievable. Your efforts and dedication were not in vain.

To Mrs Francesca Pim-Kaime and family. You saw the star even before it could begin shinning and through your sacrifice and mentorship you helped place it in the clear open sky. Thank you so much for your help with secondary school fees sponsorship. I am forever indebted.

I also dedicate this work to my daughter Edlyne Wakiri and her mum. Most of the time I was away from home, I am sorry. I hope you will one day understand that it was not in vain and that your mere presence in my life kept me going.

Lastly and most importantly I dedicate this work to God. May it be a testament to all who believe in God that His Grace and blessings does follow those who put their trust and faith on Him. When things are dark and hopeless, do your level best and pray to God to guide your steps, do not despair. Also, that there is no dream too big that it cannot be achieved.

Acknowledgements

I would like to convey my special thanks to my academic advisors Prof Ntobeko Ntusi, Prof Richard Naidoo, and Prof Sebastian Skatulla. Thank you for providing the required support to achieve the training, research mentorship, and access to the finest research resources. Thank you for also entrusting me with the running of the project and believing in my capacity, it was a learning experience for all of us. Thank you to Prof S Skatulla and Prof N Ntusi for providing the scholarship during my training. Special thanks to Prof Jennifer A Kirwan for providing mentorship, technical assistance, and guidance during the data analysis, manuscripts writing, and drafting of this thesis. Thanks to Prof Karen Sliwa for offering a scholarship for biostatistics capacity building training.

I also acknowledge Dr R Manganyi for helping obtain valve fragment samples from the Christiaan Barnard Division of Cardiothoracic Surgery, Groote Schuur Hospital, Cape Town. Subash Govender, Dr Riyaadh Roberts, and Prof Dhiren Govender for assistance with histology and histopathology diagnosis of the valve samples.

Special thanks go to A/Prof Lubbe Wiesner and A/Prof Henry Adeola for accepting to give me access to their mass spectrometry facilities and being collaborators in this research. Furthermore, my gratitude goes to Alicia Evans, and Nandi Mehlala for their technical support and training on the mass spectrometry experiments. I would also like to thank all the nurses and doctors at the Groote Schuur Hospital Cardiac clinic for their hospitality and support during patients' recruitment.

I would also like to acknowledge help from friends and colleagues at the institute; Dr. Aremu Olukayode, Dr. Evelyn Lumngwena, and Dr Mathilda Mennen to name but a few. I also thank colleagues at the Metabolomics South Africa association for professional mentorship. I acknowledge social support from friends who made my stay in Cape Town feel like home away from home, more specifically; L Mhlanti, L Andera, T Sikhakhane, Dr S Kariuki, Dr S Irungu,

and J Wambugu. Special thanks to my special friend who offered a shoulder to lean on and helped me bear the difficult circumstances during the writing of this thesis.

Lastly, I would like to convey my heartfelt gratitude to my parents Moses and Eunice Mutithu for their love, unconditional support, teaching me resilience, value for hard work and commitment to a dream, prayers, and spiritual support. I would also like to thank Edlyne and Esther, my siblings, the Maina's, and the Wakiri Sr's for supporting me through the process.

Table of contents

Abstract

Background: Cardiovascular diseases contribute to approximately 37% of the noncommunicable disease related deaths. Africa has the largest burden of rheumatic heart disease affecting about 33 million people aged below 50 years. Furthermore, the prevalence of degenerative aortic stenosis has been on the rise in sub-Saharan Africa which further complicates the management of valvular heart disease in these regions. The pathomechanisms leading to rheumatic heart disease and aortic stenosis pathology have not been fully understood.

Study aims: Therefore, in this study we aimed to use metabolomics techniques to explore the metabolic biomarkers in patients with rheumatic heart disease and degenerative aortic stenosis patients undergoing valve replacement surgery.

Methods: Whole blood samples were collected from matched rheumatic heart disease and aortic stenosis patients, as well as matched controls. Furthermore, valve tissue samples were collected from the rheumatic heart disease and aortic stenosis patients who were going through a valve replacement surgery at the Groote Schuur Hospital, Chris Barnard Division of Cardiothoracic surgery. Hematoxylin and Eosin staining was used to assess the histopathological features of the excised valve biopsies. Method development pilot experiments were performed to determine suitable parameters to use in the study. Metabolites were extracted from serum and valve tissues for untargeted metabolomics analysis using ultra-performance liquid chromatography with quadrupole time-of-flight mass spectrometry and *in situ* biomarker localisation with matrix assisted laser desorption ionisation mass spectrometry imaging. Univariate and multivariate statistical analyses were used to explore metabolic changes, associations, and investigate area under the receiver operating characteristic curves of the potential biomarkers. Bioinformatics tools were used to explore the gene-metabolite interactions of the potential serum biomarkers.

Results: The histopathological assessment explored the frequency of the following histopathological features among rheumatic heart disease and aortic stenosis patients. Aortic stenosis was predominantly associated with features of calcification 8(80.00%) as compared to rheumatic heart disease 10(25.64%), $p<0.01$. In addition, it was observed that neovascularisation was mostly associated with rheumatic heart disease 24 (61.54%) as compared to aortic stenosis 3(30.00%), $p=0.01$. Furthermore, presence of vegetations and Aschoff bodies was solely reported among valve fragments obtained from rheumatic heart disease patients. Fibrosis was a histopathological feature commonly found among patients with rheumatic heart disease 30(76.92%) and aortic stenosis 9(90.00%) undergoing valve replacement surgery. Admixed features of chronic and acute rheumatic heart disease were reported on 59-year-old woman having aortic and mitral valve replacement surgery. The method development pilot experiments established suitable extraction solvents, liquid chromatography flow rates, and the feature extraction methods. After the pilot experiments, the untargeted metabolomics study of serum samples showed seven metabolites differentially expressed in rheumatic heart disease which were independent of patients' baseline characteristics (covariates) and could differentiate rheumatic heart disease from healthy controls (AUC>0.7). Four metabolites could differentiate aortic stenosis from controls (AUC≥0.7). Of the perturbed metabolites in rheumatic heart disease and aortic stenosis, 7-HOCA and deoxycholate showed a moderate association with left ventricle ejection fraction. Furthermore, acylcarnitine and ketone bodies showed a correlation with left ventricular mass index in rheumatic heart disease and left atrial area in aortic stenosis patients. Elevated levels of cortisol were associated with the presence of valve calcification in rheumatic heart disease and aortic stenosis. Metabolomics of valve biopsies showed calcification, myxoid change, collagen deposition, and fibrosis had moderate associations with fatty acids, amino acids and derivatives, and bile acid metabolite classes in rheumatic heart disease and degenerative aortic stenosis patients. In addition, ventricular remodelling and function, aortic and mitral valve flow characteristics and measurements showed moderate associations with fatty acids, glycerophosphocholines, and amino sulphonic acid metabolites. Metabolites that correlated

with histopathological features were spatially localised on mitral and aortic valves using matrix assisted laser desorption/ionisation imaging mass spectrometry.

Conclusion: The study suggests cases of chronic rheumatic heart disease admixed with acute rheumatic fever are very rare, however they should be considered in regions with a high prevalence of rheumatic heart disease. In addition, the study highlights the importance of method development pilot experiments especially in metabolomics studies. The metabolites altered in rheumatic heart disease and degenerative aortic stenosis not only associate with cardiac remodelling but also are involved in major energetic pathways, amino acids metabolism, and inflammation regulation processes. Moreover, rheumatic heart disease patients with single or double valve replacement, and degenerative aortic stenosis patients appear to have distinct metabolic signatures on the excised valve biopsies. The study also reports metabolites that associated with histopathological and transthoracic parameters of valvular heart disease in rheumatic heart disease and degenerative aortic stenosis. In situ localisation of some of the metabolites is also reported.

List of Abbreviations

ACC/AHA	American College of Cardiology/American Heart Association
ACE	Angiotensin-converting-enzyme
ACN	Acetonitrile
AF	Atrial fibrillation
AKI	Acute kidney injury
ANCOVA	Analysis of covariance
ANOVA	Analysis of variance
AR	Aortic valve regurgitation
ARF	Acute rheumatic fever
AS	Aortic valve stenosis
AUC-ROC	Area under the receiver operating characteristic
AV mean PG	Mean aortic valve pressure gradient
AV Vmean	Mean aortic velocity
AVA	Aortic valve area
AVR	Aortic valve replacement
BAV	Bicuspid aortic valve
BCAA	Branched chain amino acids
BMI	Body mass index
BSL 2	Biosafety level 2
CAS	Calcific aortic stenosis
CCA	α-Cyano-4-hydroxycinnamic acid
CEM	Continuous electron multiplier
CI	Chemical ionization
CMR	Cardiovascular magnetic resonance
CRP	C-reactive proteins
CT	Computed tomograph
CV	Coefficient of variation
DALYs	Disability-adjusted life year
DC/AC	Direct current/alternating current
DC/RF	Direct current/radiofrequency
DDA	Data dependent acquisition
DEM	Discrete electron multipliers
DHB	2,5-Dihydroxybenzoic acid
EDV	End-diastolic volume
EF	Ejection fraction
EI	Electron impact
EMP	Endothelial microparticles
ESC/EACTS	European Society of Cardiology/European Association for Cardio-Thoracic Surgery
ESI	Electron spray ionization
FA	Formic acid
FBMN	Feature based molecular networking

FC	Fold change
FDR	False discovery rate
FFPE	Formalin-Fixed Paraffin-Embedded
FTICR	Fourier-transform ion cyclotron resonance
FTIR	Fourier Transform Infrared Spectroscopy
GAS	Group A streptococcus
GBD	Global burden of disease
GC-MS	Gas chromatography mass spectrometry
H&E	Haematoxylin and Eosin
HIV	Human immunodeficiency virus
HMDB	Human Metabolome Database
HPLC	High-performance liquid chromatography
ITO	Indium tin oxide
KEGG	Kyoto Encyclopaedia of Genes and Genomes
LA	Left atrium
LC-MS	Liquid chromatography mass spectrometry
LC-MS/MS	Liquid chromatography tandem mass spectrometry
LMIC	Low- and middle-income countries
LTQ	Linear-ion trap orbitrap
LV	Left ventricle
LVEDD	Left ventricular end-diastolic diameter
LVEDV	Left ventricular end-diastolic volume
LVEF	Left ventricular ejection fraction
LVESV	Left ventricular end-systolic volume
LVMI	Left ventricular mass index
LVPWd	Left ventricular posterior wall thickness in diastole
LysoPA	Lysophosphatidic acid
LysoPE	Lysophosphatidylethanolamine
m/z	Mass-charge ratio
MALDI	Matrix assisted laser desorption/ionization
MALDI-MSI	Matrix assisted laser desorption/ionization imaging mass spectrometry
MCCV	Monte Carlo cross validation
MG	Monoglyceride
MR	Mitral valve regurgitation
mRSD	Median relative standard deviation
MS	Mitral valve stenosis
MS/MS	Tanden mass spectrometry
MSEA	Metabolites set analysis
MV mean PG	Mean mitral valve pressure gradient
MVA	Mitral valve area
NBF	Neutral buffered formalin
NMR	Nuclear magnetic resonance
NO	Nitric oxide
OPLS-DA	Orthogonal partial least square discriminant analysis
PCA	Principal component analysis
PET	Positron emission tomography

PS	Phosphatidylserine
QC	Quality control
QTOF	Quadrupole time-of-flight
RF	Radiofrequency
RHD	Rheumatic heart disease
ROC	Receiver operating characteristic
ROI	Regions of interest
RT	Retention time
RT-qPCR	Reverse transcription-quantitative polymerase chain reaction
SA	Sinapinic Acid
SAM	Significance analysis microarrays
SAM	s-Adenosylmethionine
SBP	Systolic blood pressure
Sciex OS	Sciex operating system
SIM	Select-ion-mode
SNP	Single-nucleotide polymorphism
SNR	Signal to noise ratio
SSA	Sub-Saharan Africa
STICH	Search tool for interactions of chemicals
TAV	Tricuspid aortic valve
TAVR	Transcatheter aortic valve replacement
TB	Tuberculosis
TOF	Time-of-flight
TTE	Transthoracic echocardiogram
UHPLC	Ultra-High-Performance Liquid Chromatography
UPLC-QTOF-MS	Ultra-performance liquid chromatography with quadrupole time-of-flight mass spectrometry
VHD	Valvular heart disease
VIP scores	Variable importance in projection scores
WHF	World Heart Federation

List of figures and tables

Figures

Tables

List of publications and conferences

Publications

Mutithu DW, Aremu OO, Mokaila D, Bana T, Familusi M, Taylor L, Martin LJ, Heathfield LJ, Kirwan JA, Wiesner L, Adeola HA, Lumngwena EN, Manganyi R, Skatulla S, Naidoo R, Ntusi NAB. A study protocol to characterise pathophysiological and molecular markers of rheumatic heart disease and degenerative aortic stenosis using multiparametric cardiovascular imaging and multiomics techniques. PLoS One. 2024;19(5):e0303496. doi: 10.1371/journal.pone.0303496

Hahnle L, Mennen M, Gumedze F, **Mutithu D**, Adriaanse M, Egan D, Mazondwa S, Walters R, Appiah LT, Inofomoh F, Ogah O. Greater Disease Severity and Worse Clinical Outcomes in Patients Hospitalised with COVID-19 in Africa. Global Heart. 2024;19(1). doi: 10.5334/gh.1314

Mutithu DW, Kirwan JA, Adeola HA, Aremu OO, Lumngwena EN, Wiesner L, Skatulla S, Naidoo R, Ntusi NA. High-Throughput Metabolomics Applications in Pathogenesis and Diagnosis of Valvular Heart Disease. Reviews in Cardiovascular Medicine. 2023;24(6):169. doi: 10.31083/j.rcm2406169

Mutithu DW, Kirwan JA, Adeola HA, Wiesner L, Naidoo R, Ntusi NA. Metabolic biomarkers associated with cardiac remodeling in patients with severe rheumatic heart disease and degenerative aortic stenosis. Basic Cardiovascular Sciences Scientific Sessions of the American Heart Association. 2023. Posters P1177-2023.

Mutithu DW, Roberts R, Manganyi R, Ntusi NAB. Chronic rheumatic heart disease with recrudescence of acute rheumatic fever on histology: a case report. European Heart Journal-Case Reports. 2022 Jul 1;6(7):ytac278. doi: 10.1093/ehjcr/ytac278

Mutithu D, Wiesner L, Naidoo R, Ntusi N. P-550 Serum metabolomics provide distinctive biomarkers of metabolic processes in rheumaticheart disease and degenerative aortic stenosis. 18 Annual Conference of the Metabolomics Society. 2022. Poster P550.

Keeton R, Tincho MB, Ngomti A, Baguma R, Benede N, Suzuki A, Khan K, Cele S, Bernstein M, Karim F, Madzorera SV, Moyo-Gwete T, Mennen M, Skelem S, Adriaanse M, **Mutithu D**,...Riou C. T cell responses to SARS-CoV-2 spike cross-recognize Omicron. Nature. 2022 Mar;603(7901):488-492. doi: 10.1038/s41586-022-04460-3.

Aremu OO, Samuels P, Jermy S, Lumngwena EN, **Mutithu D**, Cupido BJ, Skatulla S, Ntusi NAB. Cardiovascular imaging modalities in the diagnosis and management of rheumatic heart disease. International Journal of Cardiology. 2021 Feb 15;325:176-85. doi: 10.1016/j.ijcard.2020.09.049

Chakafana G, **Mutithu D**, Hoevelmann J, Ntusi N, Sliwa K. Interplay of COVID-19 and cardiovascular diseases in Africa: an observational snapshot. Clinical Research in Cardiology. 2020 Dec;109:1460-8. doi: 10.1007/s00392-020-01720-y

Muhamed B, **Mutithu D**, Aremu O, Zühlke L, Sliwa K. Rheumatic fever and rheumatic heart disease: Facts and research progress in Africa. International Journal of Cardiology. 2019 Nov 15;295:48-55. doi: 10.1016/j.ijcard.2019.07.079

Conferences

Metabolomics South Africa Workshop on Metabolomics Data Processing and Analysis [North-West University, South Africa, 25 Jul 2023 – 27 Jul 2023]. Facilitator.

19th Annual International Conference of the Metabolomics Society [Niagara Falls, Canada, 18 Jun 2023 – 22 Jun 2023]. Oral presentation and session co-chair.

18th International Conference of the Metabolomics Society [Valencia Conference Center, Spain, 19 Jun 2022 – 23 Jun 2022]. Poster presentation.

Non-Communicable Disease (NCD) Research Symposium [Lord Charles Hotel, Somerset West, Cape Town, 3 Mar 2020 – 3 Mar 2020] Poster presentation on metabolomics biomarkers in RHD. Link: https://www.cebhc.co.za/research-key-outputs/ncd-research-symposium/

ACGT Proteomics and metabolomics symposium [University of Cape Town, Cape Town, 11 Dec 2019 – 14 Dec 2019] Oral presentation of an abstract based on valvular heart disease metabolomics biomarkers Workshop training: Skyline applications, SWATH data analysis Link: https://acgt.co.za/newsroom/2019/12/

45th Department of Medicine annual research days [Lecture Theatre II, New Groote Schuur Hospital. Cape Town, 1 Oct 2019 – 2 Oct 2019] Oral presentation for "Rheumatic fever and rheumatic heart disease: Facts and research progress in Africa" publication. Link: http://www.health.uct.ac.za/notice/call-abstracts-department medicine-annual-research-day-symposium.

1 INTRODUCTION AND STUDY RATIONALE[1]

1.1 Metabolic biomarkers in valvular heart disease

Valvular heart disease (VHD) is the damage to some or all of the heart valves; the most affected valves are mitral (MV), aortic (AV), and pulmonary valves (PV). Valvular heart disease in Africa are mainly composed of rheumatic, degenerative, and congenital VHD.[1] According to The Global Burden of Disease report, incidence rates of rheumatic heart disease (RHD) have been steady while the incidence rates of non-rheumatic VHD were shown to be on the rise between 1990 and 2019.[2] Central and sub-Saharan Africa is the most affected by incidences of RHD (29.40/100,000) as at 2017.[2] There is also an increase in non-rheumatic VHD incidences reported between 1990 and 2017.[2] The most prominent non-rheumatic valve disease in sub-Saharan Africa are dystrophic aortic and mitral valve disease.[2]

Repeated incidences of acute rheumatic fever (ARF) following pharyngeal infection with group A *Streptococcus* (GAS) are thought to cause RHD.[3] GAS M proteins are thought to induce autoimmune reactions against the host's cardiovascular myosin, actin, tropomyosin and laminini.[3,4] Genetic susceptibility has also been postulated to contribute towards pathogenesis of rheumatic valvulopathy.[5] Degenerative valve disease mostly present as aortic stenosis due to dystrophic calcification or mitral valve regurgitation due to myxomatous degeneration.[2] Degenerative aortic valve stenosis (AS) and myxomatous mitral valve degeneration is caused by inflammatory pathways, advanced age, smoking status, and hypertension.[2,6] In addition, NOTCH1, fibrillin-1 (FBN1), and filamin (AFLNA) genes mutations have also been associated with

[1]This chapter is based on a manuscript:
Mutithu DW, Aremu OO, Mokaila D, et al. A study protocol to characterise pathophysiological and molecular markers of rheumatic heart disease and degenerative aortic stenosis using multiparametric cardiovascular imaging and multiomics techniques. *PLoS ONE*. 2024;19(5):e0303496. doi:10.1371/journal.pone.0303496

the development of degenerative calcification and MV prolapse.[7,8] As per the WHF and 2020 ACC/AHA guidelines, the 2-D echocardiography with Doppler is the imaging gold standard for characterising RHD and degenerative valve disease patients.[9,10] However, cardiovascular magnetic resonance (CMR) is increasingly important for diagnosis in cases where echocardiographic image quality is poor or where there is discordance between clinical finding and echocardiography. CMR provides a superior complementation in the diagnosis and characterisation of VHD patients based on its excellent image quality, unlimited acquisition window, flow quantification, and tissue characterisation.[11] Despite the existing diagnostic techniques and known pathogenic mechanisms, it is challenging to diagnose asymptomatic patients and still much remains to be understood to adequately describe the pathogenesis of rheumatic and non-rheumatic VHD.[12] There is therefore a need to develop diagnostic and pathophysiological biomarkers for early diagnosis and therapeutic targets.

Current developments in biomarker research have contributed to the increased application of multi-omics techniques in health research.[13] Metabolomics is one of the multi-omics techniques that has been used to characterise metabolic profiles in the common VHD such as degenerative and congenital valvular heart disease.[14–17] Specifically, dysregulation of inflammatory processes, energy metabolism, amino acid, serotonin, and calcium metabolism have been associated with myxomatous mitral stenosis (MS) and mitral regurgitation (MR).[17,18] Additionally, formate and lactate have been reported to be potential diagnostic biomarkers of degenerative MS and MR with a high AUROC, sensitivity, and specificity.[18] Elevated levels of arachidonic acid has been shown to be a marker of worst outcome and reduced reverse remodeling post-valve replacement in patients with aortic stenosis.[15] In summary, available reports show that metabolomics techniques could be used to describe metabolic patterns to explain pathological mechanisms and diagnosis of valvular heart disease. However, only a few studies have reported the dysregulation of metabolites in rheumatic heart disease, and none have compared metabolic profiles in rheumatic

and degenerative aortic stenosis. The study intended to explore the metabolic biomarkers in rheumatic heart disease and degenerative aortic stenosis using metabolomics techniques.

1.2 Rationale for the study

Africa is sitting in a precarious position when it comes to the diagnosis and management of valvular heart disease. RHD is endemic in most regions of African continent. Additionally, Africa has an aging population that positively correlates with an upsurge of cardiovascular disease (CVD) risk factors such as diabetes, hypertension, and obesity that mostly leads to development of degenerative AS.[2] For effective diagnosis and management of RHD and degenerative AS, early diagnosis is vital and therapeutic targeting of the processes that lead to the worsening of the valvulopathy is key. Furthermore, the only proven management of VHD is valve repair or replacement surgery. In resource limited regions of the world including Africa, there is limited available diagnostic tools and access to valve replacement facilities is almost non-existent. Therefore, there is an urgent need to develop methods and techniques for early diagnosis to supplement the existing diagnostic methods.[19] Further, there is need for research on potential pathways and processes as therapeutic targets to halt progression of rheumatic and non-rheumatic VHD. Advances in multi-omics technology have shown the potential in biomarker discovery research especially on CVD. Specifically, advances in metabolomics techniques provide a suitable platform to use for biomarker discovery studies. Reported metabolomics studies on CVD have shown it to be a powerful tool for diagnosis, prognosis, and pathogenesis biomarker discovery. The role of metabolic pathways' dysregulation in the progression of disease in other diseases have been extensively researched and identified diagnostic biomarkers and therapeutic targets. Currently not much research has been done to understand the impact of metabolic dysregulation on the pathogenesis of RHD and degenerative AS.

This work was conceived in a near-perfect time and space for biomarker discovery studies in Africa. There was access to well curated biomaterials on emerging (degenerative AS) and

prevalent (RHD) and access to cutting edge technologies. The work was planned to explore circulatory biomarkers in patients with chronic RHD and degenerative AS and compared to healthy controls. The study also intended to investigate the relationship between circulatory and tissue specific biomarkers to inform the reliability of the circulatory biomarkers in describing valve metabolic changes. Further our research investigated trends of metabolic biomarkers between severe and mild VHD to establish the role of biomarkers in pathogenesis and progression of VHD.

2 LITERATURE REVIEW[2]

2.1 Introduction

2.1.1 Common types of valvular heart disease

Classification of valve disease depends on the cause of the lesions. The most common types of VHD are; degenerative valve disease, congenital valve disease, RHD and valve disease due to cardiac traumas (**Figure 2.1**). In this study the focus was to understand the pathogenesis of RHD and degenerative aortic valve disease. RHD is highly prevalent in Africa and is one of the leading causes of mortality due to CVD.[20–22] On the other hand, there has been a sharp rise in the prevalence of degenerative valve disease in Africa; probably attributable to the rising life expectancy.[23]

Degenerative valve disease

Despite not being common in LMICs, degenerative aortic valve disease contributes to more than 80% of VHD in high-income countries (HICs) and low- and middle-income countries (LMICs).[8,24] Degenerative AS mostly progress covertly and most patients become symptomatic when they have severe stenosis or regurgitation associated with shortness of breath, syncope or angina.[25] Calcific aortic valve disease is caused by the activation of myofibroblasts, osteoblast differentiation and is accelerated in congenital valve defects such as bicuspid aortic valve (BAV).[26] Further, cases of dystrophic aortic calcification have been linked to inflammation, advanced age, smoking and hypertension.[8,27] As the valve's endothelial layer loses integrity, lipid accumulation and

[2]This chapter is based on a manuscript:
Mutithu DW, Kirwan JA, Adeola HA, et al. High-throughput metabolomics applications in pathogenesis and diagnosis of valvular heart disease. *Rev. Cardiovasc. Med*. 2023;24(6):169. doi:10.31083/j.rcm2406169

migration of inflammatory cells contribute to progressive calcification, especially in calcific aortic stenosis which is shown to worsen over time (**Figure 2.2**).[27]

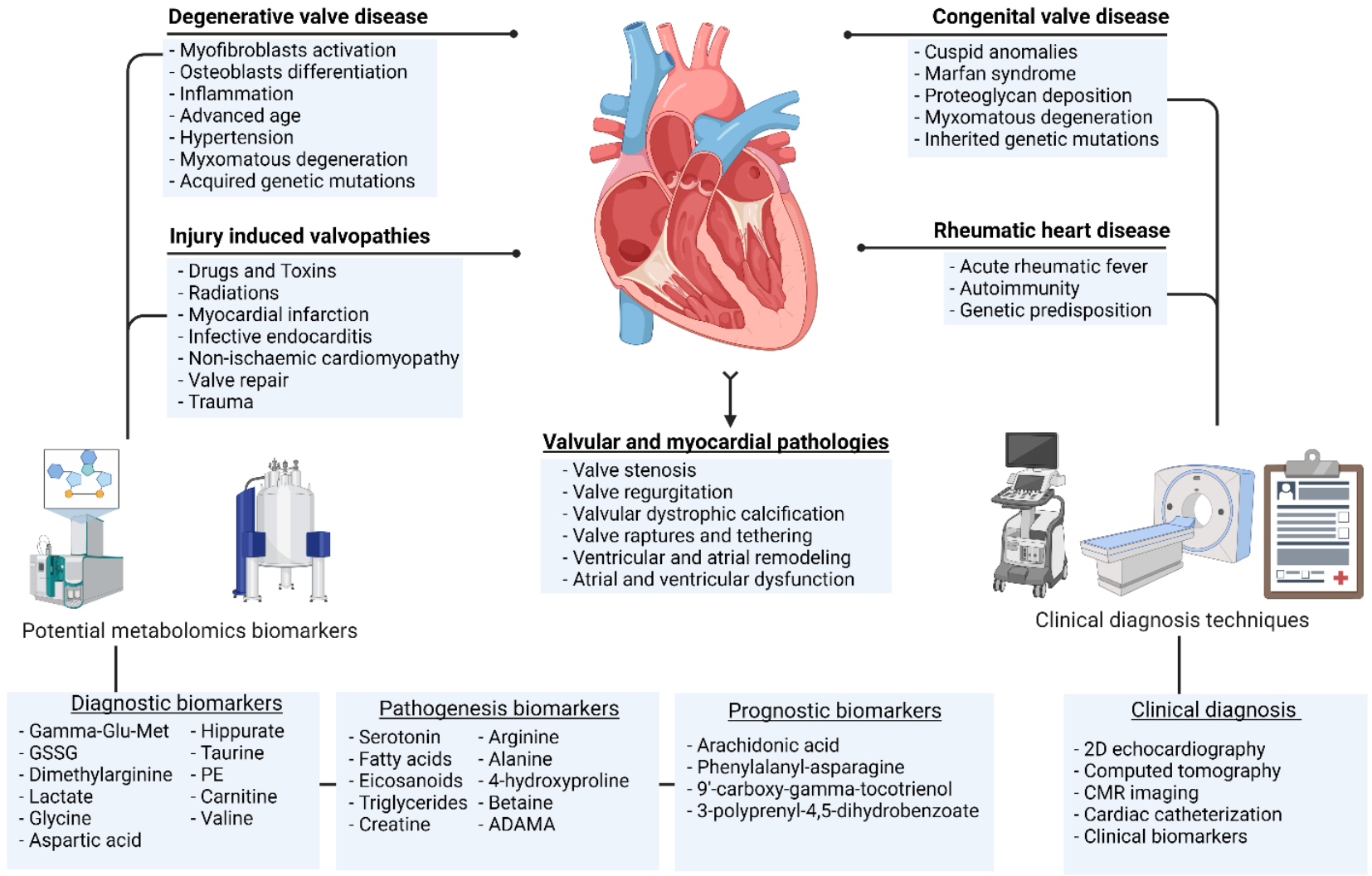


Figure 2.1. Common types of valvular heart disease showing their different aetiologies, their associated valvular and myocardial pathology, the diagnostics techniques, and the potential metabolomics biomarkers. Adapted from Mutithu DW, Kirwan JA, Adeola HA, et al. High-throughput metabolomics applications in pathogenesis and diagnosis of valvular heart disease. *Rev. Cardiovasc. Med*. 2023;24(6):169. doi:10.31083/j.rcm2406169.

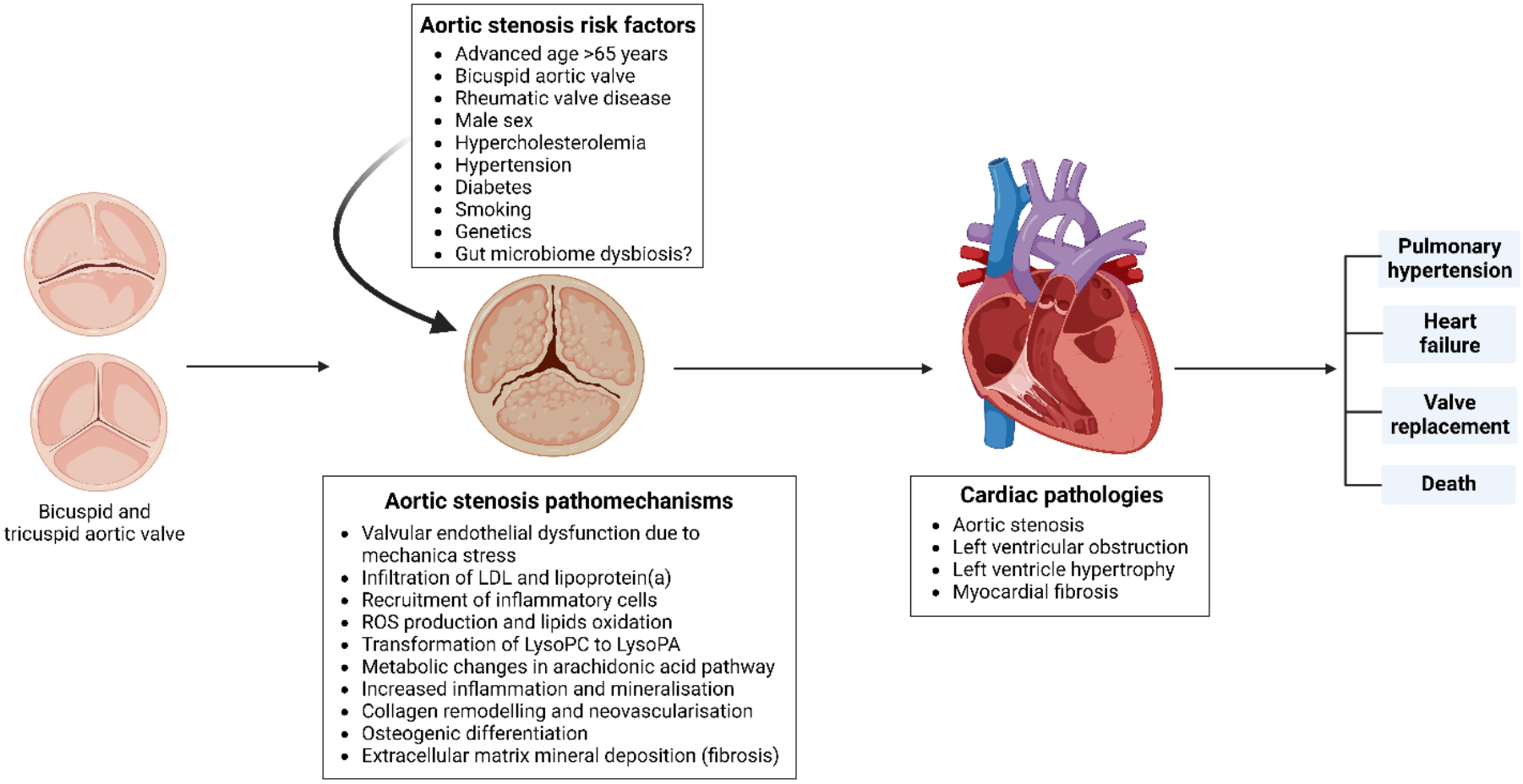


Figure 2.2. Pathogenesis of degenerative aortic stenosis. Created with Biorender.com.

Mechanical stretching of the valve apparatus leads to the deposition of glycosaminoglycans and proteoglycans by the ventricular interstitial cells which is accentuated by the genetic predispositions and environmental factors. The high sheer forces on the valves could also be a contributor to degenerative aortic valve disease since they are predominantly reported on the left side of the heart. Genetics studies have identified single nucleotide polymorphisms (SNPs) at the lipoprotein(a) locus in patients diagnosed with calcific aortic valve disease.[28] In addition, genetic perturbations in the NOTCH1, Fibrillin-1 (FBN1), Filamin (AFLNA) gene have also been associated with development of degenerative calcification.[8,29]

Degenerative aortic valve disease is mostly diagnosed when it has progressed to stenosis. Using echocardiography, it is easy to assess valvular flow dynamics and valve morphology.[25] Imaging techniques such as computed tomography (CT) or positron emission tomography (PET) have shown high resolution in non-invasive assessment of calcification at early stages.[25,30] However, CT and PET are specialised tools that require special training and are not readily available in the LMICs. Thus, there is need for a simple and cheap tools for early diagnosis of valve disease.[25] Histologically, degenerative aortic valve disease is characterised by collagen degradation, mucopolysaccharide accumulation, elastic disruption, fibrosis and myxoid change.

Rheumatic valve disease

RHD is a chronic valvular condition – a sequel of acute rheumatic fever (ARF).[31] Recent systematic reviews and meta-analysis studies have reported RHD prevalence to be 1.7% to 14.67% –most of the studies are based in low- and middle-income countries (LMICs).[32–34] In sub-Saharan Africa (SSA), the estimated prevalence of RHD is 4.6/1000 and 7.9/1000. Distressingly, it has been reported that cases of asymptomatic RHD are approximately 5 times more than symptomatic RHD (21.7/1000).[22] Globally, RHD accounts for 1,426 disability-adjusted life-years (DALYs) per 1000 individuals.[20,22] Africa is among regions with high rates of disability-adjusted life-years which negatively affects the economic capacity of affected communities.[35] The peak prevalence of RHD

is between 25-45 years which is as a result of recurrent ARF and delays in RHD detection.[36] The prevalence of RHD is twice as much among females than males and most pronounced between the ages of 20-40 years[37] (**Figure 2.3**).

Rheumatic fever occurs after infection with group A *Streptococcus* (GAS) due to autoimmune response (**Figure 2.3**). Repeated autoimmune flares of the valvular epithelial surface proteins lead to valvulitis and collagen deposition.[4] Pharyngeal infection and inflammation are followed by a latent phase of approximately two to three weeks before initial symptoms of ARF can become apparent.[4] Group A *Streptococcus* M protein is the virulence factor which aids in host infection. The M protein is shown to share similar immunogenic epitopes with cardiovascular proteins such as myosin, actin, tropomyosin and laminin.[31,38] Genetic susceptibility is an important factor in RHD pathogenesis. A significant association between RHD and human leucocyte antigens (HLA), TNF-α, IL-4, IL-10, IL-6, IL-1RaVNTR, ecNOS4, TGF-α1, ACE and MBL2 have been predominantly observed in RHD cases.[5,39]

RHD is predominantly a “left-sided” valvular disease most commonly presenting with mitral valve involvement and to a lesser extent, with aortic valve disease.[21,36] Tricuspid and the pulmonary valve is frequent. Mitral regurgitation and mitral stenosis are the commonest lesions often diagnosed from the age of 10 years, and may progress to mixed mitral valve disease by the age of 20 years.[36] Prevalence of aortic valve involvement in RHD proportionally increases with age.[40]

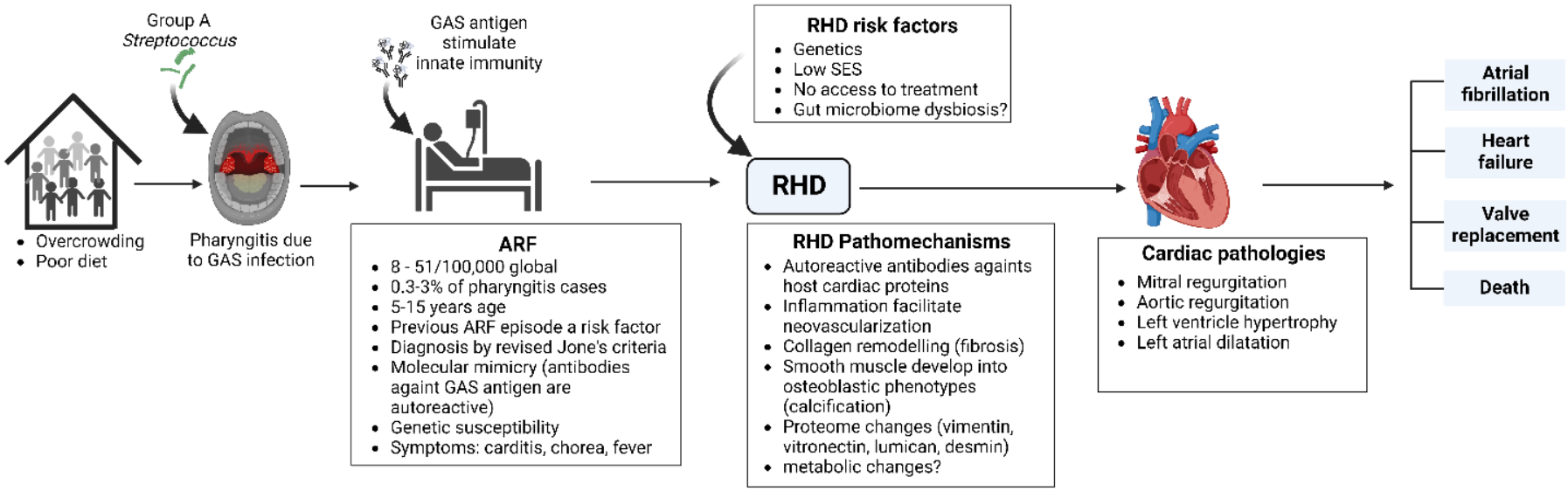


Figure 2.3. Pathogenesis of Rheumatic heart disease. Tonsillitis due to GAS infection, ARF and induction of immune reactions leading to Valvulitis and carditis. Chronic RHD characterised with Valve stenosis, regurgitation, left heart cardiomyopathies and heart failure, cardiac surgery, or mortalities. Created with Biorender.com.

ARF has limited visible damage to the valves but at times may present with thrombi on the leaflets, oedema, and some form of chronic inflammation.[41] Macroscopically, RHD presents with commissural fusion, calcification and thickened cusps.[41–43] Mitral valve involvement in RHD presents with shortened and fused chordae tendineae and scarred commissures leading to "catfish mouth" features. Rheumatic aortic valves present with fibrosis and calcification of the free edges of the cusps. Histologically, RHD is characterised by neovascularisation, focal or diffused calcification, infiltration of eosinophils, lymphocytes, or plasma cells, presence of Aschoff bodies or markers of chronic inflammation.[44–46] The Aschoff bodies are nodules composed of Anitschkow cells with a wavy nucleated cell termed as "caterpillar cells". Aschoff bodies result from inflammatory lesions that healed through fibrosis, deposition of macrophages and collagen necrosis.[31,47] Aschoff nodules are the classic markers of ARF and are mostly not observed in chronic RHD. It is mostly believed that there is limited inflammation during chronic RHD however presence of Aschoff's nodules has been reported in RHD cases suggesting recrudesce of active ARF and myocarditis.[48]

2.1.2 Diagnosis and management of valvular heart disease

Generally, VHD is diagnosed with echocardiography, through assessment of valve anatomy and the haemodynamic parameters. In LMICs where access to healthcare is limited, most valvular disease is diagnosed when the patient is symptomatic.[49,50] Transthoracic echocardiography is sufficient to diagnose VHD however, transoesophageal echocardiography, computed tomography and cardiovascular magnetic resonance (CMR) may provide additional information and may be helpful when transthoracic echocardiography is insufficient or not diagnostic.[51,52] In addition, cardiac catheterisation can also be used for a detailed valvular lesion assessment.[52]

Symptomatic management of VHD involves medical therapy, but definitive therapy is through valve repair, surgical valvular replacement, or via transcatheter aortic valve implantation. The latter options are expensive and not readily available to many patients in LMICs. Therefore, early

diagnosis and prevention of RHD is very important. Penicillin prophylaxis has been demonstrated to retard progression of RHD.[53,54]

2.1.3 Metabolomics techniques

Metabolomics is a systematic measure of intermediate small molecules and metabolism products in an organism, cell, or tissue. Metabolomics accounts for the effects of genetics, epigenomes, transcriptome, proteome, nutrition, gut microbiota, physical activities, and environmental exposures as factors affecting the organism's metabolism[55] Metabolomics holds greater potential in understanding tissue pathomechanisms.[56] Metabolomics studies measure complex sample matrices hence an array of tools and techniques have developed to study metabolic dysregulations.[57] Some of the common techniques and tools are: nuclear magnetic resonance (NMR), mass spectrometry (MS) and Fourier-Transform infrared (FTIR) spectroscopy.[58–60]

Since the study is explorative and uses complex bio-samples, MS-based metabolomics techniques were used that have higher sensitivity as compared to NMR systems.[61,62] MS identifies molecules based on variations of their molecular masses. The ions are separated based on their ion abundance and mass-to-charge ratios (*m/z*) after ionization, acceleration through vacuum and detection by analyser.[56] Generally, mass spectrometry set up is made up of a means of sample pre-separation, sample introduction, ion source, mass analyser and a detector.

Metabolites' pre-separation techniques

Samples are introduced into MS using pre-separation methods such as gas chromatography (GC)-MS, liquid chromatography (LC)-MS or direct injection techniques.[63,64] GC-MS is often used on volatile analytes that can be vaporised without decomposition. In GC the mobile phase is a carrier gas and the stationary phase is a microscopic polymer or liquid on an inert support held within a column **(Figure 2.4**).[56] LC-MS is the most common pre-separation method employed in metabolomics studies. LC setup has a liquid mobile phase and an inert solid stationary phase;

compounds are separated depending on their interaction with the two phases (**Figure 2.5**). Pre-separation techniques where compounds are separated through interaction of polar mobile phase and the non-polar stationary phase is referred to as reverse-phase chromatography.[63] Molecule separation in reverse-phase chromatography is based on hydrophobic interaction between the analytes and the stationary phase. The hydrophobic interactions are altered by gradually changing the polarity of the mobile phase from polar to organic.[63] The chromatographic gradient alters the interaction of the analytes and the stationary phase that helps separate molecules based on the type of the mobile phase and the chemical-physical characteristics of the compounds.[63,64] For example, isobaric compounds that have the same nominal mass but different molecular formula, may flow differently through the column particles for separation because they have different polarities based on the location of their functional groups.[65] Likewise, some (but not all) isomeric compounds; they have the same nominal formula i.e. the same number of atoms but are arranged differently thus giving the molecules different chemical properties and are separated based on their different interactions with the phases. Effective separation of isobaric and isomeric compounds is determined by the column's resolution. The resolution is affected by type of column particles, solvent type and flow rate, pH of the mobile phase, and column temperatures.[65] Resolution of a column can be improved by optimizing the gradient parameters of the LC setup.[65] After separation in the LC system the analytes are injected into the ion source and then into the mass analyser of the LCMS equipment.

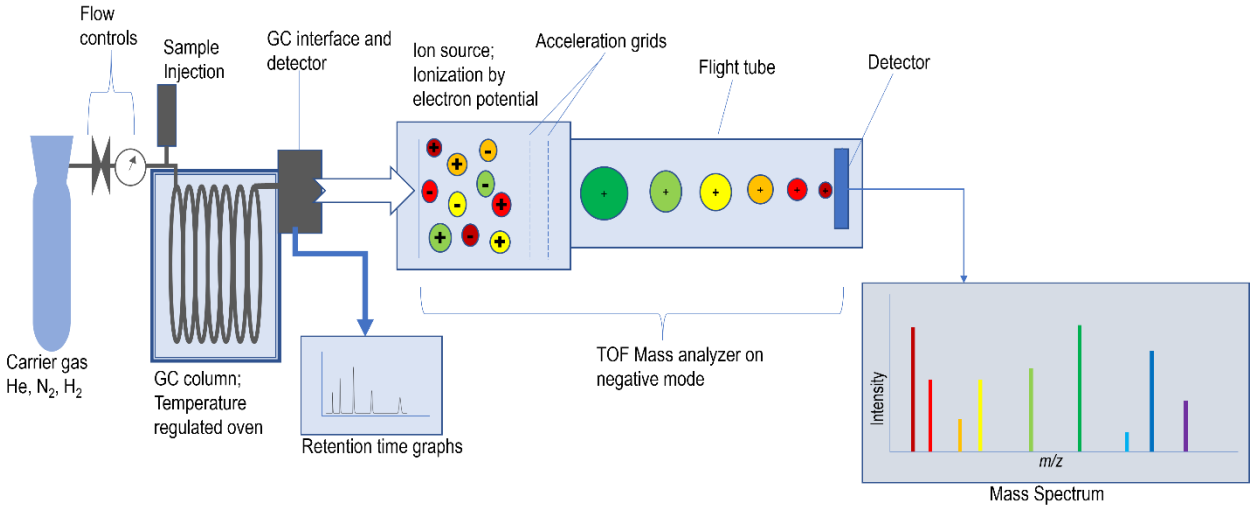


Figure 2.4. Gas chromatography coupled to a time-of-flight mass spectrometer (GCMS)

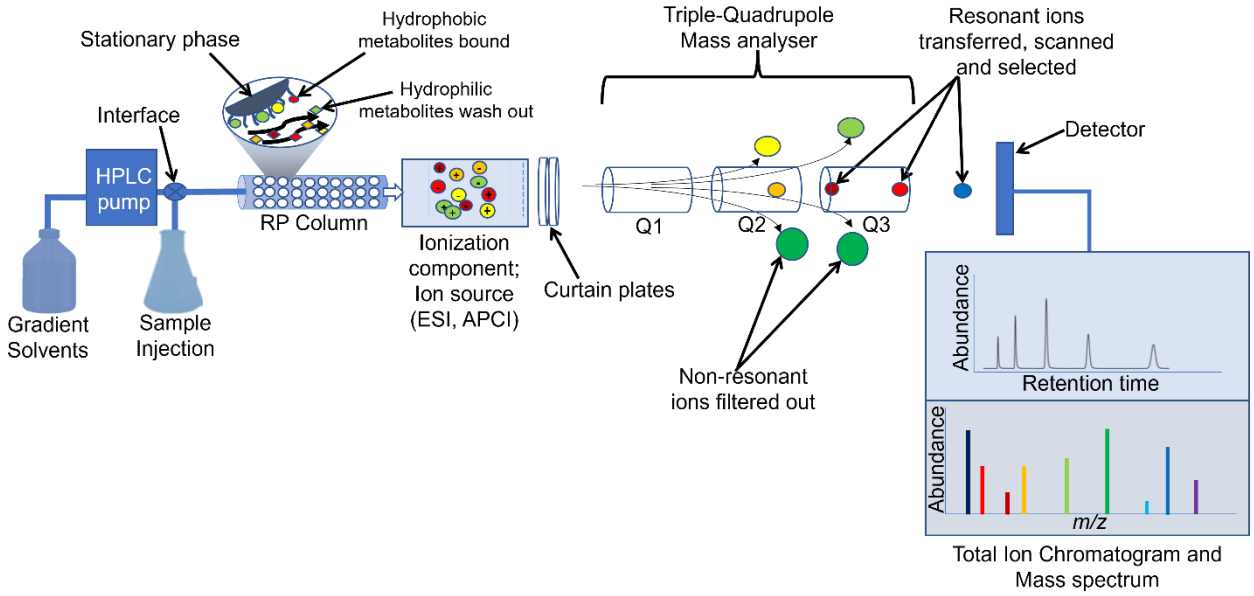


Figure 2.5. Liquid chromatography coupled to a quadrupole mass spectrometry analyser (LC-MS)

Ionization techniques

Ionisation ranges between soft and hard ionisation.[66] Hard ionisation uses very high energy levels to fragment the sample and yields many low molecular weight fragments. Soft ionisation uses low fragmentation energy and yields fewer molecular fragments. Electrospray ionisation (ESI) and

matrix assisted laser desorption ionisation (MALDI) are soft ionisation methods while electron impact (EI) is hard ionisation, and chemical ionisation (CI) is considered moderate ionisation. For ESI, the analyte is forced through a charged capillary with the help of a nebulizing gas to produce charged droplets.[67,68] Evaporation of the solvent causes the droplet size to decrease, and the ions are ejected due to charge repulsion. To remove the solution around the droplets, gas and heat are introduced into the ESI setup. Normally, small molecules emerge singly charged while larger molecules emerge with multiple charges. Electrospray ionisation is mostly coupled with LCMS techniques for analysis of peptides, proteins and non-volatile metabolites up to 6kDa.[64,66,69] ESI is affected by chemical structure of the compound, pH, and composition of the solvent. Despite ESI being the most preferred mode of ionisation due to its non-disruptive nature, its performance is impeded by ion suppression/matrix effect and is highly sensitive to the type of solvent used.[66] Ion suppression is complicated and challenging to deal with and there is currently no universal approach. However, several strategies have been incorporated into LCMS protocols aimed at keeping ion suppression to the minimum. Some of the common approaches are diluting the samples during reconstitution, protein precipitation, polar and non-polar phase extraction, switching ionisation modes, reducing ESI flow rate, optimization of column conditions, and use of column modifiers.[70] Matrix-assisted laser desorption/ionisation method is also regarded as soft ionization. Ionisation in MALDI is achieved by shining a laser beam onto a solid mixture of matrix and analytes. The matrix is meant to absorb excess ionisation energy from the laser beam to induce ionisation with minimal fragmentation.[71]

Mass analysis and detection

The mass analyser is an integral component of the mass spectrometer where charged ions are separated based on their mass differences. Acceleration of charged ions through the analyser is determined by the mass-to-charge ratio (*m/z*), voltage potential along the analyser, magnetic field and the analyser's mean free path. When the ions are exposed to an electrostatic field, the force

applied on them is explained by Coulomb's law and it explains how the charged ions will be separated based on their mass/charge (**Equation 2.1**).[72] Some of the commonly used analysers are time-of-flight (TOF), quadrupole and orbitrap.[72]

$$F = \frac{1}{4\pi\varepsilon_0} \frac{q_1 q_2}{r^2} \quad \textbf{(Equation 2.1)}$$

Where F is the electric force, $1/4\pi\varepsilon_0$ is the Coulomb constant (k), q_1 and q_2 are the ions charges, and r is the distance of separation.

Quadrupole mass analyser is also referred to as transmission mass analyser; it is made up of 4 parallel rods placed in a radial array. The opposite rods are placed at opposing direct currents (DC) and an alternating-current (AC) is superimposed at specified radiofrequencies (RF).[72] To select ions of specific *m/z*, the rods are put at specified DC/RF levels such that only the selected ion species will have a stable trajectory between the rods to reach the detector (**Figure 2.5**). Sequentially changing the DC/RF ratios allows ions of different *m/z* to be separated in space and time before reaching the detector.[68,73] Depending on the type of experiments, quadrupoles can be operated in full spectrum scanning mode by placing the DC and RF at fixed ratios; where all the ions are detected. To operate on selected-ion-mode (SIM) the DC/RF setups are altered to remove ions with unstable trajectory and only allows ions with specific m/z to be detected.[68] The quadrupoles can also work as ion-transporter when operated on RF-only mode; in this mode the ions are not preselected.[72] Advances have seen the introduction of triple quadrupole MS instruments. Triple quadrupole technology is built on the single quadrupole in that there are 3 sets of quadrupoles: Q1, Q2 (collision cell) and Q3. Depending on the type of experiment Q1 and Q3 can be operated as mass filters set on scanning or selected-ion modes. In the case of tandem mass spectrometry (MS/MS), Q1 is operated on transfer mode (data independent acquisition) or SIM (data dependent acquisition), Q2 as a collision/fragmentation cell and the daughter ions monitored in Q3 by scanning or selected-ion modes.[72,74] Quadrupoles are high-resolution

instruments with mass resolutions going down to 0.1 Da and are best suited for targeted metabolomics applications.[72]

Orbitrap analysers are based on Fourier transformation of the image currents of the ions and are regarded as a type of ion trap mass analyser.[75] Altering the electric potential and the electric fields injects and traps ions within the electrodes due to inertia that maintains their elliptical trajectories and their motions depending on their *m/z*.[68] Ions of the same *m/z* will have a harmonic axial motion around the inner core of the electrode. In the detector the ions oscillate around the spindle, and they form axial oscillations. The ions' axial oscillation frequencies are detected as image current that can further be converted to the corresponding mass-to-charge ratio. Currently, orbitrap instruments are coupled to either a direct ions source, linear-ion-trap systems (LTQ orbitraps), or quadrupoles. Orbitrap instruments are most suited for untargeted metabolomics experiments because of its high-resolution capabilities. They are reported to have an ultra-high-resolution capabilities of over 1,000,000 for m/z 300-350 at 3 seconds thus achieving sub-1ppm mass accuracy.[75,76]

Another common mass analyser is the time-of-flight (TOF) which is compatible with GC, LC, or direct injection systems. A typical TOF instrument has a grid at potential (V) for accelerating the ions, a field-free vacuum flight tube and the detector. The acceleration of the ions gives them kinetic energy which is proportional to the charge of the ions. In theory, the separation of the ions is determined by the product of their kinetic energy and the applied electrostatic energy as shown in **Equation 2.2.** To ensure ions of the same mass arrive together and the ions of different masses don't arrive simultaneously on the detector and improve the quality of spectra, ion sources are pulsed.[58,72,77,78] Time-of-flight instruments are coupled to a continuous ion source such as ESI or MALDI (**Figure 2.5**). The initial spatial spread of the ions at the initial time (t=0) determines the resolution of the TOF instrument.[79] The flight tube resembles a tube that is kept at high vacuum and at opposite electric potentials and allows ions to migrate towards the ion detector. Ion species

are differentiated based on their differences in their flight time. To correct the spatial spread, TOF instruments are installed with a reflector (ion mirror) which uses an electromagnetic gradient to turn around the ion's trajectory and lengthen their flight path. As mentioned earlier, **Equation 2.3** assumes that all the ions of the same mass receive equal kinetic energy, however in reality some ions of the same masses might receive different kinetic energies as shown in **Figure 2.6**. These ions with similar masses but different kinetic energy will have different velocities in the flight tube; the ions with less kinetic energy will fly slower. Therefore, to correct the difference, slower ions take a shorter path to turn around in the ion mirror as opposed to the faster ions. The differences in ions' turn-round paths in the reflector ensures that the ions of similar masses arrive simultaneous at the detector despite differences in their initial kinetic energies (**Figure 2.6**).[80]

$$Ve = \frac{1}{2} mv^2 \qquad \textbf{(Equation 2.4)}$$

Where e is the charge of the electrons, v is the applied voltage, and m is the mass of the ions.

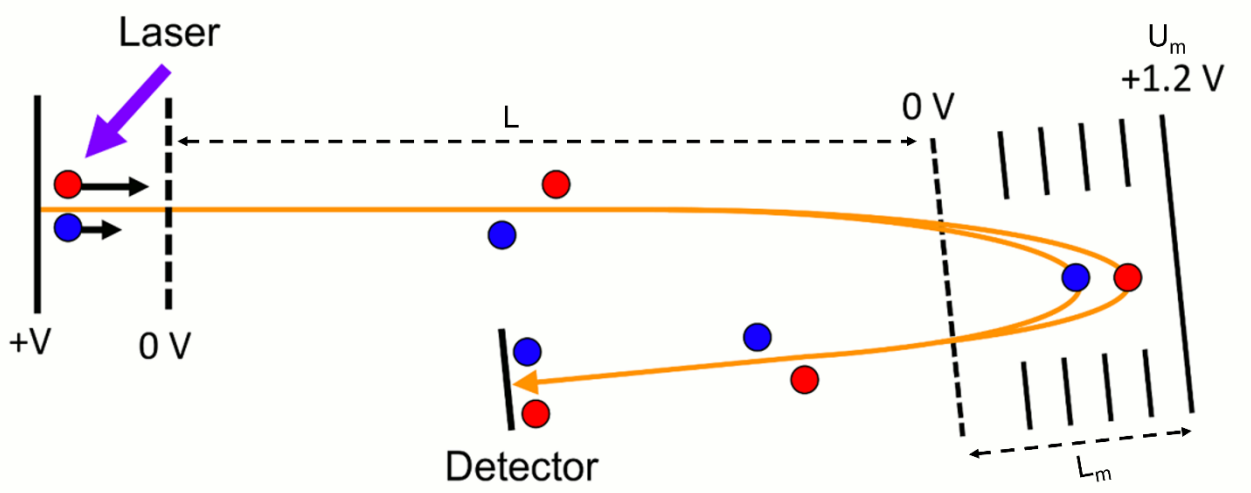


Figure 2.6. Schematic illustration of the single-stage reflectron of TOF MS. (L, field free path length; L_m, length of ion mirror; U_m, voltage applied across the ion mirror). Adapted from Hosseini S, Martinez-Chapa SO. Principles and mechanism of MALDI-ToF-MS analysis. *Fundamentals of MALDI-ToF-MS Analysis: Applications in Bio-diagnosis, Tissue Engineering and Drug Delivery.* Springer, Singapore. 2017:1-9.

Quadrupoles have been coupled to TOF technologies (QTOF). A QTOF instrument has Q1 and Q2 quadrupoles; Q1 is operated on scanning or SIM mode while Q2 is operated as an ion-

transport for MS1 experiments and a collision cell for MS/MS experiments.[78] Currently Q-TOFs are reported to have resolutions of ≤ 70,000 at m/z 200 with a mass accuracy <3ppm.[79,80] The unlimited mass range makes QTOF best suited for untargeted metabolomics. Further, multiple pulses accumulation generates enhanced spectra with low signal to noise ratio.[72,74,78]

The detector measures the ions' current as it is projected from the analyser. There are different types of detectors but most detectors work on the principle of amplifying the signal from the analyser.[72,81] Commonly used detectors are discrete electron multipliers (DEM) or continuous electron multiplier (CEM) which are based on the arrangement of the multiplication surfaces being fixed (DEM) or not fixed (CEM). Further there is the dynode type detector which works by converting the received ions into opposite charge electrons. The amplification capability of the detectors' surface gets depleted; thus, requiring frequent replacement (approximately every 2 years).[72,81]

Matrix assisted laser desorption/ionization

The process of MALDI ionisation starts with ablation and desorption of the matrix-sample mixture by the laser beams' irradiation. The matrix used in the MALDI is used to facilitate the soft ionisation by absorbing the excess the laser energy; it also charges the analyte by transferring the protons. Then, samples are either protonated or deprotonated by the ablated hot gases; the ionised molecules are then accelerated through vacuum and recorded by the ion detector.[77,82,83] MALDI is typically suited to analyse peptides, proteins, and nucleotides of up to 500kDa and can be operated on both the negative and positive modes. MALDI ionisation efficiency depends on the type of matrix and molecules being ionised. A suitable matrix should have a strong laser wavelengths absorbance, not react with analyte, be soluble in solvents, and stable in vacuum.[71] For proteins ionisation, α-cyano-4-hydroxycinnamic acid (CCA) and sinapinic acid (SA) are the most used matrices.[71,77] To ionise peptides, glycopeptides, glycolipids and metabolites below 10kDa; 2,5-dihydroxybenzoic acid (DHB) has been reported as the most suitable matrix.[71,77] The

main advantage of MALDI is that it does not require specialized pre-separation methods and can be used directly on thin tissue sections.[77,83–85] MALDI MS imaging (MALDI MSI) has been used to establish spatial localisation of biomolecules on tissue sections. The techniques' workflow includes tissue acquisition, cryosectioning and mounting on indium tin oxide (ITO) glass slides, matrix spraying, and laser scanning.[82,83,86,87] High resolution images of the mounted sections are obtained and are co-registered with the tissue sections loaded into the MALDI equipment.[84] Rasterised laser scanning is conducted around the tissue following the registered marks (teach marks). If there are zones on the sections with different pathology, they would be registered as regions of interest (ROIs).[84,87] Systematic directing of laser beams around the tissue collects spectra with *m/z* data that is then reconstructed to form the image. Using bioinformatics software and statistical methods, differential changes of *m/z* features can be established between ROIs or different tissue sections,[86,87] (**Figure 2.7**). MALDI can be combined with haematoxylin and eosin (H&E) to establish high resolution images for localizing the *m/z* features of interest reconstruction (**Figure 2.7**).[84] MALDI techniques are however limited since they do not have retention time dimension; therefore, losing resolution to distinguish isobaric *m/z* features.[78] In addition, poor reproducibility of signal intensities and ion suppression are limitations associated with MALDI MS methods.[74,78] The other important component of a MS setup is the mass analyser and detector.

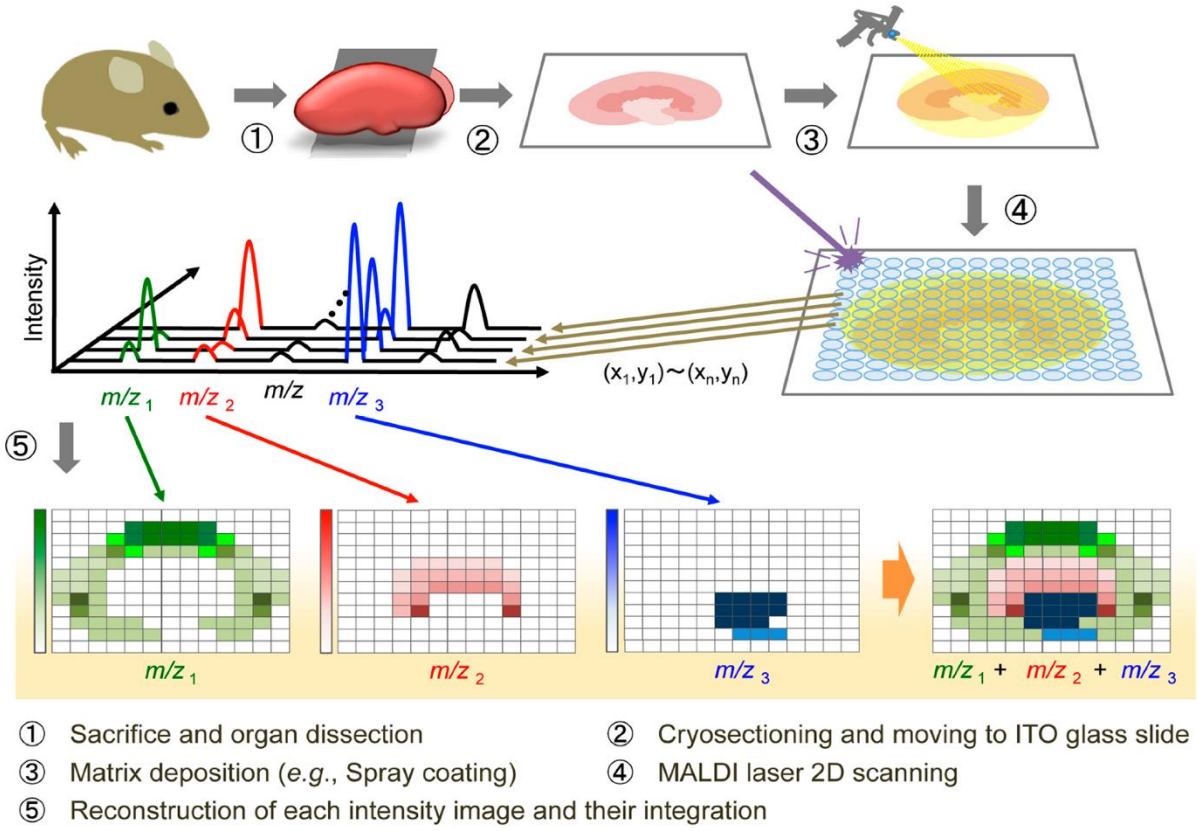


Figure 2.7. MALDI MSI Workflow. (1) Isolation of tissues. (2) Thin Sectioning of tissues and mounting on ITO slides, (3) Spraying tissues with matrix (4) laser scanning in the MALDI instrument and capturing m/z spectra at each spot (5) reconstructing the m/z image to the spatial location of where it was detected to show ROIs. Adapted from Fujimura Y, Miura D. MALDI mass spectrometry imaging for visualizing *In Situ* metabolism of endogenous metabolites and dietary phytochemicals. *Metabolites*. 2014; 4(2):319-346.

Data acquisition, pre-processing, analysis, and biomarker discovery

There are established data handling pipelines specific for targeted and untargeted MS experiments. To make LCMS, GC-MS, or MALDI-MS raw data compatible with bioinformatics tools, several pre-processing steps are required including noise removal by smoothing, missing values (zeros) filtration, and baseline correction and they affect the number and quality of the extracted features. Then retention time drifts are adjusted using spectral alignment based on quality control samples or internal standards. To ensure that intensities are comparable between samples, sample-based intensity normalization is performed to remove the systematic errors or biases and experimental variation. Finally, a matrix data type with retention time, *m/z* values, and intensities is generated using peak detection and peak picking algorithms[88] (**Figure 2.8**).

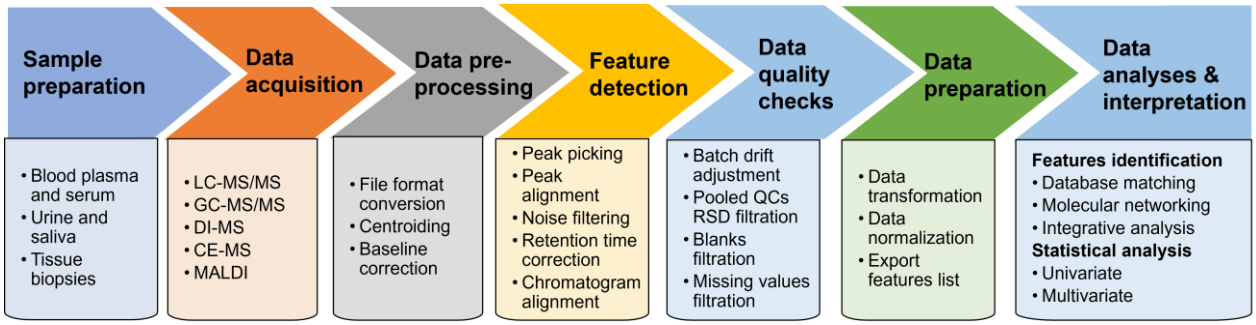


Figure 2.8. Mass spectrometry data analysis strategy for metabolomics experiments

During statistical analysis, the data matrices for each test group are compared to determine differences. This is after ensuring the differences observed are not due to batch variations, sample handling, inefficient peak picking, or equipment systematic drifts. Due to the complexity and multidimensionality of untargeted metabolomics data, dimensional reduction methods are employed to extract potential biomarkers.[89,90] Dimensionality reduction methods reduce the number of features (dimensions) while preserving the variance of the original data as much as possible to make it easy for visualisation, improve algorithms and models performance. Common dimensionality reduction methods in metabolomics are principal component analysis (PCA) (unsupervised) or orthogonal partial least square – discriminant analysis (OPLS-DA) (supervised).[91] The main difference between supervised and unsupervised dimensionality reduction methods is that the supervised method is trained on the class labels while the unsupervised method reduces the number of variables without considering the class information. Further, to extract features that are different between groups, fold change (FC) analysis based on volcano plots are also applied.[74,89,90]

Processing of MALDI MSI data is performed on several ROIs or between tissue sections by noise reduction, dimensionality reduction, and correlation analysis.[87] Comparing the significant features against metabolomics spectral libraries helps with the identification/annotation of significant *m/z* features. The identification process considers fragmentation (MSMS) patterns of the precursor molecule; isotope patterns and adducts associated with the suggested feature.[57,74,92] Development

of computerised pipelines such as open access pipelines (XCMS Online, MALDIquant, MetaboAnalyst, and MS-DIAL) or commercial pipelines (MassHunter for LCMS or SCiLS Lab for MALDI MSI) have made metabolomics data handling less laborious[86,88,93–96]

2.1.4 Application of metabolomics biomarkers in valvular heart disease

Several metabolomics studies have investigated metabolic biomarkers in VHD. Glycerophospho-N-oleoyl ethanolamine, monoglyceride and phosphatidylethanolamines have been reported as suitable risk predictors of BAV.[97] In addition, dysregulated levels of alpha-tocopherol and choline have been reported to distinguish asymptomatic bicuspid and tricuspid aortic valve individuals.[98] Furthermore, including alpha-tocopherol alongside biomarkers from other modalities including endothelial microparticles and C-reactive protein in an aortic valve morphology prediction models showed the best performance in discriminating symptomatic and asymptomatic BAV from TAV patients.[98] In summary, dysregulation of fatty acids biosynthesis pathways and metabolites with antioxidant and anti-atherogenic properties have been reported as the discriminant features between congenital valve disease and healthy participants.[97,98] Congenital and degenerative AS have been compared and dysregulation of arginine and proline metabolism pathways in a patient cohort with BAV and TAV prior and up to 7 days after transcatheter aortic valve replacement has been reported.[99] One-year post-transcatheter aortic valve replacement of the BAV patients, upregulation of arachidonic acid metabolism pathway was associated with poor recovery based on the changes in left ventricular mass index.[99]

Metabolomic biomarkers have been reported in degenerative valve disease such as calcific aortic stenosis and degenerative aortic valves. Upregulation of metabolites involved in inflammation (pyroglutamic acid and succinic acid) and ischemia (alanine) are reported as best predictors of calcific aortic stenosis using untargeted GC-MS metabolomics.[100] Further, dysregulation of plasma amino acids groups related to urea cycle, branched amino acids and indicators of (NO)-associated endothelial dysfunction have been reported in calcific aortic stenosis.[101] Long-chain acylcarnitines

have been reported as increased in patients with aortic stenosis before undergoing TAVR but were later decreased post-intervention.[102,103] Furthermore, there were decreased levels of long chain phosphatidylcholines and increased levels of amino acids and biogenic amines in the aortic stenosis patients as compared to healthy controls, but the levels were reversed 6 weeks post-TAVR.[102]

To the best of my knowledge, there are few reported metabolomics studies investigating the pathogenesis, diagnosis or prognosis of RHD. RHD has a different aetiology from degenerative and congenital valve.[4,38,49,104] Therefore, based on the reported performance of MS metabolomics techniques performance in the diagnosis of other valvular disease; metabolomics would help in understanding RHD.

Metabolomics to study pathogenesis of valvular heart disease

Few studies have studied metabolomics in RHD. Purine, glutamine, glutamate, pyrimidine, arginine, proline and linoleic metabolic pathways have been reported to be changed in RHD patients.[105] Like other severe VHD, the amino acid metabolism and energetic pathways were affected in RHD patients. Vascular endothelial cells have previously been reported to be activated by linoleic acid metabolism which might suggest it promotes inflammation in RHD patients.[105] It is known that RHD mostly affects mitral valves leading to MS or MR.[21,36] Some studies have investigated metabolic profiles in mitral valve disease such as myxomatous mitral valve stenosis and regurgitation.[17,18] Calcium metabolism, energy metabolism and inflammation are some of the biological processes that have been associated with mitral valve disease. Changes in amino acids associated with serotonin biosynthesis have also been reported in mitral valve disease which may suggest involvement of the serotonin autocrine signalling processes.[17,18] Further, the dysregulation of serotonin levels and fatty acids may explain the increased rates of depression among VHD patients.[106]

Patients with AS have been reported to have changes in metabolites of lipids metabolism, alanine pathway, and immune response.[100] Worsening of calcification have been shown to be directly proportional to inflammation; in fact, calcification is believed to follow similar processes as atherosclerosis.[107] Some strong correlations between valve pathology, VHD severity, and clinical markers and metabolic biomarkers have been widely reported.[15,101,108,109] While comparing mild-severe CAS and severe CAS, lipid metabolism and biosynthesis have strongly been associated with severe CAS.[110] To be specific, LysoPE, MG, and LysoPA showed the strongest association with CAS.[110] In addition, aortic valve replacement (AVR) has been reported to affect nitric oxide (NO) synthesis, fatty acids, and tetrahydrobiopterin metabolism.[109] In addition, the involvement of NO synthesis, tetrahydrobiopterin and oxidative stress processes in CAS further suggests pathogenesis similarity to atherosclerosis. The AVR interventions have been shown to reverse the levels of antioxidant metabolites, NO metabolism metabolites, and steroids.[109] Such reversals may suggest that they are involved in the worsening of the valve pathology, or they may represent adaptive strategies to protect the heart or body from the consequences of heart failure.

With regards to BAV, changes of metabolites mapping to glycine, serine and threonine metabolism, and taurine metabolism pathways were observed.[14] While comparing stenotic BAV and tricuspid AV with and without dilatation, changes in alpha-tocopherol and choline pathways has been reported.[98] The affected pathways suggest the role of inflammation, oxidative stress, and endothelial damage in BAV.[98] In addition, metabolic patterns observed subsequent to valve repair or replacement can be used to associate their contribution to the worsening congenital and degenerative aortic valve pathology. Valve-specific differences in changes of metabolites associated with arginine and proline metabolism pathways before-TAVR and 7 days post-TAVR have been reported. Furthermore, arachidonic acid has been associated with poor haemodynamics outcomes post-TAVR.[15]

Metabolomics for valvular heart disease diagnosis and prognosis

In RHD patients, some metabolites (caprolactam, N-acetylneuraminate, arachidonic acid, L-5-hydroxytryptophan, D-pantothenic acid, and 4-nitrophenol) can differentiate patients from healthy individuals.[105] MS and MR can be differentiated by changes in formate and lactate levels in RHD patients.[18]

Exploration of plasma and urine metabolites can tell between AS patients from healthy individuals.[111] In addition, the plasma biomarkers have been shown to be agreeable to those detected in urine after normalisation with creatinine.[111] The candidate metabolites detected in plasma and in urine showed excellent biomarker performance in differentiating aortic stenosis patients from healthy controls.[111]

Urine metabolomics of BAV patients reported glycine, hippurate, and taurine as being able to differentiate patients from healthy individuals.[14] It is easy to work with urine since it does not specialised collection methods.[14] In addition, serum glycerophospho-N-oleyl ethanolamine, monoglyceride, and phosphatidylethanolamine have been shown to differentiate BAV patients from healthy participants.[112] In calcific BAV patients, changes in lipids and lipoprotein metabolism is associated with endothelial damage and inflammation. A recent study used random forest prediction model where alpha-tocopherol predicted aortic valve morphology or dilation of the ascending aorta in BAV patients when combined with endothelial microparticles (EMPs) and C-reactive proteins (CRP).[98]

Prognostic metabolites in valvular heart disease

Metabolic biomarkers can be used to predict patients' outcomes after valve replacement surgery. Changes in amino acids, biogenic amines, and glycerophospholipids have been correlated with aortic valve pathology changes in patients with high gradient aortic stenosis after valve replacement.[113] Specifically, glycerophospholipids levels reversed post-TAVR towards healthy

control levels.[113] Further, acylcarnitine, alanine and PCs showed strong correlation with changes in left ventricular ejection fraction (LVEF), left ventricular end-diastolic diameter (LVEDD), left ventricular mass index (LVMI), and left ventricular posterior wall thickness in diastole (LVPWd) after AVR suggesting that the metabolites would predict reverse remodelling.[113] Long chain acylcarnitines have been reported as suitable predictors of LV reverse remodeling after AVR since long chain acylcarnitines (C16, C18:1, C18:2, and C18) were decreased in AS patients 24 hours post-AVR.[103] Changes in arachidonic acid metabolism pathway post-TAVR, low hemodynamics, and poor ventricular function before-TAVR are associated with mortality.[15] Arachidonic acid metabolism could also be involved in delaying ventricular reverse remodelling post-intervention. Arachidonic acid metabolism can be a suitable therapeutic target to facilitate decreased myocardial fibrosis and regained myocardial function.[15] The reported metabolic biomarker with prognostic capabilities suggests that metabolomics can give us biomarkers that may guide treatment of VHD patients.[114]

Limitations of metabolomics applications in valvular heart disease

It is always challenging to apportion causation of circulatory metabolites with cardiac pathology as some of the metabolites could be indicators or epiphenomena of metabolic disturbances in other organs. Elevated levels of phosphatidylcholine have been reported in AS while it could also be associated with metabolic processes elsewhere.[102] Further, metabolomics studies are not capable of determining functional consequences that would arise from therapeutically targeting specific metabolic biomarkers or pathways due to unknown supplementary pathways. Despite there being several biomarkers from omics studies, only a handful have found application in clinical practise due to low specificity and sensitivity.[115] To avoid pitfalls of biomarker discovery in VHD, studies should be meticulously designed. Special attention is needed in patients and healthy controls selection, sample size, samples handling, and use of other molecular techniques to validate the

identified biomarkers. Further, correlation of circulatory metabolites to tissue-specific metabolites would provide closer evidence to their application in the pathophysiology studied.

2.2 Study rationale

It is evident that VHD is highly endemic in Africa and more alarming is the high prevalence of RHD in the LMICs. RHD is one of the leading cardiovascular related causes of mortality and morbidity among individuals aged below 50 years. There is also an increasing prevalence of degenerative aortic valvular disease due to increasing life expectance in the developing countries. It is challenging to detect early progressions of the VHD using the standard diagnostic tools and the challenge is compounded in regions with limited resources. VHD proceed covertly and only presents when almost irreversible cardiac damage has occurred; therefore, there is an urgent need to develop easy and affordable diagnostic tools for early diagnosis that can easily be deployed in resource limited regions. Metabolomic biomarkers can be used in combination with the current existing diagnostic tools for screening individuals suspected to progress with VHD.

For most of patients diagnosed with VHD the only definitive intervention is valve, repair, replacement, or TAVR. However, there are no biomarkers that can be used to prognose outcomes or for risk assessment of patients who might end up with worst outcome post-intervention. Metabolic dysregulation assessment with metabolomics techniques provides a unique opportunity to understand the metabolic signatures associated with good or worse outcomes post-intervention.

2.3 Aims and objectives

Hypothesis

Metabolic profiling can diagnose and explain metabolic processes associated with RHD and degenerative aortic valvular disease progression.

Study aims

This study aimed to investigate RHD and degenerative aortic valvular disease circulatory biomarkers for fibrosis, specific valve lesions, their spatial distribution, and metabolic changes on affected tissues.

Objectives

- Describe the histological features of valve fragments from RHD and degenerative AS patients and associate them with spatial localized metabolites.
- Determine circulatory metabolic biomarkers with diagnostic capabilities and the associated pathways dysregulated in RHD, healthy controls, and degenerative AS.
- Describe tissue-specific metabolic biomarkers and related pathways dysregulated in RHD and degenerative AS that may explain the aetiology of the observed valve pathology.
- Co-localise the dysregulated metabolites on the valve tissues in patients with RHD and degenerative AS using MALDI-MSI.

3 STUDY METHODOLOGY

This chapter forms a detailed explanation of the methods used in this thesis for the work presented in the subsequent chapters.

3.1 Study design, study setting and recruitment

3.1.1 Study design

The study aimed to determine metabolic biomarkers that could classify RHD from degenerative AS, and healthy individuals (controls). The healthy participants were sex-, age-, and ethnicity-matched. Blood samples and valve biopsies from chronic RHD and AS groups were collected during a planned open-heart valve replacement surgery. Control blood samples were obtained from healthy participants. The samples were categorised into RHD and AS groups based on histopathological features observed in H&E-stained thin sections of the valve fragments and clinical diagnosis by a qualified pathologist. Circulatory and tissue-specific metabolic biomarkers were analysed using untargeted metabolomics with a LC-MS/MS QTOF system. Further, colocalisation of histopathological features with metabolic biomarkers was performed on valve biopsies using MALDI-MSI (**Figure 3.1**).

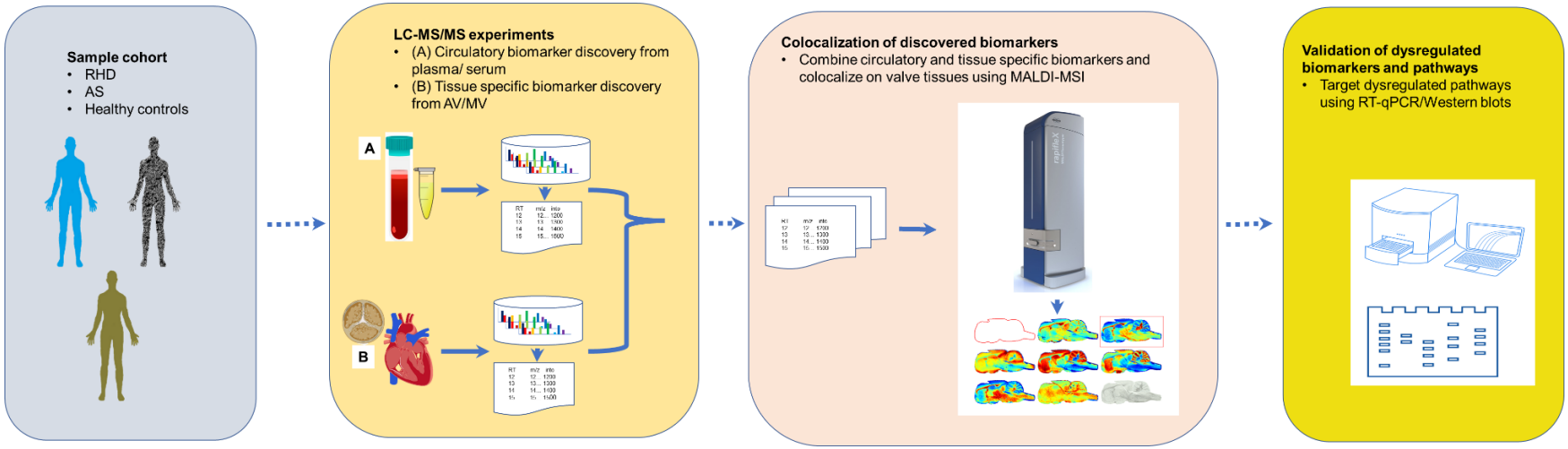


Figure 3.1 Schematic summary of experimental design. (RHD: rheumatic heart disease, AS: degenerative aortic stenosis, LC-MS/MS: tandem liquid chromatography mass spectrometry, MALDI-MSI: Matrix assisted laser desorption/ionisation mass spectrometry imaging, RTqPCR: reverse transcription quantitative real-time polymerase chain reaction)

3.1.2 Study setting

Before recruitment of participants, the study was reviewed and approved by the University of Cape Town Faculty of Health Sciences Human Research Ethics Committee (UCT HREC approval reference number 574/2018 linked to 554/2017) (**Appendix 1**). All the participants had to read and sign the study participation consent forms before being recruited into the study. The participants were recruited from the Cardiac Clinic, Groote Schuur Hospital, Cape Town, who were booked to undergo valve replacement surgery in the Christiaan Barnard Division of Cardiothoracic Surgery, Groote Schuur Hospital, Cape Town. Valve fragments and blood samples were immediately processed and stored at -80°C at the Cape Heart Institute until analysis. The experiments were conducted at the Cape Heart Institute, Department of Medicine; Division of Molecular Pathology, Department of Pathology; UCT Pharmacology and Toxicology laboratory, Department of Forensic Toxicology, and the Hair and Skin Research Laboratory, Department of Medicine, University of Cape Town.

3.1.3 Study population size and sampling

A total of 106 participants (inclusive of the cases and controls) were recruited into the study, 6 participants were later excluded after recruitment due to human immunodeficiency virus (HIV) infection, tuberculosis (TB) infection, infective endocarditis, or dilated cardiomyopathy. The patient cohort was subdivided into two subgroups i.e.; patients clinically diagnosed with RHD and those with degenerative AS based on transthoracic echocardiography (TTE) and CMR as per the World Heart Federation and American College of Cardiology/American Heart Association guidelines on the diagnosis of RHD and VHD.[9,10,116] In addition, participants who were age-, sex-, ethnicity-, and comorbidities-matched and without any cardiovascular disease were recruited as healthy controls. Patients younger than 18 years and older than 80 years, with atherosclerosis, HIV infection, cardiomyopathy, congenital valve lesions, other inflammatory conditions, previous cardiac surgery or valve repair, and unable to consent were excluded (**Table 3.1**).

Table 3.1. Inclusion and exclusion criteria

Inclusion criteria	Exclusion criteria
• Patients diagnosed with severe or rheumatic valve disease, calcific aortic valve stenosis as per Jones criteria 2015, WHF, and 2020 ACC/AHA guidelines • Patients aged 18-80 years • Patients with written informed consent	• Patients with: -autoimmune inflammatory disorders -atherosclerosis, HIV infection, cardiomyopathies, congenital valve lesions -previous cardiac surgery or valve repair -aged <18 and >80 years • Patients unable to consent

It should be noted that it is nearly impossible to determine *a priori* the sample size required for untargeted metabolomics.[117] However, this study attempted to estimate the sample size using G*Power[118] (version 3.1.9.7); the sample size was estimated to be 13 and 27 in the two groups assuming normal distribution in the population: with an accepted significance level of less than 0.05 (p value<0.05). Multivariate studies mostly focus on the number of variables that will be significantly different given a certain sample size and with less emphasis on the effect size of the differences. Since the beginning of this PhD project, other tools for determining sample size and power while considering effect size of the multivariate changes have emerged.[119–122] The allowed minimum fold change of the significant metabolites was >2 which translated to an effect size (d=1.5), this was suitable to give an expected 95% power (1-β).[123] Although the estimations were thought sufficient to predict the desired sample size for economic and planning purposes, there are inherent limitations associated with using unimodal effect size and standard deviation estimations in multidimensional metabolomics dataset.[117] Haematoxylin and eosin-stained thin sections were used for histopathological confirmation of chronic RHD or degenerative aortic valve disease. The sections were reviewed by a qualified pathologist based on the presence or absence of classical RHD features; Aschoff bodies, inflammation, neovascularisation, vegetation, hypertrophy, fibrosis, or collagen deposition.

Valve leaflets without signs of chronic inflammation, fibrosis, and neovascularisation were classified as degenerative aortic valve disease.[42,124] Periphery blood was collected from RHD, AS and healthy controls into VACUETTE® Z serum clot activator vacutainer 6ml tubes (Greiner Bio-One International GmbH, Kremsmünster, Austria) for serum isolation. In addition, periphery blood was also collected into VACUETTE® K3E K3EDTA vacutainer 9ml tubes (Greiner Bio-One International GmbH, Kremsmünster, Austria) to isolate plasma. The collected samples were immediately transported to the laboratory for further processing before storage.

3.2 Sample collection, processing, and analysis

3.2.1 Consenting and Samples collection

The study was conducted in compliance with the 2013 Helsinki declaration. Consenting participants signed the informed consent forms, and they were thus recruited into the study (**Appendix 2**). After consenting, samples were collected from the participants as per the specimen collection standard procedures. During recruitment, the patient's echocardiographic and CMR examination data, demographics (date of birth, weight, height, blood pressure, comorbidities, gender, and ethnicity), and sample processing logs were recorded in the recruitment/sampling checklist that was approved by HREC (**Appendix 3**).

The non-fasting blood samples were kept at room temperature for at least 60 minutes before processing. Plasma and serum were isolated by centrifuging the blood samples at 2,000 g for 15 minutes in a cooled centrifuge (4°C), then 500 µL aliquots of plasma and serum were obtained and stored at -80°C until analysis. Valve biopsies were collected after excision by the cardiothoracic surgeon or forensic pathologist. The valve biopsies were immediately transversely dissected perpendicular to the cut-edge (valve ring) and free edge of the cusps and aliquoted. The aliquots meant for H&E staining were immersed in 10% neutral buffered formalin (NBF) while the other aliquots were immediately snap frozen in liquid nitrogen and stored at -80°C until analysis. The H&E-staining analysis aliquots were stored in the 10% NBF until embedding in wax and the blocks were carefully archived until histology experiments.

3.2.2 H&E processing and histopathological classification

Tissues for histopathological classification were processed for FFPE and H&E staining after being fixed with 10% normal buffered formalin at room temperature. Extreme care was taken to ensure that the tissues were always covered with an adequate amount of formalin. The fixative volume should be 5-10 times of that of the tissue volume. Fixed tissues were trimmed into appropriate sizes and shapes to fit into the tissue cassettes. The Leica tissue processing machine, Leica TP 1020 (Leica microsystems Nussloch GmbH, Heidelberger Str., Nussloch, Germany) was used to process the tissues for 22 hours using a protocol for processing human tissues (**Appendix 4**).

The tissue sections were cut at 3 µM thickness with the blade angle set at 3º. An ice tray was used to cool the blocks and enable sectioning of the tissues. The sections were then picked up with forceps and floated onto a 40ºC water bath. These cut sections were mounted onto the surface of a clean microscope glass slide (Marienfeld, Lasec) and placed on a hot plate at 60ºC for 10-15 minutes to bake. The slides were then hydrated by passaging through xylene to remove wax and cleared in graded alcohols (95%, 80%, 70%), followed by a thorough wash in tap water. Thereafter the slides were stained in Mayers haematoxylin for 10 minutes, washed in tap water and placed in ammoniated water for blueing. This was followed by staining in Eosin-Phloxine for 2 minutes, then washed with water. The slides were then dehydrated by passaging through 3 stages of graded alcohol and then into 3 jars of xylene. The microscope glass cover slip (Lasec) was then mounted using Entellan® mounting medium (Merck, Darmstadt, Germany). To assess the quality of the H&E staining the slides were viewed under the microscope. The formalin fixed paraffin embedded tissue (FFPE) blocks were appropriately stored until further analysis.

A senior pathologist assisted with reviewing the H&E slides to investigate the morphological and histological features associated with RHD and degenerative AS. The histological reports were correlated with the medical history, echocardiographic reports, and cardiovascular magnetic resonance imaging results of the patients.

3.2.3 Sample processing

To prepare samples for metabolomics studies, the serum aliquots from RHD, AS, control participants were thawed in ice. From each aliquot (RHD, AS, and controls) 50 µL was obtained and pooled into a larger aliquot that formed the pooled long-term quality control (QC) sample which was re-aliquoted into 1,000 µL aliquots and stored in -80°C. The blank samples were composed of HPLC-grade water. The thawed samples and the prepared pooled QC were consequently re-thawed in ice during metabolites extraction. All the sample handlings and processing were performed in biosafety level (BSL) 2 cabinets at the Cape Heart Institute laboratories. Metabolomics analysis with LC-MS/MS were performed in batches since all the samples could not be analysed in one run.

3.2.4 Pilot experiments for metabolome extraction and analysis methods development

To determine the extraction solvents, and the column settings to be used in this study, pilot experiments were conducted using both our in-house laboratory methods and those sourced from the literature. A small section of the samples was selected to use for the pilot experiments. The solvents and HPLC-grade water were obtained from Sigma-Aldrich unless otherwise stated. To prepare extraction solvents, different solvent combinations earlier used in our laboratory were tried i.e., 100% methanol, 100% acetonitrile, a mixture of methanol and acetonitrile (1:1v/v), and a mixture of methanol and ethanol (1:1v/v). To extract the metabolites, 50 µL of the serum sample was mixed with 300 µL of each of the extraction solvent combinations. The mixtures were vortexed at high speed and incubated overnight in -20°C. The samples were then thawed and centrifuged at 13,000 g for 10 minutes. The supernatant was then transferred into 3 mL culture glass tubes and evaporated by blowing nitrogen gas into the tubes. The dried samples were reconstituted with HPLC grade water with 0.1% formic acid. The mobile phases were prepared as follows: A = HPLC grade water with 0.1% formic acid and 2mM ammonium formate, B = a mixture of 100% acetonitrile and 100% methanol (1:1v/v) with 2mM ammonium formate and 0.1% formic acid. The stationary phase was Omega

Polar C18 column (particle size 1.6 µm, pore size 100Å, length 10 cm, internal diameter 3 mm) (Phenomenex Inc., Torrance, CA, USA) that was operated at 40°C column oven temperatures. The metabolites were separated using a column gradient that was set to start at 2% B and increased to 100% B for 90 minutes. The flow rate was set at 0.7 mL/min for the entire duration of the gradient. The gradient settings used are detailed in (**Table 3.2**). The maximum pressure limits were set at 400 bar and the injection volume was set at 10 uL.

Table 3.2. The pilot liquid chromatography gradient and flow program

Time (min)	B Conc (%)	Flow (mL/min)
1.00	0.0	0.7
60	100.0	0.7
70	100.0	0.7
70.10	0.0	0.7
90.00	0.0	0.7

The triazine pesticide was used as an external standard where 10ng was mixed in 10% methanol mixed with distilled water (Sigma-Aldrich). To determine the suitable chromatographic settings, an in-house pesticide mixture was used for method development to determine the suitable settings to determine the suitable resolution of the molecules. The pesticide mix was used since it had a mixture of polar and non-polar compounds. To obtain the MS data, the information dependent acquisition method was used in the positive ionisation mode using X500R QTOF-MS/MS as detailed in the mass spectrometry sections later. The total acquisition time was 90 minutes, the total scan time was 0.9 seconds per scan, and the MS accumulation time set at 0.15 seconds and the MSMS accumulation time was 0.1 seconds. The mass range analysed was between 50 – 1,200 Da for both MS and MS/MS acquisition modes. The raw dataset files were converted to .mzXML format from the proprietary .wiff2 format using ProteoWizard (version 3.0.1908) MS-Convert method as described later. Features extraction, annotation, and pathway analysis were performed in XCMSOnline (version 2.7.2), XCMS (version 1.47.3) and CAMERA (version 1.34.0). For feature detection parameters, the matchedfilter method was used, Gaussian-like peak shape full width at half maximum (fwhm) = 30 seconds, signal-to-noise-threshold (snthresh) = 2, maximum noise

cutoff ratio (max) = 10, m/z steps for binning (step) = 0.1 m/z, steps = 2, m/z differences between bins used to classify features (mzdiff) = 0.01 Da. Retention time correction method = obiwarp, the step sizes in m/z to build the chromatogram profile (profStep) = 0.1 *m/z*. Features/peaks grouping was based on density, bandwidth filter for finding well behaved peak groups (WBPG) within a retention time (bw) = 2 seconds, *m/z* width of slices to classify features falling in the same retention time (mzwid) = 0.01 *m/z*, the fraction of samples per group where the peak must be detected to be accepted (minfrac) = 0.5 (50%), minimum number of samples per group (minsamp) = 1. Normalisation of the features was performed using a regression technique known as Locally Estimated Scatter-Plot Smoother (LOESS). The samples where the detected and accepted peaks were missing were filled using 'fill peaks' function that fills peaks based on if the peak was detected in other samples. The peak must have been present in at least 1 sample in a group to be filled as missing peak otherwise it was excluded from the data matrix. Features annotation was performed in CAMERA where the isotopes and adducts were annotated, m/z absolute error (mzabs) = 0.05 *m/z*, general ppm error (ppm) = 5 ppm, multiplier of the standard deviation (sigma) = 6, percentage of the FWHM width to use for retention time matching (perfwhm) = 0.6, maximum ion charge state (maxcharge) = 3, maximum number of expected isotopes (maxiso) = 5, general used intensity values = into. The putative identification was based on matching the experimental in-source MS^1 and MS/MS fragments to the MS^1 and MS-MS spectra in the METLIN database. The putative annotation was based on the following parameters: - ppm = 50; adducts = M+H, M+Na, M+K, M+Na-($C_{12}H_{22}O_{11}$-H_2O) (**Appendix 6**). The pilot experiments were meant to establish quantitatively a suitable extraction solvent to use in the study. It was also important to establish suitable data pre-processing, data processing, features putative annotation, and pathway analysis techniques and tools to use in the study. After optimising the settings for the equipment and the data processing parameters, then I performed metabolome extraction for the main study.

3.2.5 Study metabolome extraction from serum

A metabolite extraction step was conducted on serum from samples (RHD, AS, and controls), a pooled QC, a solvent blank and an external standard. To prepare the external standard the NIST SRM 1950 human plasma reference standard was procured (National Institute of Standards and Technology, Gaithersburg, MD, USA). The samples were allocated at random to each batch. However, in each batch there was, on average, 3 RHD samples, 2 AS samples and 3 RHD samples. During batch preparation, there was 1 pooled QC, 1 external standard, and 1 blank that were extracted together with the samples. The pilot experiments had shown that a mixture of acetonitrile and methanol (1:1v/v) performed better as the extraction solvent. Therefore, to extract metabolites, 100 µL of samples, pooled QCs and external standards were added to 400 µL of ice-cold extraction solvent. In a batch, there were 2 pooled QCs and 2 external standards that were injected alternatively every after 3 samples injections, in addition 1 blank was injected at the start and the end of each batch (**Figure 3.2**).

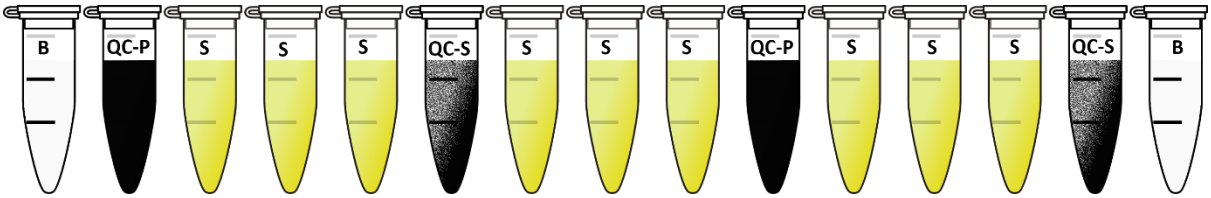


Figure 3.2. Schematic for batch design used in the LCMS/MS analysis of serum. QC-P = pooled sample, QC-S = external standard, B = Blank, S= Samples (RHD, AS, and HC)

The mixture was vortexed at high speed for 2 minutes (Labnet International, Edison, NJ, USA). The samples were then centrifuged at 14,000 g for 15 minutes (Sigma Laborzentrifugen GmbH, An der Unteren Söse, Osterode am Harz, Germany) precooled to 5°C. The supernatant was dried and reconstituted in reconstitution solvent made of HPLC-grade water mixed with 100% acetonitrile (98:2v/v) and modified with 1% formic acid or 1 mM acetic acid for positive and negative modes respectively (**Appendix 2**). The reconstituted samples were transferred into micro-inserts of 2 mL autosampler vials then loaded into the ExionLC™ AD Series UHPLC autosampler that was set at 15°C. The sampling speed of 5 µL/s and a needle stroke height of 52 mm settings were used. The rinse settings were set to rinse for 2 seconds at a flow rate of 35 µL/s with a rinsing volume of 500 µL and was set to rinse before and after aspiration.

3.2.6 Study metabolome extraction from biopsies

Next, the biopsy tissues were extracted and analysed. To obtain the pooled QC sample for the valve biopsies, 20 mg was taken from all the samples (RHD and AS), mixed and homogenised by crushing them in liquid nitrogen until it formed a uniform powder. The homogenised pooled QC sample was then re-frozen until analysis. The samples were processed for metabolites extraction in batches where each batch was composed of cases (RHD and AS), pooled biopsies QC, pooled serum QC (long-term QC), and blank samples (**Figure 3.3**). To extract metabolites from the frozen valve biopsies, 40 mg of the tissue was added to 400 µL of the ice-cold solvent made of 100% acetonitrile and 100% methanol mixed in equal proportions (1:1v/v). To prepare the blank sample, only 400 µL of the prechilled extraction sample was added and processed as the tissue biopsies. In addition, the long-term QCs were extracted by mixing 100 µL of the pooled serum QC into 400 µL pre-chilled extraction solvent. After mixing the valve biopsies with the extraction solvent they were homogenised with a prechilled Omni Bead Ruptor 24 homogenizer (Omni International, Cobb Place Blvd, Kennesaw, GA, USA) using 0.6 g of 1.4 mm ceramic beads and homogenized for three times over 20 seconds at 5,500 rpm with 30 seconds pause intervals while maintaining the temperature at 0 °C (**Appendix 5**). The extracts were centrifuged at 14,000 g for 15 minutes, the supernatant was dried and then reconstituted in the reconstitution solvent as stated earlier.

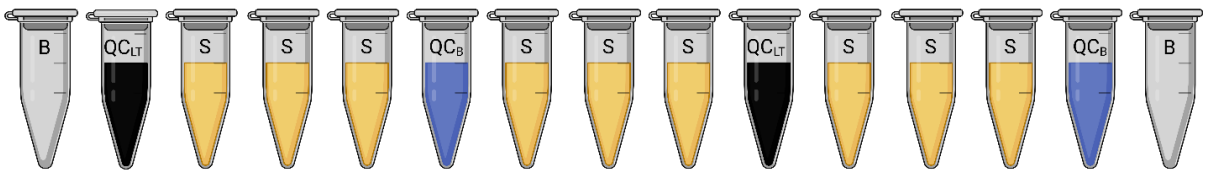


Figure 3.3. Schematic for batch design used in the LCMS/MS analysis of tissue biopsies. QCLT = long-term QC (pooled serum), QCB = pooled biopsies QC, B = Blank, S= Samples (RHD and AS). Created with BioRender.com.

3.2.7 Chromatographic and mass spectrometry data acquisition

To establish system suitability, the equipment was calibrated every day before commencing any experiments. Calibration was performed using 10 ng of the Sciex ESI positive and negative calibration solutions (Sciex, Framingham, MA, USA). The autocalibration method was used and the equipment was considered calibrated and ready for analysis when all the masses were

detected and within the preset mass window. In addition, the system was first equilibrated and purged with ethanol for 20 minutes and it was ensured that the back pressure was stable. The serum and biopsy extracts (samples, QCs, and standards) and blanks were injected systematically in all batches for extracts as stated earlier (**Figure 3.2 and Figure 3.3**). The blanks were used to later identify and subtract system contaminations while QCs were used for retention times' normalisation and batch drift adjustments. On average each batch was composed of 15 injections that run for 60 minutes which translated to a batch length of 15 hours. The samples were held in the ExionLC™ AD Series UHPLC autosampler that was cooled to 15 °C. The serum and tissue metabolite extracts for the entire project were analysed with reversed phase chromatography using ExionLC™ AD Series UHPLC system operated on Sciex OS software ver.1.4 (Sciex, Framingham, MA, USA). Separation was performed with Omega Polar C18 column with particle size 1.6 µm, pore size 100Å, length 10 cm, and internal diameter 3 mm (Phenomenex Inc., Torrance, CA, USA); operated at 40°C column oven temperatures. The mobile phase A was composed of HPLC grade water, while mobile phase B was a mixture of 100% acetonitrile and methanol (1:1v/v) both modified with 2 mM ammonium formate as final molarity and 0.1% formic acid (v/v) for positive mode. To acquire in negative mode, the mobile phase A and B was modified with 1 mM acetic acid. The mobile phases for both modes were run on a 60-minute gradient at a rate of 0.4 mL/minute. The gradient was allowed 2 minutes equilibration at 2% B, then the gradient was increased linearly for 45 minutes to 95% B, the gradient was held again for 2 minutes at 95% B before being decreased sharply to 2% B within 6 seconds, the gradient was then allowed to equilibrate at 2% B for 2 minutes before being stopped (**Table 3.3**). The autosampler injection volume was 5 µL/seconds set at 15°C with rinsing set at 2 seconds at 35 µL/seconds and the rinsing volume was 500 µL. The minimum and maximum backpressure was set at 0 bar and 400 bars, respectively.

Table 3.3. The study liquid chromatography gradient and flow program

Time (min)	B Conc (%)	Flow (mL/min)
2.00	2.0	0.4
45.00	95.0	0.4
47.00	95.0	0.4
47.10	2.0	0.4
60.00	2.0	0.4

The LC-MS/MS data was acquired in positive and negative ionisation mode on a high resolution Sciex x500R Q-TOF mass spectrometer operated with Sciex OS software ver.1.4 (Sciex, Framingham, MA, USA) with dual ESI ion source system. The data was collected using information dependent mode with a mass range between 50 – 1,200 Da with a resolving power of ≥ 40,000 (FWHM) at *m/z* 956 and mass accuracy of <5 ppm. The total duration was set at 60 minutes with a total scan time at 0.6672 seconds with approximately 5355 cycles. Ion source gas 1 was set at 40 psi; ion source gas 2 was set at 65 psi and the temperature set at 500°C. The curtain gas was set at 25 psi. To obtain the TOF-MS data, the ion spray voltage was set at 5000 V, CAD gas 7 and data obtained in profile format with a mass range of 50−1200 Da. TOFMS accumulation time was 200 milliseconds, declustering potential 70 V, spread 0 V, collision energy 10 V, collision energy spread 0 V and data acquisition rate was 4 scans/second. The TOF-MSMS data of the fragment ions was obtained in a data dependent format, the following settings were used; maximum candidate ion 7, intensity threshold at 50 counts/seconds, mass tolerance of 50 mDa, ion spray voltage 70 V, target mass range 50 – 1200 Da, accumulation time 60 milliseconds, declustering potential 70 V, declustering potential spread 0 V, collision energy 35 V, collision energy spread 15 V with Q1 unit resolution and data acquisition rate was 8 scans/second. The data was obtained in .wiff and .wiff2 formats for each batch and was stored in a secure device until processing and analysis.

3.2.8 Mass spectrometry data pre- and post-processing

After obtaining the raw data, it was further processed and analysed through a workflow as summarised in **Figure 3.4**. The data was pre-processed to convert proprietary .wiff files to open .mzML file format using an open-source MS-Covert in ProteoWizard (v 3.0.1908) software. To

reduce the size of the data files, the data was also converted from profile to centroid formats using the peak picking and the continuous wavelet transform (CWT) algorithms; the algorithm settings were MS level = 1 - ∞, minimum signal to noise ratio (min SNR) = 0.1, and the minimum peak spacing = 0.1.[125,126]

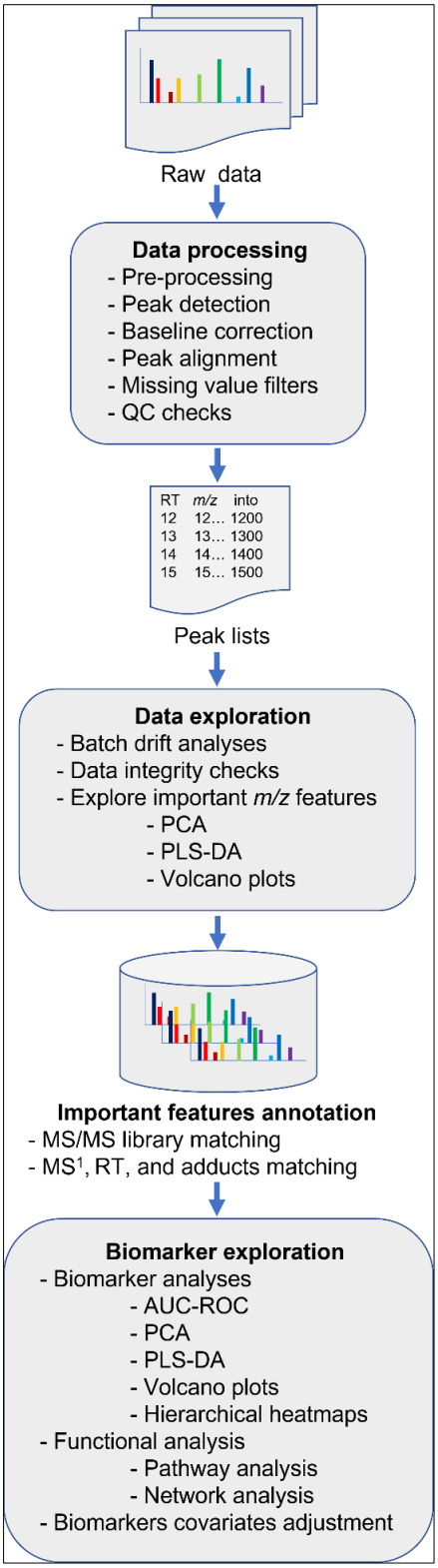


Figure 3.4. LC-MS/MS Metabolomics data processing and analysis workflow. (QC: Quality Control; m/z: Mass-charge ratio; PCA: Principal Component Analysis; PLS-DA: Partial Least Square-Discriminant Analysis; MS1: accurate mass; MS/MS: Tandem Mass spectrometry; RT: Retention time; AUC-ROC: Area under the curve-receiver Operator Curve)

After conversion, the data was further processed to extract *mz*RT features using MS-DIAL (v 4.38) which is an open-source metabolomics workflow software composed of peak detection

and spotting for feature extraction, alignment, and de-noising (**Figure 3.4).** The software was downloaded from (http://prime.psc.riken.jp/compms/msdial/main.html). The software was saved on the local computer meaning data did not need to be uploaded online. The software required pre-installation of .NET Framework 4.0, or the most recent version obtained from (https://dotnet.microsoft.com/en-us/download/dotnet-framework/net40).

The workflow started with defining the data collection parameters; the MS^1 and MS^2 tolerances to integrate the MS^1 and MS^2 scans. To execute peak detection and smoothing, the minimum peak height and mass slice width were defined (**Appendix 6**). Peak picking was conducted after retention time smoothing which was based on linear weighted moving averages guided by the defined sigma levels. The linear-weighted-moving average algorithm was deployed for smoothing of the chromatograms which considered the retention time and the accurate mass of the samples. The peak detection was performed with 3 threshold values and a backtracking method i.e., first derivative and second derivatives that used the amplitude filter and the first derivative filter values. The two values were intended to evaluate noise which was obtained by 5-point approximations.[96] To detect the peak edge the amplitude and first order derivative had to exceed the amplitude filter and the second order derivative is negative.

To generate the *m/z* features list, the algorithm considered all the data points of the base peak and assigned the scan number, retention time, the base peak *m/z*, and the intensity. To obtain the final spot also referred to as region of interest (ROI), the adjacent spots of the same *m/z* and retention time were combined and only the one with the highest intensity was picked (**Figure 3.5**).[96,127]

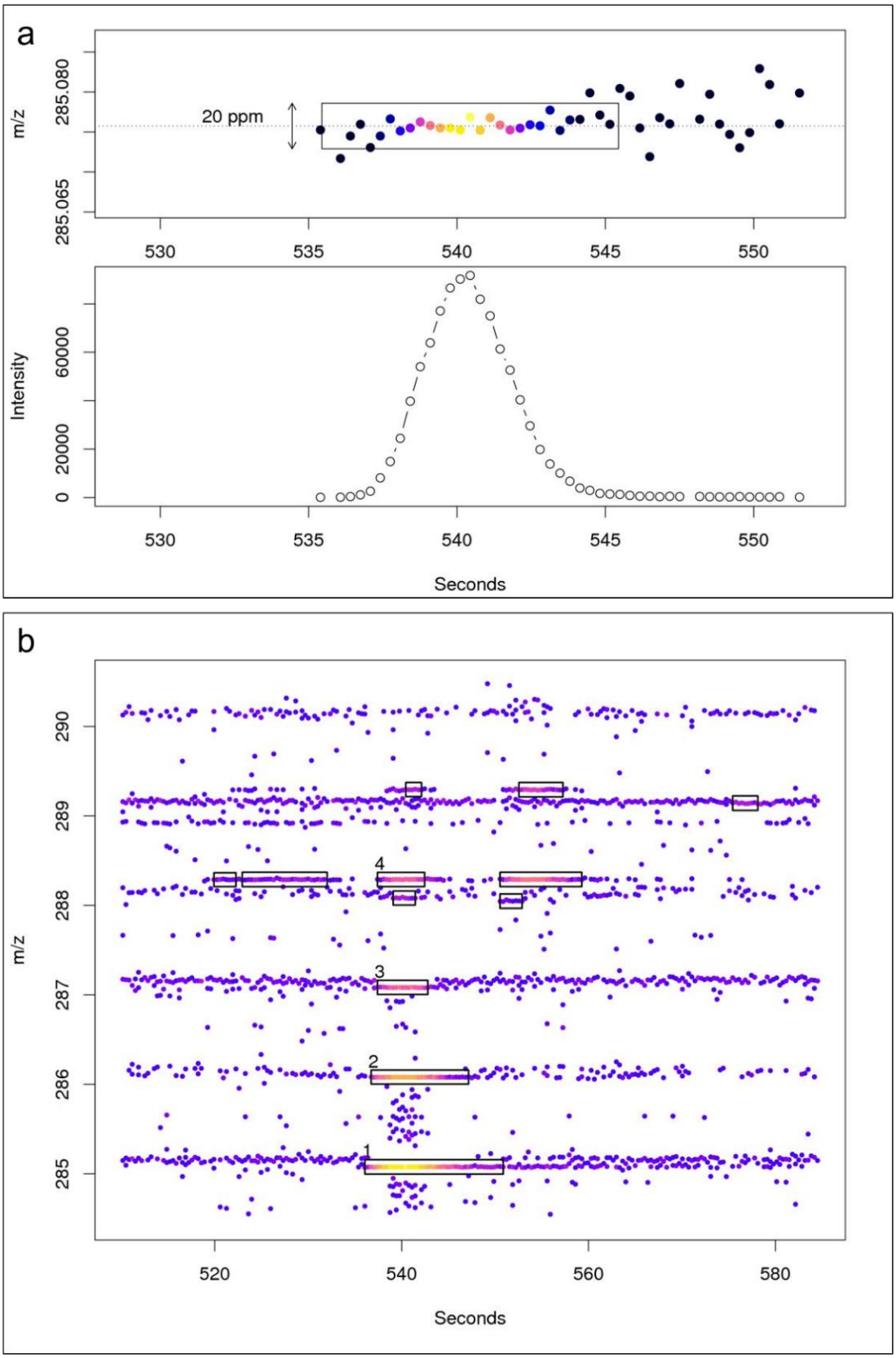


Figure 3.5. Schematic description of m/zRT features matrix extraction based on regions of interest (ROI). (a) The predefined mass error that traces the chromatogram based on the detected mass intensities to build the corresponding chromatogram. (b) The distributed regions of interest with the same retention time but at different masses extracted for further analysis. Adapted from Tautenhahn et al. Highly sensitive feature detection for high resolution LC/MS. *BMC Bioinformatics.* 2008;9(1):504. doi:10.1186/1471-2105-9-504

A very low minimum peak height was used that ensured to detect even the lowly expressed features but at the expense of data processing time and noise. Noise was removed during statistical analysis as explained in detail later. The parameters for peak picking and *m/z* features extraction can be found in **Appendix 7**.

Noise in the MSMS data was reduced using MS2Dec deconvolution algorithm that extracted MS/MS chromatogram, performed baseline correction, smoothing, extracted the model peak, and fitted the MSMS chromatogram peaks to the pre-extracted MS peaks. The algorithm required the segment value and the chromatogram band width, along with the median absolute abundance of the chromatogram that established the linear baseline correction that determined true peaks from pseudo peaks.[96]

Following MSMS deconvolution, the peak alignment algorithm was executed for peak spots alignment. Peak alignment was meant to find the warping function to synchronize all the extracted peaks to a common retention time within all the samples.[128] Peak alignment was important for downstream statistical inferences when comparing levels of potential biomarkers. The alignment was designed to correct the non-linear distortion of the chromatographic system to the individual features detected in each sample. It assumed that similar molecules would behave the same in the chromatogram and any distortions in their retention times is due to systematic variations.

The peak alignment was performed in 4 steps i.e. (1) constructing the reference table, (2) matching each sample peak table to the reference table, (3) filtration and (4) filling-in the missing values.[96] To construct the reference peak table the algorithm used the pooled QC samples as the reference file based on the defined RT and MS^1 threshold values. To assign each sample peak to the reference peak table the algorithm used the weighted preferences (RT factor and MS^1 accurate mass factor) on whether to use RT or the MS^1 while adding an extracted sample peak to an aligned peak group.

The algorithm interrogated each feature at a time and compared it with the corresponding peak extracted in the reference sample. If the queried feature could not be aligned to a feature in QCs and not found in any other samples, it was deleted to avoid unnecessary features. After successful alignment, the algorithm generated an alignment ID, average retention time, average *m/z*, and the intensities of the *m/z*RT features. Aligned features that were missing in some of the samples were filled by linear interpolation using the average retention time, average m/z and the average intensity of the preceding and succeeding detected features[128] The blank samples were analysed, and the algorithm provided an option to filter unwanted features based on the blanks. Features were removed if they had an intensity which was more than 5-fold-changes between samples and blanks. Further filtration was done based on the frequency of missing values in the aligned feature within sample groups and/or generally within the whole dataset as detailed in **Appendix 7**.

It was impossible to run all the samples in one batch thus samples were analysed across multiple batches. Batch drift analysis and correction was performed using LOWESS normalisation based on the pooled QC samples. LOWESS is a nonparametric model suitable for predicting the fitted curve. The drifts in RT and intensities were assumed to be non-linear, thus a nonparametric method was the most suitable for drift correction. During batch correction, it was important to provide the LOWESS span which is the number of data points to be considered in the predicting of the curve. It was accepted by default to use a span setting of 0.75; if a small span was used there was insufficient data points which resulted in huge variance; if too large a span was used; the regression was over-smoothed and resulted in data loss. In cases where the nonlinear regression methods were not appropriate for batch correction, linear methods (ComBat, QC-RLSC, ANCOVA, or Eigen-MS) were used as explained in **section 3.2.9** of this thesis. For more details on the parameters used in this work for data pre-processing on positive and negative modes, refer to **Appendix 7**.

3.2.9 Statistical analysis and metabolic features discovery

MetaboAnalyst was used for statistical analysis to extract important *m/z* features; the software is available at (https://www.metaboanalyst.ca/home.xhtml).[95,129] MetaboAnalyst is an open-source web-based bioinformatics platform for metabolomics data analysis. To identify discriminant metabolites the marker/peak list from MS-DIAL was prepared from the aligned features with all the QCs and samples. At first, data integrity checks were conducted i.e., batch effect analysis and missing values. In cases where the non-linear regression batch correction methods were not appropriate the QC samples were processed for batch drift analysis using batch effect correction module in MetaboAnalyst. The module used ComBat, QC-RLSC, ANCOVA, or Eigen-MS methods which were automatically selected based on the method giving the least distance between samples after correction.[89,95] In addition, the relative variation (CV) of the extracted features was analysed based on the QC samples. Missing values were also checked however in MS-DIAL, the filtration step had removed variables missing in more than 50% of sample group and 23% overall samples. The missing values that remained were filled with a small non-zero value (1/10 of the minimum peak area).[96]

Further, data normalisation was performed to remove systematic bias and ensured data consistency so that groups were comparable for easy down-stream statistical and biological analyses. The algorithm offered 3 normalisation options i.e., row-wise (sample) normalisation, data transformation, and column-wise (variable) scaling. In this work, row-wise normalisation by sum was used. Normalisation by sum is based on a mathematical model that forces the intensities of each experimental run to be equal which is similar to total ion chromatogram normalisation.[130] The data was not transformed but column-wise scaling was performed using either auto-scaling algorithm or the range scaling where the variable's intensity was centred around the mean and divided by the standard deviation or mean-centred and divided by the range of each variable, respectively. Since there is no agreement on the suitable normalisation and scaling methods, the effectiveness of the normalisation was investigated by visual

inspection of the Gaussian distribution of the normalised data.[129] The samples and variables distribution had to form a bell-shaped Gaussian curve as a sign of successful normalisation.

To discover markers associated with RHD and degenerative AS, univariate statistics tests such as student's t test, fold change, and volcano plots were used to explore differential levels of the detected *m/z*RT features. Further, multivariate statistics tests such as PCA and PLS-DA were used for dimensionality reduction to extract highly influential markers.

PCA was used for visual assessment of data sets to determine the relative differences between groups. In addition, PLS-DA was used to determine discriminant features; PLS-DA is a supervised dimensionality reduction technique that extracts variables contributing to the differences in the studied groups. Features contributing the most to the separation of groups (discriminant features) were extracted based on the weights of the PLS-DA loading scores. From the PLS-DA, important features were selected based on their variable importance in projection (VIP) scores. The VIP scores were based on the weighted sum of squares of the PLS-DA loadings. The VIP scores accounted for the y-axis variance of each component i.e., the extent of separation that the variable contributes to the separation of the groups along a specific component. The VIP was also based on the sums of the PLS-DA regression coefficients i.e., how the variables better predict separation of the groups analysed. PCA and PLS-DA despite being powerful tools for dimensionality reduction in multivariate data, they have limitations; PLS-DA specifically is based on mathematical model projections that are prone to overfitting. To ensure stability of the model, cross validation of the model was performed that measured the sum of squares between predicted data and the original data; by cross validating the sum of squares captured by the model, the algorithm provided a predictive relevance value (Q^2). PLS-DA models with very low or negative Q2 were considered overfitted and invalid to explain differences between groups.[131] The stability of the PLS-DA model was also validated using the permutation test technique.[95,129] In this work, the aim was to find features/biomarkers that were different between groups. Therefore, PLS-DA was used to

visualise general group differences, but univariate analyses was used to determine differentially expressed features/markers.

To extract discriminant features between groups, univariate 2-sided student's t test and analysis of variance (ANOVA) were used to test significance levels, while the magnitude of dysregulation was assessed using fold change (FC) which was visualized using volcano plots. P values of <0.05 were considered significant. Further, a heatmap with hierarchical clustering was used to visualize dysregulation trends. Hierarchical clustering assessed distance between variables in different groups. The algorithm assumed that all the variables were separate clusters then it bonded the clusters step-by-step into bigger clusters. In each step, it bonded each of the closest pairs to the source cluster. The process of clustering was repeated until only one cluster could be built by the next step leading to the formation of a hierarchical tree. The distance between clusters/vectors and their similarities were determined by Pearson's correlation based on the principle that the more similar the vectors/variables are the stronger is their correlation coefficient and thus the closer the distance.[132] To determine the inter-cluster distance between two clusters, cophenetic distance was computed using averaged-linkage clustering method.[95,132] When the variables are grouped together into clusters, their trends are further investigated by their differences in intensities and how they vary between study groups and color-coded for easy visualization with heatmap plots.[95]

The significant *m/z* features and those showing interesting trends from the heatmap plots were considered and added to the marker list. The selected *m/z* features were further processed for annotation by MS/MS, MS and isotope matching to publicly available human metabolome databases using GNPS feature based molecular networking[133] and CUE mass mediator tools.[134]

3.2.10 Metabolite annotations

Annotation of the extracted features was based on accurate masses, retention time, and the isotopic pattern matching with publicly available MS/MS databases such as, but not limited to:

MassBank, PubChem, METLIN, HMDB, NIST. To expand the number of identified features, annotation was also based on features' accurate mass (mass error ±10 ppm) and adducts formation ([M+H], [M+2H], [M+Na], [M+K], [M+NH4], [M+H+NH4], [2M+H]) and ([M-H], [M+Cl], [M+HCOOH-H], [M-H-H2O] for +ESI and -ESI respectively using CEU Mass Mediator (CMM) software. The software also took the retention time data, but it was not used in scoring the putative annotation of the queries. After the putative annotation, the annotated compounds were manually inspected to rule out false annotations. The two strategies of metabolites annotation met Metabolomics Standards Initiative (MSI) levels 2 and 3 classification guidelines.[135] MSI level 2 classifications required probable structure description through literature and/or library or database matching of fragmentation spectra pattern.[134,136] The MSI level 3 putative characterisation required unique matching of the parent ion data through literature search and/or libraries and databases.[134,136]

To annotate the important features, GNPS feature based molecular networking algorithm was used to annotate them to MSI level 2. The annotation was based on matching the experimental mass, isotopic ratios, MS/MS fragmentation patterns, and fragment intensities to the reference spectra libraries. The similarity scores were based on the 0-1 range where 0 = no match and 1 = perfect match. The similarity matching assumed that experimental and theoretical values followed a normal Gaussian distribution using predefined sigma value to compare expected vs observed values.

To obtain the isotope ratios matching, the reference library needed to have molecular formula where the theoretical isotopic classification could be derived using the McLaurin expansion. The MS/MS spectra similarity matching used the dot-product and reverse dot-product cosine values.[137–139] The dot product algorithm compared the experimental spectral data (vector 1) and the library spectrum data (vector 2) and the vector magnitude to obtain the dot-product similarity score. The cosine of the angle between vector 1 (dot) and vector 2 (product) was the dot-product value that indicated the measure of vectors similarities. The angle between identical vectors was 0° which translated to cosine = 1; a dot-product of 0 or 1 meant not similar

in direction or magnitude or perfect match in direction and magnitude, respectively.[137–139] Identical molecules have identical spectra with similar mass intensities therefore their vectors coincided. The reverse dot-product matching algorithm omitted peaks that did not match to the library to improve the matching scores. To minimize chances of false identification, annotation was limited to features with a minimum of 4 fragmentation peaks. In the GNPS environment, to improve the success of positive identification of unknown features, molecular networking algorithm was used.[137–139]

The molecular networking algorithm performed a pairwise spectra similarity matching where each MS^2 experimental spectra was compared to each other; similar MS^2 spectra were grouped together to form a network. The algorithm then used the *m/z* changes between neighbour spectra to build molecular networks based on the assumption that structurally related molecules would have similar MS^2 fragmentation patterns after ionization (

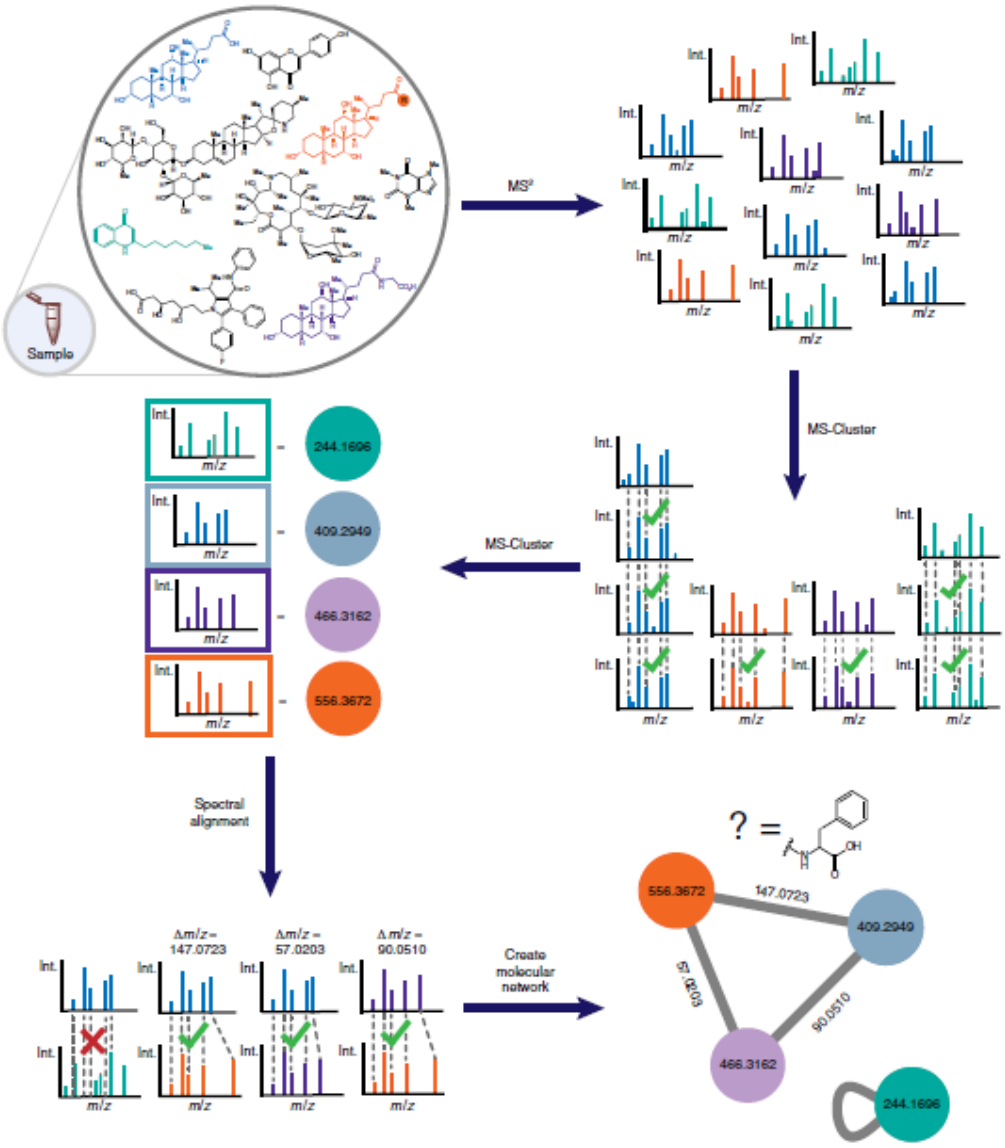


Figure 3.6).[133,137,140] The MS^2 spectra were aligned to each other in the dataset, the cosine score was assigned to describe their similarity. To create a node, identical spectra were grouped into one node or into a cluster based on the similarity scores of their ion fragments using the MS-Cluster algorithm.[133,140] Related molecules exhibit similar MS^2 fragmentation patterns, and the algorithm would use their similarities to create network edges.[133,140] After the clustering and networking, the consensus spectrum was queried against the spectra library database and assigned a putative annotation and a cosine score. With the feature based molecular networking, the unknown nodes in a cluster were tentatively annotated by interrogating the mass shifts from the putatively annotated node within a direct edge.[133,140]

The important *m/z* features that could not be annotated for MSI level 2 were considered for annotation in level 3. MSI level 3 annotation was performed with CEU Mass Mediator; a web-based metabolomics annotation software available from http://ceumass.eps.uspceu.es/.[134] The software performed experimental mass matching based on approximately 122 rules summarised into: ionization, adduct formation, relationships of the different adducts to the primary metabolite, and the retention times.[134] The software mathematical model scored the experimental masses' similarities to the reference molecules in the publicly available databases i.e., HMDB, KEGG, LipidMaps, MINE, KNApSAck and inhouse libraries.[134] The retention time and ionization matching were applicable to lipids and fatty acids while accurate mass and MS^1 composite spectra characteristics were used to annotate non-lipid metabolites. The similarity scores from the algorithm ranged from 0-2. A similarity score of 2 showed strong evidence to support the characteristic matches of the annotated metabolite, 1 meant not enough evidence to refute or accept the characteristics matched the annotated metabolite, and 0 suggested there was enough evidence to reject matching characteristics of the annotated metabolite. Further, the software provided the respective database unique identifiers of the putatively annotated metabolites.[134,141] The annotated metabolites were manually curated to ensure fully formed peaks were annotated and processed.

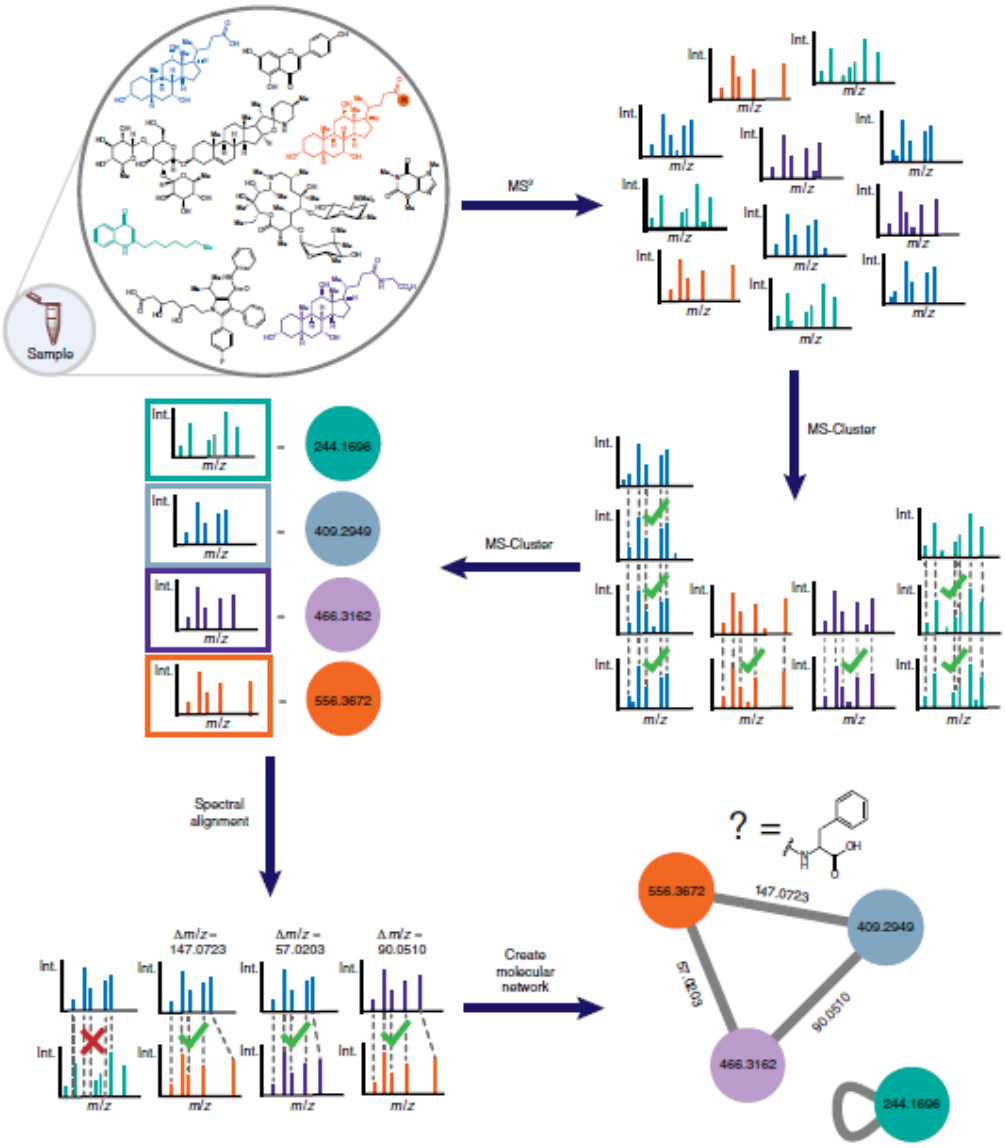


Figure 3.6. Molecular networking scheme. Starting from acquisition of the experimental spectrum to creating a molecular network. Adapted from Aron et al. Reproducible molecular networking of untargeted mass spectrometry data using GNPS. *Nat Protoc.* 2020;15(6):1954-1991.

To reduce chances of identification and annotation of spurious features, the annotated metabolites were manually curated to ensure they originated from an observed molecule. In MS-DIAL, the aligned features were evaluated to assess the accuracy of peak picking by examining extracted chromatogram. Peak picking was considered successful if there was evidence of co-elution of the extracted chromatogram in the samples. It was assumed metabolites with similar characteristics i.e., molecular weight and molecular formular tend to co-elute. The molecules could have different intensities in the different study groups but had similar peak shapes after peak picking and alignment. Use of peak shapes for curation had significantly been shown in other studies to reduce chances of false discovery in untargeted

metabolomics.[142] There were computational tools meant to filter out noisy and non-aligned features however it is difficult to design automated parameters adaptable to the variabilities of the data; I therefore resorted to manual evaluation of the peak shapes. In untargeted metabolomics there is need to come up with an intricate balance between statistical significance and biological relevance. It is important that studies do not process noisy features and that one does not remove features with biological relevance especially those with low levels.

3.2.11 Functional analysis

To determine the biological relevance of the identified biomarkers a combination of pathway enrichment analysis (over-representation) and pathway topology analysis was performed using MetaboAnalyst pathway analysis.[143] The over-representation test used a hypergeometric test to determine if a group of metabolites would be represented in each group of molecules by chance. According to metabolites set analysis (MSEA), a “set” comprises metabolites sharing a biological process i.e., those that change together in a disease condition or which are always found in the same tissue.[144] MSEA used a global test algorithm due to its flexibility in p-value estimation in small and large sample sizes calculated from Q-stats asymptotic distributions. The metabolites’ library was obtained from Kyoto Encyclopaedia of Genes and Genomes (KEGG) through computational mining or manual curation.[129,144] MetaboAnalyst MSEA is reported to have a library set of ~2,000 human metabolites from literature and public databases associated with pathways/diseases or tissues.[95] The library can be searched using unique IDs or common names of the query metabolites. The query data was formatted to have metabolites’ unique identifiers. The pathway analysis function used globaltest and GlobalANCOVA algorithms for pathway topology analysis using the degree of centrality or betweenness of functionally related metabolites.[143] The degree of centrality measures the number of connections that one metabolite has with other metabolites while betweenness centrality measures the shortest path going through the metabolites of interest. The output from the analysis showed a graphical representation based on the p-value and the pathway

impact value of specific metabolite sets. Important pathways were those with the lowest p-values and the highest impact.

To further elucidate the relevance of the identified metabolites, I conducted gene-metabolite interaction using Metabridge (https://www.metabridge.org/) and network analysis module in MetaboAnalyst.[95,145] A list of annotated metabolites was processed into corresponding HMDB or KEGG IDs. The list was uploaded into the Metabridge where the IDs were mapped to the corresponding human KEGG IDs, further the software identified all the reactions associated with the submitted set of metabolites. Further, the genes that coded for the enzymes corresponding to the mapped pathways were generated, tabulated, and downloaded.[145] The list with the metabolites, and human gene names was used to perform gene-metabolite interaction networking to visualize the interactions of the submitted metabolites and genes using the KEGG orthologs (KOs).[95] The metabolites and genes referenced for mapping were obtained from STICH ('search tool for interactions of chemicals') with high confidence interactions. The algorithm mapped the metabolites and genes (seeds) onto networks with direct neighbours; the networks were further built to create large subnetworks. The generated subnetworks could then be viewed using the visualization panel. Important nodes (metabolites and/or genes) were determined based on their position in the network. Nodes that were centrally located in a network were more important than the marginally positioned nodes. Important metabolites and genes were tabulated based on the global structure by ranking them based on their node degree and node betweenness which referred to the number of links a node had with other nodes and the number of shortest paths that went through a node, respectively. To derive functional insights from the gene-metabolite interactions, the important nodes were selected, and enrichment analysis conducted with the gene ontologies (GO) and the biological processes (BP) associated with the set of the selected nodes. Hypergeometric tests were used to determine the significance levels of the associated GO:BP queries.[95] The mapped GOs, metabolites and the associated biological processes were exported in tabular format with their degree, betweenness centrality, and p-values.

3.2.12 Diagnostic biomarkers analysis using AUC-ROC

I sought to test the capabilities of the significant metabolites to differentiate between RHD, HC, and degenerative AS. The potential diagnostic biomarkers were analysed using area under receiver operator characteristics (AUC-ROC) model.[146] The ROC had been used for diagnostic, prognostic, and predictive analysis of metabolites investigated for application in biomarker studies.[14,98,109,112,147] The ROC analysis was considered robust since its performance was not affected by the sample size, prevalence of the variables nor the normality of distributions. The ROC model used a set of biomarkers to accurately predict the sample groups as compared to the actual groups of the sampled population. The model performance was based on the sensitivity (proportion of actual positive cases classified as positive) and the specificity (actual negative cases classified as negative) of the diagnostic metabolites set obtained from AUC-ROC. The AUC was a single metric that represented the probability of the metabolic biomarker to precisely classify a randomly selected sample into the disease group than it would a healthy sample. The AUC range between 1 (perfect classification) to 0 (poor performance); 0.5 AUC indicate the classification is by chance.[146] The AUC-ROC was classified as either excellent (1-0.9), good (0.9-0.8), fair (0.8-0.7), poor (0.7-0.6) and fail (0.6-0.5).[146] From the ROC curve I also determined the cut-off points (optimal points) which was a balance between specificity and sensitivity of the diagnostic biomarker. To increase the confidence of the reported AUC-ROC, the associated 95% confidence interval calculated using 500 bootstrap replications was also reported. To increase the prediction capabilities of the diagnostic biomarkers a multivariate ROC analysis where a combination of the most promising markers was combined into a diagnostic panel with single AUC-ROC. The multivariate ROC models were created using linear support vector machine (SVM) algorithm models based on Monte Carlo sampling for cross validation and permutation for model's significance testing and nested cross validation (CV) to determine its generalizability.[129] To further estimate the robustness of the diagnostic biomarker, their significance levels were adjusted with some of the baseline clinical parameters using linear regression model with metadata table analysis

module in MetaboAnalyst.[129] The features were considered robust if their significance was generally not affected by the covariates after adjustment.

3.2.13 Spatial localisation of dysregulated metabolites on valve leaflets

The candidate metabolites were further investigated to determine their co-localisation with pathological features on the diseased valves using MALDI-MSI. Thin cryosections (5 µm) of frozen valve fragments were thaw-mounted on ITO and microscope slides for MALDI-MSI and H&E staining, respectively. The slides were dried in a desiccator for about 15 minutes. The mounted sections were protected from direct sunlight by covering the slides with aluminium foil. Diseased and non-diseased regions within the H&E-stained slides were highlighted as ROI by a pathologist. The mounted ITO slides were sprayed with 40 mg/ml 2,5-dihydrobenzoic acid (DHB) matrix that was dissolved in a 1:1 mixture of methanol: distilled water (v/v) with 0.1% formic acid. After thoroughly mixing the matrix by sonication, 12 layers of the matrix were sprayed on the ITO slides using the HTX TM-Sprayer (HTX Technologies, Chapel Hill, NC, USA); refer to **Appendix 8** for further details. After drying the sprayed ITO slides, the empty space between the two tissues was cleared of matrix with methanol; the space was used to load the calibration standards. To calibrate the MALDI TOF equipment for low molecular weight experiments, a mix of the matrices was used; DHB, HCCA (a-Cyano-4-hydroxycinnamic acid), and SA (sinapic acid). The slide images were obtained by scanning and processing the images with CyberViewX and the image uploaded to FlexImaging 5.1 software. The MALDI-MSI data was acquired using rapifleX® MALDI Tissuetyper® with flexControl 4.0 and flexImaging 5.1 (Bruker Scientific LLC, Billerica, MA, USA), in positive reflector mode over an *m/z* range of 50-1,200 Da at 50 µm raster width. The spectra were accumulated from 200 laser shots at 10kHz frequency. The statistical analysis and data visualisation was done with SCiLS^{TM} Lab 2020a. The m/z features were then co-localized with the histopathological features, and their detection levels compared between the ROIs. PCA multivariate analysis was used to determine features that were discriminately expressed between the compared ROIs. The discriminant features

were used to generate ion maps that showed the differential spatial changes of the ions at different ROIs. For further details on the MALDI-MSI protocol, refer to **Appendix 8**.

3.3 Ethics consideration and data dissemination

3.3.1 Ethics approval

The study was approved by the Faculty of Health Science, Human Research Ethics Committee, University of Cape Town (HREC REF:574/2018) as a study linked to (HREC REF 061/2018) (**Appendix 1**). A written and duly signed informed consent were obtained from all participants; thumb-printed consents duly signed by witnesses for the case of illiterate participants were also obtained. The study strictly adhered to the principles of Helsinki 2013 declaration.

3.3.2 Data dissemination

The clinical data of the participants were anonymised with a unique identifier, stored in a password-locked database and the source files stored in a secure locker. The samples were stored as per approved standard operating procedures and good clinical practice guidelines. A data transfer agreement was processed before any data could be transferred to a third party. Primary and secondary findings from the study were to be published in peer reviewed journals.

3.3.3 Data management plan (DMP)

A data management plan was developed using the UCT template and archived at UCT's DMP archive (https://dmp.lib.uct.ac.za/plans/1689) (**Appendix 9).**

4 CLINICAL AND HISTOPATHOLOGICAL CHARACTERISTICS OF RHEUMATIC HEART DISEASE AND AORTIC STENOSIS HEART VALVES[3]

4.1 Introduction

According to the GBD, in the SSA regions, there was minimal change in age-standardised RHD prevalence between 1990 and 2017 while the prevalence of non-RHD showed a general upwards trend.[2,148] SSA is among the regions with the highest rates of disability-adjusted life-years due to RHD which negatively affects the economic strengths of affected communities.[49] Furthermore, ARF affects about 8 to 51 per 100,000 individuals globally and mostly affects children aged between 5-15 years old.[149] There are cases of ARF among adults; however it is almost non-existent in persons older than 45 years.[149,150] As highlighted in chapter 2 of this thesis, there are histological features that are commonly associated with RHD and degenerative AS. To effectively analyse H&E slides, all forms of artefacts should be avoided. The common artefacts are poor orientation, calcification, autolysis, shrinkage, freeze artefacts, or cracking during sectioning.[151,152] Water carryover is another common source of artefact that causes pink haze on the slides and affects the cytoplasmic counter staining.[151–153] This study aimed to explore the demographics and clinical parameters of patients with RHD and aortic stenosis undergoing valve replacement and to describe the histopathological features of valve biopsies obtained during the valve replacement surgery.

[3]This chapter is based on the article:
Mutithu D. W., Roberts R., Manganyi R., Ntusi N. A. B. (2022). Chronic rheumatic heart disease with recrudescence of acute rheumatic fever on histology: a case report. *Eur Hear J - Case Reports*, 6(7), ytac278. doi:10.1093/ehjcr/ytac278

4.2 Materials and Methods

4.2.1 Clinical data and sample collection

The participants were recruited into the study as described in section 3.1.2 of this thesis. During recruitment, demographics data and vitals were collected from the participants. In addition, CMR analysis was performed to obtain detailed CMR cardiac performance parameters of structure, function, haemodynamics, and tissue characteristics. Further, the participants' medical files were carefully reviewed to obtain their clinical history i.e., CVD history and diagnosis, presence of comorbidities, medications, and echocardiography parameters. The data were collected and stored for the entire duration of the study in a password locked computer dedicated for the study. Valve biopsies were collected and processed as detailed in section 3.2.

4.2.2 Tissue fixing and embedding.

The tissues for histological examination were fixed in formalin immediately after excision. They were then processed in a tissue processor overnight (**Appendix 4**) with Leica tissue processing machine, Leica TP 1020 (Leica microsystems Nußloch GmbH, Nußloch, Germany), followed by embedding in paraffin wax the following day using the Leica EG1140H embedder and the Leica EG1140C chiller plate (Leica microsystems Nußloch GmbH, Nußloch, Germany) was used to cool and harden the wax tissue blocks (**Appendix 4**). The blocks were stored in appropriate conditions until required for sectioning and H&E staining.

4.2.3 Sectioning and H&E staining

Thin sections were obtained and processed as described in section 3.2.2, the H&E processing and histology methods. The slides were carefully stored in microscope slide boxes awaiting reviewing by a pathologist.

4.2.4 Histopathological assessment of FFPE valve biopsies

To assess the quality of the H&E staining the slides were first pre-viewed under the light microscope. The pre-viewing of the slides ensured the staining contrast was sufficient, no artefacts, and no tissue folding. A senior pathologist reviewed the H&E slides to investigate the morphological and histological features associated with RHD and degenerative AS. The histological reports were interpreted alongside with the clinical history, echocardiographic reports, and cardiovascular magnetic resonance imaging results. The histological reports were used to supplement classification of the VHD patients into those of probable rheumatic or non-rheumatic aetiologies.

4.2.5 Statistical analysis

Comparison of the clinical, echocardiographic, CMR and histological variables was performed using one-way analysis of variance (ANOVA) for parametric variables while Kruskal-Wallis Test was used for non-parametric tests. To compare the categorical variables, chi-square or Fisher exact tests were used to analyse the distribution of the variables among the groups. The statistical analysis was performed in R (version 4.2.0) and R studio (version 2022.07.1+554).

4.3 Results

4.3.1 Study recruitment and clinical data analyses

Out of the 100 participants included in the study, 32 participants were excluded in this chapter due to their age and comorbidities mismatch. The baseline characteristics of the enrolled participants is summarised in **Table 4.1**. A total of 68 participants were included into this chapter: RHD n=39; AS n=10; and age-, sex-, race-, and comorbidities-matched controls n=19. The valve samples analysed in this chapter were not obtained from healthy controls. Of the 39 participants diagnosed as having RHD, 32 had valve replacement while 9 participants diagnosed as having AS had the aortic valve replaced. RHD patients were generally younger (44.7 ±13.7 years) as compared to AS and controls who were 64.2 ±12.8 and 52.1 ±8.11 years

old, respectively (p<0.001). To ensure consistency in the downstream analyses, the participants were matched based on their comorbidities i.e., hypertension, diabetes, and body mass index (BMI). This study observed elevated mean systolic blood pressure (SBP) (143.0 ±23.1 mmHg) among AS patients as compared to RHD (121.0 ±24.2 mmHg) and controls (138.0 ±23.7 mmHg), p=0.01. There were however no differences in distribution of participants known to be hypertensive and diabetic among the participants included into the study (p>0.05). Furthermore, AS patients were observed to have elevated BMI 33.6 ±7.58 kg/m^2 as compared to RHD (27.4 ±6.07 kg/m^2) and controls 30.0 ±4.07 kg/m^2 (p=0.01). RHD patients had mildly abnormal global systolic cardiac function while controls and AS patients had normal systolic cardiac function; mean LVEF RHD 44.8 ±14.7%, AS 57.0 ±4.07%, and controls 56.1 ±12.1% (p=0.02). The AS and RHD patients had elevated left ventricular end-systolic volumes as compared to healthy controls (p=0.05). RHD and AS patients had dilated left atria (44.8 ±11.4 and 30.3 ±4.59 cm^2) compared to controls 24.2 ±5.57 cm^2 (p<0.001). Degenerative AS patients had greater left ventricular mass compared to RHD patients and controls (p=0.01). The valve lesions analyses and grading demonstrated that AS patients undergoing valve replacement all had severe aortic stenosis. In addition, among the AS patients some had mild and moderate mitral regurgitation (10% and 30%, respectively). Among the RHD patients that needed valve replacement, 64.10% had severe mitral stenosis and 35.90% had severe mitral regurgitation (p<0.001). Moreover, 43.59% RHD patients in the study had atrial fibrillation while none of the AS patients had atrial fibrillation.

Table 4.1. Baseline characteristics of the study participants

	RHD N=39	**Degenerative AS N=10**	**Controls N=19**	**P-value**
Age (years), mean(±σ)	44.7 ±13.7	64.2 ±12.8	52.1 ±8.11	<0.001[b,c]
Sex Female, n(%) Male, n(%)	 27(69.23) 12(30.77)	 6(60) 4(40)	 12(63.15) 7(36.84)	 0.82
Race Black, n(%) Mixed, n(%) White, n(%)	 11(28.20) 28(71.8) 0(0.00)	 1(10.00) 6(60.00) 3(30.00)	 4(21.05) 13(68.42) 2(10.51)	 0.03
BMI (kg/m^2), mean (±σ)	27.4 ±6.07	33.6 ±7.58	30.0 ±4.07	0.01[c]

SBP (mmHg), mean (±σ)	121.0 ±24.2	143.0 ±23.1	138.0 ±23.7	0.01[c]
LVEF %, mean (±σ)	44.8 ±14.7	57.0 ±4.07	56.1 ±12.1	0.02[c]
LVEDV (ml), mean (±σ)	174.0 ±49.1	199.0 ±61.2	151 ±26.7	0.25
LVESV (ml), mean (±σ)	93.9 ±27.5	94.4 ±34.5	65.2 ±13.3	0.05
LVMI (g/m^2), mean (±σ)	63.8 ±25.8	96.6 ±41.3	48.6 ±12.9	0.01[b, c]
LA Area (cm^2), mean (±σ)	44.8 ±11.4	30.3 ±4.59	24.2 ±5.57	<0.001[a, c]
Hypertensive, n(%)	14(35.90)	8(80.00)	7(36.84)	0.07
Diabetic, n(%)	3(7.69)	1(10.00)	1(5.26)	1
Smoker				
Current, n(%)	11(28.21)	2(20.00)	5(26.32)	0.65
Ex-Smoker, n(%)	1(2.56)	1(10.00)	0(0)	
Dyslipidemia, n(%)	9(23.08)	6(60.00)	4(21.05)	0.09
Histopathology				
Aschoff bodies, n(%)	2(5.13)	0	-	1
Calcification, n(%)	10(25.64)	8(80.00)	-	<0.01
Collagen deposition, n(%)	18(46.15)	4(40.00)	-	0.46
Fibrosis, n(%)	30(76.92)	9(90.00)	-	-
Inflammations, n(%)	13(33.33)	2(20.00)	-	0.44
Myxoid change, n(%)	15(38.46)	2(20.00)	-	0.25
Neovascularization, n(%)	24(61.54)	3(30.00)	-	0.01
Vegetations, n(%)	2(5.12)	0(0)	-	1
Fibrin deposition, n(%)	6(15.38)	3(30.00)	-	0.41
Valve lesions grading				
Aortic regurgitation				
Mild, n(%)	9(23.08)	2(20.00)	-	0.37
Moderate, n(%)	9(23.08)	1(10.00)	-	
Severe, n(%)	6(15.38)	0(0)	-	
Aortic stenosis				
Mild, n(%)	2(5.13)	0(0)	-	<0.001
Moderate, n(%)	2(5.13)	0(0)	-	
Severe, n(%)	6(15.38)	10(100)	-	
Mitral Regurgitation				
Mild, n(%)	8(20.51)	2(20.00)	-	<0.001
Moderate, n(%)	12(30.77)	1(10.00)	-	
Severe, n(%)	14(35.90)	0(0)	-	
Mitral Stenosis				
Mild, n(%)	1(2.56)	0(0)	-	<0.001
Moderate, n(%)	7(17.95)	0(0)	-	
Severe, n(%)	25(64.10)	0(0)	-	
Atrial fibrillation, n(%)	17(43.59)	0(0)	-	<0.01
Medications				
Penicillin, n(%)	17(43.59)	0	-	<0.01
Anti-hypertensive meds, n(%)	9(23.08)	5(50.00)	-	0.12
Diuretics, n(%)	24(61.54)	5(50)	-	0.44
Anti-diabetic meds, n(%)	3(7.69)	1(10.00)	-	1
Beta Blockers, n(%)	28(71.80)	4(40.00)	-	0.03
Anticoagulants, n(%)	22(56.41)	3(30.00)	-	0.13
Cardiac Glycosides, n(%)	3(7.69)	0(0)	-	1
Statins, n(%)	8(20.51)	5(50)	-	0.10
Medical intervention				
SOC, n(%)	7(17.95)	1(10.00)	-	-
Valve replacement, n(%)	32(82.05)	9(90.00)	-	

RHD, rheumatic heart disease; Degenerative AS, Degenerative aortic stenosis disease; SBP, systolic blood pressure; BMI, body mass index; LVEF, left ventricle ejection fraction; LVEDV, left ventricle end-diastolic volume; LVESV, left ventricle end-systolic volume; LVMI left ventricle myocardium mass index; LA, left atrium; SOC,

standard of care treatment. ANOVA Post hoc analysis with FDR adjusted P-value <0.05; a, RHD vs controls; b, AS vs controls; c, RHD vs AS.

4.3.2 Histopathological analyses of RHD and degenerative aortic stenosis

Valve fragments from 32 RHD and 9 AS patients were processed and analysed to determine the prominent histopathological features associated with the valvular heart disease and summarised in **Table 4.2**. On histology, AS was predominantly associated with features of calcification 8(80.00%) as compared to RHD 10(25.64%), p<0.01. In addition, it was observed that neovascularisation was mostly associated with RHD 24 (61.54%) as compared to AS 3(30.00%), p=0.01. Furthermore, presence of vegetations and Aschoff bodies was solely reported among valve fragments obtained from RHD patients. Fibrosis was a histopathological feature commonly found among patients with RHD 30(76.92%) and AS 9(90.00%) undergoing valve replacement surgery.

Table 4.2. Distribution of histopathological features observed in valve biopsies obtained from RHD and AS patients

	RHD N=32	**Degenerative AS N=9**	**P-Value**
Aschoff bodies, n(%)	2(5.13)	0(0)	1
Calcification, n(%)	10(25.64)	8(80.00)	<0.01
Collagen deposition, n(%)	18(46.15)	4(40.00)	0.46
Fibrosis, n(%)	30(76.92)	9(90.00)	1
Inflammations, n(%)	13(33.33)	2(20.00)	0.44
Myxoid change, n(%)	15(38.46)	2(20.00)	0.25
Neovascularization, n(%)	24(61.54)	3(30.00)	0.01
Vegetations, n(%)	2(5.12)	0(0)	1
Fibrin deposition, n(%)	6(15.38)	3(30.00)	0.41

The assessment of the gross morphology of the isolated valves showed valves that had thickened cusps and chordae tendineae due to severe fibrosis or calcification (**Figure 4.1**).

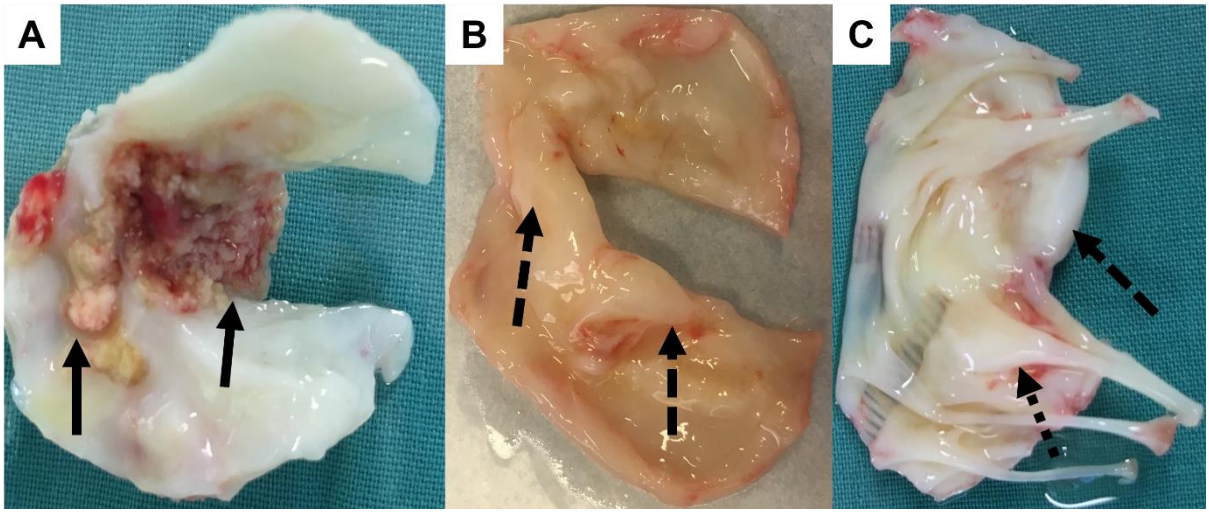


Figure 4.1. Macroscopic pathologies of the excised aortic and mitral valves. (A) Severe calcification on aortic valve, solid arrows show areas of severe dystrophic calcification. (B) Severely thickened aortic valve cusps due to fibrosis (dashed arrows). (C) Anterior view of thickened mitral valve and chordae tendineae (dotted arrow).

In addition, the study incidentally observed a case of a 59-year-old woman who was of sober habits and a 10-year history of poorly managed type 2 diabetes and hypertension. She had no history of acute rheumatic fever, prior myocarditis, or other known structural heart disease. She had presented with acutely decompensated heart failure. She was diagnosed with RHD complicated by severe aortic valve regurgitation and severe mitral valve regurgitation. She was put on heart failure medication and referred for double valve replacement surgery. She underwent open-heart surgery and had aortic valve and mitral valve replacement with bioprosthetic tissue valves. The histopathological assessment of the H&E sections from the patient's aortic and mitral valve biopsies showed features of acute-on-chronic and acute valvulitis. The mitral valve showed moderate-to-severe fibrosis of the valvular media, scattered stromal histiocytes. In addition, there was evidence of foci neovascularisation with characteristic thick-walled vessels (**Figure 4.2**). There was however no evidence of inflammation or infective endocarditis.

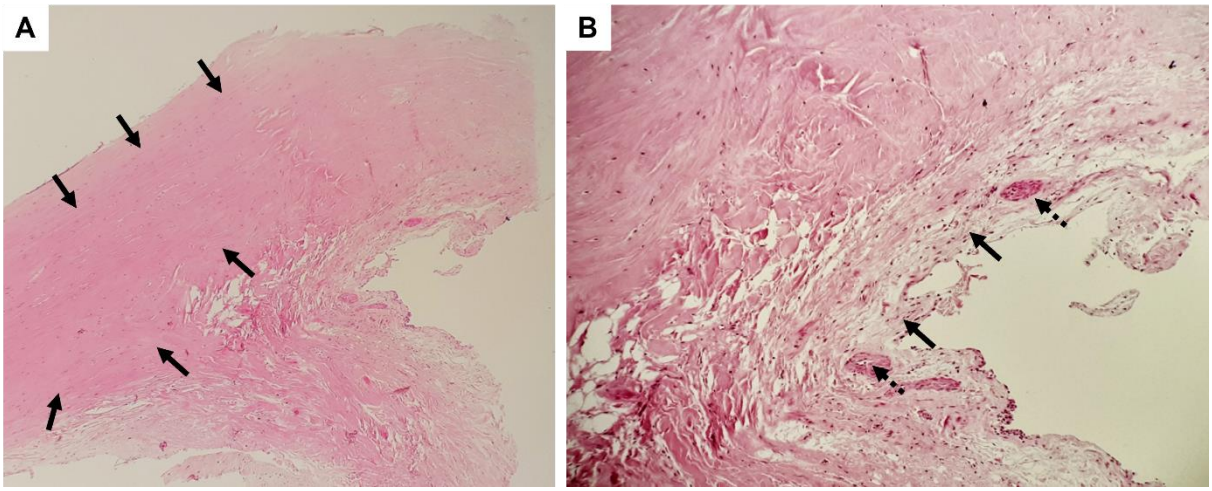


Figure 4.2. Haematoxylin and eosin staining of mitral valves. (A) severe fibrosis (solid arrows) (*magnification 40x*), (B) Scattered stromal histiocytes (solid arrows) and foci of neovascularization with thick-walled vessels (dotted arrows) (*magnification 200x*).

By contrast, the aortic valve showed features of acute on chronic valvulitis such as stromal neovascularisation, moderate-severe stromal fibrosis and presence of macrophages and plasma cells. In addition, there was evidence of fibrin deposition and localised lymphocytes. Additionally, there was discrete foci of acute rheumatic valvulitis with Aschoff bodies containing Anitschkow cells and central fibrinoid necrosis (**Figure 4.3**). The patient's treatment was adjusted to include antidiuretics, metformin, anticoagulants, aspirin, penicillin and other antibiotics, statins, and beta-blockers. The patient continued doing well and continued to attend her post-surgery check-ups.

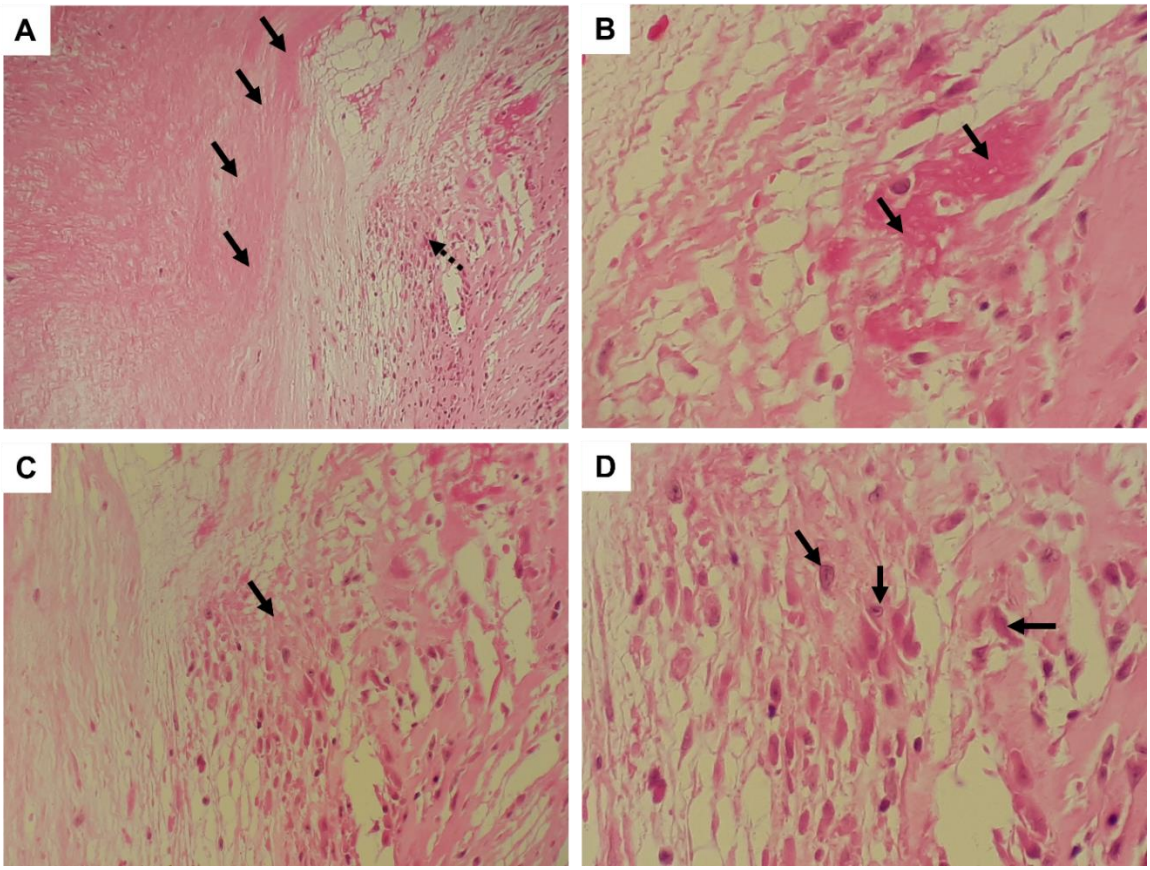


Figure 4.3. Haematoxylin and eosin staining of aortic valve. (A) Hyalinisation (solid arrows) and Aschoff bodies (dotted arrow) (*magnification 100x*), (B) fibrinoid necrosis (*magnification 400x*), (C) Aschoff body (magnification 200x), (D) Aschoff body with characteristic Anitschkow cells (*magnification 400x*).

4.4 Discussion

In this study the demographics and clinical parameters of the RHD and the aortic stenosis patients was described. Further the echocardiographic and CMR parameters that explained the cardiac performance of the recruited participants was highlighted. The histopathological parameters of the isolated valve biopsies highlighted the features associated with chronic RHD and aortic stenosis. Coincidentally, the study also observed a rare phenomenon of acute rheumatic valvulitis in a participant with chronic RHD valve.[154]

AS and RHD are mostly asymptomatic at the early stages of the pathology and often patients are diagnosed when the disease is advanced.[116,155,156] AS is prevalent in high-income countries and mostly affects patients with advanced age as opposed to RHD that is prevalent in LMICs and becomes symptomatic at a relatively young age.[155,157] In addition, RHD has been reported to be more highly prevalent among females than males.[116,158] Similarly, RHD patients recruited

into the study were generally younger than the aortic stenosis patients and were predominantly female. The number of AS participants having AVR and recruited into the study was disproportional to the number of RHD participants, reflecting local demography.

Chronic RHD and severe AS affect the left side of the heart. AS causes the narrowing of the aortic valve orifice causing pressure overload in the LV which is compensated by LV hypertrophy leading to increased LV mass.[8,159,160] In as much as LV hypertrophy causes reduced LV volume, the response is not universal; some patients have reduced ventricular volume due to thickened walls while others may have reduced volume due to reduced radius. AS may be associated with reduced diastolic volume, reduced ejection volume, impaired coronary blood-flow reserve and increased left atrial (LA) area.[8,159,160] RHD is almost predominantly a left side disease however there are rare cases of right-side valve disease. It mostly presents with mitral valve disease; most often mitral valve stenosis or regurgitation and aortic valve disease is less common.[116,161] In this study we observed RHD patients as having enlarged LA area more than the AS patients since the RHD patients were mostly complicated with mitral stenosis, mitral regurgitation, and aortic stenosis. In addition, the LV mass was increased more in AS than it was in RHD since the LV volume overload could have been compensated by LV hypertrophy. The compensatory mechanisms of responding to LV volume and pressure overload leads to both AS and RHD patients having normal LV ejection fractions.

The classical histological features associated with RHD are - presence of Aschoff bodies, neovascularisation, myxoid change, signs of inflammation, presence of vegetations, hypertrophy, and presence of fibrosis and collagen deposition.[162] Degenerative AS is expected to show fibrosis, collagen deposition, neovascularisation, calcification and infiltration of inflammatory cells.[163] In this study, features of chronicity were observed in valve biopsies from RHD and AS patients. Calcification is mostly reported in aortic stenosis.[163] In this study, calcification was predominantly found in degenerative AS patients however there were some cases of calcification among RHD –a finding that has been reported elsewhere.[162] Further, it was evident that neo-vascularisation was predominant but not exclusively among RHD since

there some cases that showed neo-vascularisation among degenerative AS patients.[162,163] The presence of neovascularisation is thought to promote calcification; in this study it could have contributed to the calcification cases among RHD.[163] As expected, fibrosis was a common histological feature among RHD and degenerative AS.[162,163] Since the chronic inflammation results from long-term injury to the tissue, it is often characterised by signs of immune response, phagocytosis, necrosis, and repair. Hence the length of chronic inflammation would be indicated by the amount of fibrosis. Further the kind of immune cells to be found depends on the kind of antigen. Agents producing cytokines are characterised with macrophages while cytotoxic lymphocytes inducing agents would show presence of lymphocytes.[164] The incidental finding of acute-on-chronic histological features showed presence of lymphocytes, macrophages and plasma cells; the finding of Aschoff bodies with characteristic Anichkov cells was the hallmark for acute inflammation.[48,162] In addition, unlike in many other reports, the features of acute rheumatic fever in this study were found on aortic valve as opposed to mitral valve.[48,162,163] The presence of acute valvulitis on aortic valve supports the fact that rheumatic fever and rheumatic heart disease lesions does also affect the aortic valves.[3,48,162,165] Cases of ARF are unlikely and almost non-existent in patients who are older than 45 years, therefore finding a 59-year-old-woman with acute valvulitis was unexpected.[154] The findings therefore suggested that there could be cases of ARF that goes unnoticed, especially in RHD endemic regions.

4.5 Limitations

The study reported on the histopathological features of RHD and degenerative AS. However, the patients included in the study were at the advanced stages of VHD and therefore the features of pathology were not fully representative. Furthermore, due to the limited number of samples the study could not achieve the intended statistical power.

4.6 Conclusion

RHD is endemic in SSA. Many patients require valve replacement surgery at a fairly young age. On the other hand, AS prevalence is on the rise and the patients also require valve replacement surgery. RHD and AS have different aetiologies and risk factors and therefore have different histopathological features. In this chapter, we describe the different imaging, macroscopic and histological findings in RHD and AS. We also report on a rare finding of chronic RHD admixed with acute rheumatic valvulitis. The study has therefore shown that in high-risk areas, consideration of ARF in the differential diagnosis of VHD should not be ruled out despite not meeting the criteria for definite ARF diagnosis as per revised Jones criteria 2015.

4.7 Contributions and acknowledgement

The valve samples were obtained from patients having valve replacement surgery at the Christiaan Barnard Division of Cardiothoracic Surgery, Groote Schuur Hospital, Cape Town after recruitment and consenting by D Mutithu assisted by Dr R Manganyi, O Aremu, and Dr E Lumngwena. The clinical data was collected and curated by D Mutithu. The valve fragments were provided by Dr R Manganyi after excision during the surgery. The samples for histology were processed by D Mutithu assisted by S Govender. The H&E slides were reviewed by Prof D Govender. Data analysis was done by D Mutithu.

5 CIRCULATORY METABOLIC BIOMARKERS IN CHRONIC DEGENERATIVE AORTIC STENOSIS AND RHEUMATIC HEART DISEASE[4]

5.1 Introduction

VHD are commonly classified based on their aetiology. In this study we focused on understanding the pathogenesis of chronic degenerative AS and RHD. The global burden of RHD and non-RHD VHD stands at 17.16/100,000 and 401.69/100,000 for ages between 5 to 70 years old persons, respectively.[2,148] Further, degenerative AS is highly prevalent in HICs countries but also the prevalence is rising in LMICs where it mostly presents with calcific AS.[2] RHD is among the leading causes of cardiovascular related mortalities and morbidities in LMICSs.[2,49] SSA is among those regions with the highest rates of disability-adjusted life-years due to RHD, which negatively affects the social economic standings of affected communities.[49]

Metabolomics has found application in biomarker discovery in CVD to identify target molecules for therapies and markers of early diagnosis and outcome prediction.[56,64,166] LC-MS with soft ionisation e.g. ESI and TOF mass analyser is the most common metabolomics technique used in discovery studies owing to its wide coverage of the molecular mass range.[79,80] Advancements in analytical technology have led to development of bioinformatics tools that aid high throughput processing, statistical analysis, and identification of potential metabolic biomarkers.[86,88,93–96] Metabolomics biomarkers associated with pathogenesis and diagnosis of VHD have been reported.[18,97–103,167,168] Little to no studies have reported metabolomics biomarkers associated with pathogenesis or diagnosis of RHD.[19]

[4]This chapter is based on a manuscript:
Mutithu, D. W., Kirwan, J. A., Adeola, H. A., Aremu, O. O., Lumngwena, E. N., Familusi, M., ... & Ntusi, N. A. (2023). *Serum metabolic profiling in rheumatic heart disease and degenerative aortic stenosis* [Manuscript submitted for publication]. Department of Medicine, University of Cape Town, Cape Town, South Africa.

In this study, patients were classified into chronic degenerative AS or RHD based on their clinical diagnosis supported by echocardiography and histological assessment of the valve fragments. This study explored serum metabolomics biomarkers associated with RHD and degenerative AS and contrasted it with healthy individuals. The study aimed to identify serum metabolic biomarkers and pathways associated with RHD and degenerative AS using LC-MS metabolomics. In addition, the study also aimed at characterising metabolic biomarkers with diagnostic capabilities to differentiate RHD, degenerative AS, from healthy individuals. Finally, logistic regression analysis was employed to assess covariation between circulatory biomarkers and predefined clinical parameters.

5.2 Materials and Methods

5.2.1 Sample collection and processing

Study participants were recruited as described in section 3.1, the study design and recruitment. Blood samples was collected from consenting RHD and degenerative AS as described in Chapter 3.2.3. In addition, blood samples were obtained from age-, sex-, ethnicity-, and comorbidity-matched participants. Blood samples were processed as detailed in section 3.2.3, sample processing methods. Briefly, serum was isolated by centrifugation at 2000 x g for 10 minutes in a cooled centrifuge to 4°C. Aliquots of isolated serum were immediately stored in -80°C freezer until analysis. Sample processing was performed in level 2 biosafety (BSL-2) cabinet. Patients were included or excluded from the study as described in the inclusion and exclusion criteria described in section 3.1.3, the study population sampling.

5.2.2 Methods development pilot metabolome extraction and analysis

As explained in section 3.2.4, the pilot experiments methods; some samples were processed to pilot the suitability of metabolites extraction solvents. Briefly, different solvent combinations were tested i.e., 100% methanol, 100% acetonitrile, a mixture of 100% methanol and 100% acetonitrile (1:1v/v), and a mixture of 100% methanol and 100% ethanol. The serum samples

were processed as detailed in section 3.2.4 and reconstituted with HPLC grade water with 0.1% formic acid. The samples were then separated using LC using Omega Polar C18 column with particle size 1.6 µm, pore size 100Å, length 10 cm, and internal diameter 3 mm (Phenomenex Inc., Torrance, CA, USA) with the mobile phase A being HPLC-grade water with 0.1% formic acid and 2mM ammonium formate and the mobile phase B being a mixture of 100% acetonitrile and 100% methanol (1:1 v/v) with 2mM ammonium formate and 0.1% formic acid. Two gradient settings were tested to determine the one suitable for metabolome extracted as detailed in **Table 5.1**. The maximum pressure limits were set at 400 bar and the injection volume was set at 10 uL. To determine the system suitability, Sciex Calibration standards were used to calibrate the equipment (Sciex, Framingham, MA, USA). To set up the suitable flow rate and data acquisition rate, triazine pesticide was used in place of the samples (Sigma-Aldrich). The mass spectrometry data was obtained in the positive ionization using full scan mode and scanning masses between 50 – 1,200 Da using X500R QTOF-MS/MS equipment operated with Sciex OS software ver.1.4 (Sciex, Framingham, MA, USA). The raw data was pre-processed and converted to .mzML from .wiff2 using MS-Convert in ProteoWizard (version 3.0.1908). Furthermore, features extraction, annotation, and pathway analysis were performed in XCMSOnline (version 2.7.2), XCMS (version 1.47.3), and CAMERA (version 1.34.0), respectively (**Appendix 6**). The pilot statistical analyses were performed in MetaboAnalyst 5.0 (https://www.metaboanalyst.ca/).

Table 5.1. The pilot liquid chromatography gradient and flow program

Time (min)[1]	B Conc (%)[1]	Flow (mL/min)[1]	Time (min)[2]	B Conc (%)[2]	Flow (mL/min)[2]
1.00	0.0	0.7	2.00	2.0	0.4
60	100.0	0.7	45.00	95.0	0.4
70	100.0	0.7	47.00	95.0	0.4
70.10	0.0	0.7	47.10	2.0	0.4
90.00	0.0	0.7	60.00	2.0	0.4

5.2.3 Study metabolome extraction and analysis

Study population

VHD patients were recruited as described in section 3.1, the study design. Patients were further classified into RHD and degenerative AS groups as described in section 3.1.3, the study population sampling. In all cases, histopathological assessment of the excised aortic valve fragments was done. Non-fasting venous blood samples were collected from VHD patients (n=39: 20 RHD, 10 degenerative AS) and from age-, sex-, ethnicity-, and comorbidity-matched individuals without VHD as controls (n=19) as described in section 3.1.3. The blood samples were processed to isolate serum as stated in section 3.2.3, sample processing methods.

Metabolites extraction

For more details on the study of metabolites extraction methods, refer to section 3.2.5. Briefly, sera from RHD, AS and healthy participants were processed for analysis in batches. Each batch was composed of 15 samples/injections to avoid long waiting times before samples were injected. Each batch was prepared by randomly selecting 3 RHD, 3 AS, 3 healthy controls, 2 pooled quality controls (QCs), 2 external standards, and 2 blanks. The QC samples were prepared by pooling 50 µL of serum from each sample into 1 aliquot which was divided into 4 aliquots of 1.5 mL to avoid multiple freeze-thaw cycles. In addition, a long term reference standard (LTRS) was prepared by reconstituting the standard (NIST SRM 1950 metabolites in human plasma standards) (Sigma-Aldrich, St. Louis, MO, United States) with an equal proportion of HPLC-grade water (1:1 v/v) and was divided into 4 aliquots of 1.5 mL each as was reported elsewhere.[169] The prepared QCs and LTRS were stored in -80°C freezer until analysis.

To extract the metabolites, 100 µL of the samples, QCs and LTRS were added to 400 µL of ice-cold extraction solvent. The extraction solvent was made by mixing 100% acetonitrile with 100% methanol (9:1v/v). The mixture of the samples and the extraction solvent was vortexed

(Labnet International, Edison, NJ, USA) and centrifuged at 14,000 x g for 15 minutes with Sigma 1-14K centrifuge (Sigma Laborzentrifugen GmbH, An der Unteren Söse, Osterode am Harz, Germany) precooled to 5°C. The supernatant was dried using a nitrogen gas concentrator in the glass tubes (Globe Scientific Inc, Mahwah, NJ, USA). It was then reconstituted in 150 µL reconstitution solvent made of HPLC-grade water mixed with 100% acetonitrile (98:2 v/v) and modified with 1% formic acid or 1 mM acetic acid in the positive or negative ionization modes, respectively. The reconstituted extracts were then centrifuged for 5 minutes at 13,000 x g to remove undissolved particles. The reconstituted samples were then transferred into the low-recovery-volume flat bottom inserts placed in 2 mL vials ready for loading into the ExionLC™ AD Series HPLC system that was set at 15°C (Sciex, Framingham, MA, USA).

LC-MS/MS data acquisition

The mobile phase (A) was 100% HPLC-grade water while mobile phase (B) was a mixture of 100% acetonitrile and 100% methanol at 1:1 (v/v). The mobile phases were modified with 2 mM ammonium formate and 0.1% formic acid (Sigma-Aldrich) or with 2 mM acetic acid for the positive or the negative modes, respectively. The 60 mins chromatography gradient tried in pilot experiments was used as indicated in **Table 5.2**.

Table 5.2. The study liquid chromatography gradient and flow program

Time (min)	B Conc (%)	Flow (mL/min)
2.00	2.0	0.4
45.00	95.0	0.4
47.00	95.0	0.4
47.10	2.0	0.4
60.00	2.0	0.4

High resolution MS/MS data was acquired in positive and negative ionisation modes as detailed in section 3.2.7, the MS data acquisition methods. The data was obtained in proprietary .wiff and .wiff2 formats and was stored in secure devices until processing and analysis.

Data processing and statistical analysis

To extract, align and adjust batch drift of the *m/z* features, the raw positive and negative mode LC-MS/MS data was processed with MS-Covert in ProteoWizard 3.0.1908.[170] The *m/z* features were obtained using the data processing pipeline composed of peak picking, baseline correction, peak alignment, filtration, and QC checks that was performed in MS-DIAL 4.38 (**Appendix 7**).[96] Batch drift adjustment was done with MetaboAnalyst batch drift correction algorithm based on QC samples that were injected every 5 samples in all batches. Spurious *m/z* features were filtered out by removing background noise based on the blank samples (maximum sample/blank<5-fold change). In addition, *m/z* features were further filtered by exclusion based on missing values in the experimental samples; features were excluded if they had more than 25% or 45% missing values in at least one group in ESI+ and ESI- modes, respectively. To obtain the significant features, data were first normalised based on their intensities between samples (row-wise normalisation) based on the sum intensities in each sample, and further normalised by adjusting the intensities of each feature (column-wise normalisation) by mean centring and dividing by the square root of standard deviation of each feature using MetaboAnalyst 5.0 web tool (https://www.metaboanalyst.ca/home.xhtml).[171] To obtain potentially important features between groups, univariate analysis of variance (ANOVA), student's t test, and volcano plot analysis was used to obtain potentially important features between groups. In addition, multivariate PCA and PLS-DA analysis was used to visualise group differences. Chemical identification of the selected features was done by matching their MS and MS/MS data to publicly available spectral libraries (HMDB, ChEBI, MassBank, etc.) with GNPS 28.2[137] and SIRIUS 4.5.3[172]. Important features without MS/MS data were matched to public databases based on their accurate mass, retention time, and adduct formation patterns using CEU Mass Mediator available at (http://ceumass.eps.uspceu.es/).[141] Important features with and without MS/MS data were annotated into MSI levels 2 and 3, respectively.[136,173] The annotated metabolites were further screened using univariate statistics tests. To describe the global metabolic patterns between

groups, multivariate chemometrics tools such as hierarchical clustering, PCA and PLS-DA were used. The *m/z* features and metabolites with a P-value <0.05 and log2 fold change (FC)>1 or <1 was considered significantly changed.

To assess the reliability of the important *m/z* features, manual curation of the extracted ion chromatograms was done in MS-DIAL by inspecting the peak shapes. The curation is time consuming therefore this was applied to the statistically important features before pathway enrichment analysis. During manual curation of the annotated compounds, the peak shape, intensity and the retention time matching with the library match and the standard molecules were checked. In addition, the fragments matching to the MS2 library to confirm the identification were also checked. Isobaric elution was also considered, i.e. molecules eluting at different times and used the MS DIAL compound search to determine the libraries match to the identification.

5.2.4 Functional analysis of significant metabolites

To explore the functional relevance of the annotated metabolites, the pathway analysis algorithm in MetaboAnalyst 5.0 (https://www.metaboanalyst.ca/) was used for pathway enrichment and topology analysis. The HMDB and KEGG identifiers of the annotated metabolites were uploaded and embedded in human pathway library for pathway analysis. The mapped pathways were visualised with scatter plots while testing the significance level, the enrichment was analysed using the hypergeometric test, while the topology analysis was done using relative-betweenness centrality on the *Homo sapiens* KEGG pathway library.[171] Significantly mapped pathways (p≤0.05) were further explored to determine the profile of the mapped metabolites using univariate analysis. To describe the genes and biological processes associated with the annotated metabolites between groups, gene-metabolite interaction analyses were performed using MetaBridge (https://www.metabridge.org/) and MetaboAnalyst 5.0 (https://www.metaboanalyst.ca/).[145,171] The HMDB identifiers of the annotated metabolites were uploaded and mapped to the KEGG identifiers where genes associated with the mapped metabolites are obtained.[145] The interaction of the genes and the

metabolites was analysed using network exploration algorithm. Integrative network analysis and visualisation was conducted using network analysis function in MetaboAnalyst, important genes and metabolites were selected based on their degree centrality (degree>2) and betweenness; important biological processes were selected based on their significance levels of being associated with the queried genes and metabolites (P-value<0.05).

5.2.5 Specificity and sensitivity analyses of dysregulated metabolites

To determine the specificity and sensitivity of the dysregulated metabolites to differentiate between RHD, AS and healthy controls, receiver operator curve model (ROC) analyses was used in MetaboAnalyst 5.0 (https://www.metaboanalyst.ca/).[171] A multivariate ROC model was constructed using Monte Carlo cross validation (MCCV) where important metabolites were used to build a classification model using two thirds of the samples, which was then validated on the remaining third of samples. Metabolites with potential predictive capabilities were those with an area under the curve (AUC) ≥0.8, a sensitivity and specificity ≥ 0.70 and 70% frequency of being selected when building the cross-validation model. To determine the robustness of the discriminant biomarkers, the effect of clinical metadata (covariates) on their significance level was assessed by analysing the correlation of the p-values with and without covariate adjustments using multivariate linear regression models.[171]

5.3 Results

5.3.1 Methods development pilot experiments outcomes

Liquid chromatography flow rate settings

To determine the suitable chromatogram method, the pesticides standard was used to explore the chromatographic settings. Based on the abilities to resolve peaks of the late-eluting compounds, the 60 minutes gradient with 0.4 mL/min had the better resolution as compared to 90 minutes gradient with 0.7 mL/min (**Figure 5.1 & Figure 5.2**).

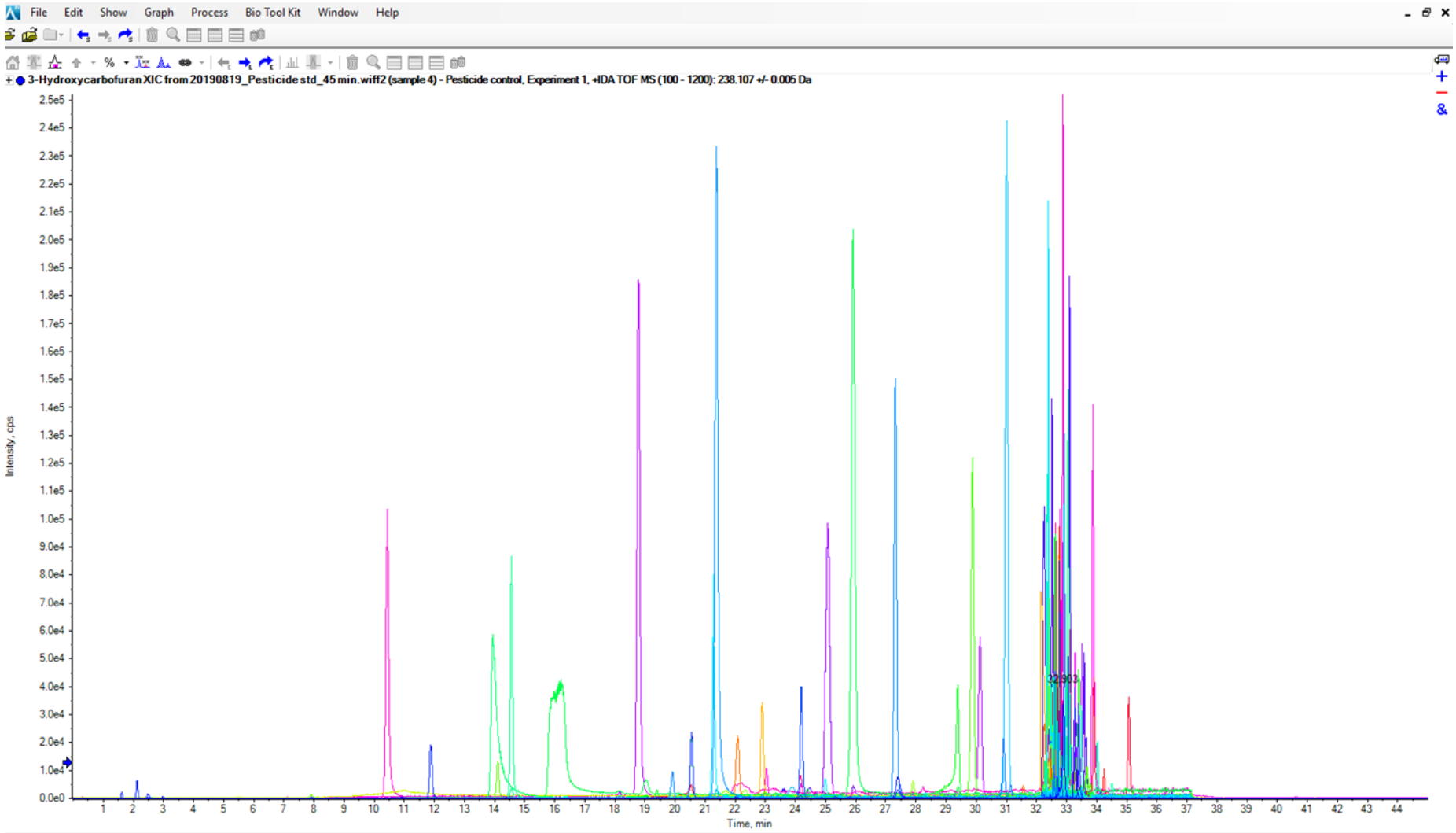


Figure 5.1. Extracted ion chromatogram of 3-hydroxycarbofuran (238.105 Da) from a 90-minute gradient.

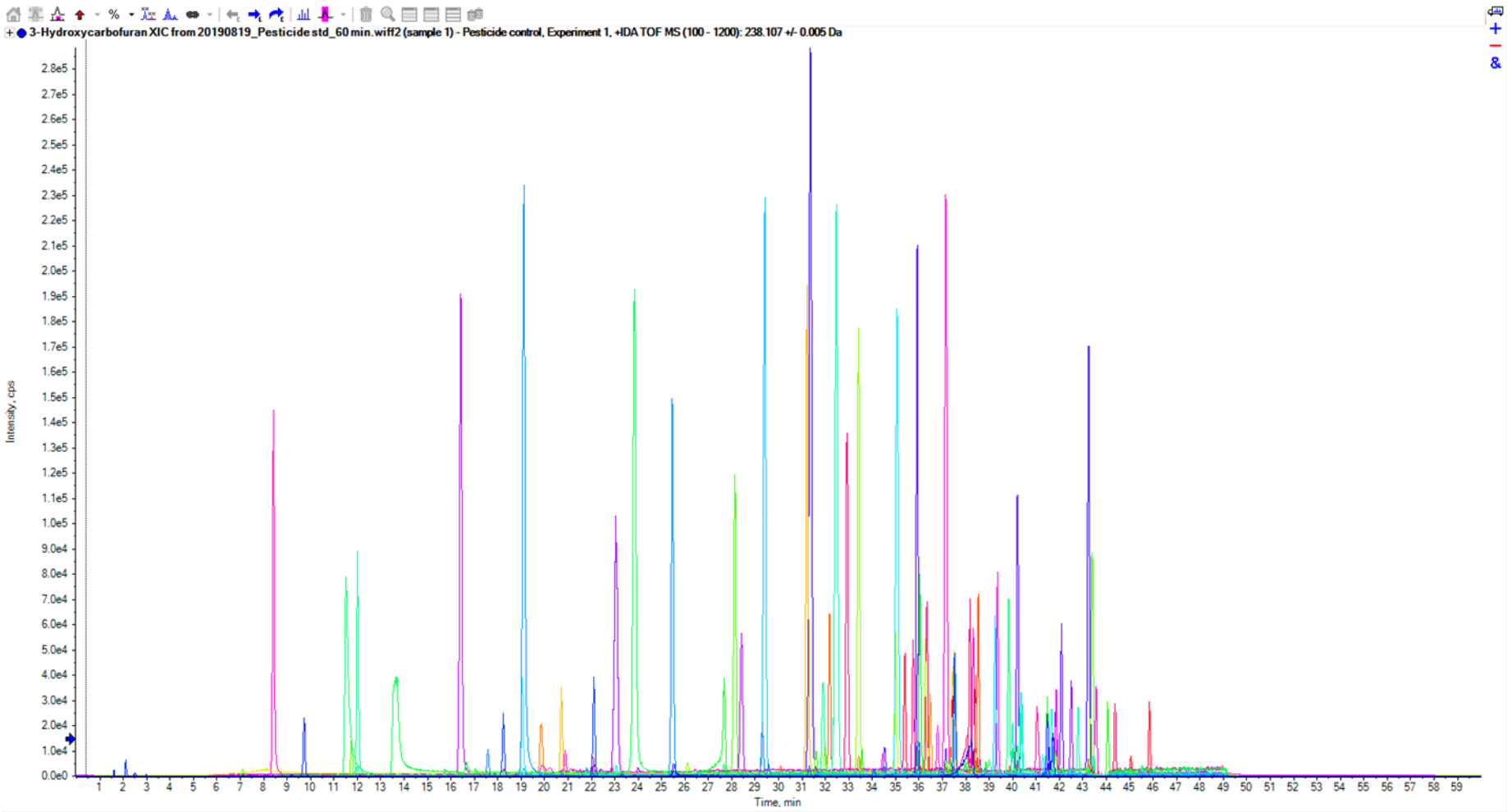


Figure 5.2. Extracted ion chromatogram of 3-hydroxycarbofuran (238.105 Da) from a 60-minute gradient.

Choosing the suitable extraction solvent

To determine the most effective extraction solvent to extract the polar and the non-polar part of the metabolome, the extracted masses from the solvent combinations were logically analysed with Venn diagram to determine the relationship of the extracted features. The findings showed that the combination of acetonitrile and methanol had the most unique features detected (**Figure 5.3 & Figure 5.4**). The pilot studies had shown that ACN had better protein precipitation capabilities for metabolome extraction. Previously reported studies have shown that adding methanol into the crushing solvent increased the extractability of the polar and nonpolar metabolites based on the relative standard deviation of the extracted features[174] Furthermore, to reduce protein binding of the metabolites, 1% FA is added as a modifier; however it is known to decrease solubility of metabolites that are more hydrophobic.[175] The mixture of acetonitrile and methanol were therefore adopted as the crushing solvent for the entire study; the solvent was mixed with formic acid or acetic acid for efficient metabolome extraction and protein precipitation. The data was acquired in 12 batches in positive and negative modes; in each batch there were 15 samples/injections. Each batch was prepared by randomly selecting 3 RHD, 3 AS, 3 healthy controls, 2 pooled quality controls (QCs), 2 external standards, and 2 blanks.

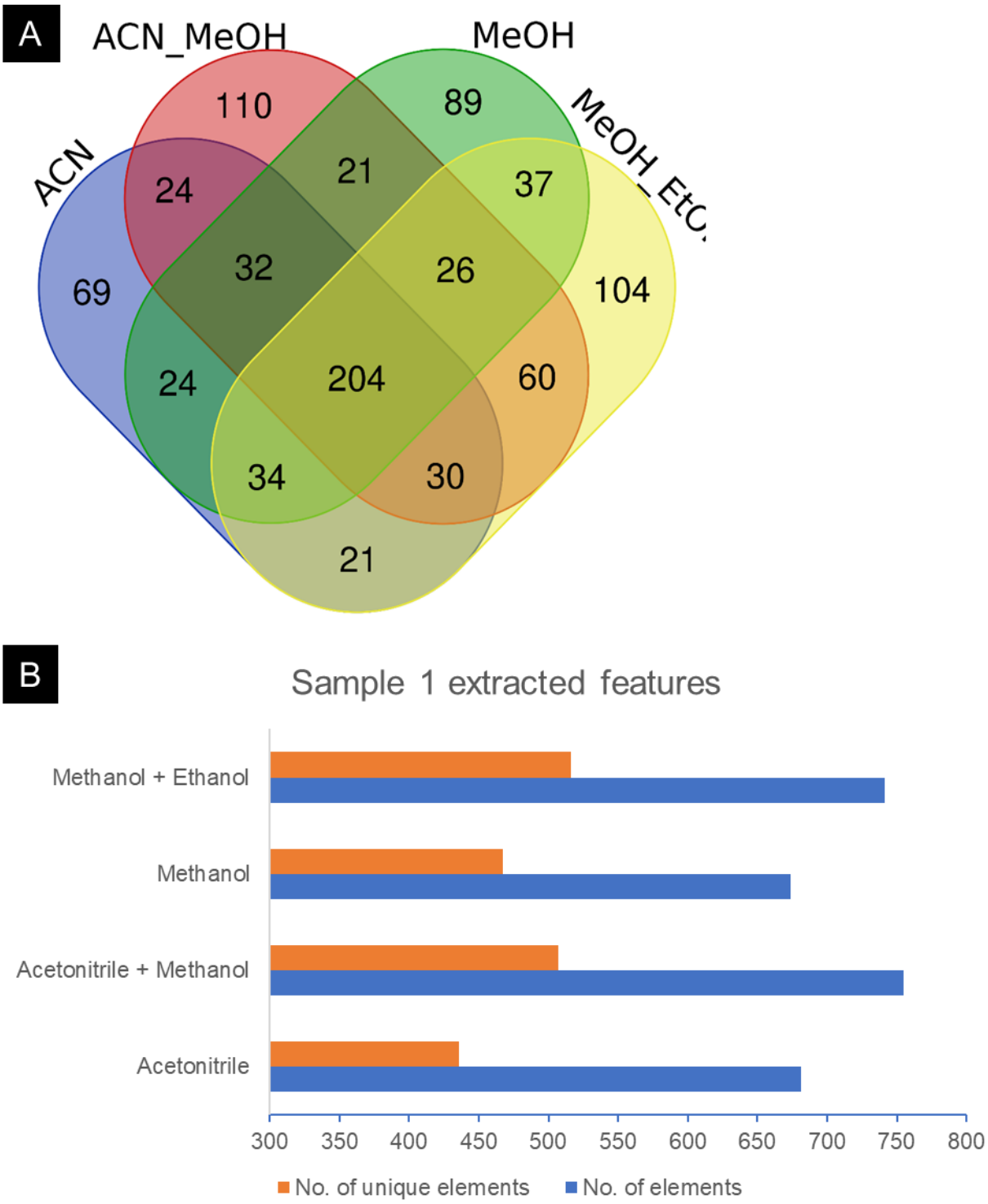


Figure 5.3. Exploration of extracted features from the 2 samples using 4 solvent combinations. The Venn diagrams showing the logical relations of the shared and unique features between different extraction methods in sample 1 (A) and (B). It also shows the distribution of the detected and unique features for each of the extraction solvent combination in Sample 1.

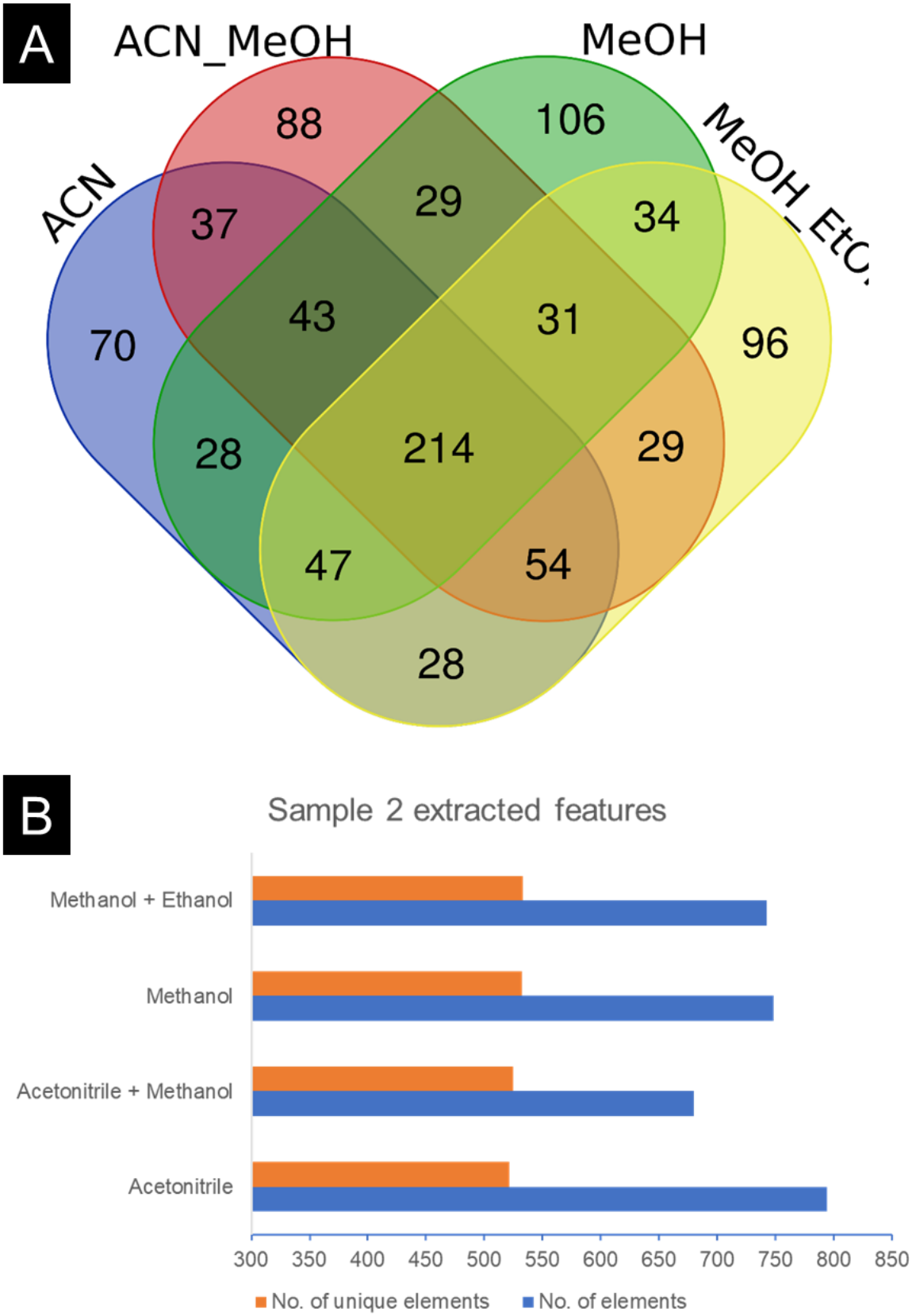


Figure 5.4. Exploration of extracted features from the 2 samples using 4 solvent combinations. The Venn diagrams showing the logical relations of the shared and unique features between different extraction methods in Sample 2 (A). It also shows the distribution of the detected and unique features for each of the extraction solvent combination in Sample 2 (B).

Choosing the peak picking and features extraction methods

To determine the suitable peak picking and features extraction algorithms and methods, the XCMS-online and the MS-DIAL software were used. The XCMS-online was web-based therefore it was available only if I had internet access. The software also required minimal technical knowledge to load the raw files, in put the parameters and download the results. Since the software is online based the storage capacity was limited and in case there was need to process bigger file sizes there could have been problems with storage. Furthermore, there was limited functionality to investigate the quality of the extracted features. MS-DIAL was downloaded and operated within a local computer; therefore, it would be readily available without requirement for internet access. MS-DIAL offered flexible functionality in that I could inspect the quality of the extracted features. MS-DIAL offered several noise filtration parameters therefore I had an option to tweak several parameters to ensure I obtained quality extracted features. The challenge with MS-DIAL was when processing huge data files approximately 60 gigabytes of data, it required a computer with a physical memory capacity of at least 16 gigabytes.

5.3.2 Study experiments biomarker discovery and exploratory analyses

After the pilot experiments the methods and parameters to use in the study were compiled. The methods and parameters comprised the entire metabolomics pipeline as further described in **Appendix 5** and **7**. The results from the study experiments are described below.

Peak detection, extraction, and feature repeatability analysis

Initially there were 41,738 and 31,772 features detected in the ESI+ and ESI- modes, respectively. After peak picking, alignment and initial filtering, 3,692 (ESI+) and 2,864 (ESI-) features (*m/z*, retention time, and intensity) were extracted. The median relative standard deviation (mRSD) of the pooled quality control samples (pooled QC) for negative mode and positive modes was 29.99% and 35,85% respectively. The RSD cut-off point for acceptable

repeatability of the extracted features was <30% in pooled QC RSD (**Figure 5.5A-B**). Following batch drift analysis and correction based on the pooled QCs, the PCA plots of before and after batch adjustment did not have a significant difference for positive and negative ionisation modes (**Figure 5.6 & Figure 5.9**). In addition, the singular value decomposition analysis did not show an obvious systematic trend that could be attributed to bias (**Figure 5.7 & Figure 5.10**). Furthermore, the distance between original data and normalised sample data did not show significant reduction in distance (**Figure 5.8 & Figure 5.11**). After batch normalisation and filtration based on the quality control samples RSD cut-off (≤30%), 1,847 (ESI+) and 1,064 (ESI-) features were extracted and used for annotation and statistical analysis. The *m/z* features perturbation patterns in the ESI+ and ESI- modes between AS, RHD and controls were explored using univariate volcano plot analysis based on the fold change and significance level. The analysis showed *m/z* features changed differently between the groups in both the positive and negative ionisation modes. Therefore, univariate analysis was used to obtain the significant features after annotation. The m/z features were annotated using the GNPS molecular networking technique.

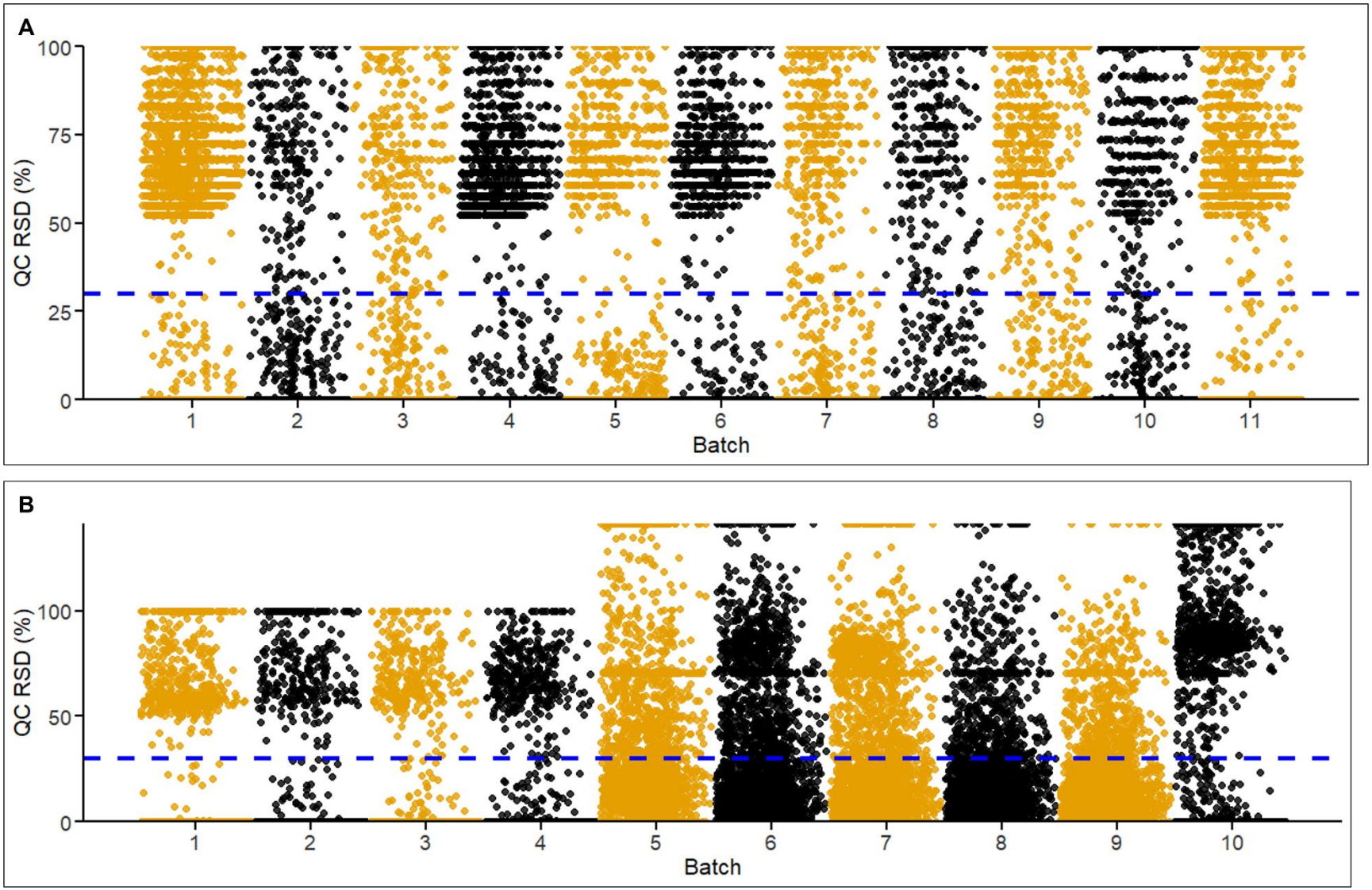


Figure 5.5. The pooled quality control relative standard deviations distribution per batch for (A) positive ionization mode and (B) negative ionization mode. The dotted line shows the <30% RSD cut-off points for features filtration with acceptable repeatability.

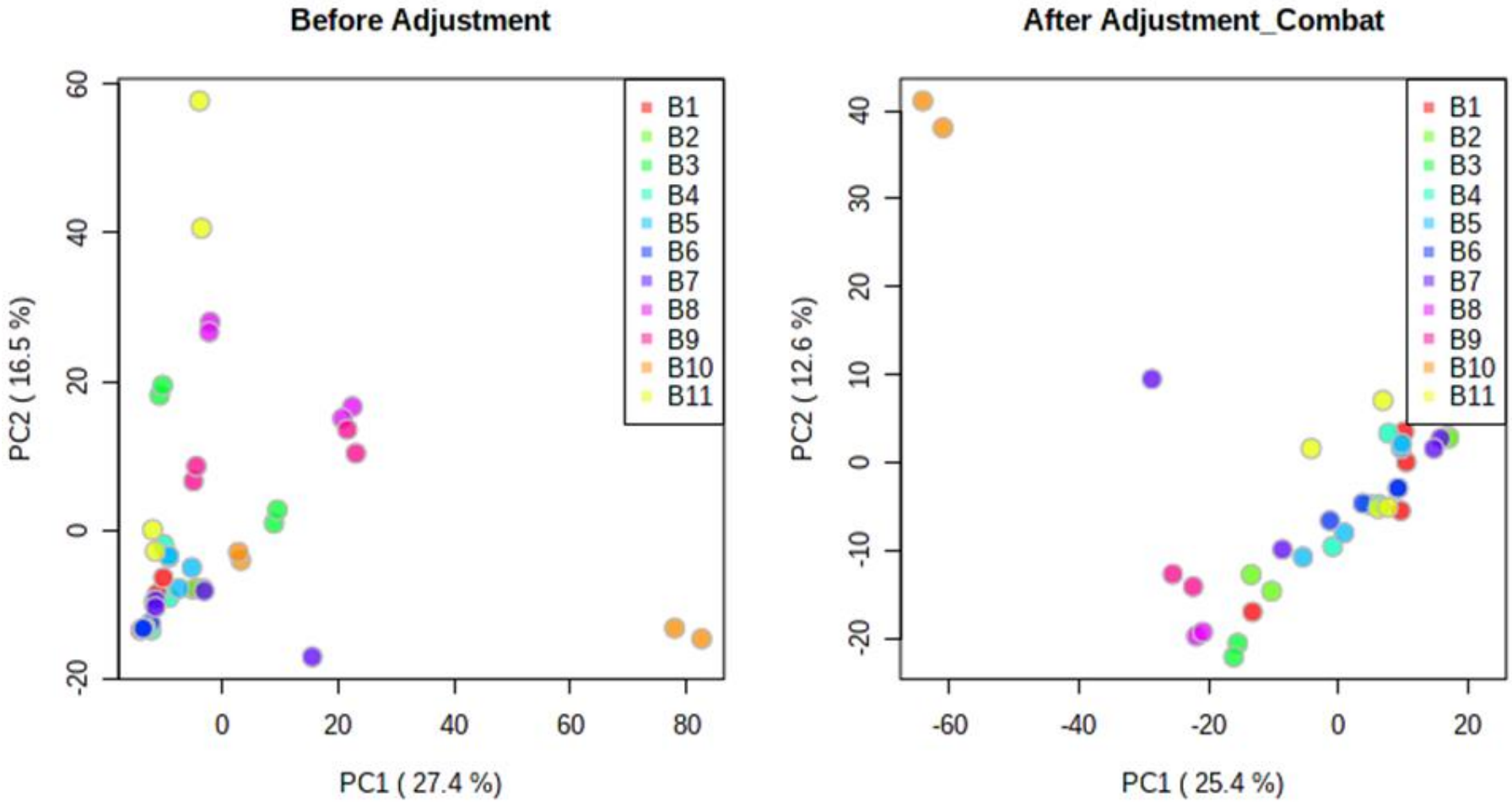


Figure 5.6. Batch drift analysis and adjustment in the positive ionisation mode (ESI+). PCA plot showing sample distribution before and after batch correction using ComBat method.

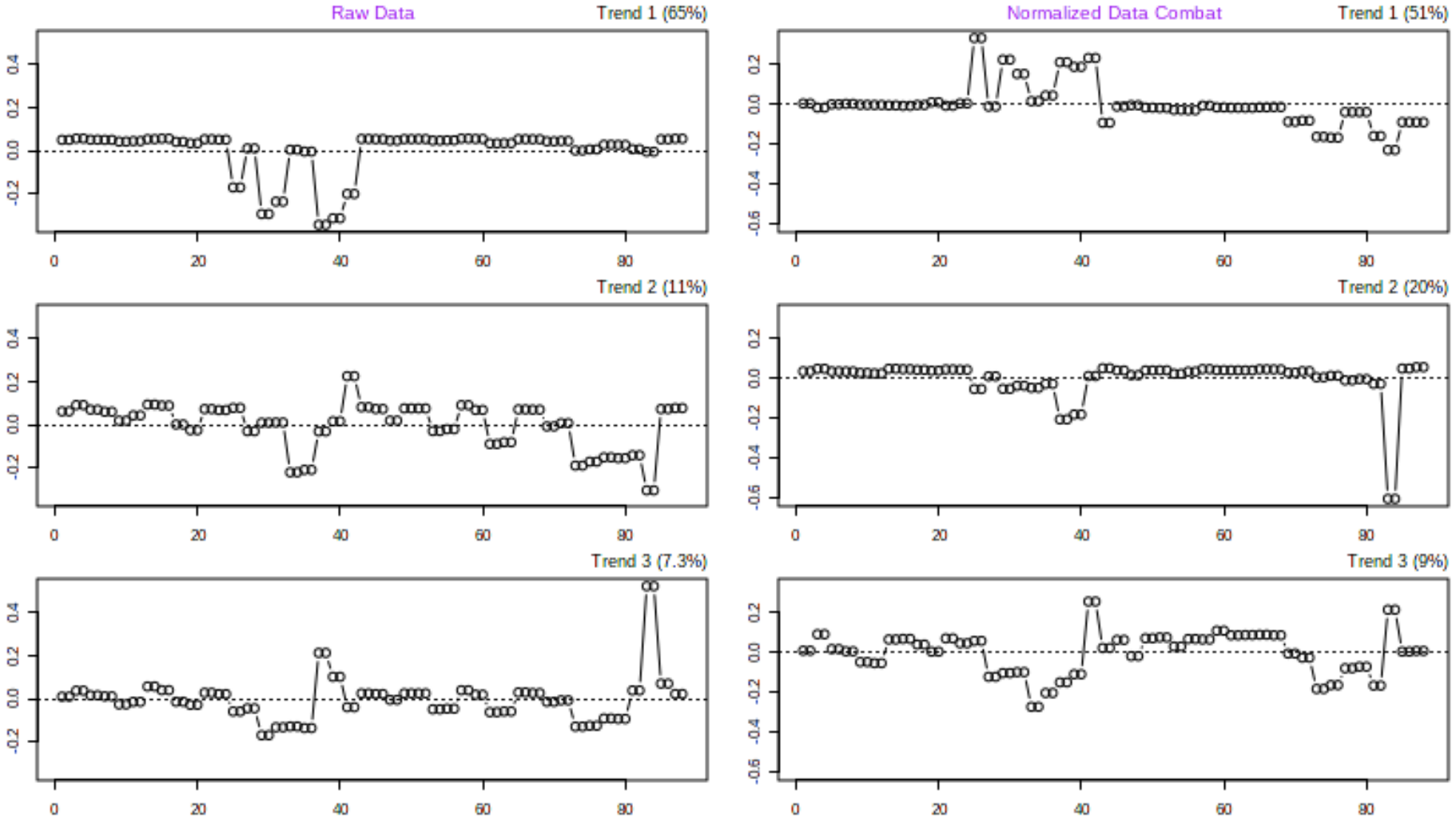


Figure 5.7. Batch drift analysis and adjustment in the positive ionisation mode (ESI+). The singular value decomposition analysis of systematic trends attributable to bias before and after normalization using ComBat method.

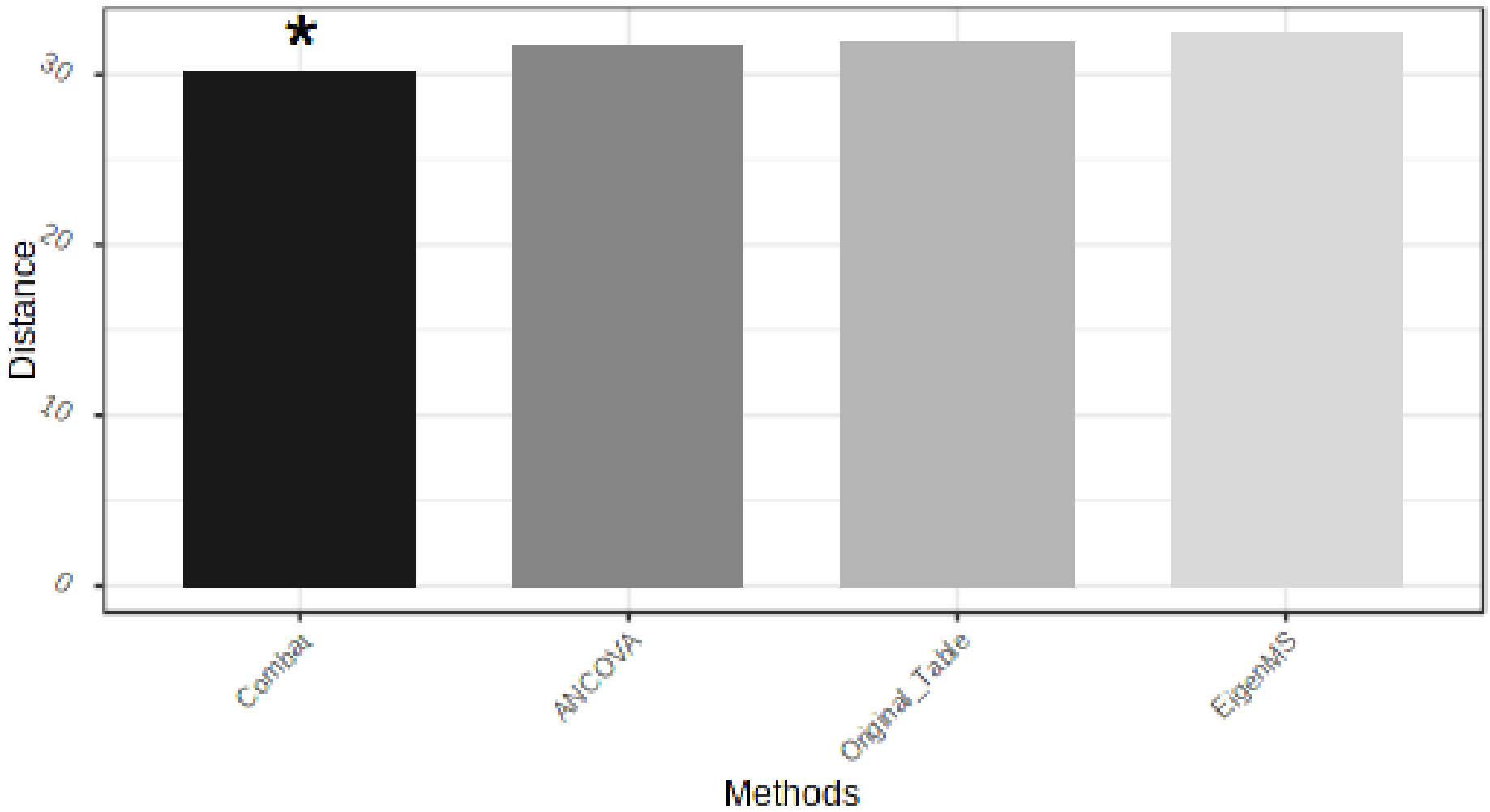


Figure 5.8. Batch drift analysis and adjustment in the positive ionisation mode (ESI+). Inter-batch distance analysis on the normalized and the original data (*shows the normalisation method with the least distance between the sample batches).

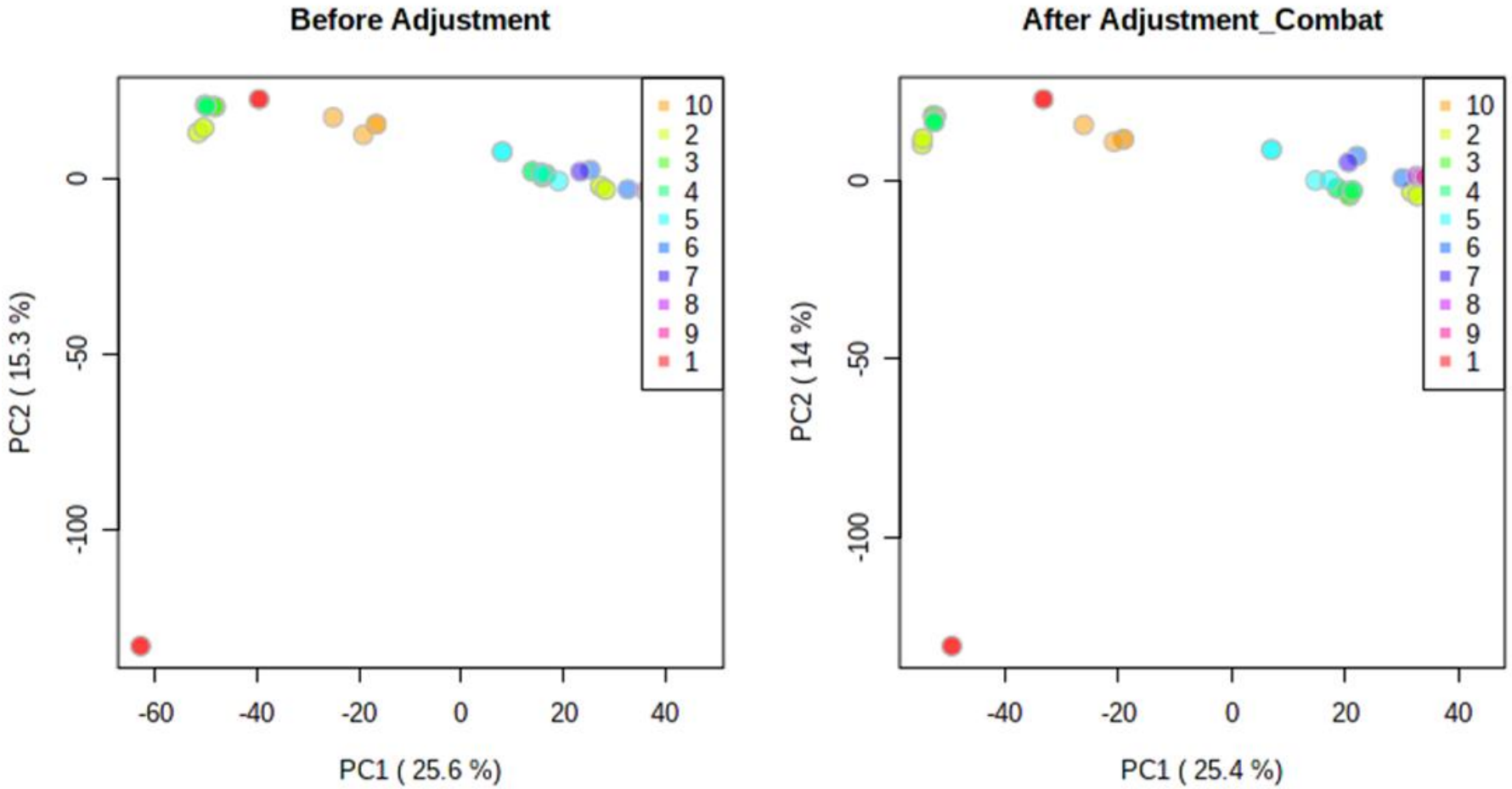


Figure 5.9. Batch drift analysis and adjustment in the negative ionisation mode (ESI-). PCA plot showing sample distribution before and after batch correction using ComBat method.

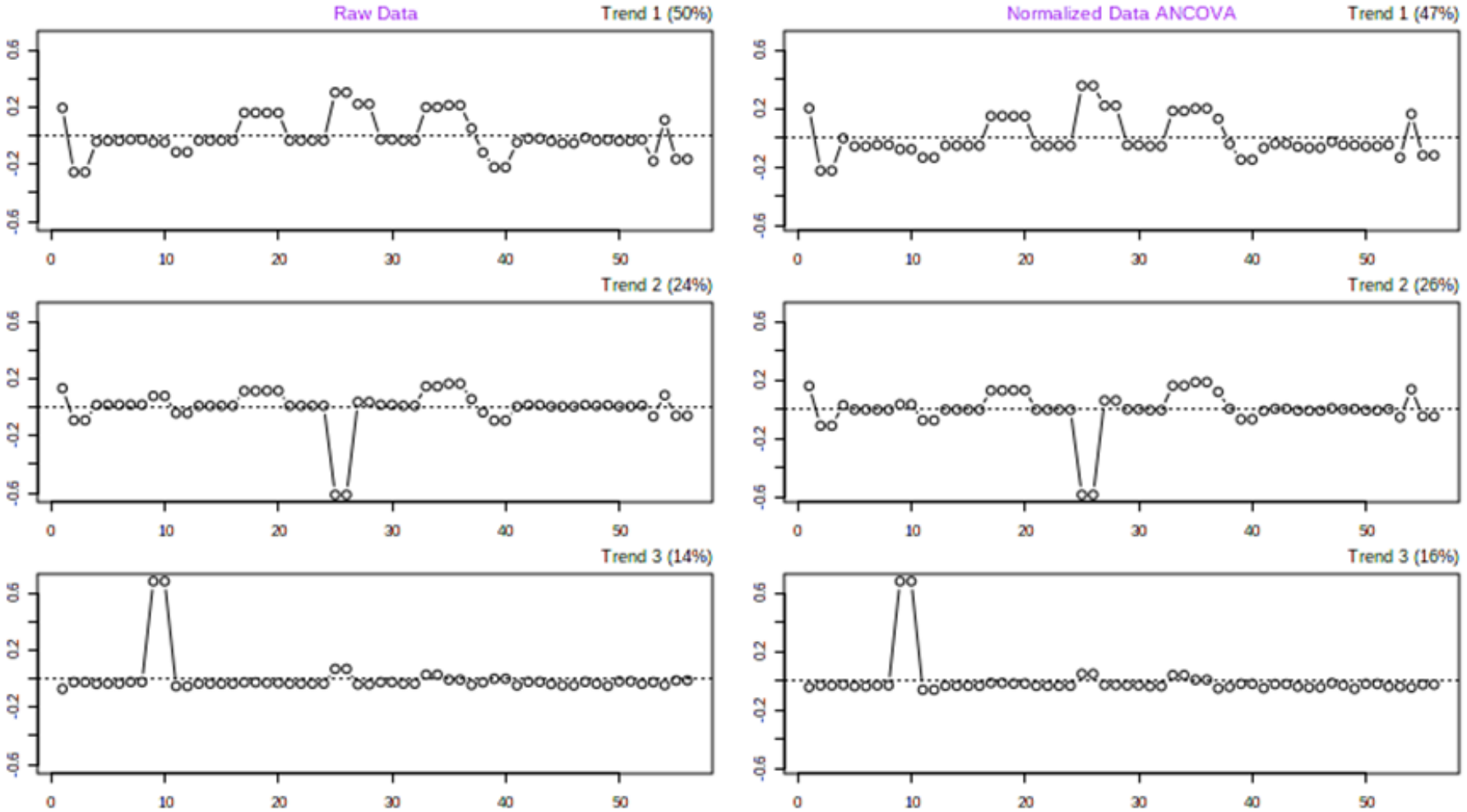


Figure 5.10. Batch drift analysis and adjustment in the negative ionisation mode (ESI-). The singular value decomposition analysis for systematic trends attributable to bias before and after normalization using ANCOVA method.

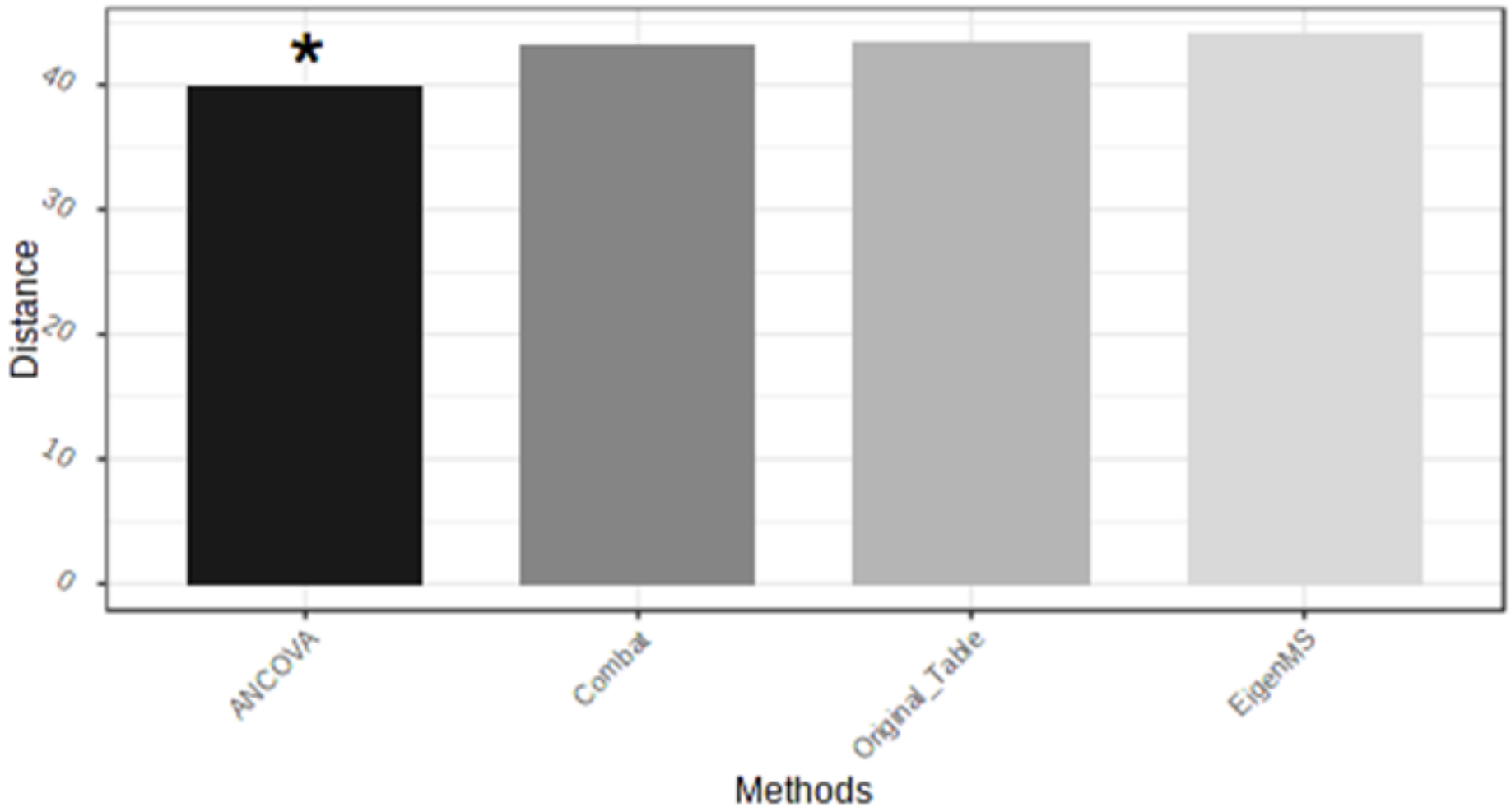


Figure 5.11. Batch drift analysis and adjustment in the negative ionisation mode (ESI-). Inter-batch distance analysis on the normalized and the original data (*shows the normalisation method with the least distance between the sample batches).

Extracted ion chromatogram qualitative analysis and features annotation

While using ANOVA, 36 metabolites were annotated to MSI level 2 and 3 and were significantly perturbed ($p<0.05$) between RHD, AS and controls **Table 5.3**. The PLS-DA model analysis of the annotated metabolites reported a low model cross validation predictive relevance score ($Q2=0.49$), which is not unusual in human studies where highly variable data and many confounding factors complicates interpretation. Permutation testing of the model showed a between and within sums of squares (B/W) ratio between the permutations and the original non-permuted data which was significantly different (empirical p-value =0.042) (**Figure 5.12 & Figure 5.13**). However, the use of B/W ratios to determine differences different more than 2 groups has been shown not to be the most feasible method.[176,177] Therefore, univariate statistics testing was used to determine the significant *m/z* features. Prior to considering a feature as successfully annotated and as potential biomarker, the extracted ion chromatogram peak shapes of the important features were qualitatively analysed. The extracted peaks were assessed based on the background noise, gaussian peak shape and the peak alignment from different samples (**Figure 5.14 & Figure 5.15**). Only the annotated metabolites that passed the quality checks were considered as detected and potential biomarkers. The important *m/z* features were annotated based on the molecular formular, predicted molecular structure and accurate mass, and matched to the public metabolic spectra databases (**Figure 5.16 & Figure 5.17**). Furthermore, the features that did not match to any known molecules were further annotated based on their molecular similarity to annotated features using the GNPS feature-based molecular networking (FBMN) algorithm (**Figure 5.18 - Figure 5.20**).

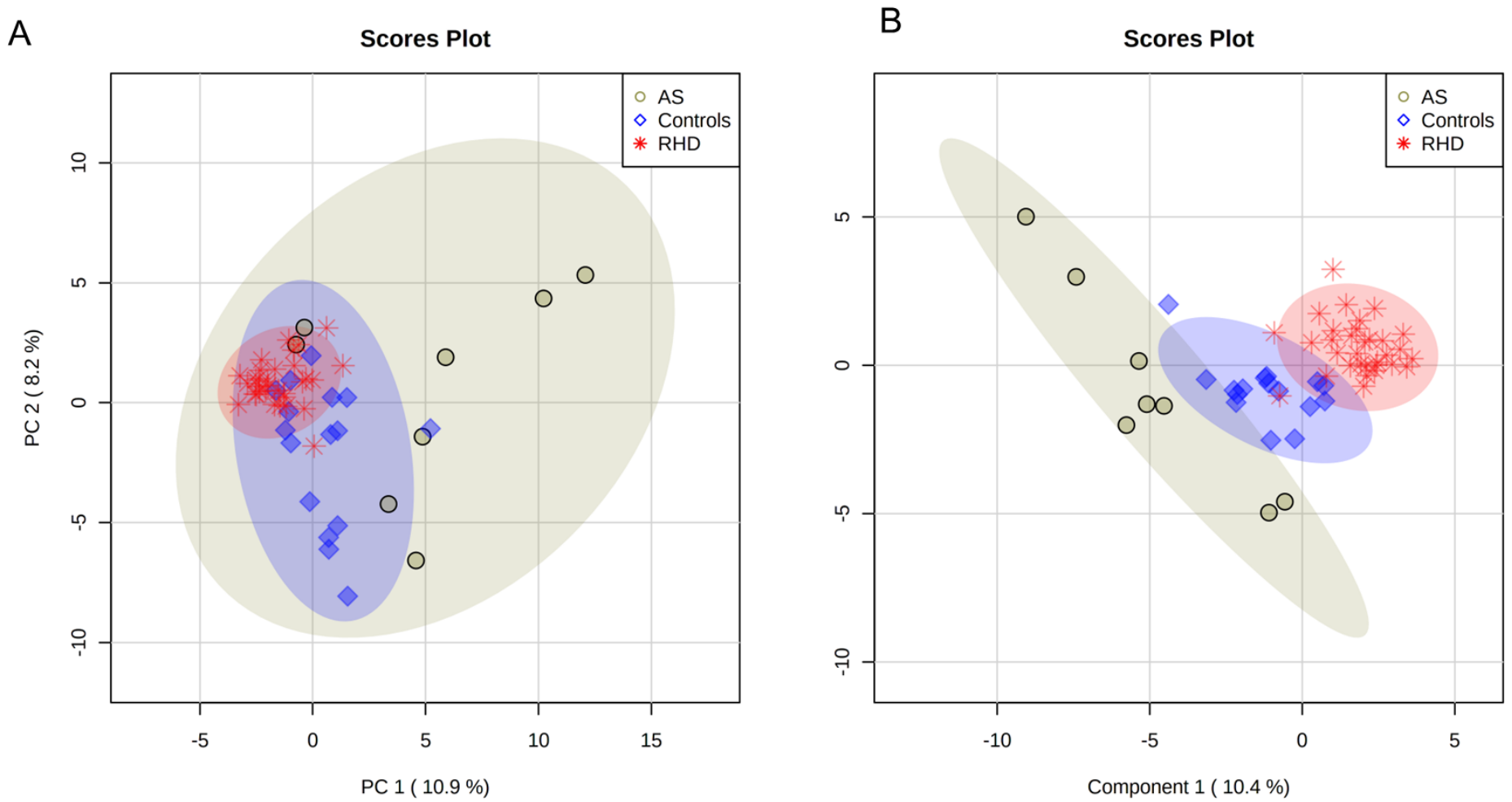


Figure 5.12. Exploratory analysis of metabolites dysregulated between RHD, AS and controls. (A) PCA scores plot, (B) PLS-DA scores plot.

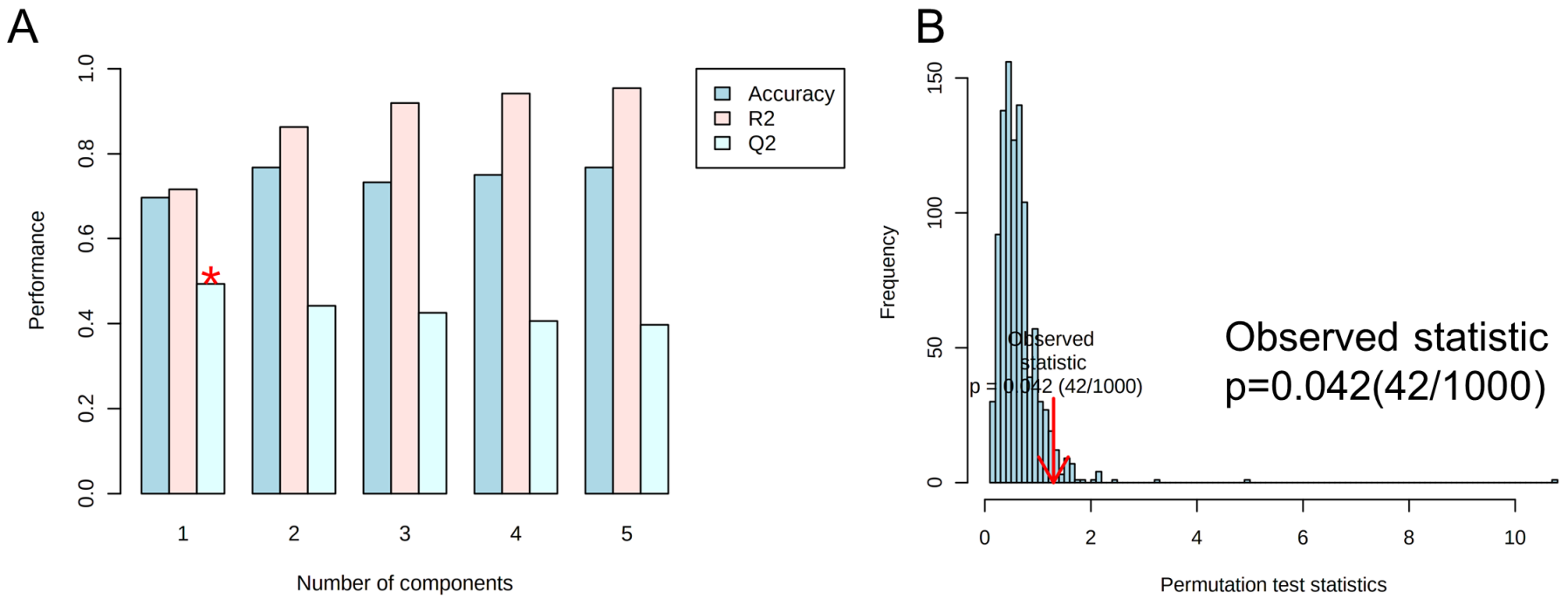


Figure 5.13. Exploratory analysis of metabolites dysregulated between RHD, AS and controls. (**A**) PLS-DA scores cross validation based on component 1; R2=0.72 and Q2=0.49, (**B**) permutations test analysis based on 1,000 permutations (empirical p = 0.042).

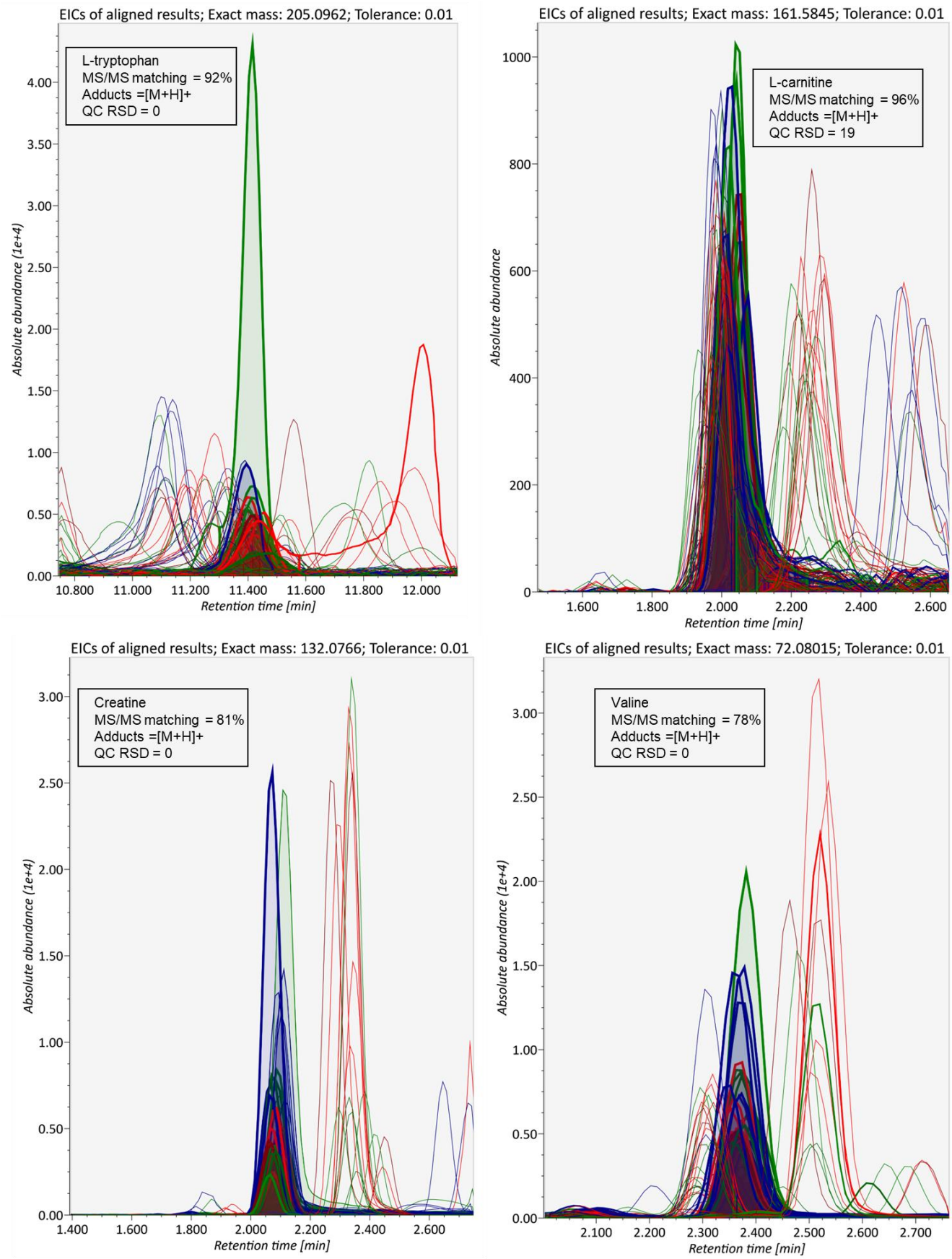


Figure 5.14. Positive and negative ionisation modes extracted ion chromatograms (EIC) of some of the aligned and annotated m/z features that were found to be important metabolite biomarkers.

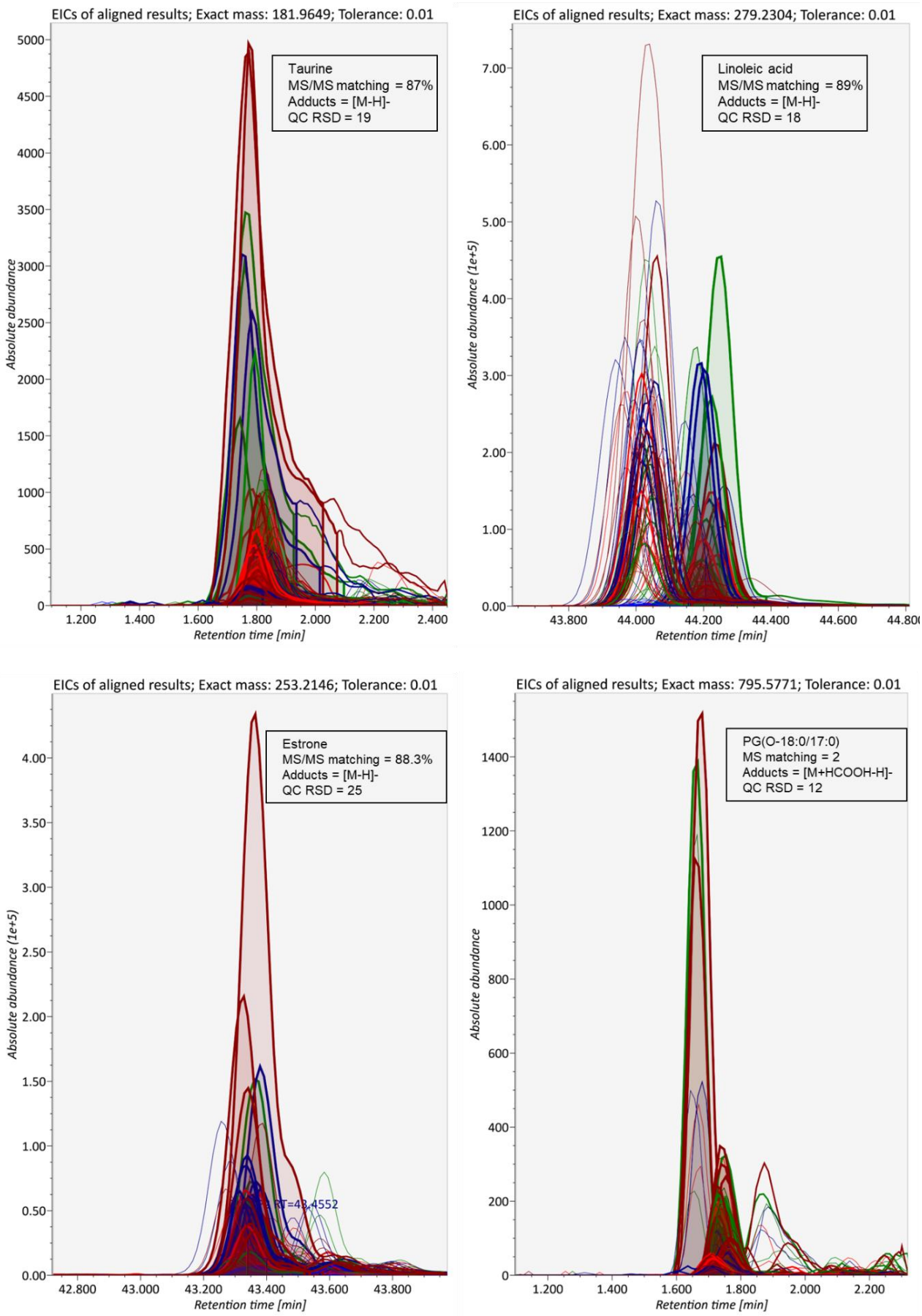


Figure 5.15. Positive and negative ionisation modes extracted ion chromatograms (EIC) of some of the aligned and annotated m/z features that were found to be important metabolite biomarkers.

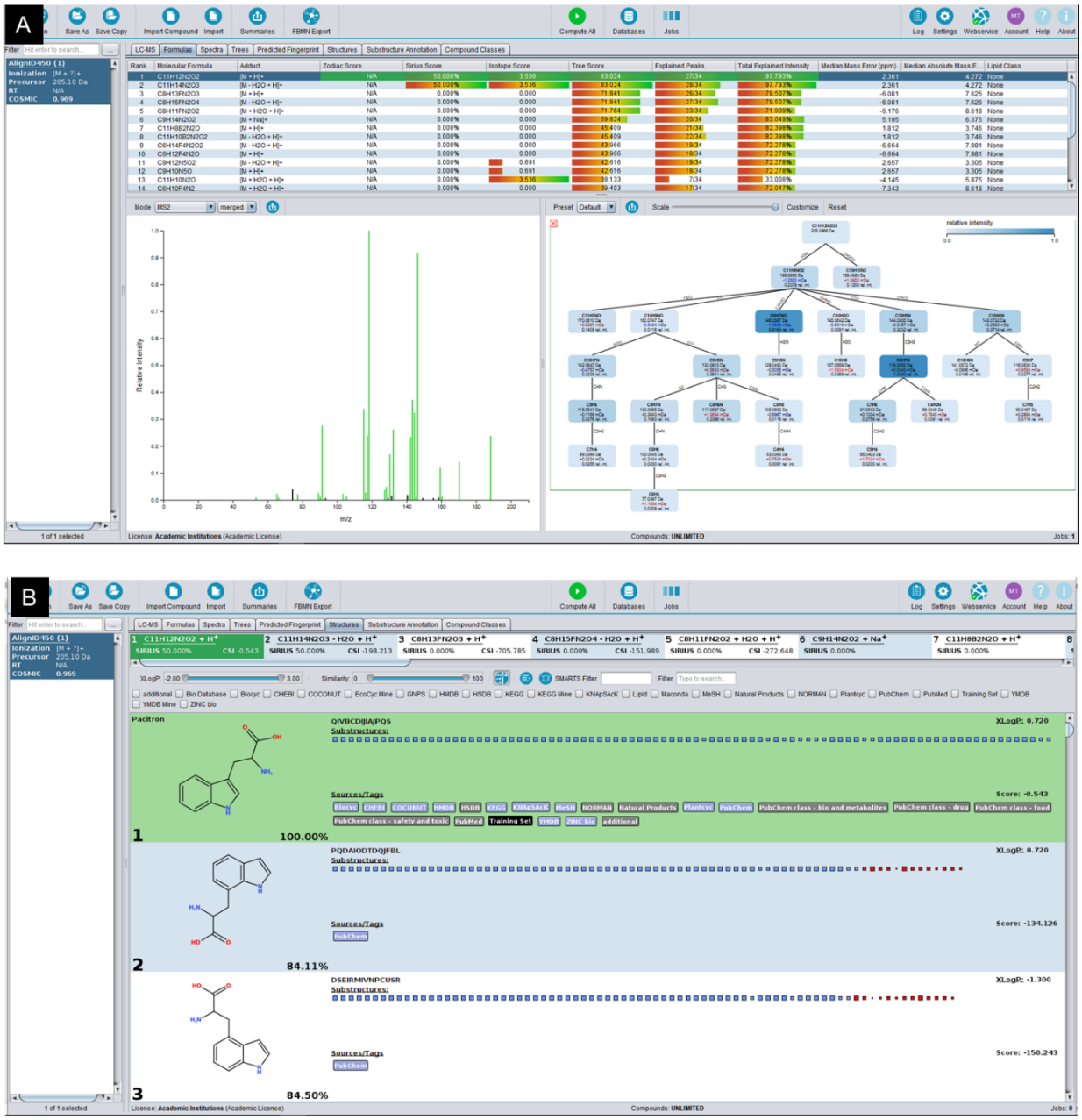


Figure 5.16. Successful annotation process of important *m/z* features based on the MS and the MS/MS fragments using SIRIUS version 5.8.6. The AlignID450 where *m/z* = 205.09657 was the precursor ion and the associated ions in the spectra, (A) MS and MS/MS fragments predicted molecular formular, the predicted molecular structure prediction tree and the predicted adducts. (B) The database matching of the predicted molecular formular and the predicted structure.

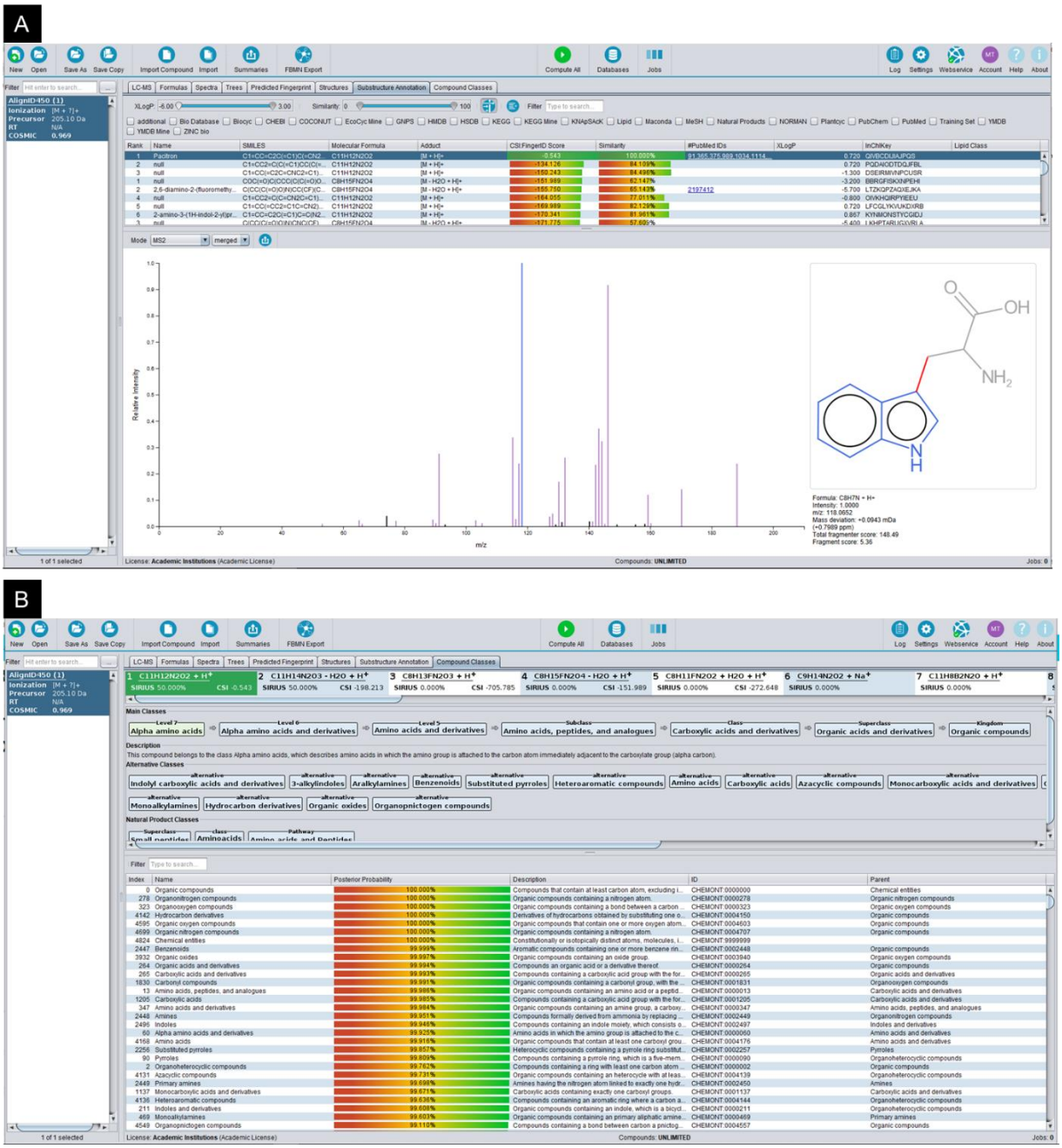


Figure 5.17. Successful annotation process of important *m/z* features based on the MS and the MS/MS fragments using SIRIUS version 5.8.6. The AlignID450 where *m/z* = 205.09657 was the precursor ion and the associated ions in the spectra, (A) Substructure annotation based on the queried MS/MS fragments of the annotated *m/z* feature. (B) The annotated compound classification using Classifire and Natural products prediction algorithms. The *m/z* feature was annotated to L-tryptophan with 100% spectra matching score and was annotated into MSI confidence level 2.

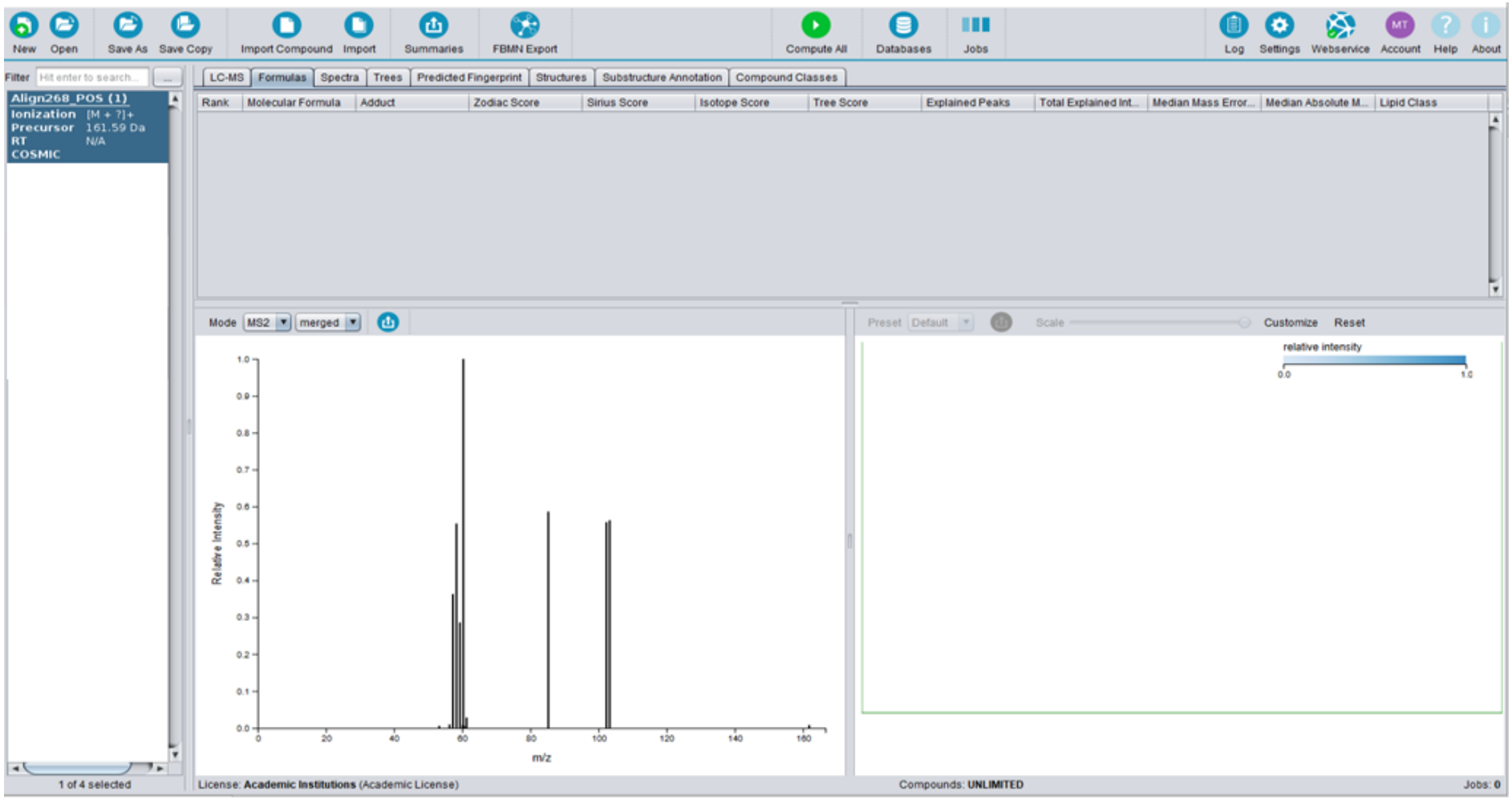


Figure 5.18. Unsuccessful example of the annotation process of important *m/z* features based on the MS and the MS/MS fragments using SIRIUS version 5.8.6. Each peak is assigned a unique peak number assigned by Sirius? In this example, MS-DIAL alignment number 268 in the positive ionisation mode (Align 288 POS) equated to a spectra with *m/z* = 161.58698 was the precursor ion and the associated fragments in the spectra. Using the Align command for peak 268 compiled the MS and MS/MS fragments associated with this peak to build a predicted molecular formula. This feature could not be associated to any known biological molecule. It was not annotated. To ensure reliability of the annotations, the SIRIUS annotations were compared to the annotations suggested by other annotation tools.

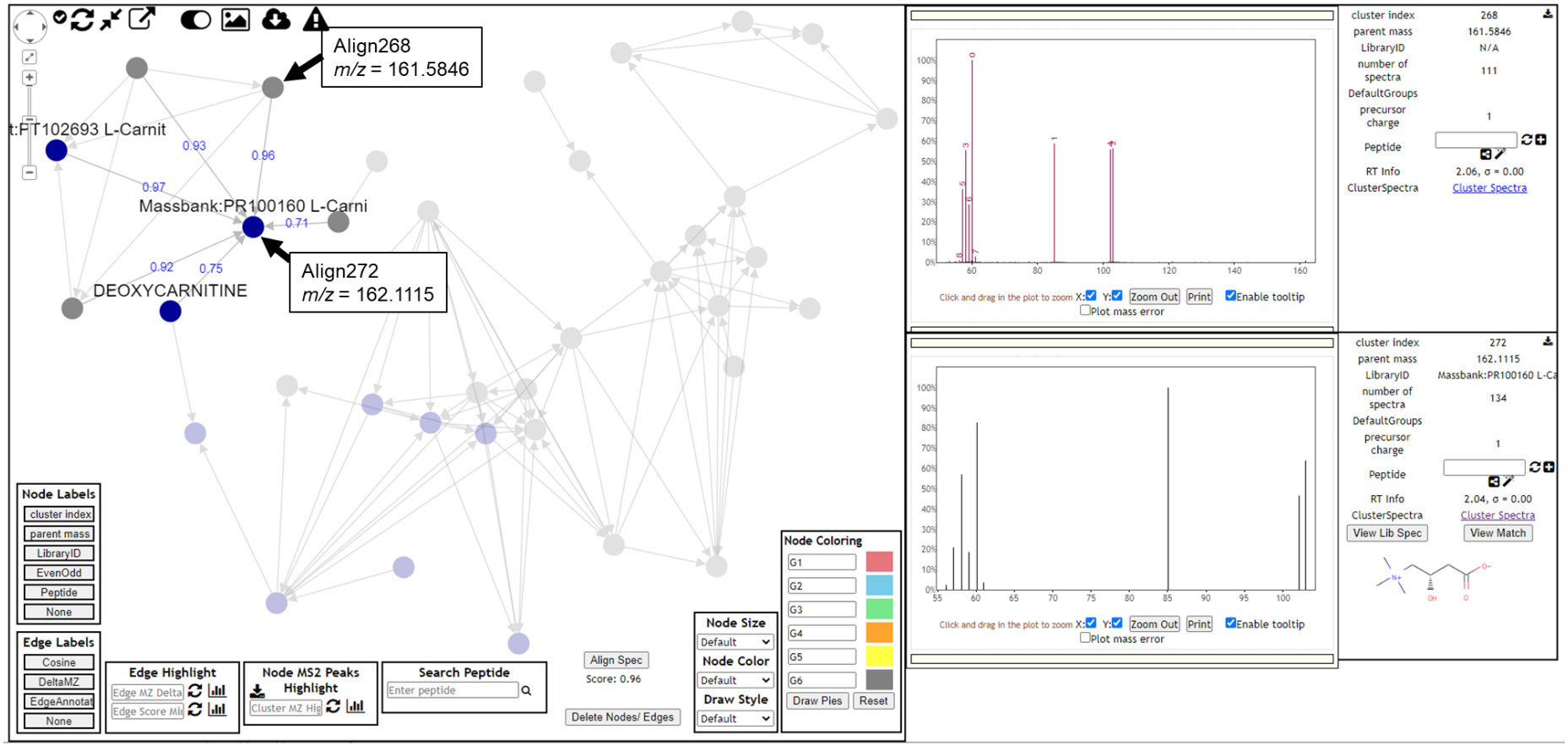


Figure 5.19. Annotation process of the important features based on the MS and MS/MS matching using the GNPS feature based molecular networking (FBMN). Feature Align268 *m/z* 161.5846 was annotated based on the molecular similarity to feature Align272 *m/z* 162.1115 which was annotated to L-carnitine.

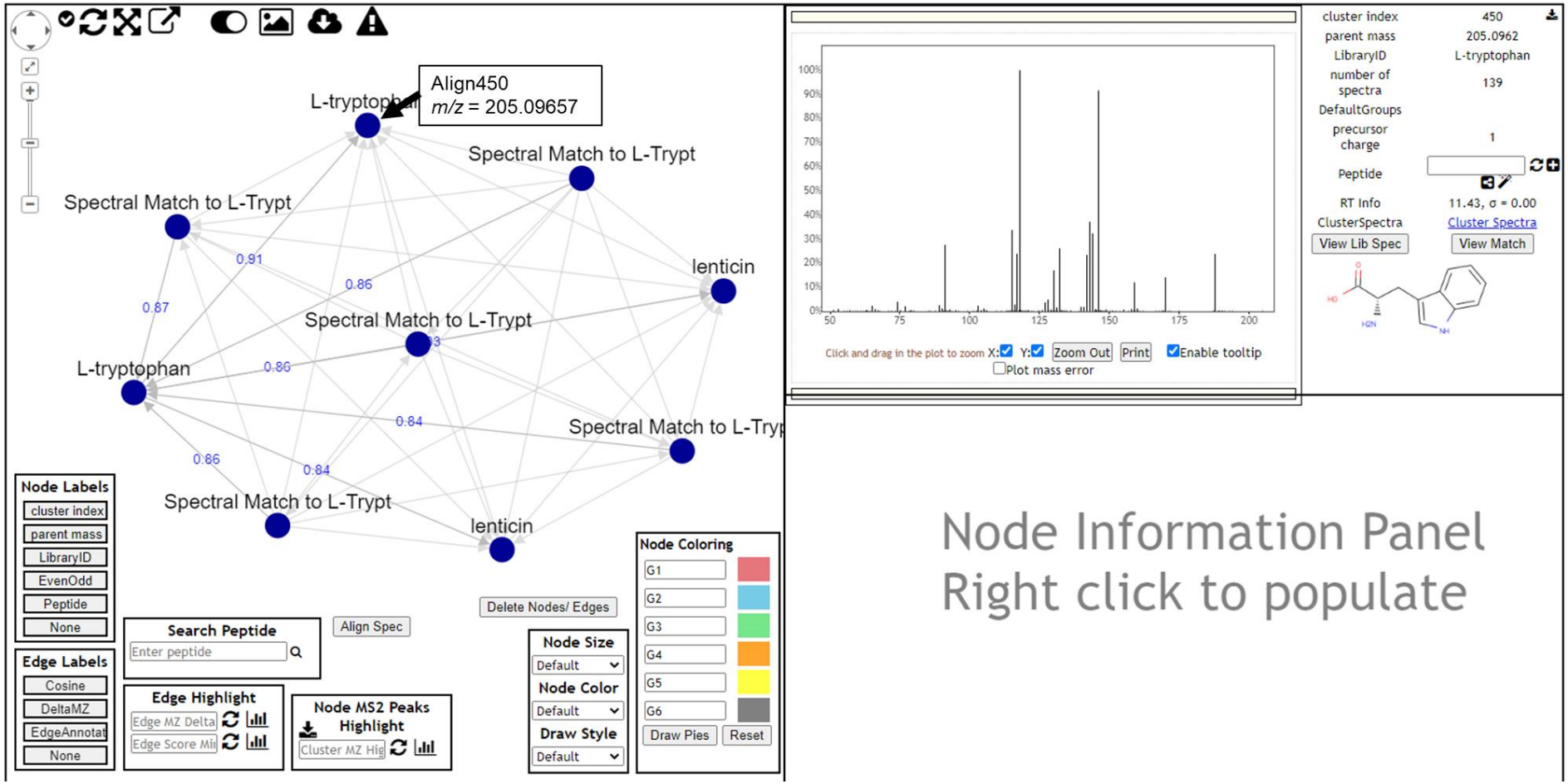


Figure 5.20. Annotation process of the important features based on the MS and MS/MS matching using the GNPS feature based molecular networking (FBMN). Feature Align450 *m/z* 250.09657 was also annotated to L-tryptophan based on the FBMN.

Table 5.3. Summary of dysregulated metabolites between Rheumatic heart disease (RHD), aortic stenosis (AS) patients, and controls.

Molecule name	RT	*m/z*	HMDB ID	MSI confidence level	$CV_{(QC)}$	ANOVA raw p.value	FDR adjusted p value	Post Hoc (Tukey's HSD p<0.05)
tryptophan	11.43	205.0962	HMDB0000929	2	0	<0.001	0.004	Controls-AS; RHD-AS
Phenylacetylglutamine	13.72	265.1174	HMDB0006344	2	20	<0.001	0.004	Controls-AS; RHD-AS
PE(20:4)	41.18	502.2923	HMDB0011517	3	23	<0.001	0.004	Controls-AS; RHD-AS; RHD-Controls
PE(22:6)	41.50	526.2926	HMDB0011526	3	29	<0.001	0.005	RHD-AS; RHD-Controls
Indole-3-acetic acid	20.96	130.0642	HMDB0302181	2	30	<0.001	0.016	Controls-AS; RHD-AS
Deoxycholate	40.59	357.2784	HMDB0000626	2	26	0.001	0.016	Controls-AS; RHD-AS
Carnitine	2.06	161.5846	HMDB0000062	2	19	0.002	0.020	Controls-AS; RHD-AS
Cortisol	26.12	363.2144	HMDB0000063	2	30	0.002	0.022	Controls-AS; RHD-AS
4-hydroxy-2-oxo-Heptanedioic acid	2.32	191.0545	C05601	3	19	0.002	0.022	Controls-AS; RHD-AS
3-Formylindole	19.94	146.0596	HMDB0029737	2	0	0.004	0.030	RHD-Controls
PA(8:0/13:0)	34.59	512.3346	HMDB0115484	3	17	0.006	0.038	Controls-AS; RHD-AS
Crotonic acid	1.85	85.0298	HMDB0010720	3	19	0.006	0.038	Controls-AS; RHD-Controls
2,4,7-Decatrienoic acid	1.69	187.0725	HMDB0035235	3	17	0.007	0.038	Controls-AS; RHD-AS
Taurine	1.78	181.9649	HMDB0000251	2	19	0.007	0.038	Controls-AS; RHD-Controls
PS(16:0/16:0)	1.75	734.4912	HMDB0000614	3	21	0.007	0.038	Controls-AS; RHD-AS
Propionylcarnitine	6.23	218.1391	HMDB0000824	2	0	0.008	0.038	Controls-AS; RHD-AS
MG(0:0/20:3)	40.23	398.3282	HMDB0011545	3	16	0.008	0.039	RHD-Controls
10-nitro-9-octadecenoic acid	30.71	328.2467	HMDB0062737	3	17	0.009	0.040	Controls-AS; RHD-AS
Butyrylcarnitine	9.61	232.1534	HMDB0002013	2	16	0.010	0.041	Controls-AS; RHD-AS
DG(a-15:0/i-12:0/0:0)	47.20	521.4223	HMDB0093443	3	16	0.011	0.047	Controls-AS; RHD-AS

LysoPC(18:1)	42.11	566.3257	HMDB0002815	2	16	0.012	0.048	RHD-Controls
3-oxo-4-pentenoic acid	1.84	97.0282	LMFA01060166	3	23	0.017	0.062	RHD-Controls
Methyl-2-decene-4,6,8-triynoate	1.65	190.9122	HMDB0033765	3	17	0.017	0.062	RHD-AS
TG(15:0)	1.63	810.7893	HMDB0043801	3	30	0.018	0.062	RHD-Controls
Stoloniferone L	39.57	445.3328	LMST01031098	3	26	0.021	0.067	RHD-Controls
Tetradecanoylcarnitine	37.37	394.2950	HMDB0005066	3	30	0.025	0.076	RHD-Controls
2,10-dihydroxy-4,6,8-decatriynoic acid	1.81	97.02925	LMFA01050235	3	30	0.030	0.088	RHD-Controls
Cyclo(-prolyl-valyl)	12.84	197.1278	HMDB0240493	2	16	0.030	0.088	RHD-Controls
LysoPE(0:0)	42.44	528.3096	HMDB0011494	3	26	0.033	0.092	RHD-Controls
PE(P-16:0)	1.75	738.5738	HMDB0011356	3	21	0.035	0.093	Controls-AS
Cer(d18:0)	1.71	696.6908	HMDB0011764	3	16	0.038	0.097	Controls-AS; RHD-AS
2-Amino-4-hydroxy-3-methylpentanoic acid	1.98	146.0822	HMDB0029449	3	4	0.046	0.101	Controls-AS; RHD-AS
2-Hydroxy-4-(methylthio)butanoic acid	2.44	133.0320	HMDB0037115	3	26	0.047	0.101	Controls-AS; RHD-AS
Lyso PC (16:1)	41.79	494.3221	HMDB0010383	2	16	0.047	0.101	RHD-Controls
Ursodeoxycholic acid	40.19	357.2795	HMDB0000946	2	26	0.047	0.101	RHD-Controls
PC(P-18:0)	1.71	796.6284	HMDB0011252	3	23	0.049	0.101	Controls-AS; RHD-AS

*p value adjustment for multiple testing done with Benjamin-Hochberg false discovery rate (FDR) adjustment method.

5.3.3 Functional analysis of the perturbed metabolites

Pathway analysis of metabolites changed between RHD and healthy controls

Univariate chemometrics analysis of the annotated metabolites between RHD and controls indicated 22 annotated metabolites as significantly altered (log2FC>2, p<0.05) **Table 5.4**.

Table 5.4. Summary of metabolites significantly changed between RHD and healthy controls indicating their corresponding pooled QC RSD (CV_{QC}), fold changes (FC and log2FC), crude and FDR adjusted p value.

Metabolite	CV_{QC}	FC	Log2(FC)	P Value	FDR adj. p
PE(20:4)	23	0.21	-2.28	0.001	0.022
3-Formylindole	0	0.08	-3.59	0.001	0.022
PE(22:6)	29	0.21	-2.27	0.002	0.022
Tryptophan	0	0.27	-1.86	0.003	0.029
MG(0:0/20:3)	16	5.25	2.39	0.006	0.046
LysoPC(16:1)	16	0.31	-1.71	0.010	0.046
Taurine	19	2.15	1.10	0.011	0.046
LysoPC(18:1/0:0)	16	4.46	2.16	0.013	0.046
2,6-nonadienoic acid	24	3.75	1.91	0.013	0.046
Ursodeoxycholic acid	26	0.07	-3.84	0.014	0.046
Crotonic acid	19	0.48	-1.05	0.017	0.048
LysoPE(0:0/22:5)	26	0.46	-1.13	0.017	0.048
Cyclo(prolyl-valyl)	16	0.42	-1.24	0.019	0.048
2,4-dihydroxy-butanoic acid	22	0.15	-2.78	0.020	0.048
Hexanoyl-Carnitine	17	0.21	-2.28	0.024	0.049
13-Docosenamide	16	0.15	-2.70	0.025	0.049
Valine	0	2.67	1.42	0.027	0.049
Tetradecanoylcarnitine	30	0.41	-1.30	0.028	0.049
Cer(d18:0/h26:0)	16	0.31	-1.69	0.029	0.049
PC(18:4/P-18:1)	28	0.27	-1.89	0.030	0.049
Inosine	23	0.37	-1.45	0.038	0.059
Butyrylcarnitine	30	2.33	1.22	0.044	0.063

Hierarchical clustering analysis of the annotated metabolites showed distinct perturbation patterns in RHD patients and healthy controls which may suggest a difference in metabolic processes (**Figure 5.21)**. After enrichment analysis (**Figure 5.22)** some of the altered metabolites were impactfully and significantly mapped to linoleic acid metabolism (FDR adj. p<0.02). Furthermore, glycerophospholipid metabolism and taurine and hypotaurine metabolism pathways were impactfully mapped. However, these were not significant after multiple testing adjustment (FDR adj. p>0.05). In details, linoleic acid, and phosphatidylcholine (PC) were mapped to linoleic acid metabolism pathway and were elevated in RHD compared

to healthy controls (FC>2, p<0.05) (**Figure 5.23A)**. In addition, Taurine was mapped to taurine and hypotaurine metabolism pathway and was increased in controls compared to RHD patients (**Figure 5.23B**). PC and LysoPC were mapped to glycerophospholipid metabolism and were significantly changed between RHD and controls (**Figure 5.23C)**.

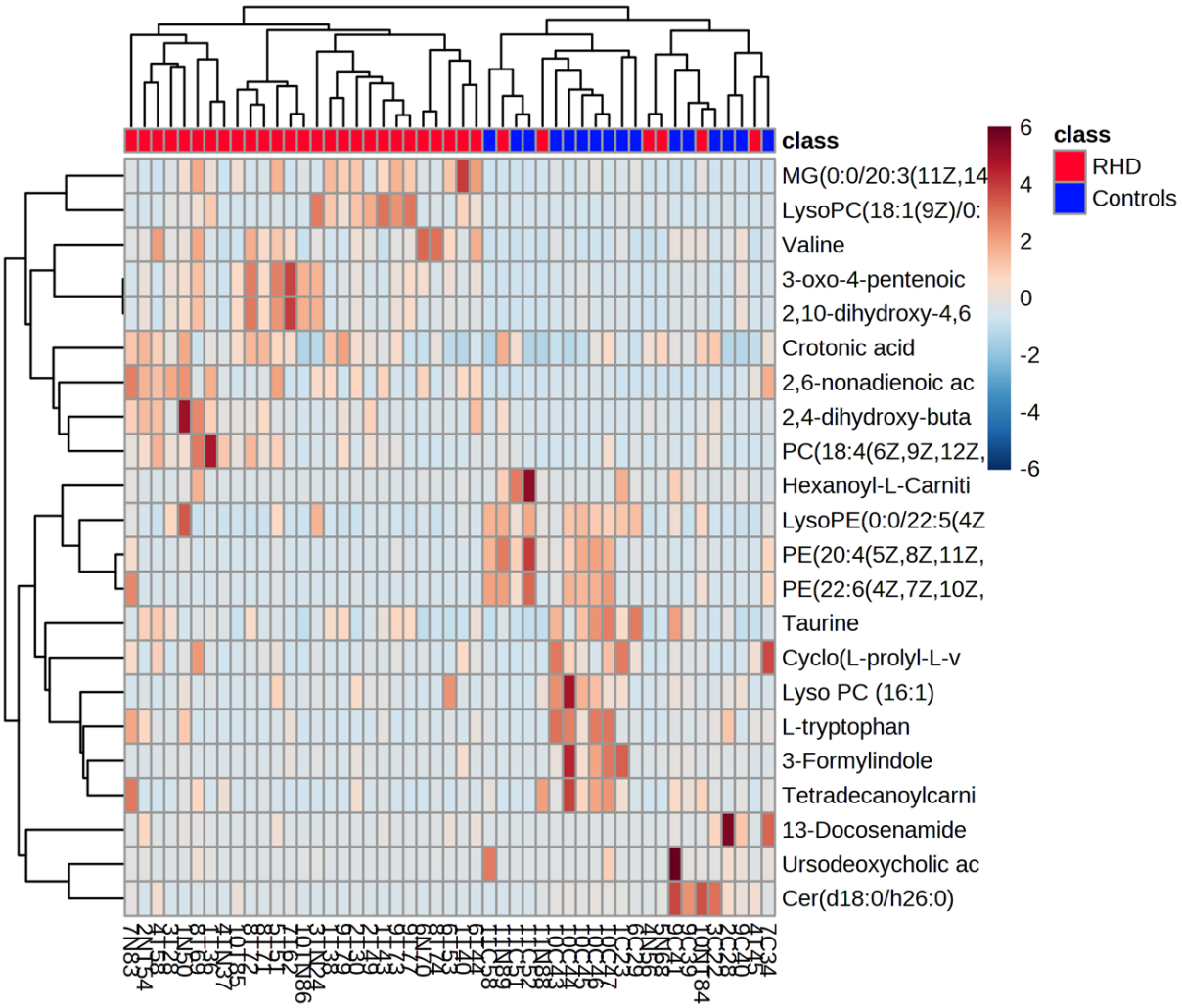


Figure 5.21. Exploration of dysregulated metabolites in rheumatic heart disease. Heatmap with hierarchical clustering of samples based on significantly changed metabolites (p<0.05).

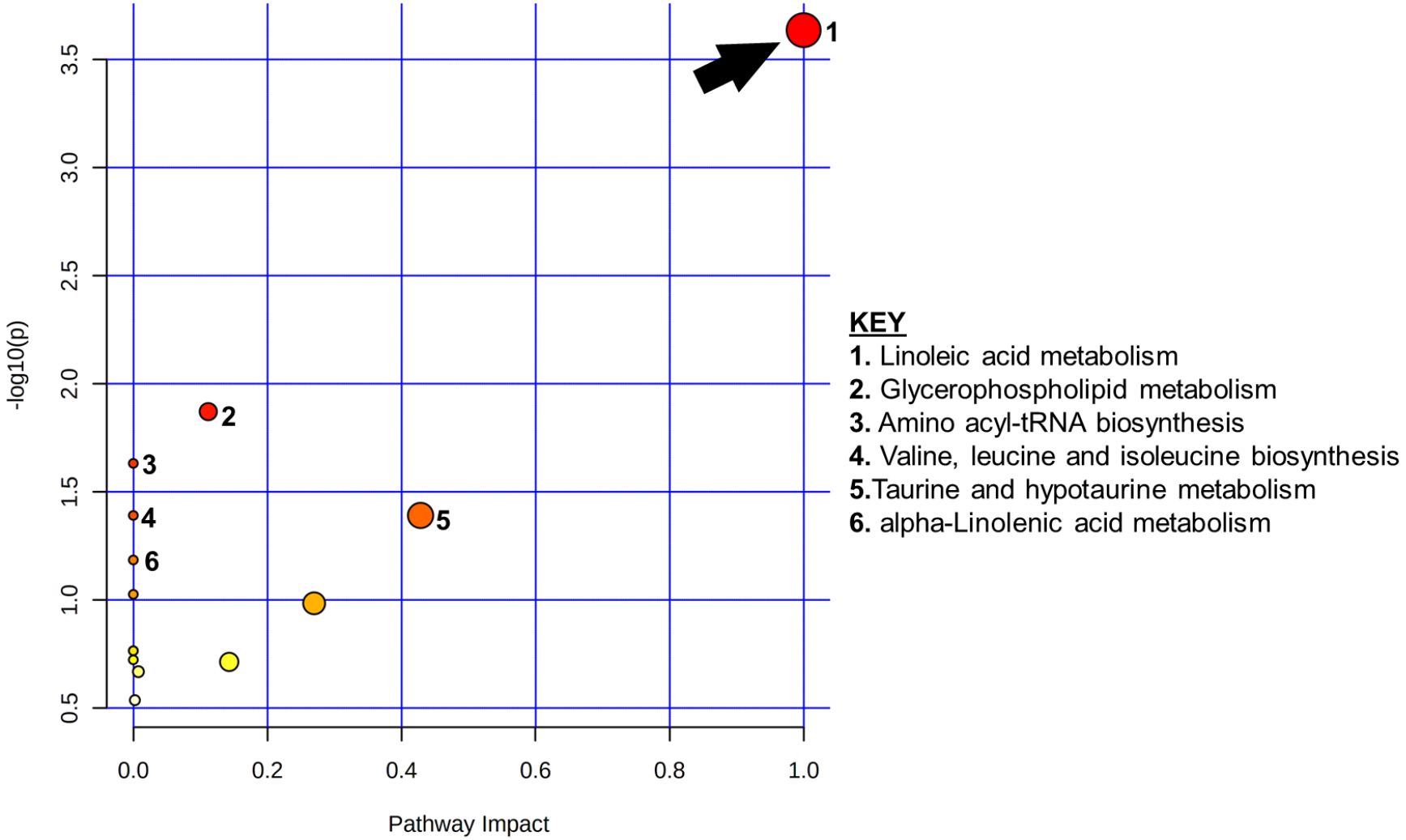


Figure 5.22. Exploration of dysregulated metabolites in rheumatic heart disease. Pathway enrichment analysis map of the enriched pathways.

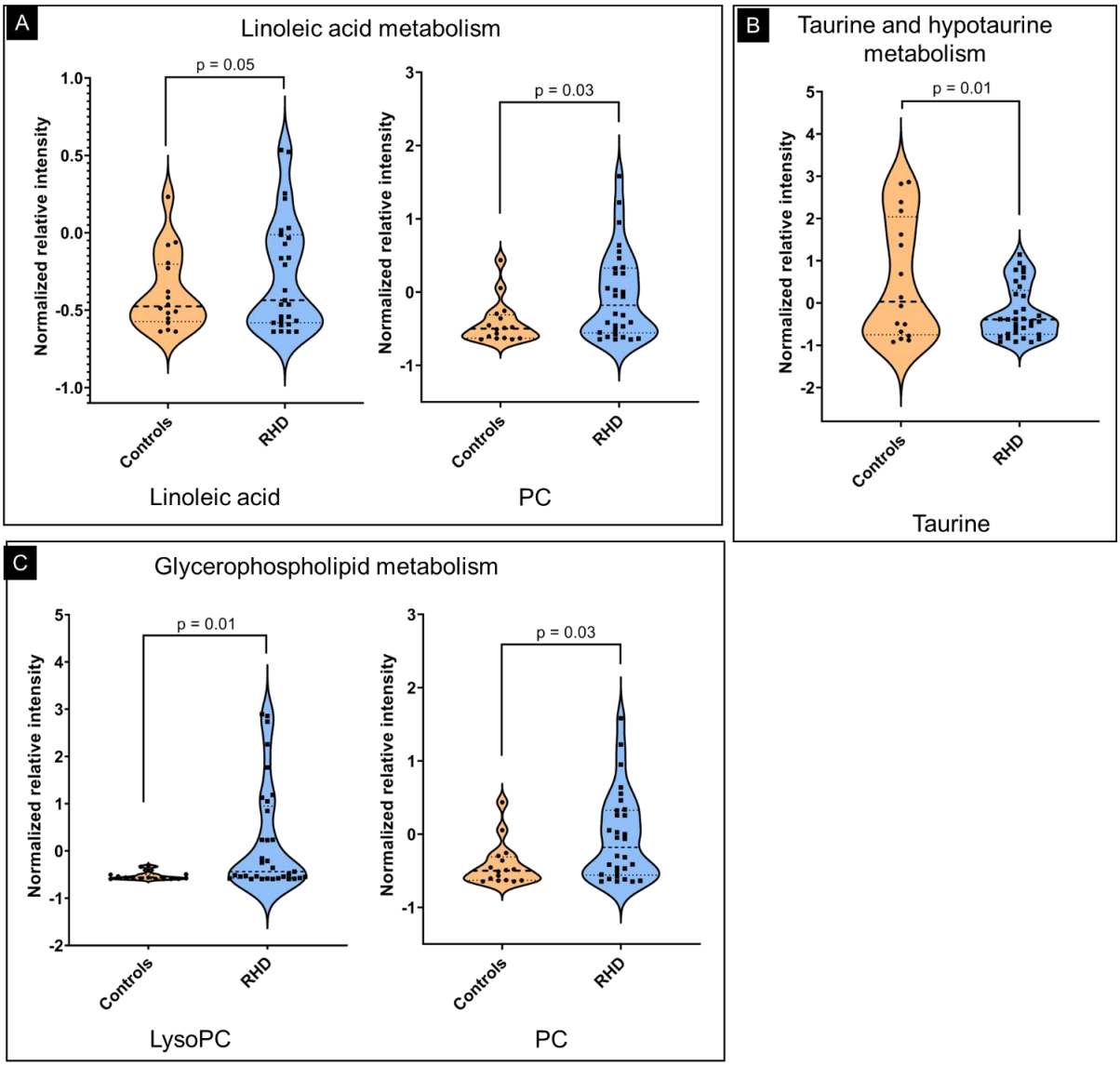


Figure 5.23. Exploration of dysregulated metabolites in rheumatic heart disease. Perturbation patterns of the metabolites mapped to (A) linoleic acid metabolism, (B) taurine hypotaurine metabolism pathways, and (C) glycerophospholipid metabolism.

Pathway analysis of metabolites changed between AS and healthy controls

After comparing the metabolites altered between degenerative AS and healthy controls, 15 metabolites were found to be significantly changed (Log2FC>2, p<0.05). The significant metabolites are summarised in **Table 5.5** showing the corresponding pooled QC RSD, fold changes and the raw and FDR adjusted p values.

Table 5.5. Summary of metabolites significantly changed between AS and healthy controls indicating their corresponding pooled QC RSD (CV_{QC}), fold changes (FC and log2FC), crude and FDR adjusted p value.

Metabolite	CV_{QC}	FC	Log2(FC)	P Value	FDR adj. p
PE(P-16:0/22:1)	21	4.15	2.05	0.002	0.056
PA(8:0/13:0)	17	9.82	3.30	0.007	0.093
Crotonic acid	19	2.25	1.17	0.010	0.093
Undecylenic acid	30	3.41	1.77	0.015	0.093
Palmitoylcarnitine	16	4.27	2.10	0.020	0.093
Phenylacetylglutamine	20	22.25	4.48	0.022	0.093
Isocaproic acid	26	6.00	2.58	0.023	0.093
Taurine	19	0.22	-2.19	0.024	0.093
Hexadec-2-enoyl carnitine	16	2.19	1.13	0.028	0.093
Butyrylcarnitine	30	2.16	1.11	0.030	0.093
PI(15:1/22:6)	23	0.09	-3.48	0.035	0.093
Ornithine	16	3.07	1.62	0.038	0.093
2,4-Hexadienyl acetate	26	2.44	1.28	0.042	0.093
PC(18:4/P-18:1)	28	4.55	2.19	0.043	0.093
2-Hydroxy-4-(methylthio)butanoic acid	26	7.80	2.96	0.046	0.093

The hierarchical clustering of the perturbed metabolites showed a pattern of metabolites generally increased or decreased in the degenerative aortic stenosis compared to healthy controls (**Figure 5.24)**. The functional analysis of the perturbed metabolites showed that the metabolites were impactfully mapped to glycerophospholipid metabolism pathway (FDR adj. p = 0.03) **Figure 5.25**. In addition, phosphatidylinositol signalling system and taurine and hypotaurine metabolism pathways were also impactfully mapped, however were not significant after multiple testing adjustment (FDR adj. p>0.05). Phosphatidylcholine and phosphatidate were increased in AS while phosphatidylinositol was decreased as compared to controls and the three metabolites were mapped to glycerophospholipid metabolism pathway (**Figure 5.26**). Furthermore, phosphatidate and phosphatidylinositol were mapped to phosphatidylinositol signalling system (**Figure 5.27A**). Taurine was decreased in AS compared to controls and was associated with taurine and hypotaurine metabolism pathway (**Figure 5.27B**)

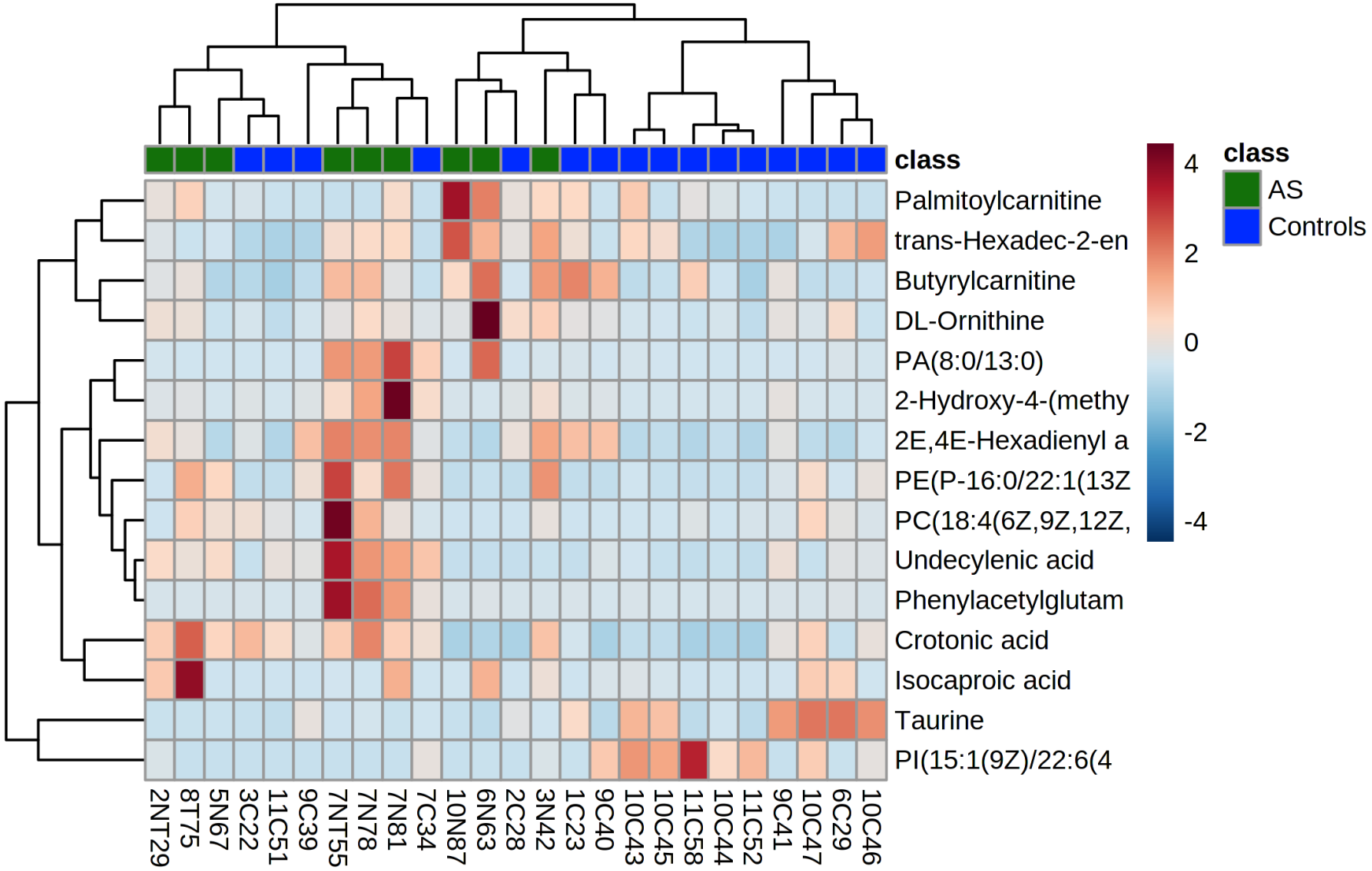


Figure 5.24. Exploration and functional analysis of perturbed metabolites in degenerative aortic stenosis. Heat map with hierarchical clustering of significantly changed metabolites in AS and controls.

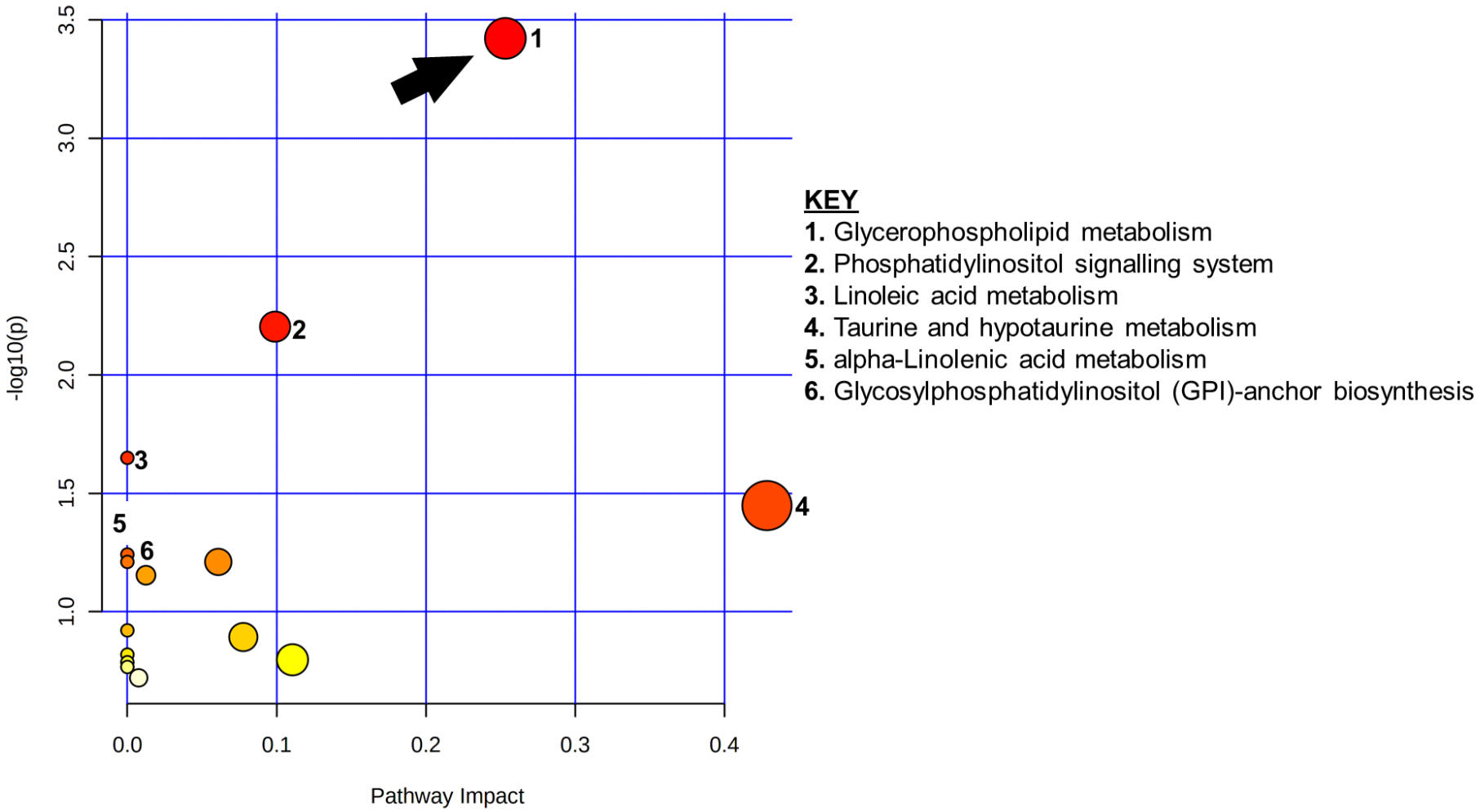


Figure 5.25. Exploration and functional analysis of perturbed metabolites in degenerative aortic stenosis. Pathway enrichment analysis map of the enriched pathways.

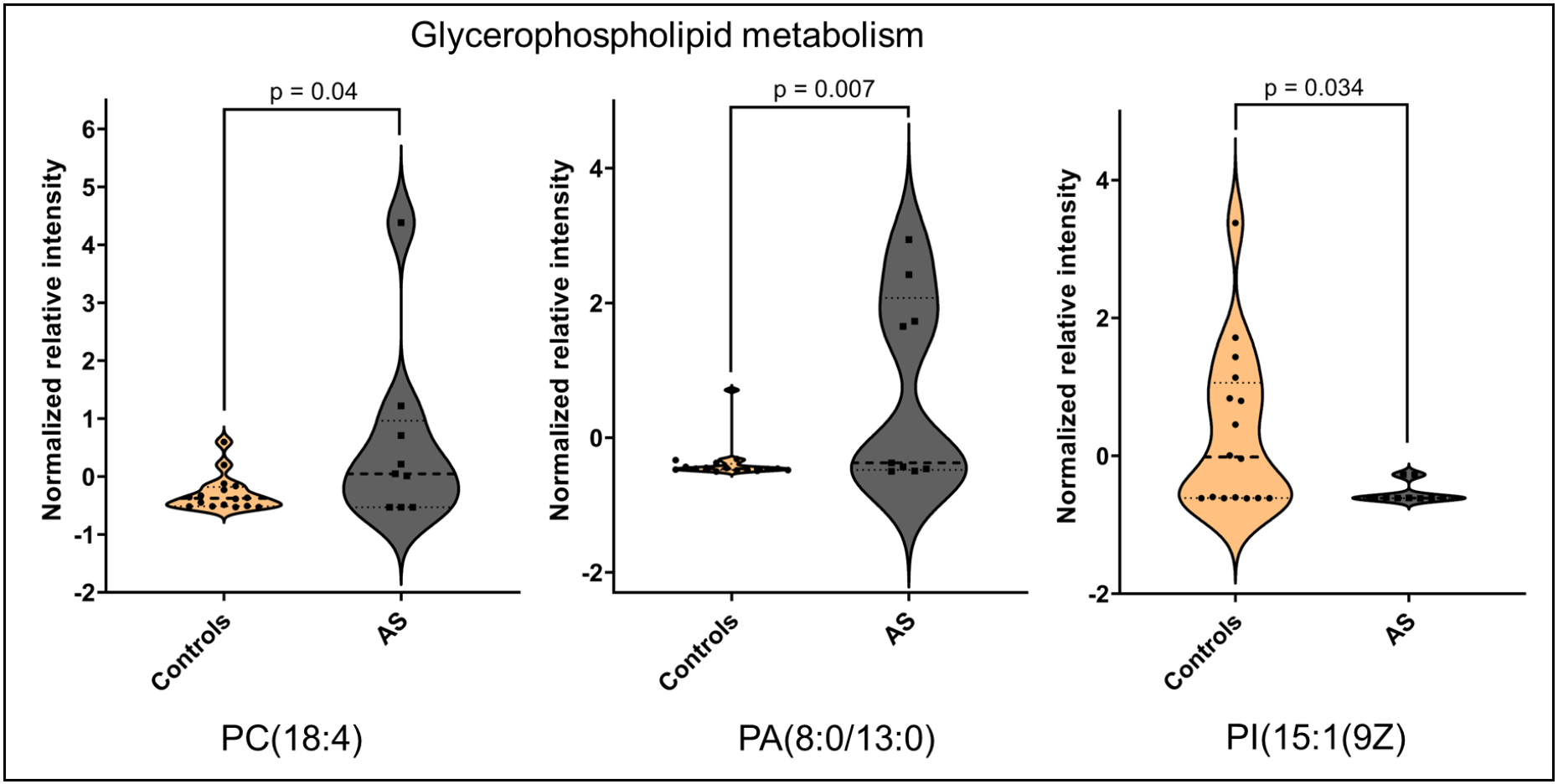


Figure 5.26. Exploration and functional analysis of perturbed metabolites in degenerative aortic stenosis. Perturbation patterns of the metabolites mapped to glycerophospholipid metabolism pathway.

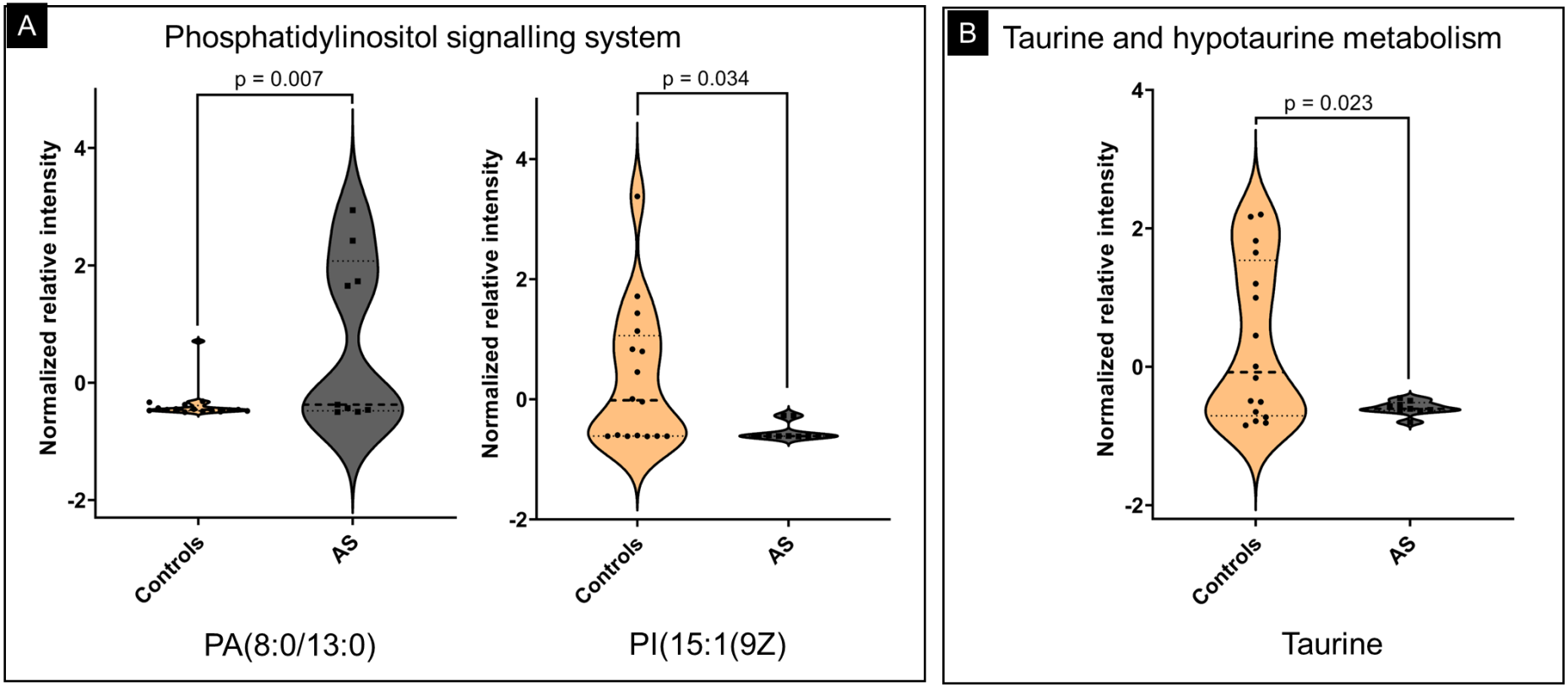


Figure 5.27. Exploration and functional analysis of perturbed metabolites in degenerative aortic stenosis. Perturbation patterns of the metabolites mapped to (A) phosphatidylinositol signalling system and (B) taurine hypotaurine metabolism pathways.

Pathway analysis of metabolites changed between RHD and AS

While comparing the metabolites detected between RHD and AS, 27 metabolites were significantly changed. The significantly changed metabolites are summarised in **Table 5.6**, showing their corresponding pooled QC RSD, fold changes, raw p value and FDR adjusted p value.

Table 5.6. Summary of metabolites significantly changed between AS and RHD indicating their corresponding pooled QC RSD (CV_{QC}), fold changes (FC and log2FC), crude and FDR adjusted p value.

Metabolite	CV_{QC}	FC	Log2(FC)	P Value	FDR adj. p
PE(20:4/0:0)	23	0.20	-2.30	<0.001	0.006
DG(a-15:0/i-12:0/0:0)	16	0.20	-2.30	0.002	0.030
Cortisol	30	0.31	-1.67	0.003	0.030
PE(22:6/0:0)	29	0.17	-2.59	0.003	0.030
Deoxycholate	26	0.05	-4.41	0.005	0.030
4-hydroxy-2-oxo-heptanedioic acid	19	0.24	-2.08	0.006	0.030
7alpha-hydroxy-3-oxo-4-cholestenoic acid	16	0.26	-1.95	0.007	0.033
Butyrylcarnitine	16	0.16	-2.62	0.008	0.033
PS(16:0/16:0)	21	0.32	-1.65	0.009	0.033
Valine	17	0.39	-1.35	0.012	0.037
10-nitro-9-octadecenoic acid	17	0.35	-1.50	0.013	0.037
Propionylcarnitine	0	0.26	-1.92	0.013	0.037
Cer(d18:1/22:1)	18	0.33	-1.59	0.014	0.038
Inosine	23	0.15	-2.73	0.017	0.042
PC(21:0/18:1)	26	0.24	-2.06	0.018	0.043
Ursodeoxycholic acid	26	0.07	-3.88	0.019	0.043
TG(15:0/o-18:0/15:0)	30	0.36	-1.46	0.020	0.043
TG(18:0/i-22:0/i-24:0)	25	0.18	-2.50	0.020	0.043
Acrylic acid	29	0.31	-1.69	0.027	0.052
3-Formylindole	0	0.26	-1.96	0.027	0.052
Acetoacetic acid	19	0.42	-1.25	0.030	0.055
Methyl-2-decene-4,6,8-triynoate	17	0.46	-1.11	0.031	0.055
PS(14:1/18:4)	29	0.18	-2.45	0.036	0.060
Cer(d18:0/h26:0)	16	0.04	-4.57	0.042	0.067
Creatine	0	0.11	-3.17	0.045	0.067
Indolepropionic acid	23	0.31	-1.67	0.049	0.067
5-Acetamidovalerate	16	0.27	-1.90	0.050	0.067

The hierarchical clustering of the perturbed metabolites between RHD and degenerative AS did not show a clear clustering of the studied groups However, there was a group of metabolites that were generally elevated together (**Figure 5.28**). The perturbed metabolites were impactfully mapped to glycerophospholipid metabolism, and to synthesis and degradation of ketone bodies however their mapping was not statistically significant after

adjusting for multiple testing (**Figure 5.29**). Phosphatidylcholine, lysophosphatidylcholine and phosphatidylserine were generally elevated in degenerative AS compared to RHD patients (**Figure 5.30**). Valine and acetoacetate were mapped to valine, leucine and isoleucine degradation pathway while acetoacetate was also associated with synthesis and degradation of ketone bodies (**Figure 5.31A & B**). Valine and acetoacetate were generally elevated in degenerative AS compared to RHD patients. The changed metabolites mapped to different pathways apart from the glycerophospholipid and linoleic acid metabolism pathways which were highlighted in both cardiac conditions (**Figure 5.32**).

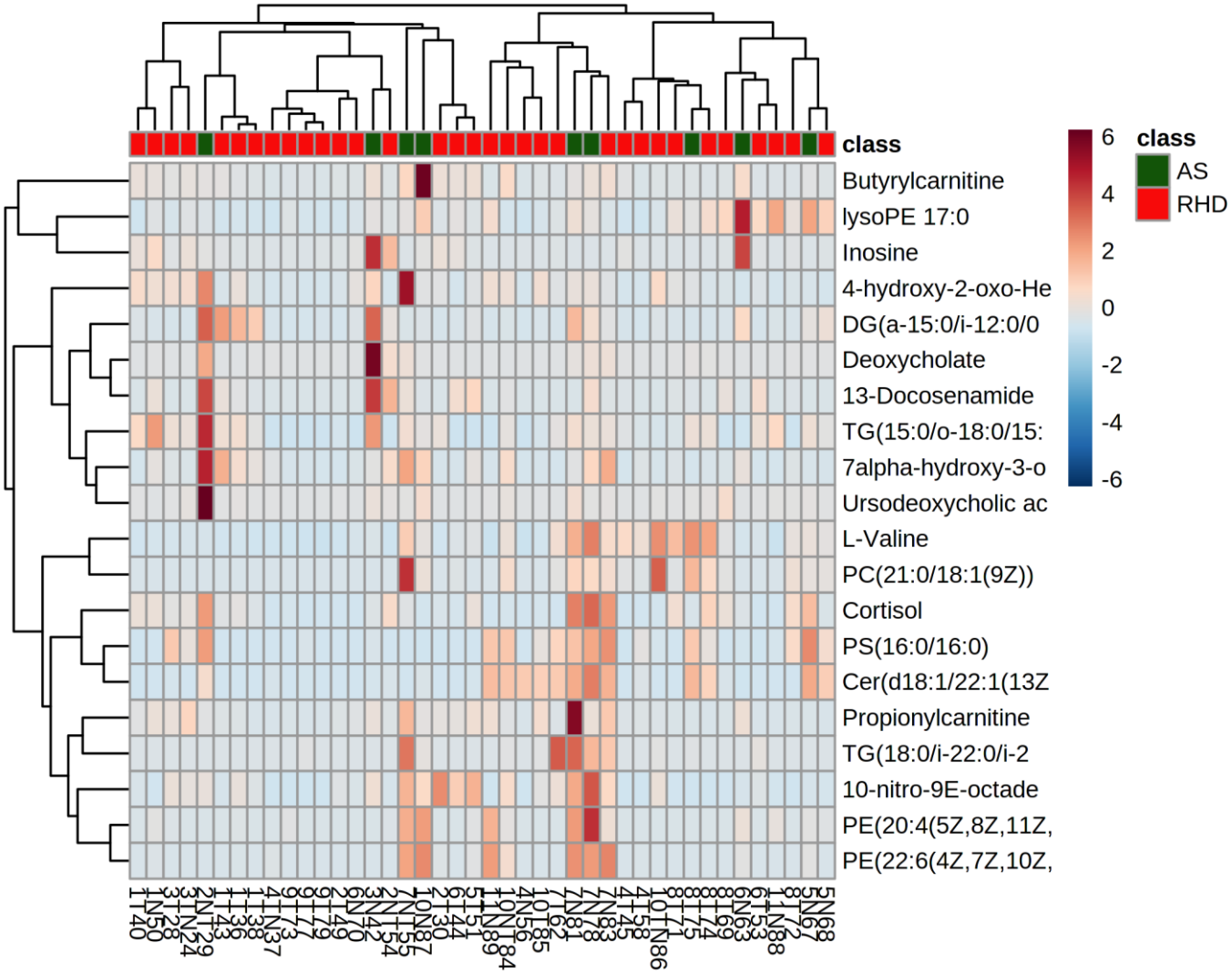


Figure 5.28. Exploration and functional analysis of perturbed metabolites in degenerative aortic stenosis. Heat map with hierarchical clustering of significantly changed metabolites in RHD and AS.

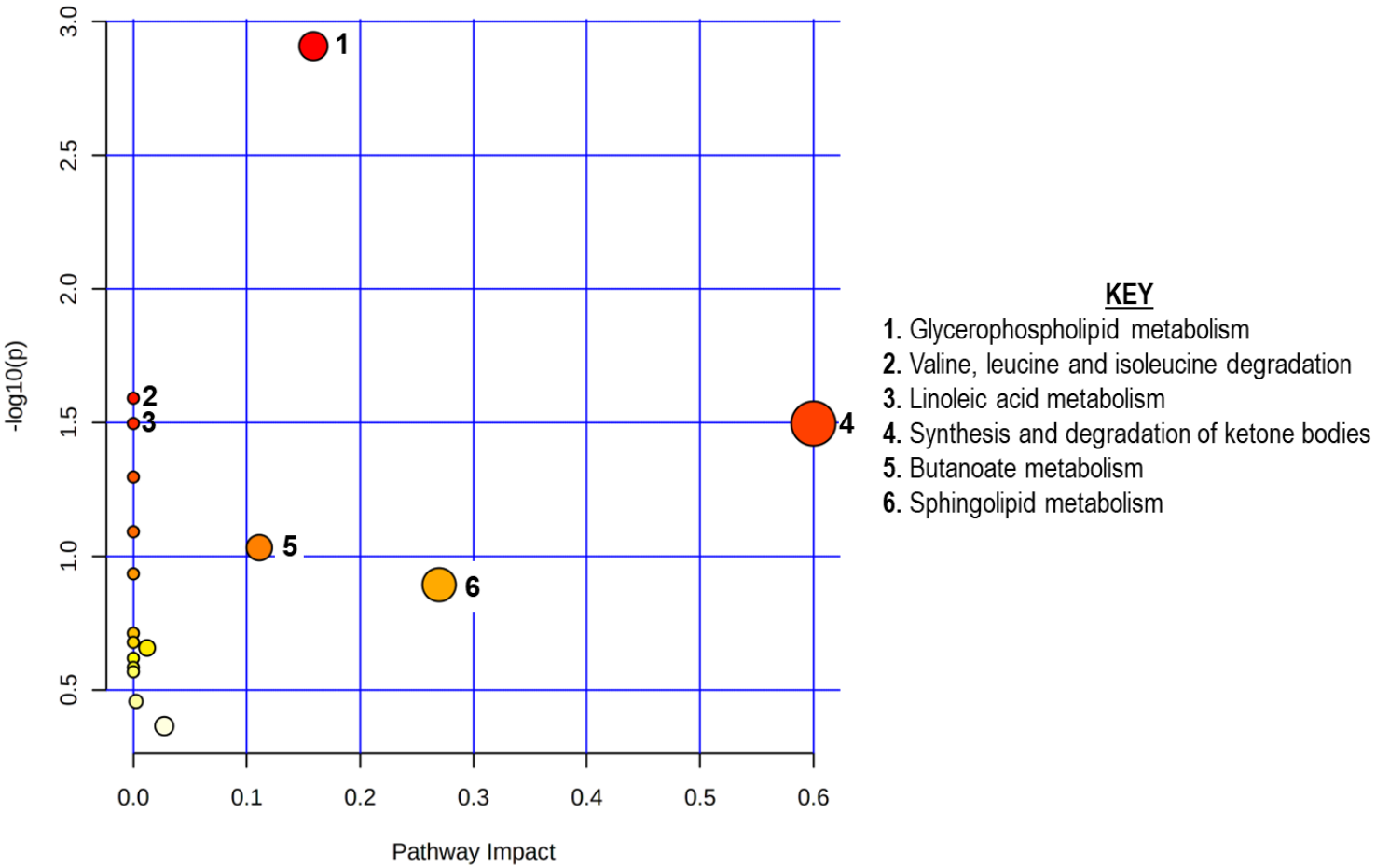


Figure 5.29. Exploration and functional analysis of perturbed metabolites in degenerative aortic stenosis. Pathway enrichment analysis map of the enriched pathways.

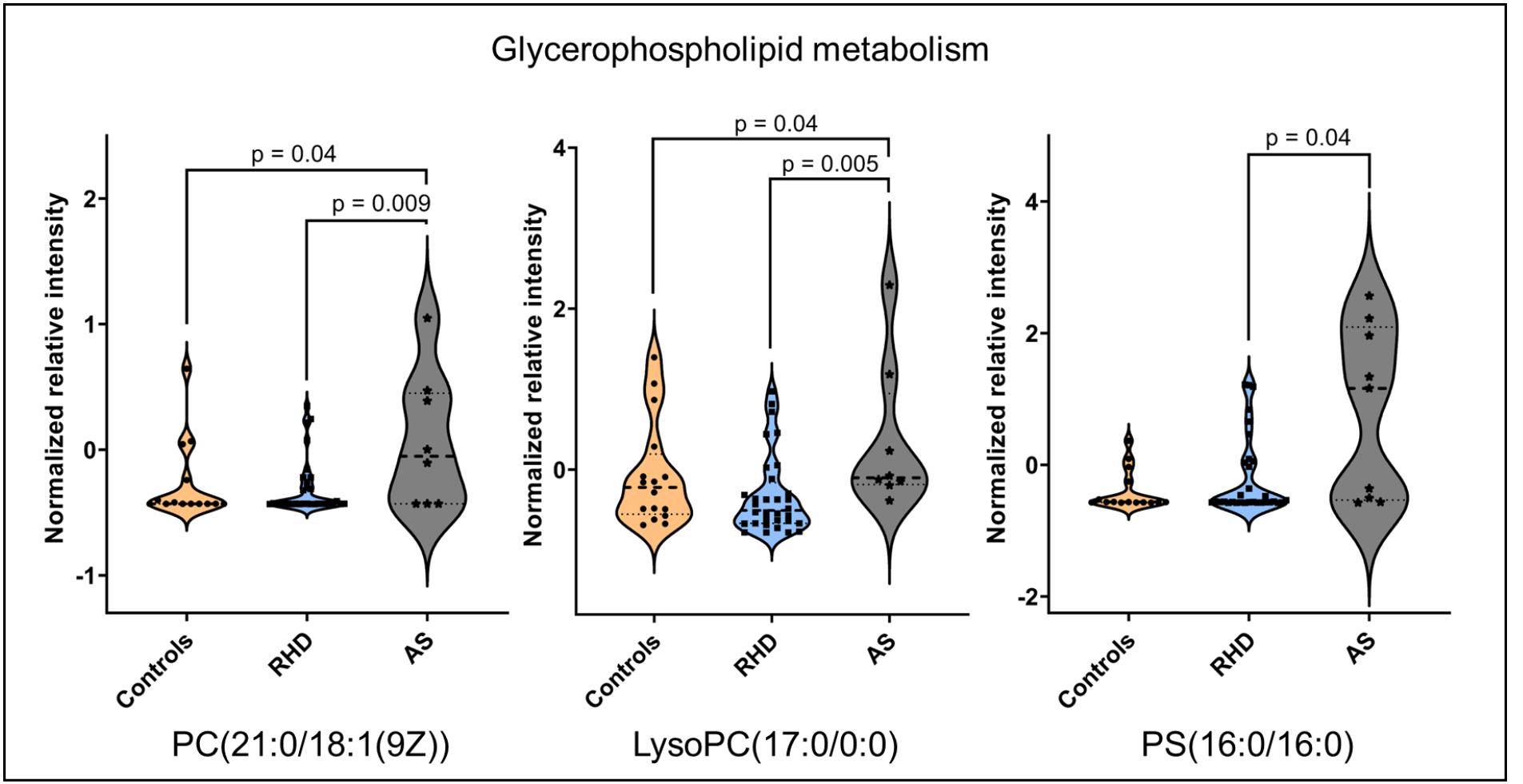


Figure 5.30. Exploration and functional analysis of perturbed metabolites in degenerative aortic stenosis. Perturbation patterns of the metabolites mapped to glycerophospholipid metabolism pathway.

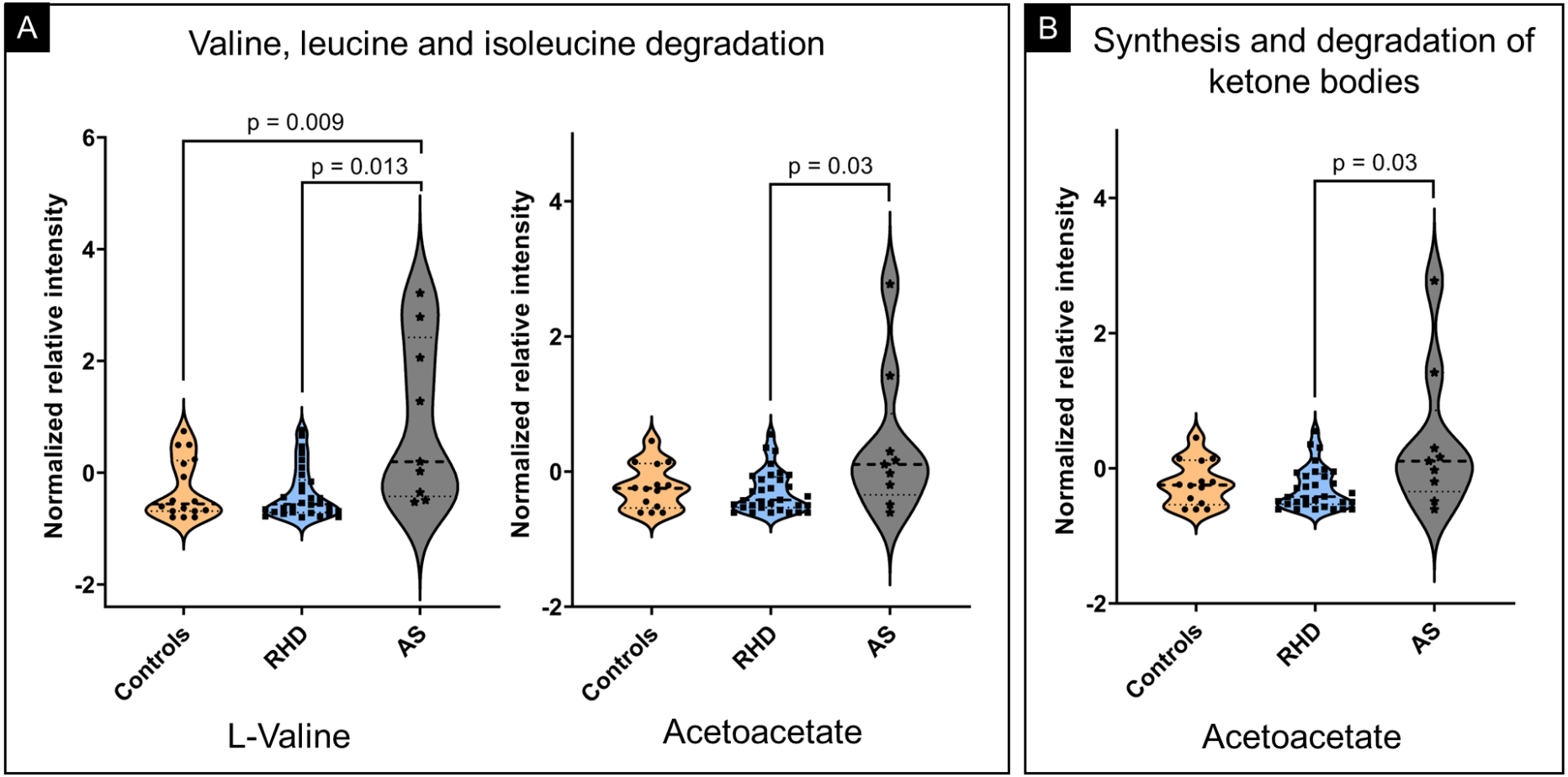


Figure 5.31. Exploration and functional analysis of perturbed metabolites in degenerative aortic stenosis. Perturbation patterns of the metabolites mapped to (A) valine, leucine and isoleucine degradation and (B) synthesis and degradation of ketone bodies pathways.

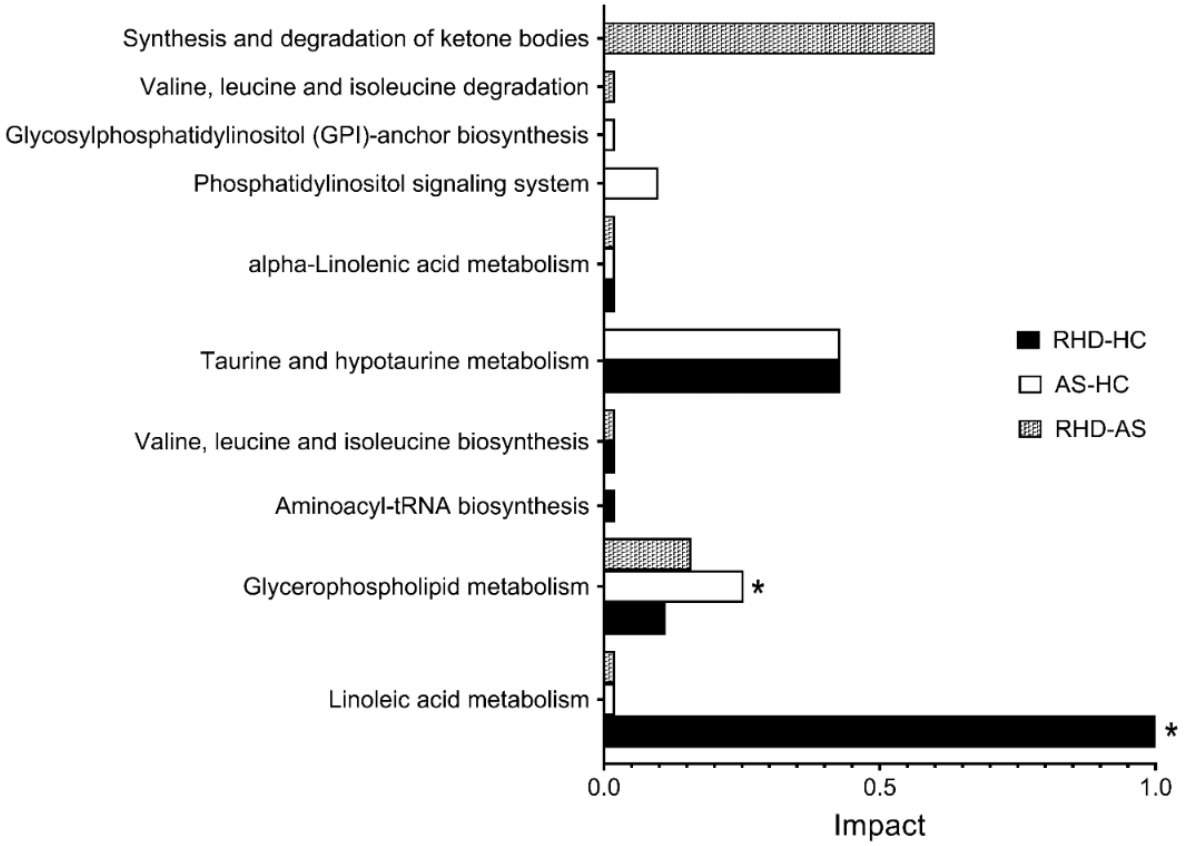


Figure 5.32. Top 5 pathways mapped by metabolites changed between RHD, AS, and healthy controls, *significantly mapped pathways (FDR adjusted p value <0.05).

Integrative gene-metabolites interactions analysis of metabolites changed between RHD, AS and controls

To further explore the functional relevance of the dysregulated metabolites in RHD, AS, and controls, gene-metabolite interaction was performed based on theoretical data of gene-metabolite connections. Important genes (circles) and metabolites (squares) were selected based on the links between them (degree) and the number of shortest paths going through the nodes (betweenness centrality) (**Figure 5.33**).

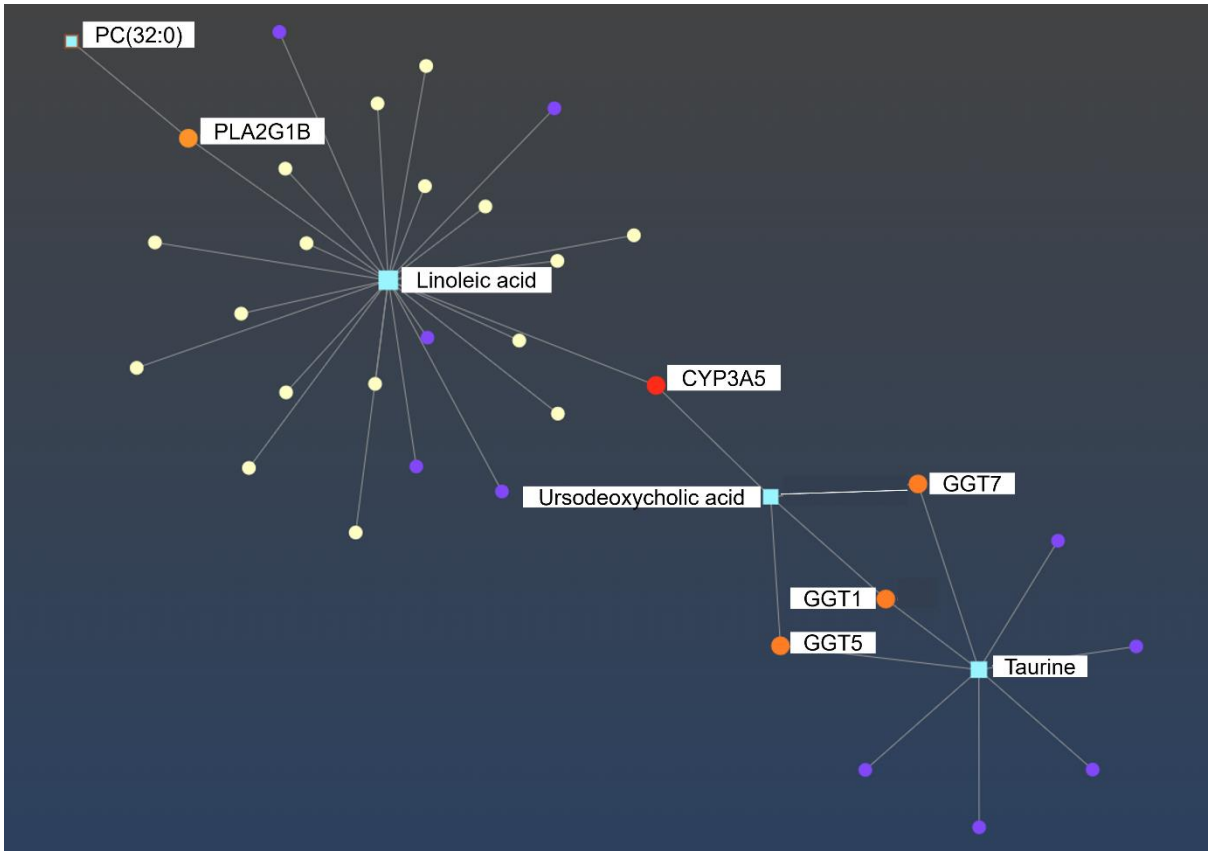


Figure 5.33. Network exploration of Gene-metabolite interactions of dysregulated metabolites in RHD. Important nodes are annotated; genes (circles) and metabolites (squares) with 2 or more links to neighbouring nodes (degree≥2).

Based on the univariate analyses, phosphatidylcholine, linoleic acid, ursodeoxycholic acid and taurine were depicted as some of the important metabolites and were centrally positioned in the gene-metabolite interaction network **Table 5.7**. The metabolites measured in the study were correlated to all the reported genes in literature that have been shown to control expression of enzymes involved in the pathways associated to the metabolites. The correlation intended to show genes and metabolites that would be important in the perturbations of the reported biomarkers. The findings indicate that, CYP3A5, GGT1, GGT7, GGT5, and PLA2G1B were some of the genes that showed a strong interaction with the important metabolites based on their number of interconnections and high number of short paths between nodes (**Figure 5.34 and Table 5.7**). The highlighted genes are potential targets for further experiments to validate the perturbed metabolites in RHD patients.

Table 5.7. RHD gene-metabolites interaction summary showing the degree centrality and betweenness of the metabolites and the interacting genes.

ID	Label	Degree	Betweenness
C01595	Linoleic acid	24	539
C00245	Taurine	8	161.5
C07880	Ursodeoxycholic acid	4	235.5
1577	CYP3A5	2	250
2678	GGT1	2	54
2686	GGT7	2	54
2687	GGT5	2	54
5319	PLA2G1B	2	34
C00157	PC(16:0/16:0)	1	0

The integrative analysis of the metabolites dysregulated between AS and healthy controls formed 2 subnetworks from the interacting genes (**Figure 5.34 and Table 5.8**). Bilirubin, taurine, and hydrocortisone were some of the main metabolites forming networks with GGT1, GGT5, GGT7 and GAD1.

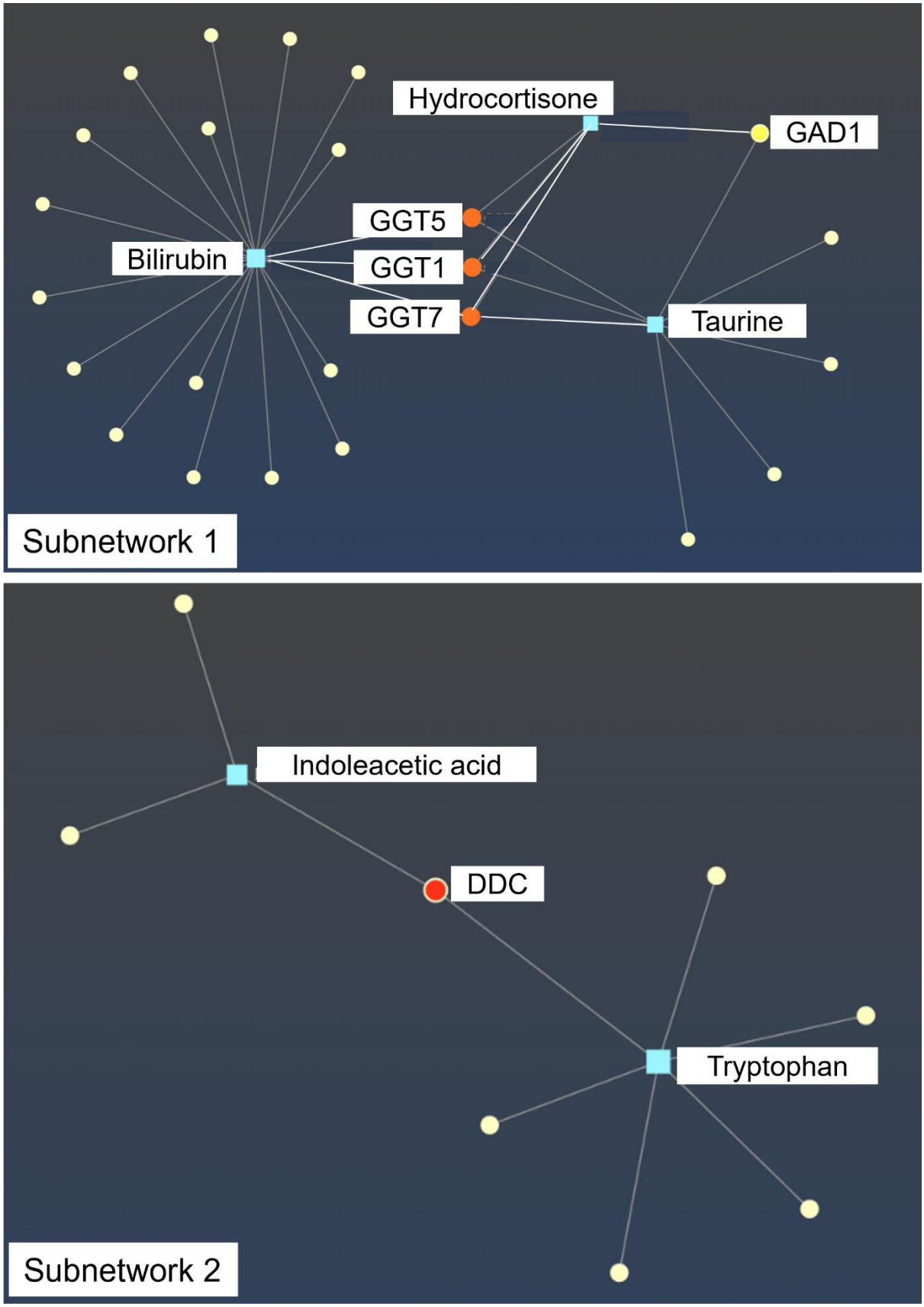


Figure 5.34. Network exploration of Gene-metabolite interactions of dysregulated metabolites in AS. Important nodes are annotated; genes (circles) and metabolites (squares) with 2 or more links to neighbouring nodes (degree≥2).

Furthermore, the subnetwork 2 showed indoleacetic acid and L-Tryptophan as one of the important metabolites interacting with DDC gene in the network (**Figure 5.34 and Table 5.8**). The DDC gene controls expression of aromatic L-amino acid decarboxylase (AADC) which catalyses biosynthesis of dopamine and serotonin and metabolism of L-tryptophan to tryptamine. Serotonin has been reported to play an important role in the control of human cardiac contractile function, in addition metabolism of amino acids have been associated with proinflammatory processes in atherosclerosis.[178,179]

Table 5.8. Aortic stenosis gene-metabolites interaction summary showing the degree centrality and betweenness of the metabolites and the interacting genes for subnetwork 1 and 2.

ID	Label	Degree	Betweenness
Subnetwork 1			
C00486	Bilirubin	20	307
C00245	Taurine	8	109.5
C00735	Hydrocortisone	4	11.5
2678	GGT1	3	43.25
2686	GGT7	3	43.25
2687	GGT5	3	43.25
2571	GAD1	2	1.25
Subnetwork 2			
C00078	Tryptophan	6	30
C00954	Indoleacetic acid	3	15
1644	DDC	2	18

The metabolites that were perturbed between AS and RHD were analysed to determine their metabolite-gene interactions profile (**Figure 5.35**). The genes for metabolite-gene interactions were obtained from the MetaCyc database.

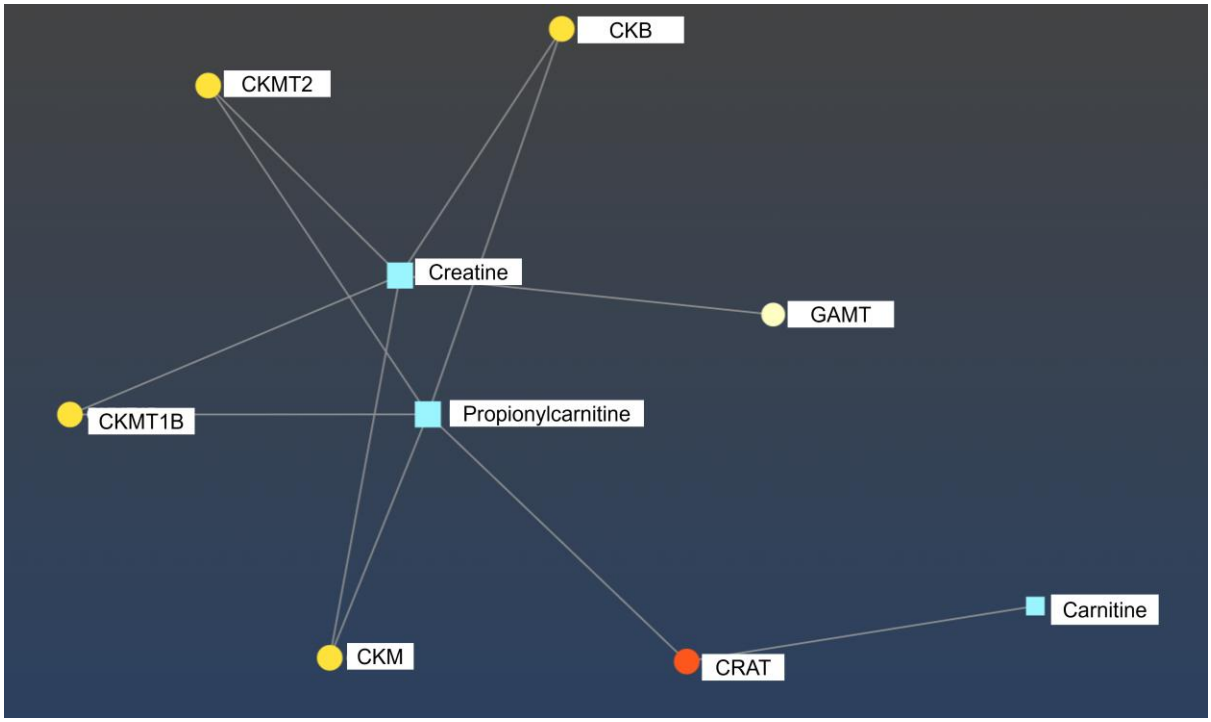


Figure 5.35. Network exploration of Gene-metabolite interactions of dysregulated metabolites between AS and RHD. Important nodes are annotated; genes (circles) and metabolites (squares) with links to neighbouring nodes.

The network analysis indicated propionylcarnitine, creatine and carnitine as some of the important metabolites with interactions with genes CRAT, CKM, CKMT1B, CKMT2, CKB and GAMT (**Figure 5.35 and Table 5.9**).

Table 5.9. Aortic stenosis and RHD gene-metabolites interaction summary showing the degree centrality and betweenness of the metabolites and the interacting genes.

ID	Label	Degree	Betweenness
C03017	Propionylcarnitine	5	15
C00300	Creatine	5	10
1384	CRAT	2	7
1158	CKM	2	1.5
1160	CKMT2	2	1.5
1152	CKB	2	1.5
1159	CKMT1B	2	1.5
2593	GAMT	1	0
C00318	Carnitine	1	0

5.3.4 Discriminant analysis of perturbed metabolites in AS, RHD and controls while adjusting for covariates

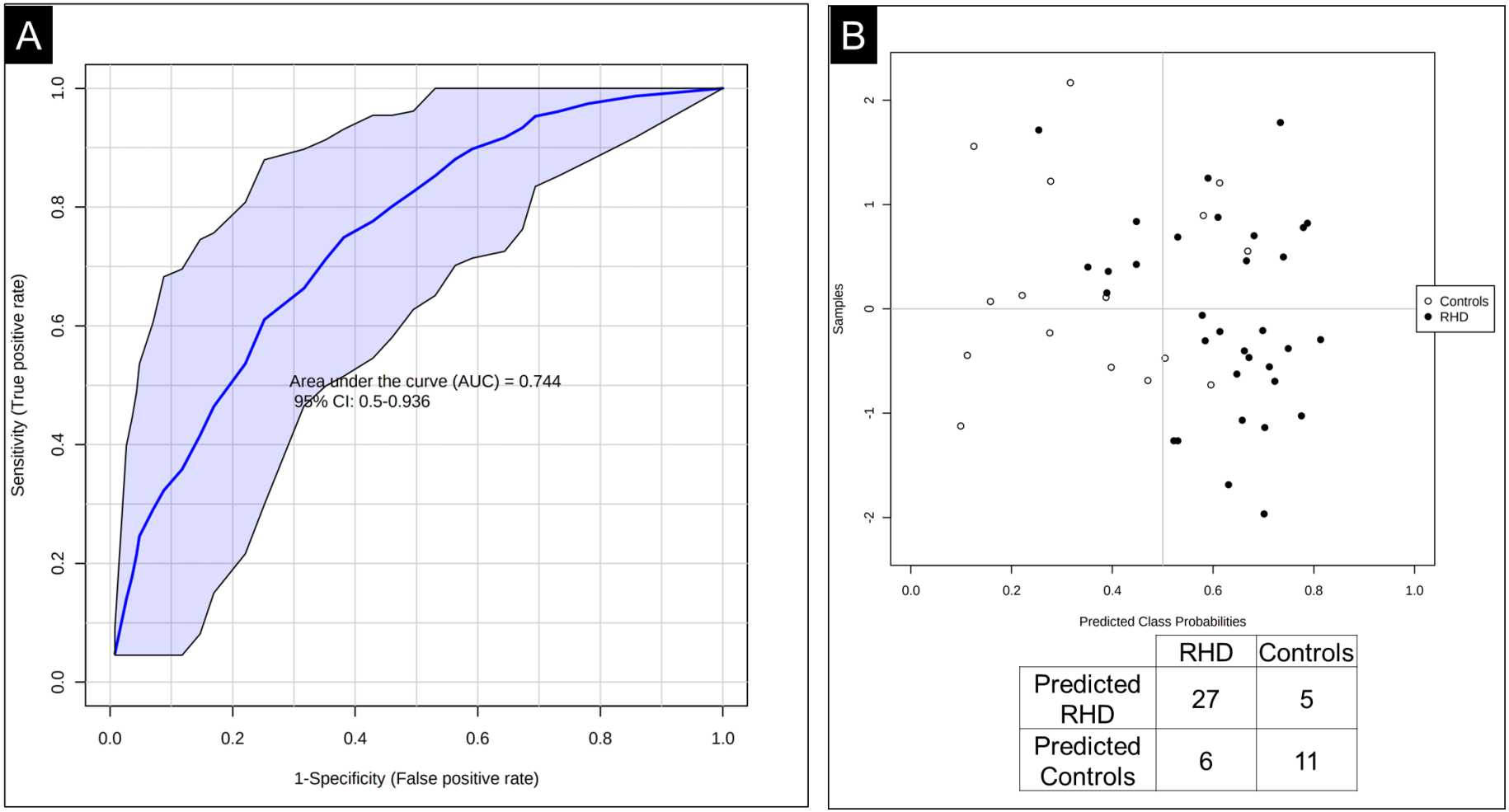
Rheumatic heart disease and healthy controls

To obtain metabolites capable of differentiating RHD from healthy controls, a univariate ROC-AUC analysis showed 14 metabolites with good predictive capabilities. Of the 14 metabolites with discriminant capabilities, 7 metabolites were generally not affected by sex, age, race,

BMI, hypertension, diabetes, or the batch effects (covariates adj. p <0.05) **Table 5.10**. Furthermore, a multivariate ROC-AUC analysis showed a group of 5 metabolites as having the best discriminative model, AUC= 0.74 [0.5-0.94] **Figure 5.36A**. Based on the predicted class probabilities analysis of the 5-metabolites model, it had a good sensitivity, 0.82[0.69-0.95] and specificity 0.69[0.46-0.91] **Figure 5.36B**. Ursodeoxycholic acid, Ceramide(18:0/26:0), and dihydroxybutanoic acid were robust predictive biomarkers of RHD and controls since they had >60% chance of being selected when building the model and were generally not affected by the covariates (**Figure 5.37 and Table 5.10**).

Table 5.10. Serum metabolites significantly changed and capable of differentiating RHD from controls and their significance levels re-tested after adjusting for covariates (sex, age, race, BMI, hypertension, diabetes, and batch effects) using multiple linear regression model, (*metabolites that remained significant after covariates adjustments and with AUC ≥0.7).

Metabolite	p value	FDR adj. p	Covariates adj. p	AUC
PE(20:4/0:0)	**0.001**	**0.022**	**0.036**	**0.76***
3-Formylindole	**0.001**	**0.022**	**0.003**	**0.74***
PE(22:6/0:0)	0.002	0.022	0.077	0.71
Tryptophan	**0.003**	**0.029**	**0.033**	**0.74***
MG(0:0/20:3/0:0)	**0.006**	**0.046**	**0.018**	**0.74***
LysoPC(18:1/0:0)	0.013	0.046	0.101	0.74
2,6-nonadienoic acid	0.013	0.046	0.081	0.73
Ursodeoxycholic acid	**0.014**	**0.046**	**0.031**	**0.76***
Crotonic acid	0.017	0.048	0.087	0.71
LysoPE(0:0/22:5)	0.017	0.048	0.056	0.73
2,4-dihydroxy-butanoic acid	**0.020**	**0.048**	**0.023**	**0.82***
Tetradecanoylcarnitine	0.028	0.049	0.073	0.78
Cer(d18:0/h26:0)	**0.029**	**0.049**	**0.008**	**0.76***
PC(18:4/P-18:1)	0.030	0.049	0.077	0.71



	RHD	Controls
Predicted RHD	27	5
Predicted Controls	6	11

Figure 5.36. Discriminatory metabolites between RHD and HC. (A) multivariate area under - receiver operator curve (AUC-ROC) model based on 5 metabolites. (B) The diagnostic model predicted class probabilities classification of the RHD and controls.

.

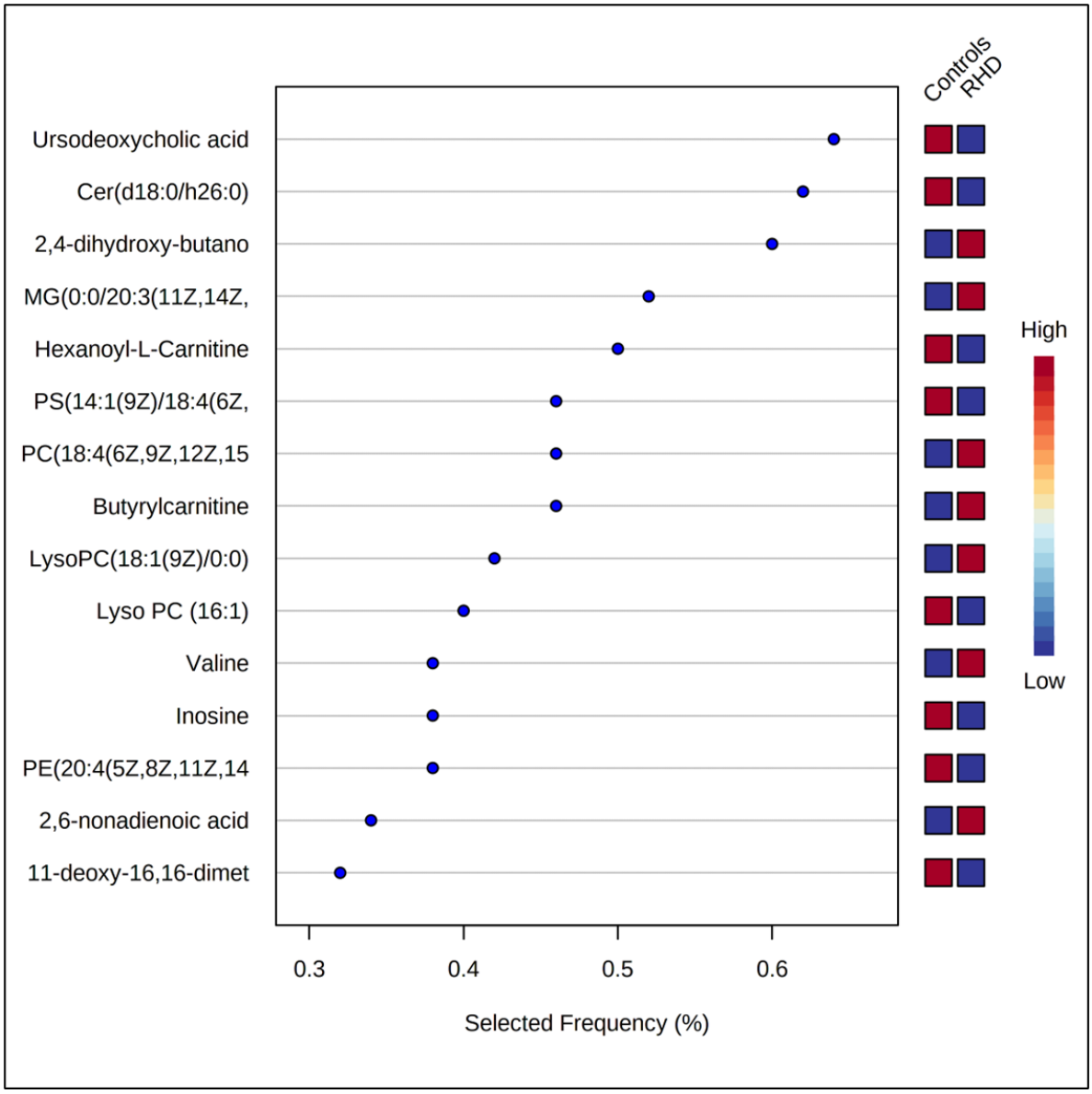


Figure 5.37. Discriminatory metabolites between RHD and HC. 15 metabolites with best predictive capabilities based on their frequency of selection to build the AUC-ROC multivariate model.

Degenerative aortic stenosis and healthy controls

In addition, metabolites capable of discriminating AS from controls were analysed where 14 metabolites were potential diagnostic biomarkers (ROC-AUC>0.7) **Table 5.11**. Of the 14 discriminatory metabolites, 4 metabolites were shown to be robust biomarkers that remained significantly different after adjusting for sex, age, race, BMI, hypertension, diabetes, and batch effects (covariates adj. p<0.05) **Table 5.11**. A multivariate diagnostic analysis showed that a set of 10 metabolites had the best capability of forming a diagnostic model (AUC=0.88) **Figure 5.38A**. Furthermore, the model's predicted class probabilities for classifying degenerative AS

and controls had very good predictive capabilities as shown by the sensitivity 1[1-1] and specificity 0.75[0.54-0.96] **Figure 5.38B**. Of the 10 metabolites forming the multivariate predictive model, PE(P-16:0/22:11) palmitoylcarnitine, isocaproic acid, taurine and lysoPC were best model performers based on their frequency of selection to form the predictive model (**Figure 5.39**). Besides their performance into the predictive model, palmitoylcarnitine, isocaproic acid, and taurine had also shown to be robust discriminating biomarkers after adjusting for the covariates (**Table 5.11**).

Table 5.11. Serum metabolites significantly changed and capable of differentiating AS from controls and their significance levels re-tested after adjusting for covariates (sex, age, race, BMI, hypertension, diabetes, and batch effects) using multiple linear regression model, (*metabolites that remained significant after covariates adjustments and with AUC ≥0.7).

Metabolite	p value	FDR adj. p	Covariates adj. p	AUC
PE(P-16:0/22:1)	0.002	0.056	0.513	0.79
PA(8:0/13:0)	0.007	0.093	0.109	0.67
Crotonic acid	0.010	0.093	0.179	0.78
Undecylenic acid	0.015	0.093	0.514	0.74
Palmitoylcarnitine	**0.020**	**0.093**	**0.009**	**0.74***
Phenylacetylglutamine	0.022	0.093	0.485	0.65
Isocaproic acid	**0.023**	**0.093**	**0.007**	**0.78***
Taurine	**0.024**	**0.093**	**0.006**	**0.68***
trans-Hexadec-2-enoyl carnitine	0.028	0.093	0.571	0.78
Butyrylcarnitine	0.030	0.093	0.141	0.77
PI(15:1/22:6)	0.035	0.093	0.428	0.75
Ornithine	**0.038**	**0.093**	**0.008**	**0.81***
2,4-Hexadienyl acetate	0.042	0.093	0.880	0.72
2-Hydroxy-4-(methylthio)butanoic acid	0.046	0.093	0.701	0.72

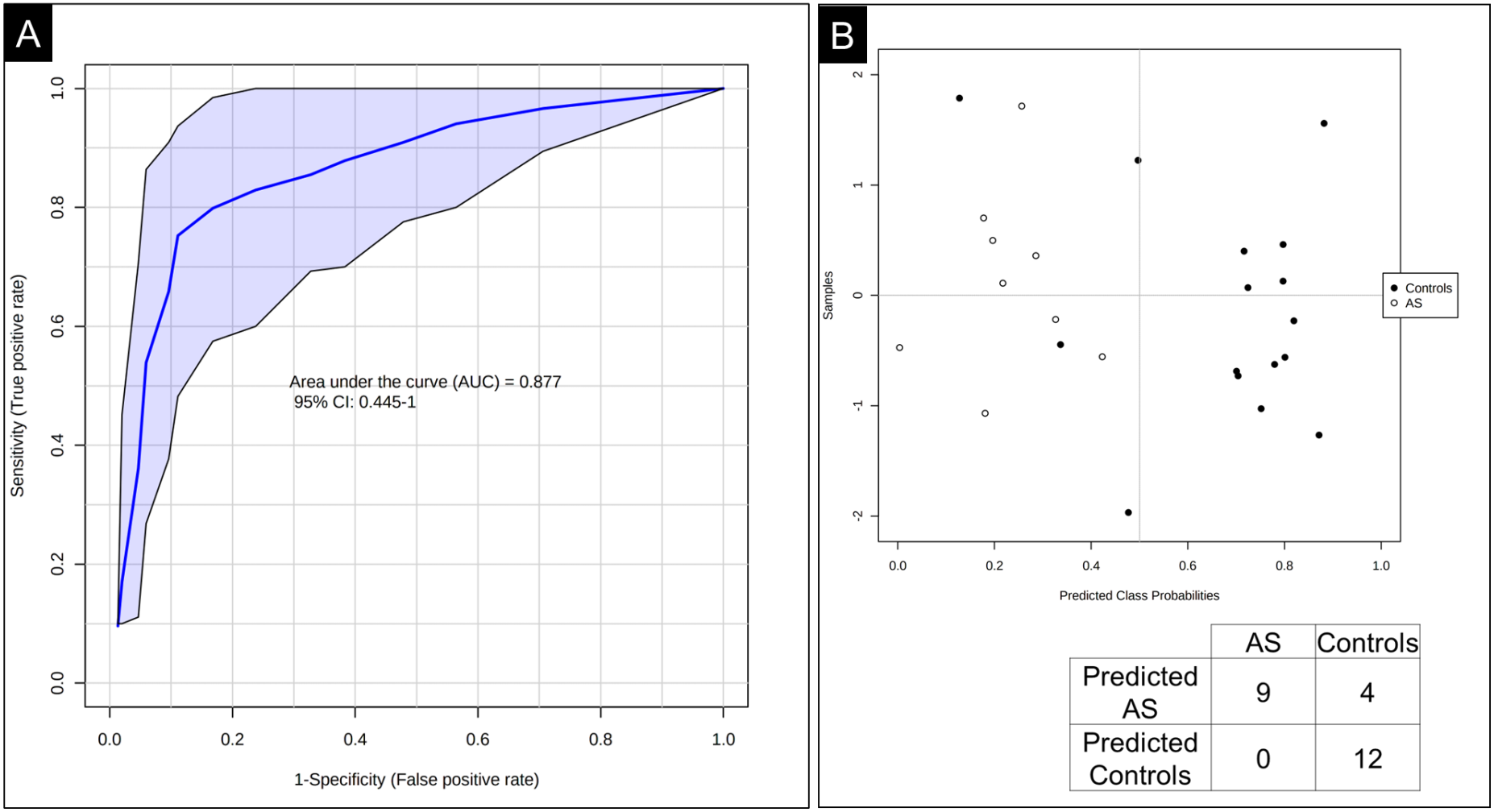


Figure 5.38. Discriminatory metabolites between AS and HC. (A) multivariate area under - receiver operator curve (AUC-ROC) model based on 10 metabolites. (B) The diagnostic model predicted class probabilities classification of the AS and controls.

.

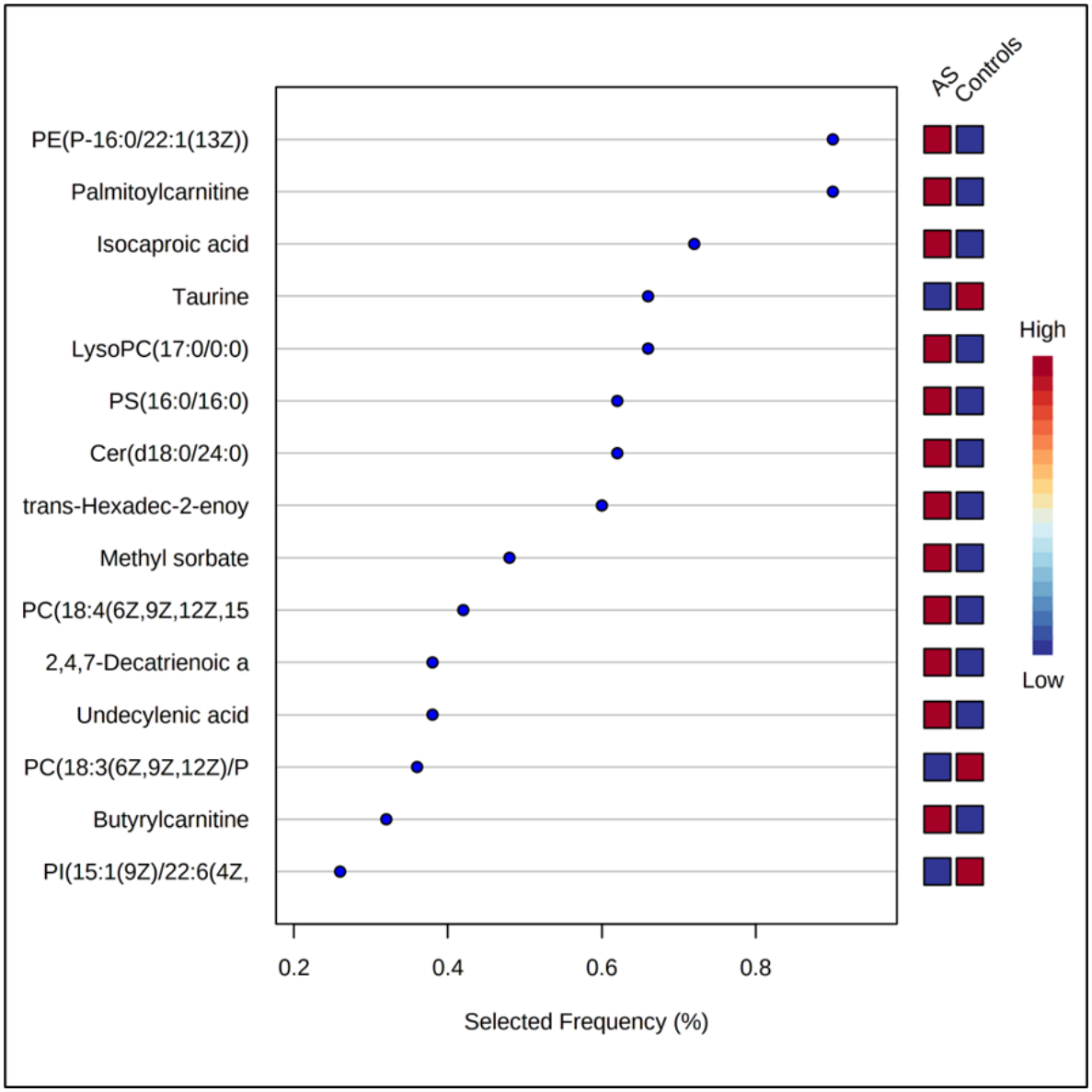


Figure 5.39. Discriminatory metabolites between AS and HC. 15 metabolites with best predictive capabilities based on their frequency of selection to build the AUC-ROC multivariate model.

Degenerative aortic stenosis and rheumatic heart disease

To investigate whether the metabolites would differentiate between degenerative AS and RHD, the metabolites changed between the two groups were analysed with univariate and multivariate AUC-ROC analysis. The univariate discriminative analysis indicated 20 metabolites as being potential biomarkers to differentiate AS from RHD patients. Of the 20 metabolites, 8 metabolites remained significantly different after adjusting for age, sex, race, BMI, hypertension, diabetes, and batch effects (covariates adj. p <0.05) **Table 5.12**. The multivariate ROC-AUC analysis showed a set of 10 metabolites as having the capabilities of

classifying the RHD and AS however the model had a large confidence interval AUC = 0.76, [95% CI 0.32 – 1] **Figure 5.40A**. A large confidence interval indicates that the AUC have lower accurate inference in the group classification and cannot be reliable in differentiating the groups. In future, similar studies should include large sample sizes to validate the reported biomarkers. Despite the limitation, based on the predicted class probabilities, the model had a very good sensitivity 0.82, [95% CI 0.69-0.95] and specificity 0.89, [95% CI 0.68 – 1] **Figure 5.40B**. Of the 10 metabolites selected to build the predictive model, acetylcarnitine, PS(32:0), lysoPE(17:0), and butyrylcarnitine were among the best predictors based on their frequency of selection to build the predictive model (**Figure 5.41**).

Table 5.12. Summary of metabolites significantly changed and capable of differentiating AS from RHD and their significance levels re-tested after adjusting for covariates (sex, age, race, BMI, hypertension, diabetes, and batch effects) using multiple linear regression model, (*metabolites that remained significant after covariates adjustments and with AUC ≥0.7).

Metabolite	p value	FDR adj. p	Covariates adj. p	AUC
PE(20:4/0:0)	**<0.001**	**0.006**	**0.008**	**0.78***
DG(a-15:0/i-12:0/0:0)	**0.002**	**0.030**	**0.044**	**0.81***
Cortisol	**0.003**	**0.030**	**0.014**	**0.74***
PE(22:6/0:0)	0.003	0.030	0.313	0.76
Deoxycholate	0.005	0.030	0.208	0.85
7alpha-hydroxy-3-oxo-4-cholestenoic acid	**0.007**	**0.033**	**0.048**	**0.75***
Butyrylcarnitine	0.008	0.033	0.593	0.82
PS(32:0)	**0.009**	**0.033**	**0.038**	**0.72***
Valine	**0.012**	**0.037**	**0.005**	**0.77***
10-nitro-9-octadecenoic acid	0.013	0.037	0.291	0.72
Propionylcarnitine	0.013	0.037	0.355	0.73
Cer(d18:1/22:1)	0.014	0.038	0.138	0.71
PC(21:0/18:1)	0.018	0.043	0.057	0.78
TG(15:0/o-18:0/15:0)	0.020	0.043	0.148	0.79
TG(18:0/i-22:0/i-24:0)	0.020	0.043	0.093	0.79
3-Formylindole	0.027	0.052	0.979	0.72
Acetoacetic acid	**0.030**	**0.055**	**0.004**	**0.73***
Methyl-2-decene-4,6,8-triynoate	0.031	0.055	0.201	0.78
Cer(d18:0/h26:0)	0.042	0.067	0.138	0.78
Creatine	**0.045**	**0.067**	**0.002**	**0.75***

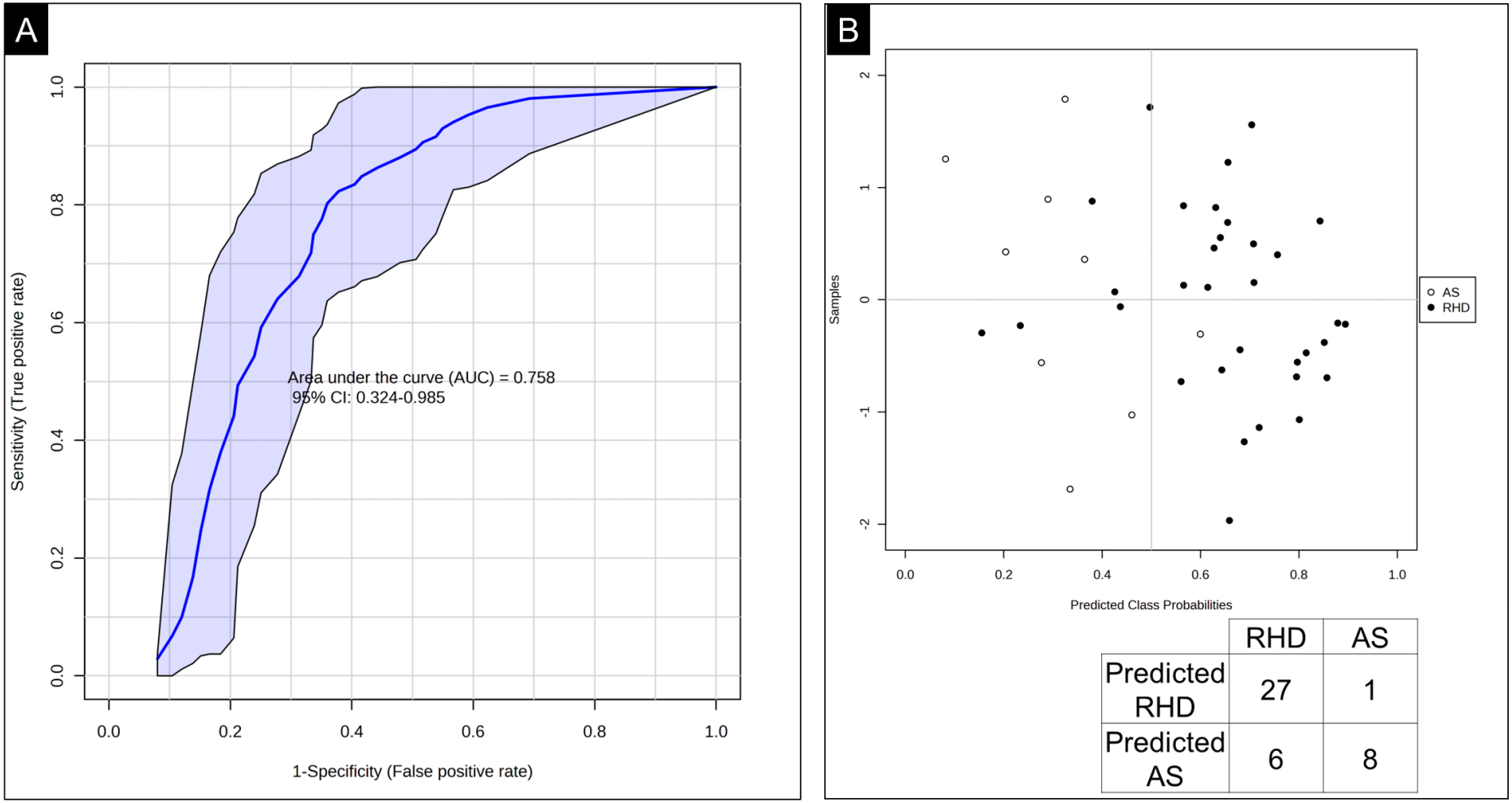


Figure 5.40. Discriminatory metabolites between AS and RHD. (A) multivariate area under - receiver operator curve (AUC-ROC) model based on 10 metabolites. (B) The diagnostic model predicted class probabilities classification of the AS and RHD.

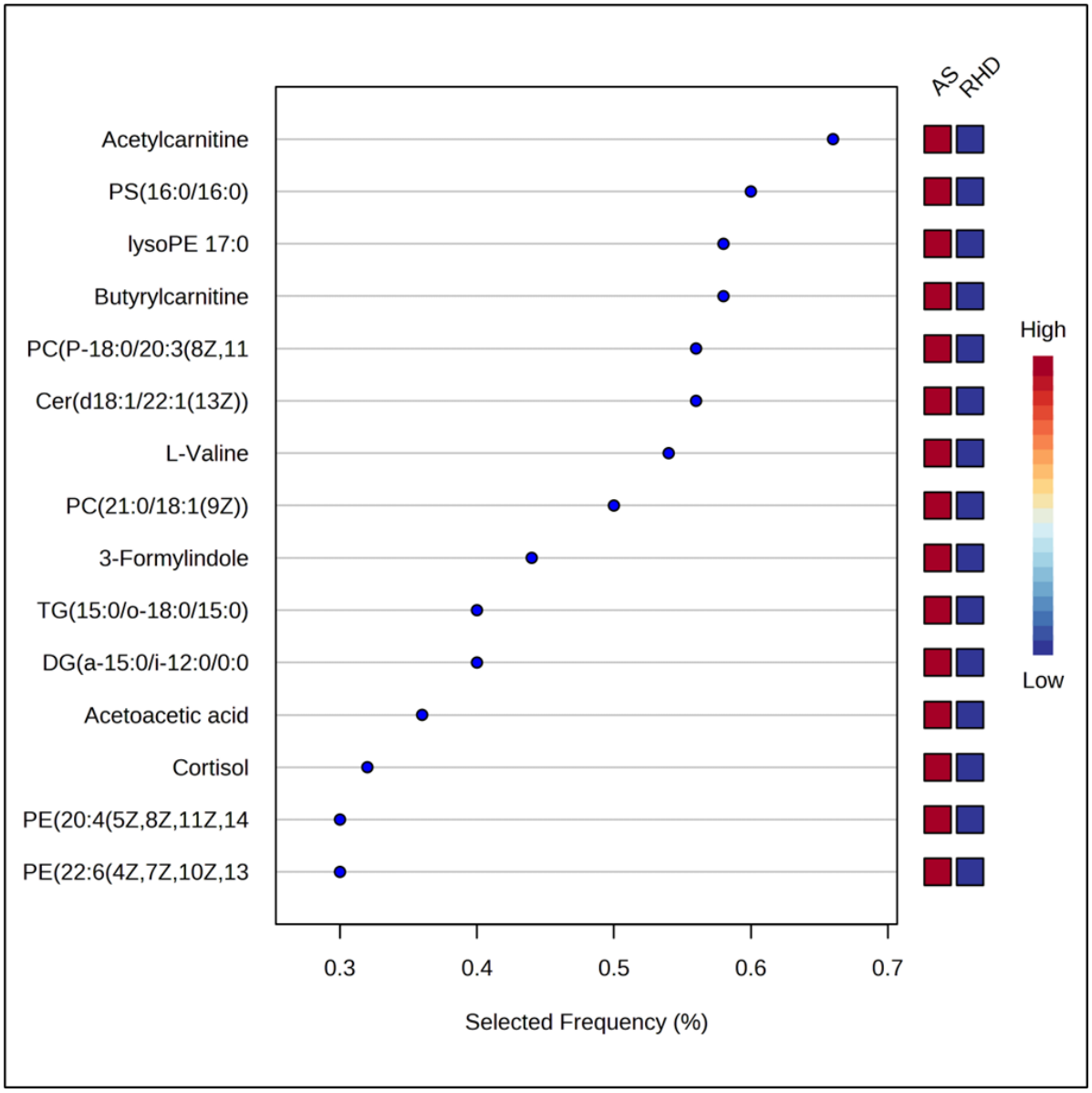


Figure 5.41. Discriminatory metabolites between AS and RHD. 15 metabolites with best predictive capabilities based on their frequency of selection to build the multivariate AUC_ROC model.

5.3.5 Correlation of cardiac remodeling parameters with changed metabolites in RHD and degenerative AS

Some of the metabolites changed between RHD and AS showed association with the CMR parameters of cardiac remodeling (LVMI, LA area and LVEF). Of the 26 metabolites different between RHD and AS, 7 metabolites showed a significant association with the CMR cardiac remodeling parameters while adjusting for covariates (**Table 5.13**).

Table 5.13. Summary of metabolites significantly changed between RHD and degenerative AS and their Pearson's correlation analyses with LVEF, LVMI and LA area after adjusting for sex, age, race, BMI, hypertension, diabetes, and batch effects.

Metabolite	p value	FDR adj. p	Covariates adj. p	LVEF (r)	p value	LVMI (r)	p value	LA area (r)	p value
PE (20:4/0:0)	**<0.001**	**0.006**	**0.008**	0.24	0.162	-0.10	0.575	**-0.48**	**0.005**
PE (22:6/0:0)	0.003	0.030	0.312	0.22	0.202	-0.19	0.278	**-0.34**	**0.047**
Acrylic acid	**0.027**	**0.052**	**0.001**	0.30	0.083	**0.35**	**0.039**	-0.22	0.196
3-Formylindole	0.027	0.050	0.979	**-0.37**	**0.028**	-0.25	0.145	0.05	0.759
PS (14:1/18:1)	0.036	0.060	0.653	**-0.41**	**0.015**	-0.02	0.887	0.08	0.637
Creatine	**0.045**	**0.067**	**0.002**	0.31	0.071	**0.34**	**0.045**	-0.20	0.253
Indolepropionic acid	0.049	0.067	0.231	-0.20	0.255	**-0.33**	**0.053**	0.16	0.357

Specifically, LVMI showed a weak positive association with acrylic acid (r= 0.35, p = 0.039) and creatine (r = 0.34, p = 0.045), but there was a negative association with indolepropionic acid (r = -0.33, p = 0.053). In addition, LVEF was negatively associated with the 3-formylindole (r = -0.37, p = 0.028) and PS (14:1/18:1) (r = -0.41, p = 0.015). The LA area showed a negative association with the phosphatidylserine molecular species metabolites**.** The described correlations are summarised is scatter plots represented in **Figure 5.42 to Figure 5.45**.

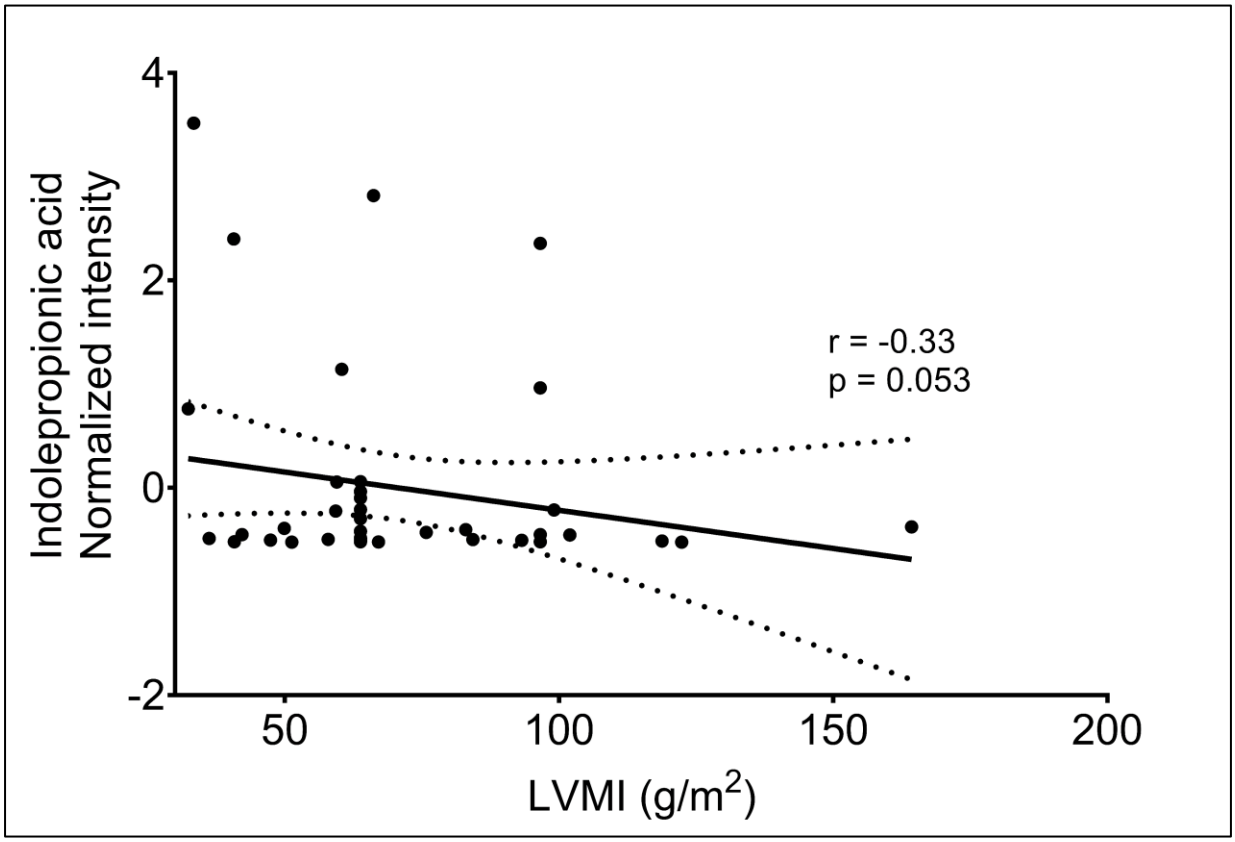


Figure 5.42. Correlation of the left ventricular mass index (LVMI) with indolepropionic acid in RHD and degenerative AS while adjusting for sex, age, race, BMI, hypertension, diabetes, and batch effects.

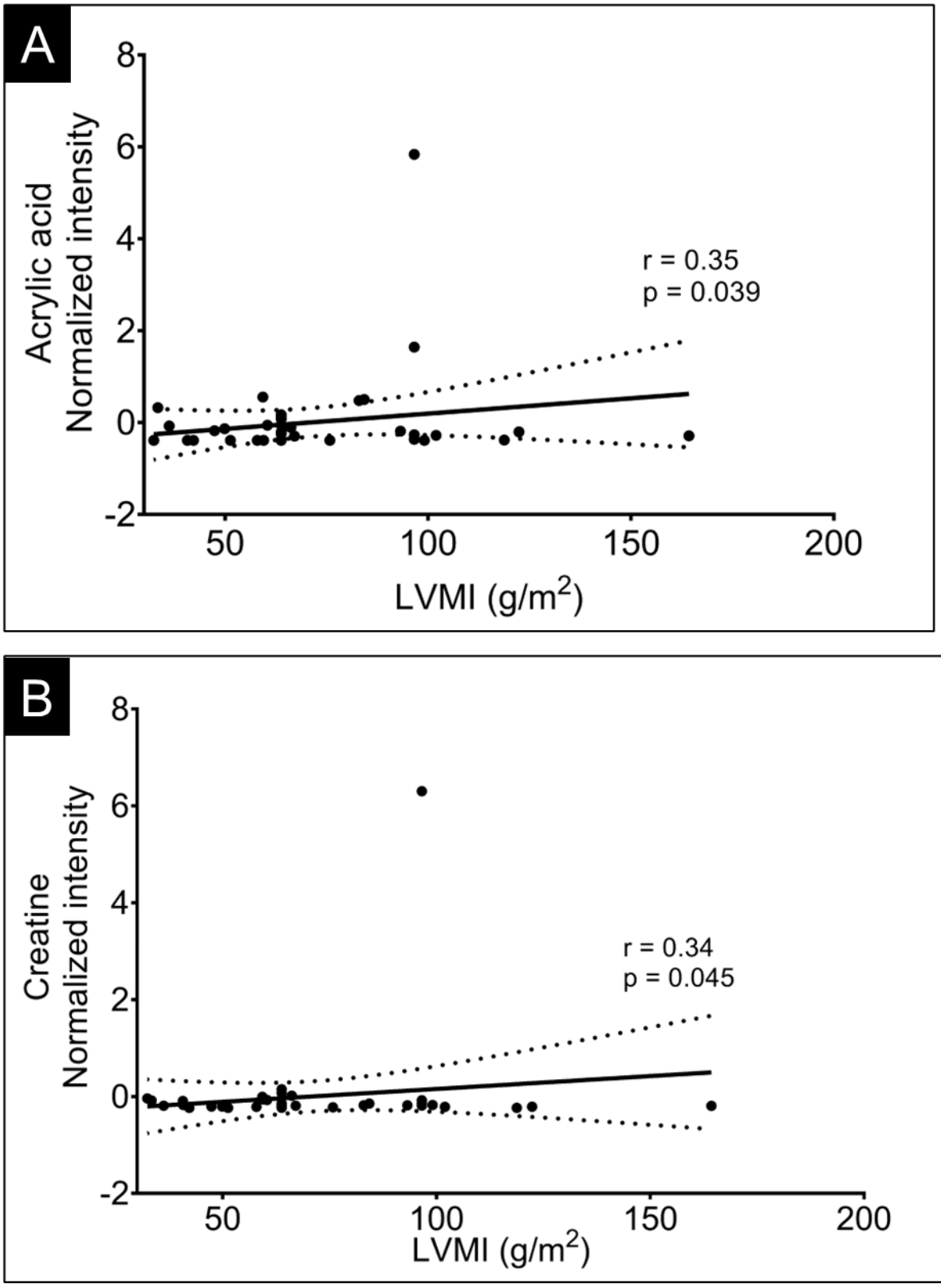


Figure 5.43. Correlation of LVMI with (A) acrylic acid and (B) creatine in RHD and degenerative AS while adjusting for sex, age, race, BMI, hypertension, diabetes, and batch effects.

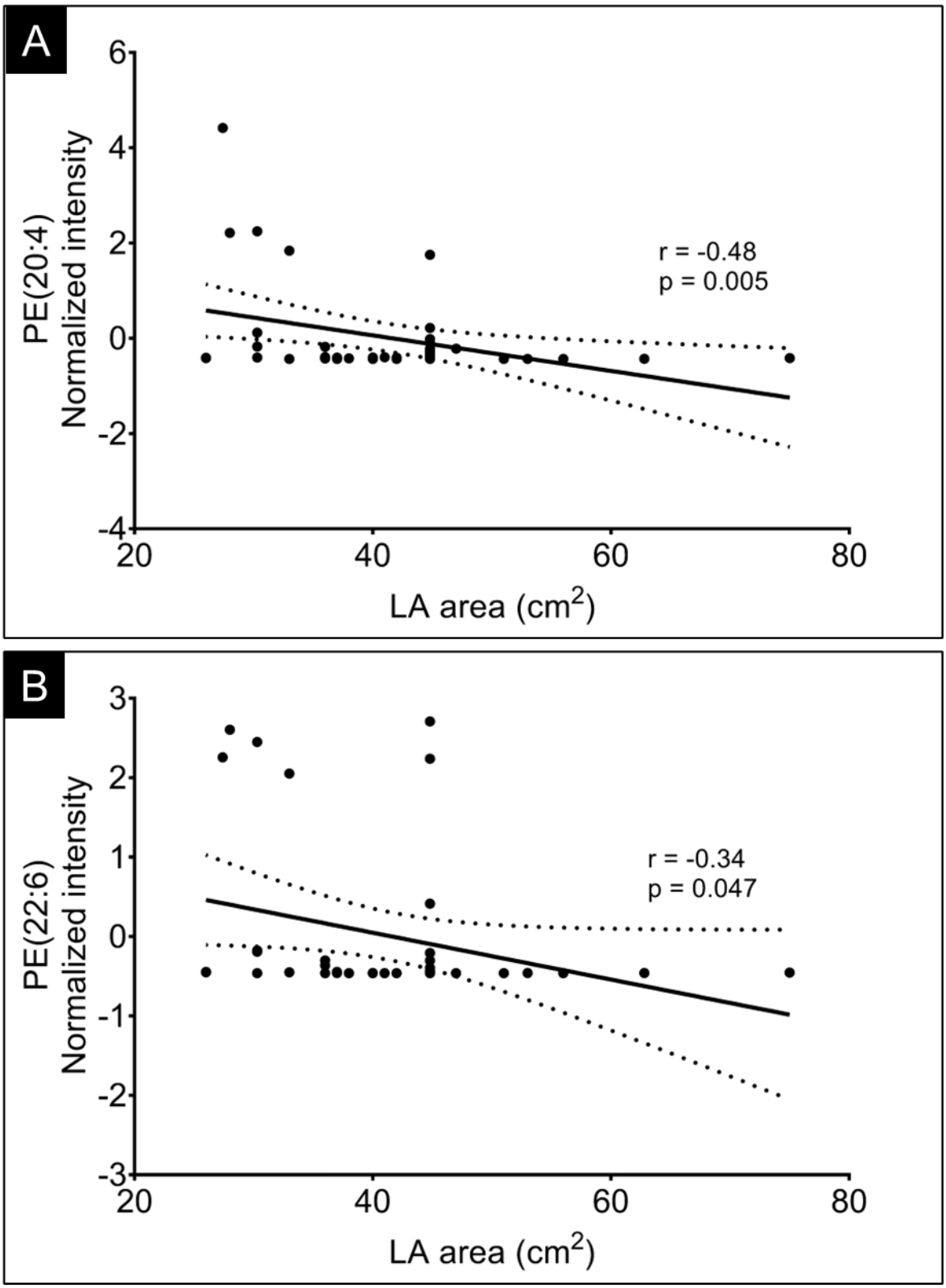


Figure 5.44. Correlation of LA area with (A) phosphatidylethanolamine (20:4) and (B) phosphatidylethanolamine (22:6) in RHD and degenerative AS while adjusting for sex, age, race, BMI, hypertension, diabetes, and batch effects.

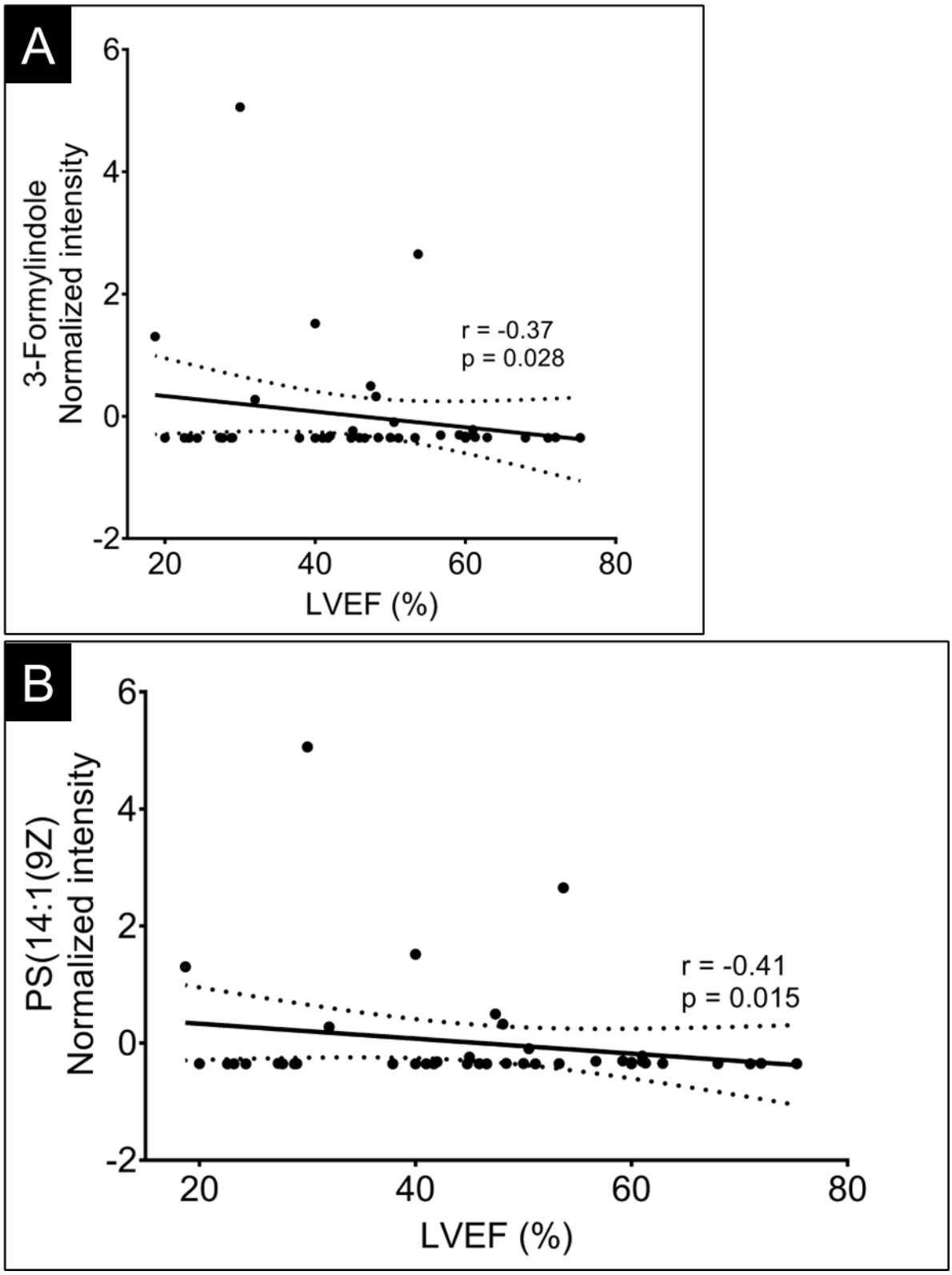


Figure 5.45. Correlation of left ventricle ejection fraction (LVEF) with (A) 3-formylindole and (B) phosphatidylserine (14:1) in RHD and degenerative AS while adjusting for sex, age, race, BMI, hypertension, diabetes, and batch effects.

5.4 Discussion

To the best of my knowledge this is the first study exploring metabolic markers in RHD and comparing it with markers in degenerative AS. In this study untargeted metabolomics was successfully applied in determining differential metabolites with biological relevance and their associated genes in patients with severe RHD, degenerative AS against controls. Further, markers with excellent ROC-AUC, specificity, and sensitivity were described as potential diagnostic markers to differentiate RHD from degenerative AS and controls. Linoleic acid metabolism was suggested to be associated with RHD when compared to controls. The metabolites dysregulated between degenerative aortic stenosis and controls and between RHD and AS were associated with glycerophospholipid metabolism pathway. In addition, the study suggested that clinical baseline characteristics affect reliability of potential diagnostic biomarkers but there exist robust markers that are stable to effects of clinical parameters.

Patients with VHD due to chronic RHD require valve replacement surgery as it is the only reliable remedy. However, in SSA the lifesaving services are not readily accessible.[180] RHD patients tends to require valve replacement at a fairly young age as compared to degenerative AS[157] and the symptoms only presents when there is severe damage of the valves. Furthermore, there is no therapeutic target to halt progression of valve damage. Derangement of linoleic acid metabolism and the arachidonic acid metabolism was found when patients with RHD were compared to the controls. There were no metabolomics studies on RHD to validate these findings however, linoleic acid is a precursor to arachidonic acid metabolism which plays an important role in mediation of inflammation.[181] Furthermore, phosphatidylcholines can be metabolised as a source of arachidonic and linoleic acid and specific PCs were elevated in RHD patients as compared to controls.[181] Linoleic acid on the other hand is a vital polyunsaturated fatty acid which is a precursor for arachidonic acid which is released upon tissue injury or irritation and promotes elevated inflammation.[181] Chronic RHD is a sequela of repeated acute rheumatic fever due to infection with Group A Streptococcus that induces an autoimmune response to cardiac proteins.[182] There is evidence of continued assault to the

tissue though the stimulation of CD4+ T cells that recognise the M protein and the cardiac proteins and elevated C-reactive proteins in chronic RHD.[182–184] The reported findings on continued inflammation in chronic RHD corroborates this study's findings where anti-inflammatory metabolites were decreased while pro-inflammatory metabolites were elevated in RHD patients compared to controls. Valine, leucine and isoleucine biosynthesis and degradation pathways were shown to be altered in patients with RHD. Specifically, valine was elevated in RHD while tiglyl-CoA, a by-product of isoleucine degradation was decreased. Dysregulation of branched chain amino acids (BCAA) has also been reported in myxomatous valve disease and calcific aortic stenosis.[18,101] The observed elevation of branched chain amino acids (BCCAs) in RHD may suggest dysregulated BCAA catabolic processes in RHD and elevation of inflammatory metabolic processes. Accumulation of BCAAs have been shown to cause a chronic induction of mTOR activity and was associated with cardiac hypertrophy and suppression of autophagy.[185] Further, changes in the BCAA metabolism and catabolism are a hallmark for metabolic shifts in a failing heart.[18,186] Patients with severe RHD often present with mitral valve disease, reduced LVEF, and dilated LA with complications of cardiac hypertrophy, embolic stroke, and heart failure.[187] The gene-metabolite interaction showed metabolites (linoleic acid, cholic acid, ursodeoxycholic acid, taurine, and beta-D-glucose) and genes (*SOAT1* and *LCT*) as being associated in a network. The gene *SOAT1* codes for sterol O-acyltransferase that controls transport of cholesterol.[188,189] Cholic acid and ursodeoxycholic acid are bile acids made from cholesterol[190], taurine is involved in the bile acids conjugation and is shown to influence the cholesterol uptake.[191] Some of the sterol esters are made from sterols and requires the acetyl-CoA from the glycolytic pathway.[192]

Lipid accumulation, calcification and inflammation are central to be the aetiopathogenesis of degenerative AS.[193] Advanced age, elevated BMI, dyslipidaemia, and hypertension are some of the risk factors associated with aetiology of degenerative AS. We observed dysregulation of glycerophospholipid and glycerolipid metabolism pathways among degenerative AS patients as compared to controls. Dysregulation of phospholipids, lysophosphatidic acid and

antioxidants linked to lipid metabolism have previously been reported in calcific AS.[109,194,195] Phosphatidylcholine, phosphatidic acid, 5,6-DHET were reported as elevated in degenerative AS patients compared to the controls. The phospholipids and eicosanoids have been reported in calcific AS valve and have been associated with inflammation.[15,196,197] Eicosanoids are elevated in AS patients with minimal LV reverse remodelling post-aortic valve replacement.[15] Taken together, the findings suggest dysregulation of inflammatory processes in degenerative AS patients. We observed degenerative AS patients to have increased left ventricular myocardium mass index indicating LV hypertrophy. Previous studies have reported the profound shifts in metabolism of strained/hypertrophied hearts from fatty acids oxidation to glucose utilisation.[198–201] The findings may suggest derangement in energy metabolism processes in patients with degenerative AS.

RHD and degenerative AS have distinct aetiopathogenesis yet both present with valve stenosis; predominantly due to fibrosis in RHD and calcification in degenerative AS however there are isolated cases of calcification in RHD.[202,203] Whether calcification in RHD follows similar processes to that seen in atherosclerosis and degenerative AS remains to be understood.[3] Metabolic profiles in degenerative AS have extensively been studied[15,99–101,109], yet there are no studies that have reported on the metabolic processes in RHD. Our study showed the dysregulated metabolites as mostly associated with glycerophospholipid metabolism where phosphatidylcholine, phosphatidylserine, and lysophosphatidylethanolamine tended to be elevated in AS compared to RHD and controls. The phospholipids especially PC are abundant membrane lipids and form the monolayer of the HDL particles, that directly controls cholesterol transport.[204] Further, the dysregulation of the PC and the lysoPC may affect free fatty acid metabolism which has been associated with proinflammation and formation of atherosclerotic plaques.[205,206] From the histopathological assessment, features of inflammation were mostly associated with RHD while degenerative AS was associated with fibrosis and collagen deposition. PCs has been reported to be reduced in cases of myocardial infarction compared to coronary artery disease cases with

atherosclerotic plaques.[205] Accumulation of glycerophospholipid species metabolites are protective against oxidative stress due to their antioxidant characteristics and reduced levels may indicate heightened oxidative stress levels.[205] There has been a strong relationship between phospholipids and oxidised phospholipids and the extent of calcification in cardiovascular diseases. However, the exact mechanism is not well understood. In other studies, PCs has shown contradicting trends depending on the PC moiety; a strong correlation with calcification was however reported.[207,208] LysoPC has been shown to mediate calcification through the activation of NF-kB/IL-6/BMP pathways in human valve interstitial cells.[196] The phospholipids dysregulation indicates shifts in energy metabolism through the production of acetyl-CoA that leads to ATP production through the TCA cycle.[209] Failing hearts are shown to shift their preference from fatty acids oxidation to glycolytic pathways which would explain the accumulation in AS of the phospholipids that are precursors to acetyl-CoA synthesis from the glycerophospholipid metabolism pathway.[209] The suggested changes in fatty acids and energy metabolism was further reinforced by the gene-metabolite interaction findings that showed enrichment to fatty acid biosynthesis and amino acid metabolic processes. The fatty acids and amino acid metabolic processes derangements have been reported before mostly in AS.[18,99–101,109]

As per the WHF and ACC/AHA guidelines, RHD and AS is characterised through patient's symptoms, valve anatomy, the hemodynamics, and the blood biomarkers.[9,10,116] However, it is difficult to detect VHD in asymptomatic patients and the diagnosis requires specialised personnel and equipment which are not always readily available especially in resource limited regions.[116,210] Metabolomics studies have proposed potential diagnostic biomarkers in cardiovascular diseases especially calcific AS.[14,15,18,109,211]; Formate and lactate are potential diagnostic biomarkers in differentiating patients with mitral valve stenosis and regurgitation from controls.[18] Further, a panel of diagnostic metabolites associated with NO metabolism have been reported to distinguish severe AS patients from controls.[109] To differentiate RHD from the controls, a panel of 10 metabolites showed very good sensitivity and specificity. Of

the 10 metabolites, 13-docosenamide and cholic acid were the best performers. There were no studies to contrast these findings however, 13-docosenamide was reported elevated in a metabolomics study on patients with psoriasis vulgaris which is a condition characterised with chronic inflammatory cells infiltration.[212] Cholic acid on the other hand is a primary bile acid made from cholesterol catabolism pathway for excretion. Decreased levels of bile acid excretion have been associated with cases of atherosclerosis in patients with coronary artery disease.[213,214] Cholic acid was reported elevated in controls as compared to RHD patients. To differentiate AS from controls, 3-oxo-5S-amino-hexanoic acid, palmitoylcarnitine and acylcarnitine were best performers and they were all elevated in AS. Amino-oxohexanoate is an intermediate in lysine degradation pathway; lysine is important in calcium absorption. Lysine is elevated in AS patients compared to controls.[101] Acylcarnitines and palmitoylcarnitines play an essential role in fatty acids transport into the mitochondria during beta-oxidation. Like our findings, acylcarnitine was reported elevated in AS and could differentiate calcific AS patients from controls.[109] To the best of my knowledge no other studies have explored diagnostic biomarkers between RHD and AS. A panel of 10 metabolites demonstrated very good sensitivity and specificity to differentiate RHD from AS. Phosphatidylserine (PS), lysophosphatidylethanolamine (lysoPE) and valine were the best biomarkers elevated in AS compared to RHD. Valine has been reported elevated in AS than controls.[101] On the other hand, PS and lysoPE are phospholipids that are metabolised to lysophosphatidic acid which was reported as a distinctive marker in controls and calcific AS and was elevated in AS than controls.[109] The PCA and PLS-DA features exploration analysis showed AS to be similar to controls suggesting RHD had greater perturbations in metabolic processes. To determine the effect of the comorbidities on the robustness of the diagnostic biomarkers, the metadata adjustment analysis indicated that comorbidities influence the significance levels of the diagnostic biomarkers. It is therefore important to ensure that the recruited participants in metabolomics studies are matched on their comorbidities.

RHD mainly presents with pure mitral regurgitation or plus aortic regurgitation while degenerative AS primarily presents with aortic stenosis.[10,157,215] Consequently, these leads to the RHD patients presenting with decreased LVEF and hypertrophied LV; while AS presents with normal LVEF, hypertrophied LV and dilated LA.[216–218] Of the 7 metabolites that showed significant correlation with LVEF, LVMI and LA area, 3 metabolites (PE (20:4/0:0), acrylic acid, and creatine) were different between RHD and AS and their differences were largely independent of the covariates. Creatine showed correlation with LVMI, and it is known to be a key source of immediate energy by rapidly synthesising ATP.[219] Elevated levels of creatine have been reported elsewhere to be associated with increased LV systolic pressure.[219] In this study there was a positive correlation between creatine and LVMI which may suggest that as the valve pathology worsened the heart needed to pump harder and therefore requiring a quick supply of ATP through increased levels of creatine. In addition, acrylic acid showed a positive correlation with LV mass index. Previous studies on the synthetic acrylic acid showed cardioprotective capabilities by suppressing the inflammatory and cell apoptosis responses.[220] As to whether the acrylic acid detected in this study had been used to facilitate cardiovascular drugs delivery as polyacrylic acid platform[221] is beyond the scope of this study. In this study, elevated LA correlated with decreased levels of PE (20:4/0:0). In human liver, phosphatidylethanolamine is methylated into phosphatidylcholine.[222] Through gut microbiome mediated processes phosphatidylcholine is converted to trimethylamine (TMA) that is associated with heightened formation of foam cells, atherosclerosis, and calcification.[223,224] In this study, there could have been an increased uptake of phosphatidylethanolamine used to form phosphatidylcholine that might have led to calcifications reported in degenerative AS patients who had dilated left atrium.

5.5 Limitations

Based on the study design and ethical issues, this study could not determine whether identified potential biomarkers were causal or only associated with the diseases studied. Therefore, further longitudinal studies using *in vitro* or *in vivo* models are warranted to determine the

causality effect of the reported metabolites and pathways. With regards to the integration analysis of the gene-metabolite interactions there is a limitation since the data-driven integration depends on the data available in the databases. Furthermore, tissue specific were not readily available especially the genetics and metabolomics data. The limited depth of the integration data makes it challenging to interpret and draw conclusions on the findings. The features reported here were annotated to MSI confidence levels 2 and 3; future studies should consider using pure standards of the reported features to quantify the magnitude of the changed metabolites. Furthermore, this was an explorative study from one centre. Therefore, multicentre studies with larger cohorts and multiple analytic platforms are required to validate the described potential diagnostic metabolic biomarkers and pathways.

5.6 Concluding remark

In conclusion, the current study is the first to describe metabolic biomarkers expressed in poorly studied valvular heart disease (rheumatic heart disease and aortic stenosis) in SSA.

Method development pilot experiments

From the pilot experiments, we observed that the choice of the extraction solvents is dependent on the samples being analysed and the metabolome being targeted. In addition, we have demonstrated that it is important to perform pilot experiments to optimise untargeted metabolomics pipeline parameters since the conditions are not universal.

Study experiments

The study experiments and observations described untargeted metabolomics biomarkers that distinguished RHD and AS from healthy controls and described metabolic pathways affected in VHD patients requiring valve replacement surgery. The results suggested that perturbation of several metabolites in RHD and AS associated with energy metabolism, amino acids regulation, and immune response regulation. Linoleic acid metabolism and glycerophospholipid metabolic pathways were some of the key pathways affected in RHD and AS patients. Furthermore, we report on metabolites that are associated with valve calcification,

cardiac function, and altered pathological parameters in RHD and AS patients. We believe these findings should allow for planning of larger longitudinal cohort studies to allow targeted assessment of the protective and causal mechanisms of the reported metabolic biomarkers.

5.7 Contributions and acknowledgements

The blood samples were collected from patients undergoing valve replacement surgery at the Christiaan Barnard Division of Cardiothoracic Surgery, Groote Schuur Hospital, Cape Town after recruitment and consenting by D Mutithu assisted by Dr R Manganyi, O Aremu, and Dr E Lumngwena. The healthy control samples were obtained from participants without VHD recruited at the University of Cape Town. The healthy controls were recruited and consented by D Mutithu assisted by O Aremu, and Dr E Lumngwena. The clinical data was collected and curated by D Mutithu and reviewed by Prof N Ntusi. The samples were processed by D Mutithu. The LC-MS/MS experiments were performed by D Mutithu assisted by A Evans. The data was processed and statistically analysed by D Mutithu and reviewed by Prof J Kirwan.

6 TISSUE SPECIFIC METABOLIC BIOMARKERS IN RHEUMATIC HEART DISEASE AND DEGENERATIVE AORTIC STENOSIS PATIENTS[5]

6.1 Introduction

In SSA, the leading causes for VHD patients requiring valve replacement are RHD, BAV, and degenerative AS.[225,226] There exist methods for diagnosing RHD or degenerative AS however, at present there is no methods that can be used for early detection and predict disease progression.

Metabolomics reports on the changes in the metabolome within a biological system. It has been shown that metabolomics is valuable in understanding the diseases processes and pathomechanisms due to its proximity to the organism's physiology.[19,227] Metabolomics techniques exist used to perform spatial localisation of metabolic biomarkers which is akin to molecular tissue typing to investigate the microenvironments on the studied tissues.[228–230] RHD and degenerative AS have distinct pathophysiological features with the RHD dominated by myxoid change, and collagenation while degenerative AS mostly has features of fibrosis and calcification.[154,231] At the microenvironment level, the physiological changes are preceded by changes in metabolome of the affected cells. It is therefore expected that different regions of pathology would have different metabolic signatures.[230]

Most of the biomarkers reported in metabolomics have not found use at the bedside as diagnostic or prognostic biomarkers.[232] One of the approaches to ensuring that the identified

[5]This chapter is based on a manuscript:
Mutithu, D. W., Kirwan, J. A., Adeola, H. A., Aremu, O. O., Lumngwena, E. N., Familusi, M., ... & Ntusi, N. A. (2024). *Major energetic metabolites associated with echocardiographic and histological features of aortic and mitral valve disease* [Unpublished manuscript]. Department of Medicine, University of Cape Town, Cape Town, South Africa.

biomarkers find their way to the bedside is conducting correlation assessments with the clinical parameters. Several studies have explored the correlations of the metabolic biomarkers with the clinical parameters in cardiovascular diseases.[19,233–235] However, few studies have reported on the metabolic biomarkers detected on the heart valves in RHD patients with single or double valve replacement and degenerative AS patients. Even fewer studies have reported on the spatial localisation of the potential metabolic biomarkers detected in the tissue's homogenates.[228,229,236]

Therefore, in this chapter we aimed to determine the tissue specific biomarkers detected in the homogenates of aortic and mitral valves. The samples were obtained from RHD patients having single (mitral valve only) or double (mitral and aortic valves) valve replacement and degenerative AS patients (aortic valve only). The biomarkers were analysed using untargeted metabolomics approach with UPLC-QTOF-MS. The detected biomarkers were correlated to the TTE and CMR parameters of valvular heart disease.

6.2 Methods

6.2.1 Study population

The VHD patients described in chapter 4.3.1 were used in this chapter. However, metabolites were extracted from valve biopsies obtained from patients who had valve replacement surgery. RHD and AS patients scheduled to undergo valve replacement surgery at the Division of Cardiothoracic Surgery were recruited from the Cardiac Clinic Groote Schuur Hospital, Cape Town. Patients were classified as either RHD or degenerative AS based on TTE and CMR as previously reported. The samples were divided into RHD with single valve replacement (RHD 1 - mitral valve only); RHD patients with mitral and aortic valve replacement (RHD 2 - mitral and aortic valve); and degenerative AS - aortic valve only. The study was conducted in accordance with the principles of the Declaration of Helsinki; and was reviewed and approved as described in sections 3.1.1 and 3.2.1. The inclusion and exclusion criteria were as described in section 3.2.1.

6.2.2 Sampling and processing

Valve fragments were collected from RHD and AS patients while undergoing valve replacement surgery at the Division of Cardiothoracic Surgery. The valve biopsies were collected as detailed in section 3.2. During the analysis, valve aliquots were thawed on ice. The valve biopsies quality control (pooled QC) samples were prepared as described in section 3.2.6. The pooled QCs were then re-frozen together with the samples until analysis. The metabolites were extracted from the valves as detailed in section 3.2.6, metabolites extraction methods. After extraction, the samples were injected onto the UPLC column as detailed in section 3.2.7. The mobile phase A was composed of HPLC grade water and mobile phase B was composed of a mixture of 100% acetonitrile and methanol (1:1, v/v) and modified with 2 mM ammonium formate and 0.1% formic acid. The mobile phase was run at a 60-minute gradient at a rate of 0.4 mL/min and autosampler set as detailed in section 3.2.7. Data was acquired in positive ionisation mode on a Sciex X500R Q-TOF mass spectrometer operated with Sciex OS software ver.1.4 (Sciex, USA) with dual ESI ion source system. After calibration, data was collected using data dependent acquisition (DDA) mode as detailed in section 3.2.7.

6.2.3 Data processing of LC-MS/MS data and statistical analysis

Raw positive mode LC-MS/MS data were processed with MS-Convert in ProteoWizard 3.0.1908 [126] and MS-DIAL 4.38 [96] as detailed in section 3.2.8. The specific parameters for this chapter are described in details hereafter. The peaks were detected and aligned using MS-DIAL where a peak was discarded if it included more than 30% missing values. In addition, peaks were excluded if they were missing in more than 30% of the samples per group in ESI+ mode. Batch drift adjustment was performed using the pooled long-term QC samples that were injected every 5 samples in all batches, including one pooled QC injected at the beginning and end of each batch. The raw data was pre-processed, aligned to the pooled QC samples, and the batches normalised with the LOESS algorithm based on the QC samples. Data was filtered by removing background noise based on the blank samples (maximum sample/blank<5-fold change). The missing (zero) values were then replaced with 1/10 of the minimum peak height

for that *m/z* feature across all samples. The extracted features were further filtered based on the technical reproducibility of individual features: features were removed where the quality controls' (QC) relative standard deviation (RSD) was greater than 30%. There was a significant variation between the analysed batches therefore we performed a within-batch features filtration. Features were selected if they met the <30% QC RSD cutoff within the batch. The features meeting the selection criteria were then amalgamated and they were selected if they were present in more than 60% of the analysed batches. After selection, the amalgamated qualifying features were then adjusted for batch drift using feature-wise ANCOVA batch correction. To obtain the significant features, data was first normalised per sample based on the sum intensities in each sample, before transformation per feature by mean centering and dividing by the standard deviation (autoscaling) of each feature using MetaboAnalyst 5.0.[237] To identify potentially important features between groups, univariate ANOVA analysis was used, while multidimensional reduction (PCA and PLS-DA) analyses were used to visualise group differences. Chemical annotation of the selected features was done by matching their mass spectra and product ion mass spectra to publicly available spectral libraries (HMDB, PubChem, ChEBI, MassBank, etc.) with GNPS 28.2.[140] A similarity score above 70% between the queried features and the library molecular formular, structure and adducts was considered during annotation. Important features lacking product ion mass spectra data were matched to public databases based on their accurate mass, retention time, and adduct formation patterns using CEU Mass Mediator available at http://ceumass.eps.uspceu.es/.[134] Important features with and without product ion mass spectra data were putatively annotated at Metabolomics Society Initiative (MSI) levels 2 and 3 of identification, respectively. Furthermore, correlation analyses between perturbed metabolites and the histopathological features, TTE valve parameters for morphology and flow dynamics were performed while adjusting for covariates. The *m/z* features and metabolites with a p-value <0.05 were considered significantly altered.

6.2.4 Functional analysis of significant metabolites

To explore the functional relevance of the annotated significant metabolites, pathway enrichment and topology analysis was performed in MetaboAnalyst 5.0.[237] The rationale behind functional analysis is explained in section 3.2.11. The HMDB and KEGG identifiers of the annotated metabolites were uploaded and mapped to the KEGG human pathway library for pathway analysis. The mapped pathways were visualised with scatter plots while testing the significance level, the enrichment was analysed using the hypergeometric test, while the topology analysis was done using relative-betweenness centrality on the *Homo sapiens* KEGG pathway library.[237] Pathways with a false discovery rate (FDR) adjusted p-value (q) <0.05 were considered significantly overrepresented.

6.3 Results

6.3.1 Baseline characteristics of the study samples

The baseline characteristics of the study participants are summarised in **Table 6.1**. The analysed samples were grouped into samples that were obtained from RHD patients who had mitral valves replaced (RHD MV) n=22, and RHD patients who had both the mitral and aortic valves replaced (RHD MV & AV) n = 19, and from patients with degenerative AS (Degenerative AS AV) n=11. Patients with degenerative AS were generally older than the RHD patient groups ($p = 0.003$). The BMI of the participants was significantly different between the three groups ($p = 0.011$); the RHD participants were clinically classified as overweight and the AS patients clinically obese (RHD-MV =27.7±7.34, RHD-MV & AV=26.6±5.14, degenerative AS-AV = 34.2±7.22 kg/m^2, respectively). The groups did not have differences in their cardiac function as shown by LVEF, LVEDV, and LVESV not being significantly different among the included groups. The RHD patients with MV replacement and the degenerative AS patients had elevated left ventricular mass index as compared to the rest of the patients. However, RHD patients with mitral and aortic valve replacement or mitral valve replacement only had elevated LA area as compared to degenerative AS patients ($p=0.016$). From the flow dynamics (AV

mean PG, MV mean PG, AV Vmean, and mitral valve velocities) it was evident that the participants had severe valve lesions ($p < 0.05$) **Table 6.1.**

Table 6.1. Baseline characteristics of the study participants.

	RHD (MV) N=22	**RHD (MV & AV) N=19**	**Degenerative AS (AV) N=11**	**p-value**
Age (years), mean(±σ)	43.1±14.8	52.4±12.7	60.5±12.3	0.003[a]
Sex				0.732
Female, n(%)	17(77.27)	13(68.42)	7(63.64)	
Male, n(%)	5(22.73)	6(31.58)	4(36.36)	
Race				0.283
Black, n(%)	6(27.27)	5(26.31)	1(9.09)	
Mixed, n(%)	16(72.73)	13(68.42)	8(72.73)	
White, n(%)	0(0)	1(5.26)	2(18.18)	
BMI (kg/m^2), mean (±σ)	27.7±7.34	26.6±5.14	34.2±7.22	0.011[b,c]
SBP(mmHg), mean (±σ)	118±19.9	128±23.7	136±27	0.094
LVEF %, mean (±σ)	44.6±13.7	42.8±16.2	52.4±13.1	0.212
LVEDV (ml), mean (±σ)	158±48.9	175±53.1	215±67.9	0.128
LVESV (ml), mean (±σ)	87.9±24.28	95.1±33.8	104±38.5	0.579
LVMI (g/m^2), mean (±σ)	50.9±13.2	74.5±35.5	99.8±37.8	0.005[b]
LA Area (cm^2), mean (±σ)	41.8±10.3	48.4±11.7	31.9±5.71	0.016[c]
Aortic root (mm), mean (±σ)	26.1±4.84	25.7±6.82	27.6±7.98	0.787
LVOT Vmax (m/s), mean (±σ)	1.01±0.3	1.0±0.17	1.03±0.33	0.890
AV mean PG (mmHg), mean (±σ)	6.36±3.69	24.9±19.6	53.4±21.8	<0.001[a,b,c]
AV Vmean (m/s), mean (±σ)	1.14±0.34	2.28±1.03	3.25±0.91	<0.001[a,b,c]
AVA $_{VTI}$ (cm^2), mean (±σ)	1.6±0.28	1.46±0.88	0.9±0.2	0.400
MV mean PG (mmHg)	11.0±6.43	10.1±4.36	1.25±0.96	0.009[b,c]
MVA $_{PHT}$ (cm^2)	1.35±1.61	1.24±0.50	4±2.40	0.037[b,c]
MV DT (msec)	679±319	551±307	201±139	0.006[b,c]
MV E Vmax (m/s)	1.90±0.58	1.80±0.46	0.86±0.36	<0.001[b,c]
MV A Vmax (m/s)	1.85±0.40	1.18±0.71	0.92±0.41	<0.005[a,b,]
MV E/A ratio	1.03±0.23	1.34±0.72	1.08±0.85	0.572
MV E/E' ratio	31.7±16.5	26.0±18.3	13.2±3.92	0.070
Hypertensive, n(%)	5(22.73)	6(31.58)	6(54.55)	0.221
Diabetic, n(%)	0(0)	1(5.26)	2(18.18)	0.132
Smoker				0.418
Current, n(%)	6(27.27)	4(21.05)	1(9.09)	
Ex-Smoker, n(%)	0(0)	3(15.79)	1(9.09)	
Dyslipidemia, n(%)	2(9.09)	3(15.79)	4(36.36)	0.235
Histopathology				
Aschoff bodies, n(%)	1(4.5)	1(5.3)	0	1
Calcification, n(%)	2(9.09)	8(42.11)	9(81.82)	<0.001
Collagen deposition, n(%)	11(50.00)	11(57.89)	4(36.36)	0.454
Fibrosis, n(%)	18(81.82)	17(89.47)	11(100.00)	1
Inflammations, n(%)	5(22.73)	10(52.63)	2(18.18)	0.126
Myxoid change, n(%)	8(36.36)	10(52.63)	2(18.18)	0.180
Neovascularization, n(%)	15(68.18)	13(68.42)	3(27.27)	0.010
Vegetations, n(%)	0	3(15.79)	0	0.114
Fibrin deposition, n(%)	2(9.09)	5(26.32)	3(27.27)	0.507
Valve lesions grading				
Aortic regurgitation				<0.001
Mild, n(%)	5(22.73)	4(21.05)	2(18.18)	
Moderate, n(%)	1(4.55)	7(36.84)	1(9.09)	
Severe, n(%)	0	8(42.11)	2(18.18)	
Aortic stenosis				<0.001

Mild, n(%) Moderate, n(%) Severe, n(%)	1(4.55) 0 0	0 2(10.53) 6	0 0 10(90.91)	
Mitral Regurgitation Mild, n(%) Moderate, n(%) Severe, n(%)	 6(27.27) 5(22.73) 9(40.91)	 4(21.05) 7(36.84) 6(31.58)	 0 1(9.09) 3(27.27)	0.009
Mitral Stenosis Mild, n(%) Moderate, n(%) Severe, n(%)	 1(4.55) 2(9.09) 15(68.18)	 0 4(21.05) 10(52.63)	 0 0 0	<0.001

Parametric values are expressed as mean± standard deviation and assessed using parametric student t test and one way ANOVA statistical methods. Categorical variables are represented as n(%), and either a chi square test or Fisher exact was used to test the null hypothesis. RHD (MV), rheumatic heart disease with mitral valve replacement; RHD (MV & AV), rheumatic heart disease with mitral and aortic valves replacement, Degenerative AS (AV), Degenerative aortic stenosis disease with aortic valve replacement; MV, mitral valve; AV, aortic valve; AVA, aortic valve area; MVA, mitral valve area; SBP, systolic blood pressure; BMI, body mass index; LVEF, left ventricle ejection fraction; LVEDV, left ventricle end-diastolic volume; LVESV, left ventricle end-systolic volume; LVOT; left ventricle outflow tract, LVMI, left ventricular mass index; LA, left atrium. ANOVA Post hoc analysis with fdr adjusted P-value <0.05; a, RHD-MV vs RHD-MV & AV; b, RHD-MV vs degenerative AS; c, RHD-MV & AV vs degenerative AS.

6.3.2 Feature extraction and data quality analysis

After data acquisition, 3538 features were detected in the positive ionisation mode, of which, 507 features were extracted after missing values filtration. The samples were analysed in 6 batches. The frequency of the missing values was also analysed in the pooled QCs and all the samples. The frequency of the missing values was random and was not affected by the batches (**Figure 6.1 and Figure 6.2**). To ensure consistency of the batches the batches were normalised using LOESS algorithm where the width of the moving window was set at (span α = 0.7). After the batch normalisation, there was a significant batch differences observed (**Figure 6.3**).

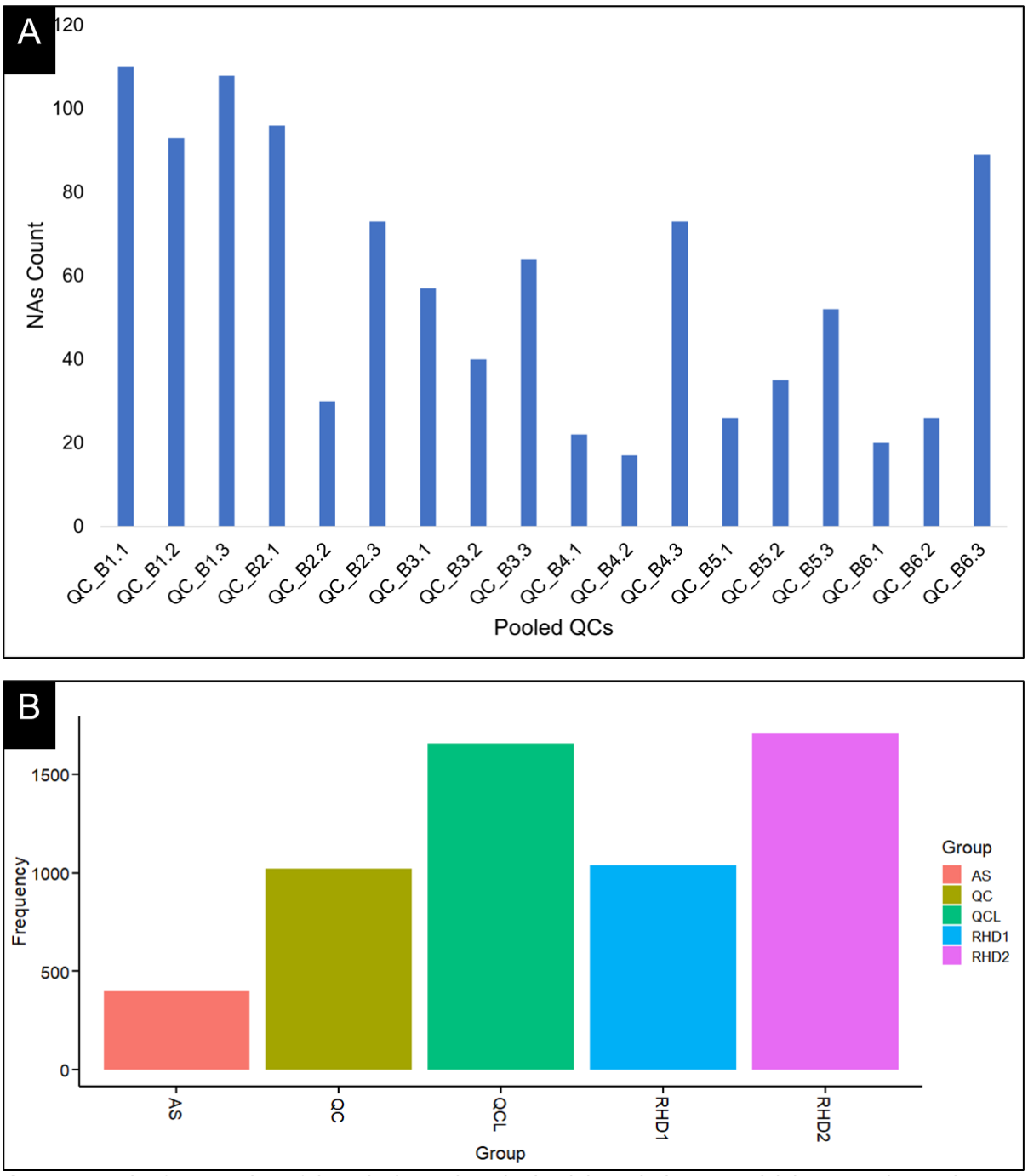


Figure 6.1. The frequencies of the missing values in the (A) pooled QCs and (B) grouped samples.

.

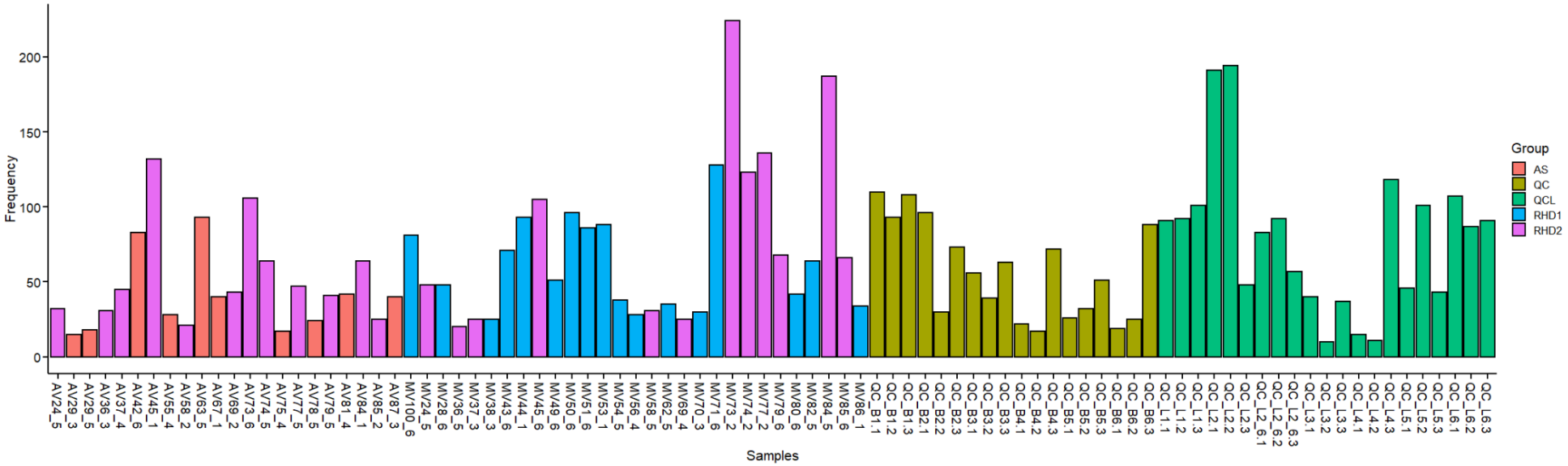


Figure 6.2. The frequencies of the missing values in individual samples analysed

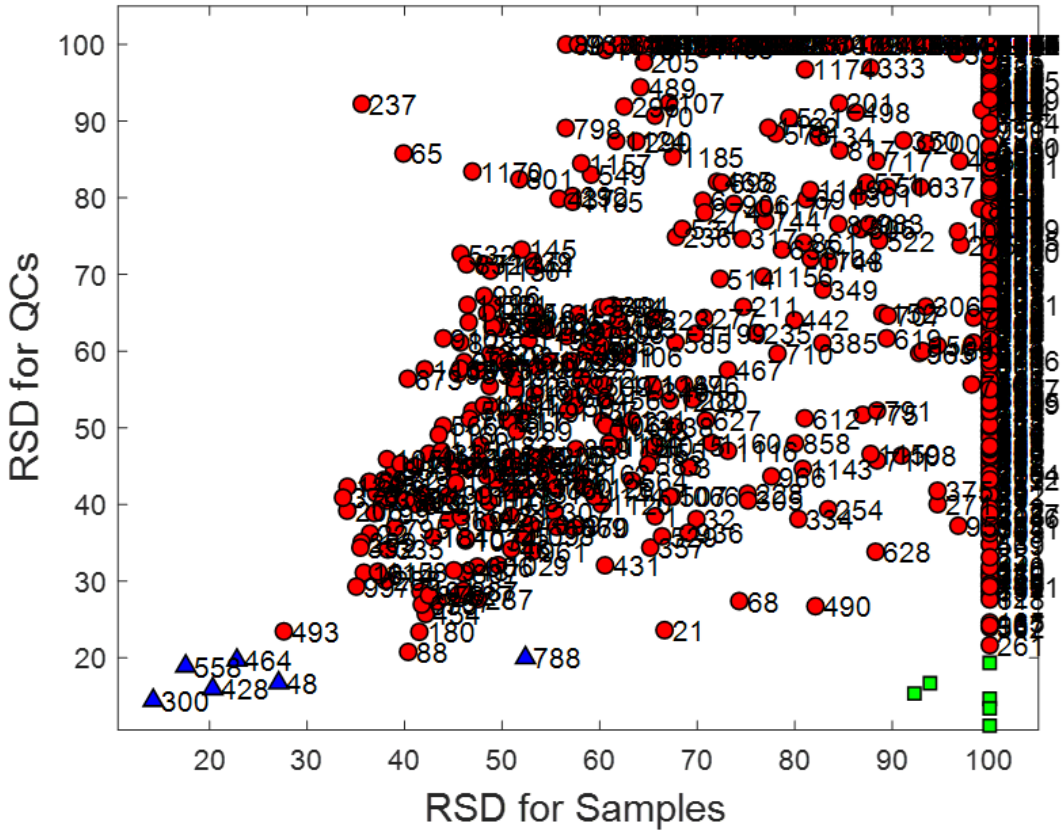


Figure 6.3. The scatter plot showing the distributions of the QCs' and the samples' relative standard deviation distribution. Triangles are samples and QCs with RSD <30%, the circles are samples and QCs with RSD>30%

After batch amalgamation, 219 features were considered usable (raw QC mRSD = 42.68%) and were taken further for batch correction using ANCOVA ((batch corrected QC mRSD = 13.26%) (**Figure 6.4 – Figure 6.6**).

.

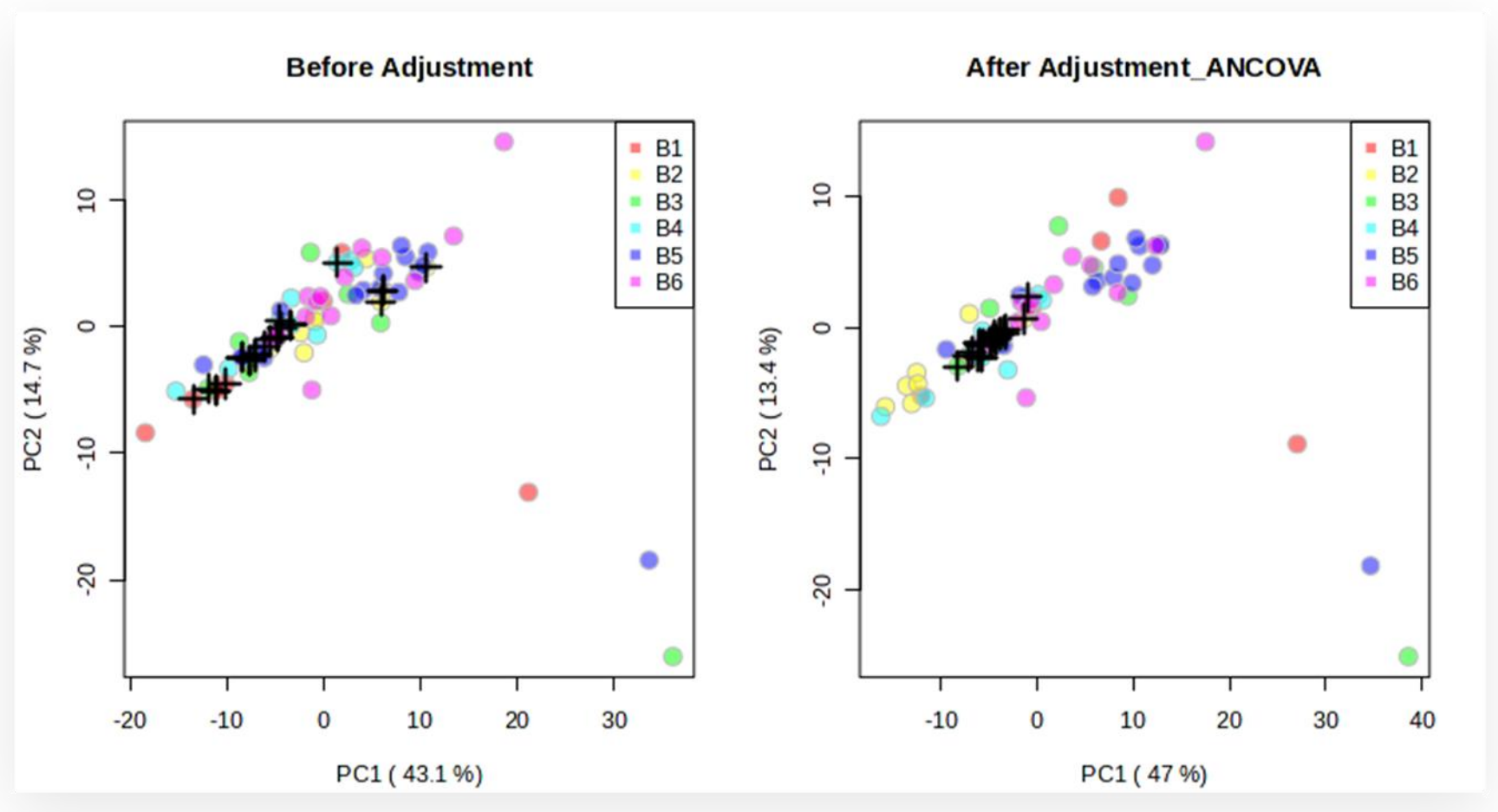


Figure 6.4. Batch drift and batch correction analysis. PCA plot showing sample distribution before and after batch correction.

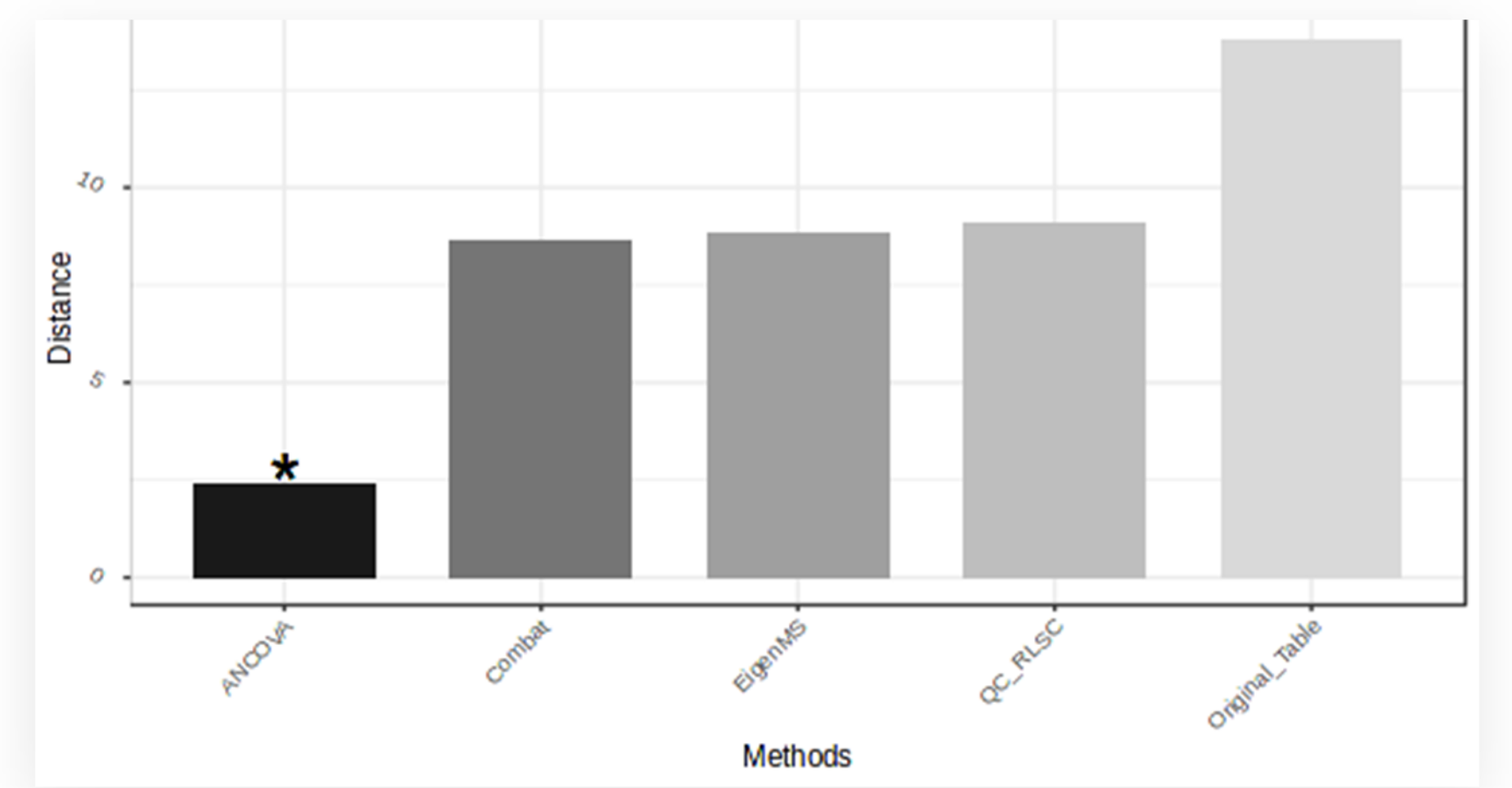


Figure 6.5. Batch drift and batch correction analysis. Inter-batch distance analysis on the batch adjusted and the original data (*shows the normalisation method with the least distance between the batches).

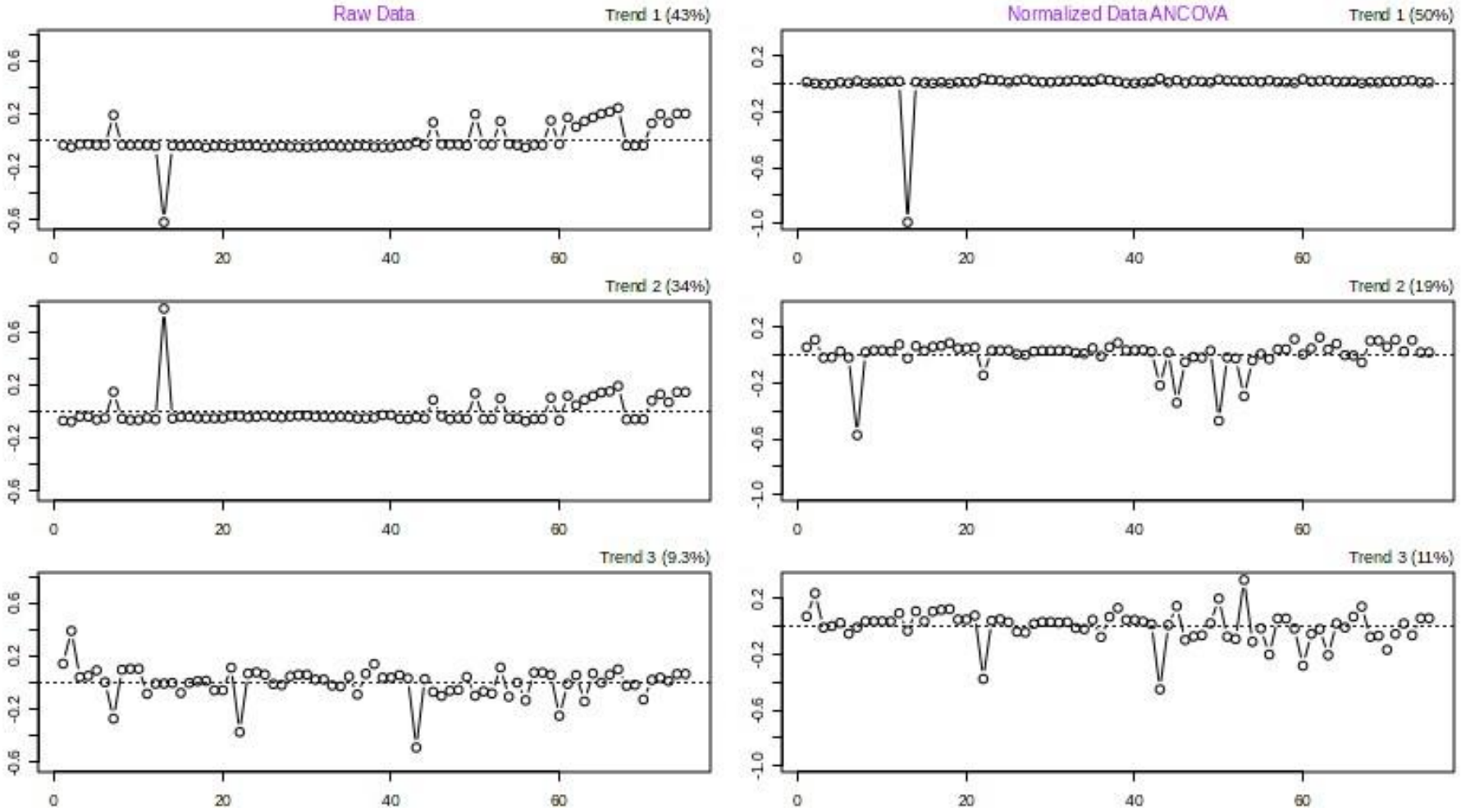


Figure 6.6. Batch drift and batch correction analysis. The singular value decomposition analysis of systematic trends attributable to bias before and after normalization.

The group-wise analyses observed 10 metabolites being significantly different between RHD and degenerative AS patients (**Figure 6.7**). Furthermore, uracil and stearoylethanolamide were different between rheumatic heart diseases patients with single valve replacement and those that required double valve replacement (**Figure 6.8**). Due to the small sample size, none of these metabolites passed false discovery testing. The multivariate analysis showed that the altered biomarkers differentiated between RHD1, RHD2 and AS, and the valve origins from the 3 valvular heart disease groups analysed (**Figure 6.9A & B**).

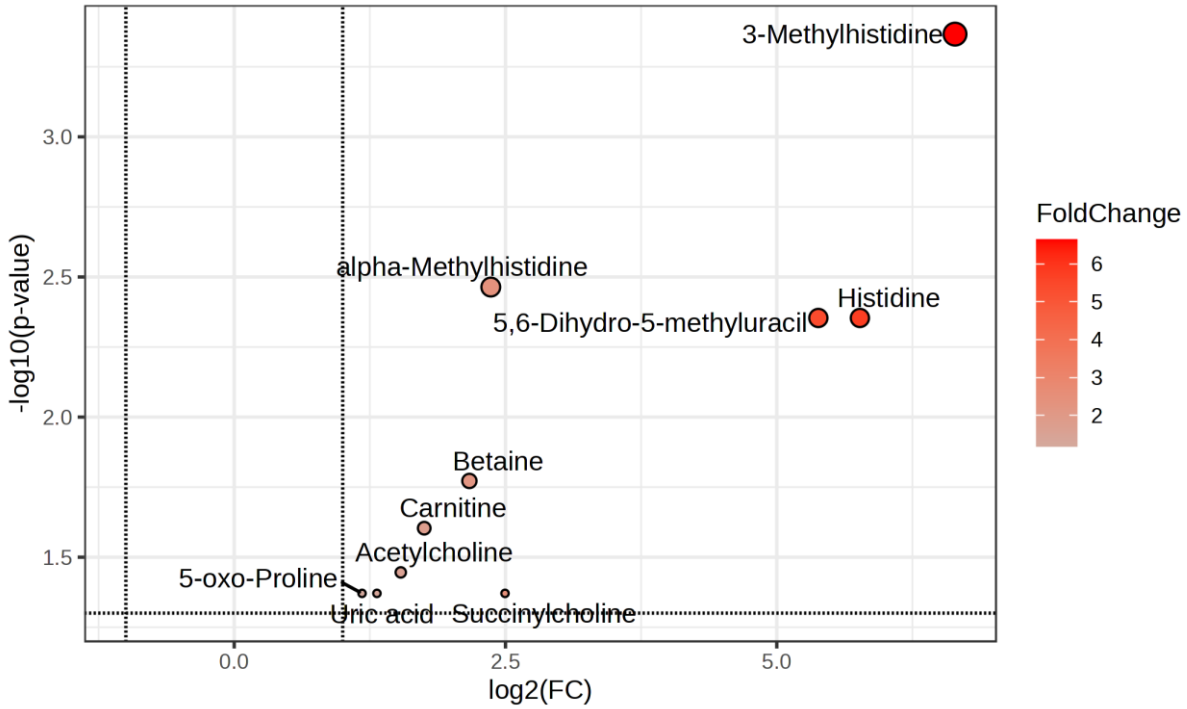


Figure 6.7. Metabolites differentially expressed between rheumatic heart disease and degenerative artic stenosis. Volcano plot of metabolites significantly changed between the aortic valves of rheumatic heart disease patients and degenerative aortic stenosis patients ($p<0.05$, $FC>2$).

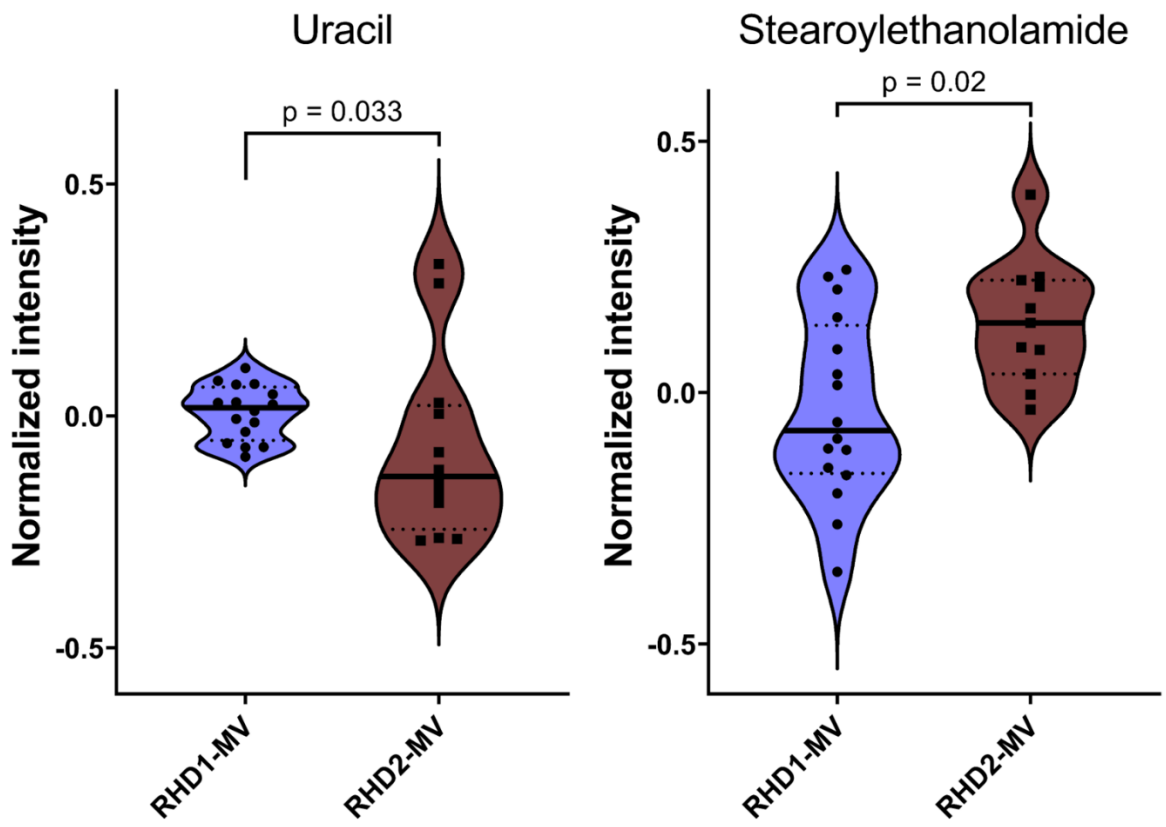


Figure 6.8. Metabolites differentially expressed between rheumatic heart disease and degenerative artic stenosis. Metabolites differentially expressed in mitral valves obtained from rheumatic heart disease patients undergoing single valve or double valve replacement.

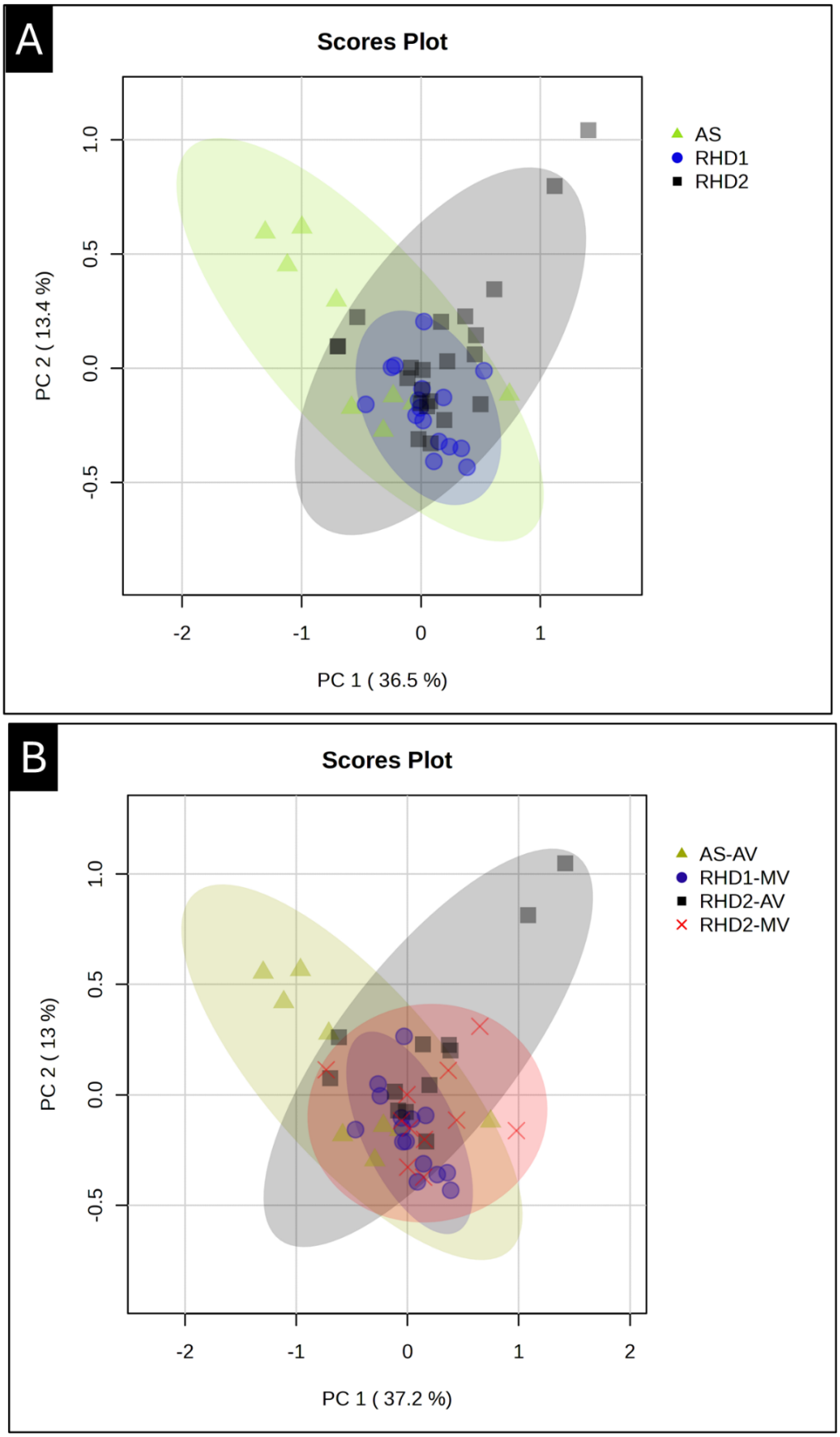


Figure 6.9. Metabolites differentially expressed between rheumatic heart disease and degenerative artic stenosis. PCA plots of variables significant by univariate statistics to visualize biomarkers differentiating (A) RHD1, RHD2, and AS and (B) the valve origins/types from the 3 valvular heart disease groups studied.

To determine the biomarkers different among the groups, univariate biomarker analysis was used to explore the biomarkers. Upon batch amalgamation, 219 features were considered potential and reliable biomarkers, 89 features were successfully annotated at MSI level 2 or 3 using spectral matching and molecular formula prediction techniques. Of the annotated features, 27 successfully annotated potential biomarkers were taken further for statistical analysis after peak shape filtering and removal of duplicates. When comparing RHD1, RHD2, and degenerative AS, 10 metabolites were expressed differently between the groups and were mostly amino acids and their derivatives (**Figure 6.10 – Figure 6.12**). However, RHD2-AV and RHD2-MV are paired, and this bias should be considered when interpreting the ANOVA results.

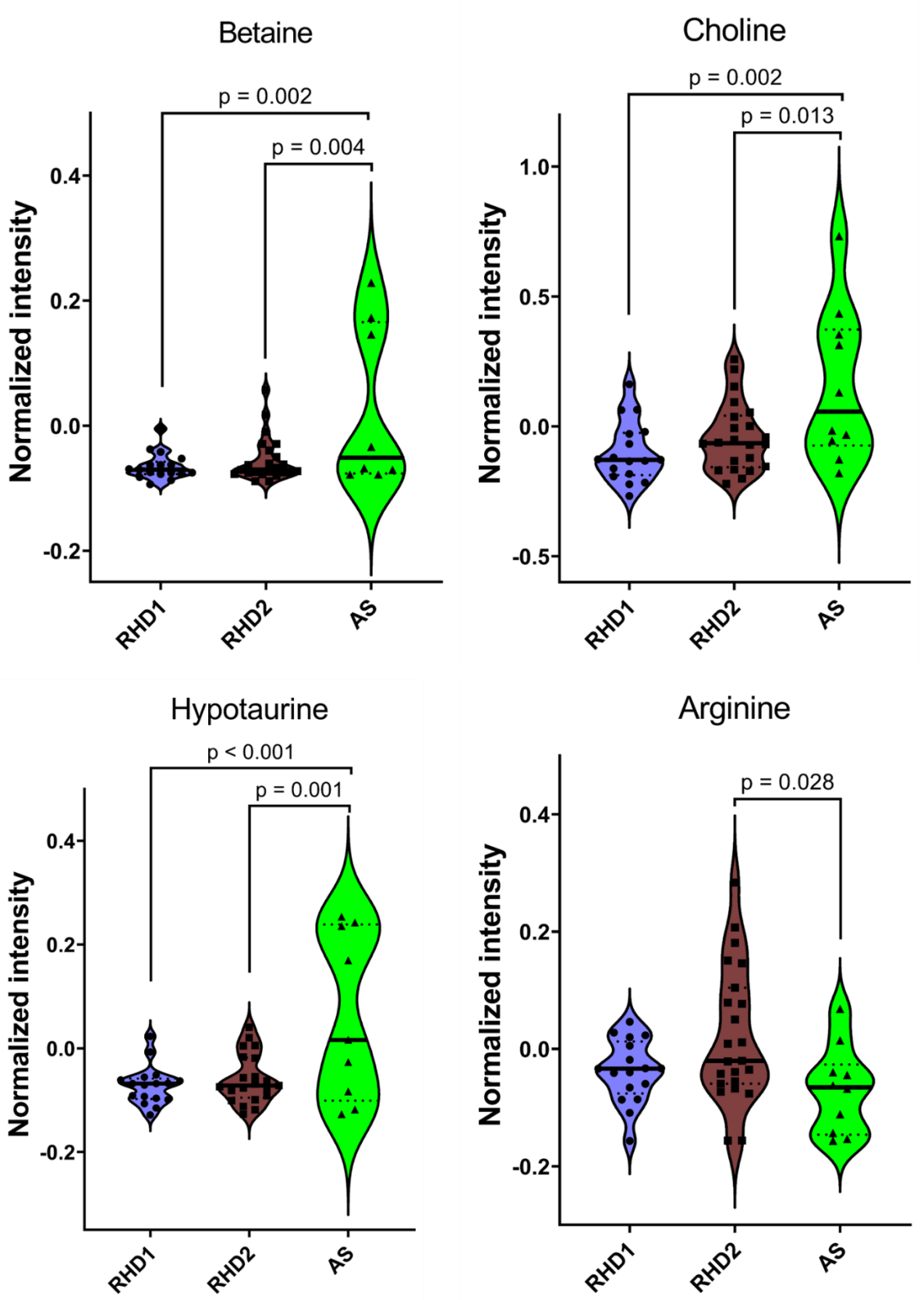


Figure 6.10. Metabolic biomarkers exploration. The metabolites (hypotaurine, choline, and betaine) that were elevated in degenerative AS compared to RHD patients undergoing single valve replacement or double valve replacement, and arginine that was not different between RHD1 and degenerative AS.

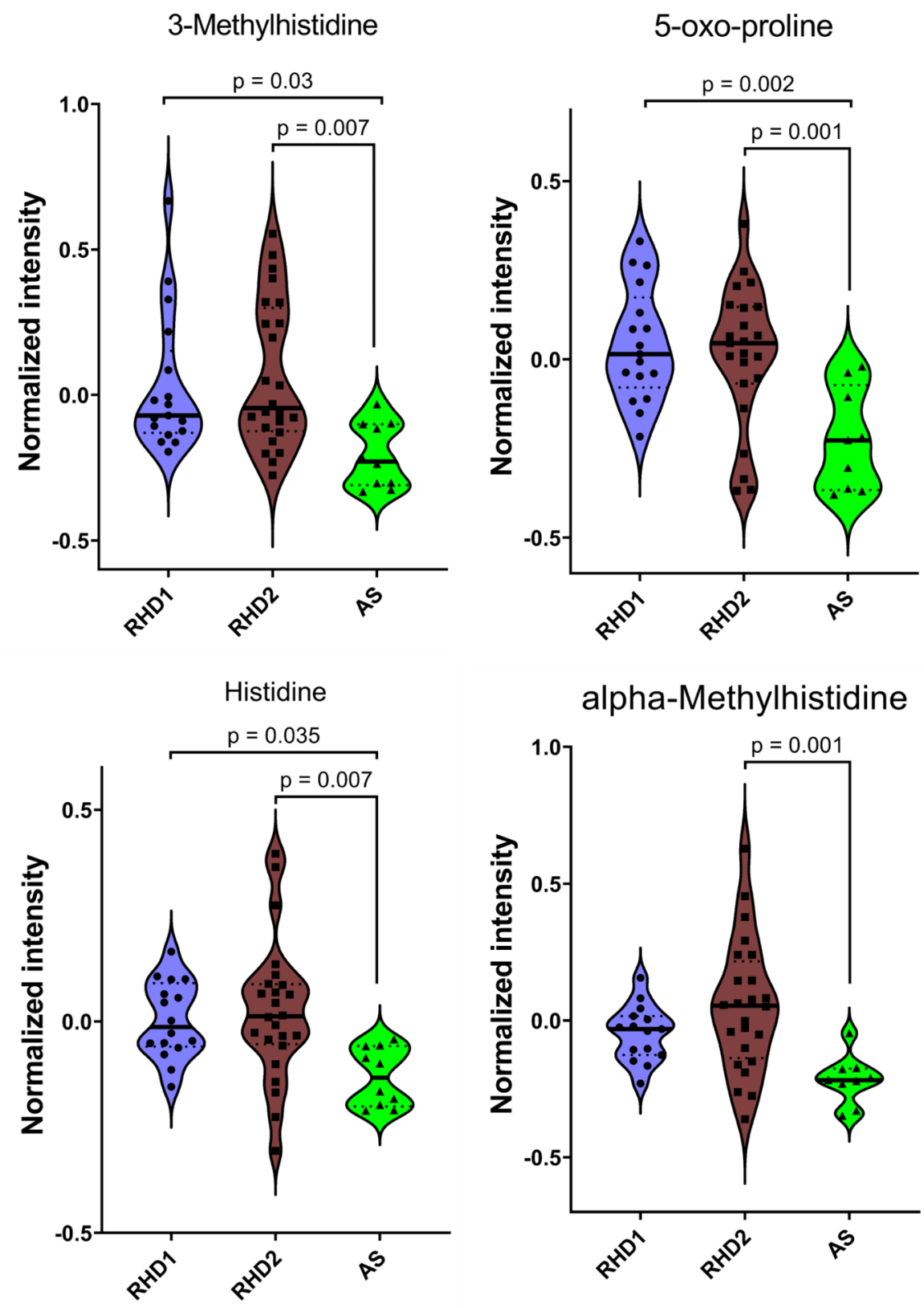


Figure 6.11. Metabolic biomarkers exploration. The metabolites that were elevated in RHD patients undergoing single (RHD1) valve replacement or double valve (RHD2) replacement compared to degenerative AS.

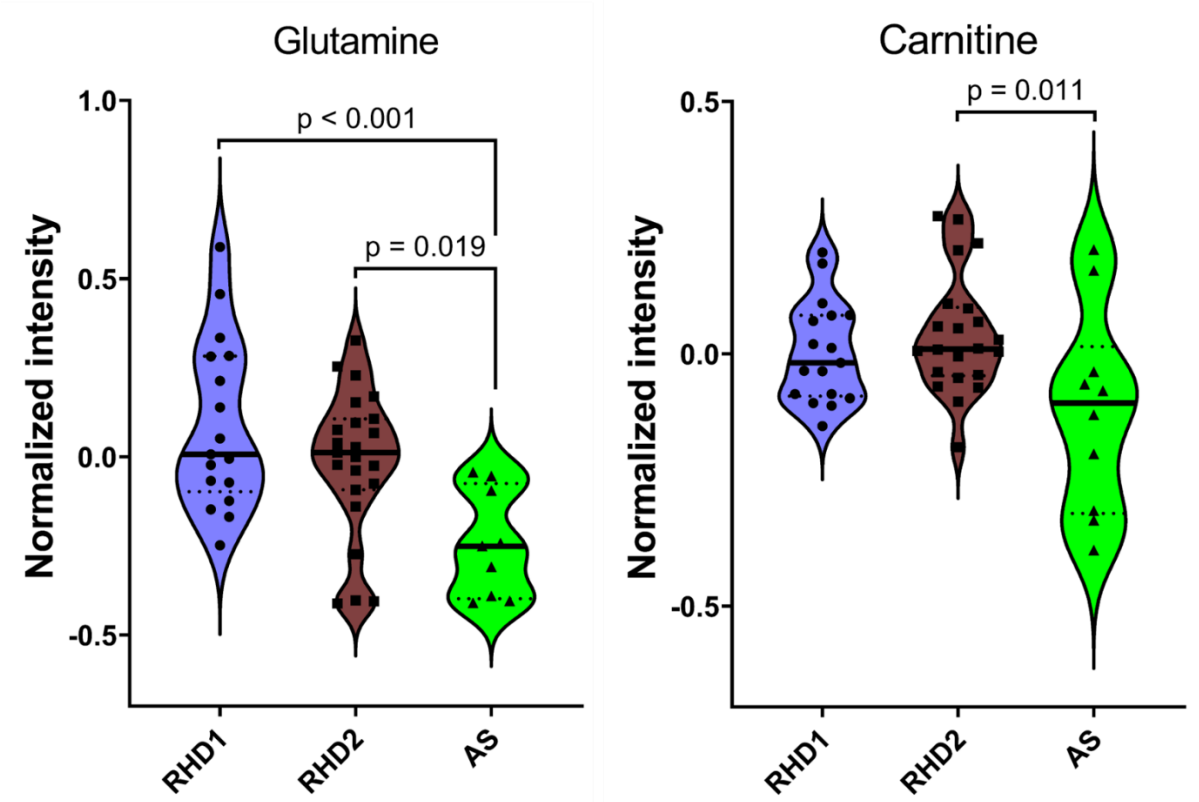


Figure 6.12. Metabolic biomarkers exploration. The metabolites that were elevated in RHD patients undergoing single valve replacement (RHD1) and double valve replacement (RHD2) and the degenerative AS.

6.3.3 Choline, carboxylic acids, and fatty acyls associated with histological and echocardiographic parameters of aortic and mitral valve pathology.

The annotated metabolites showed association with myxoid change, fibrosis, collagen deposition, and calcification which are some of the key histopathological features observed in RHD and degenerative AS. Calcification in the aortic or mitral valves significantly correlated with 13 metabolites, while myoid change showed correlation with 8 metabolites, fibrosis correlated with 1 metabolite after adjusting for the covariates (sex, age, race, BMI, hypertension, diabetes, and batch effect) **Table 6.2**. Choline (r = -0.52, p = 0.0003), propionic acid (r = 0.42, p = 0.004), histidine-betaxanthin (r = -0.4, p = 0.006) and hexadecanedioic acid (r = 0.41, p = 0.006), showed moderate correlations with histological evidence of calcification on the mitral and aortic valves after adjusting for covariates (**Table 6.2**). Furthermore, choline (r = 0.35, p = 0.019), glutamine (r = 0.43, p = 0.003), and proline (0.38, p = 0.003 showed moderate correlations with histological features of myxoid change on the studied heart valves.

Presence of fibrosis showed correlations with propionic acid (r = -0.36, p = 0.015) while adjusting for the covariates (**Table 6.2**).

Table 6.2. Histopathological features of the mitral and aortic valves correlating with tissue specific metabolites after adjusting for age, sex, race, BMI, hypertension, diabetes, and batch effect. (r= Pearsons correlation coefficient)

	Calcification	**p-value**	**Myxoid Change**	**p-value**	**Fibrosis**	**p-value**	**Collagen deposition**	**p-value**
Choline	**-0.52**	**<0.001**	**0.35**	**0.019**	-0.06	0.684	-0.18	0.240
Propionic acid	**0.42**	**0.004**	-0.07	0.638	**-0.36**	**0.015**	0.24	0.112
Histidine-betaxanthin	**-0.41**	**0.006**	0.28	0.062	0.02	0.912	-0.09	0.569
hexadecanedioic acid	**-0.41**	**0.006**	0.27	0.073	-0.07	0.652	-0.03	0.830
Triacylglycerol (16:0-17)	**-0.39**	**0.009**	0.10	0.529	-0.17	0.263	-0.03	0.866
Taurine	**-0.38**	**0.011**	0.13	0.410	-0.01	0.924	-0.11	0.490
Taurocholic acid	**0.38**	**0.012**	**-0.33**	**0.027**	-0.07	0.673	-0.13	0.418
D-Arabitol	**-0.37**	**0.013**	0.17	0.261	-0.28	0.062	0.02	0.893
Deoxycholic acid	**-0.33**	**0.030**	0.01	0.929	0.07	0.647	-0.05	0.753
MG(16:0/0:0/0:0)	**0.33**	**0.031**	-0.24	0.124	0.06	0.681	-0.16	0.294
Proline betaine	**0.31**	**0.04**	-0.20	0.204	-0.08	0.596	0.11	0.470
alpha-Methylhistidine	**0.30**	**0.047**	-0.23	0.131	-0.10	0.526	0.06	0.706
Triacylglycerol (17:1-18)	**0.30**	**0.049**	**-0.30**	**0.047**	-0.13	0.401	-0.06	0.711
Glutamine	-0.29	0.058	**0.43**	**0.003**	-0.01	0.931	0.07	0.660
Proline	-0.29	0.059	**0.38**	**0.011**	0.02	0.914	-0.06	0.681
Isoleucine	-0.23	0.137	0.29	0.056	-0.01	0.929	-0.01	0.971
Pyruvic acid	-0.22	0.156	**0.30**	**0.047**	-0.16	0.299	0.02	0.878
N-Methylglutamate	-0.10	0.511	**0.35**	**0.019**	0.24	0.115	0.11	0.492
Phosphatidylinositol	0.09	0.554	**0.30**	**0.046**	-0.08	0.625	0.19	0.206

TTE parameters showed associations with the annotated metabolites after adjusting for the covariates (sex, age, race, BMI, hypertension, diabetes, and batch effect). The mean aortic velocity (AV Vmean) and the mean aortic pressure gradient (AV mean PG) showed a negative but weak association with triacylglycerol (r = -0.39 and -0.36), taurine (r = -0.36 and -0.35), deoxycholic acid (r = -0.36 and -0.35), and histidine betaxanthin (r = -0.35 and -0.33) $p<0.05$ (**Table 6.3**). In addition, the mitral valve area (MVA_{PHT}) showed association with deoxycholic acid (r = -0.30), guanine (r = 0.31), phosphatidylinositol (r = 0.41), and creatine (r = -0.29) while the mitral valve maximum velocity (MV A Vmax) showed associations with triacylglycerol (r = 0.30), proline (r = 0.32), and methionine (r = 0.34) $p < 0.05$ (**Table 6.3**). Parameters of cardiac function and strain also showed associations with the annotated metabolites. The left ventricle end-diastolic volume (LVEDV) associated with propionic acid (r = 0.3), proline (r = -0.3), choline (r = -0.31), and histidine (r = -0.30), while the LV end-systolic volume (LVESV) showed

association with propionic acid (r = 0.3), proline (r = -0.30), histidine (r = -0.37), and glutamine (r = 0.33). The left ventricle ejection fraction (LVEF) showed association with triacylglycerol (r = -0.3) while left ventricle mass index (LVMI) associated with triacylglycerol (r = -0.3), taurine (r = -0.32), deoxycholic acid (r = -0.31), propionic acid (r = 0.29), and tryptophan (r = 0.3) $p<0.05$ (**Table 6.3**). The left atrial area (LA area) showed association with 3-methylhistidine (r = 0.31, p = 0.04).

Table 6.3. The correlation coefficients of metabolites that associated with haemodynamic and ventricular function parameters in patients with rheumatic heart disease and degenerative aortic stenosis. The correlation coefficients are adjusted for age, sex, race, BMI, hypertension, diabetes, and batch effect. (r= Pearsons correlation coefficient)

	AV Vmean(m/s)	**p-value**	**AV mean PG(mmHg)**	**p-value**	**LA Area (cm^2)**	**p-value**	**LVEDV (ml)**	**p-value**	**LVEF (%)**	**p-value**	**LVESV (ml)**	**p-value**	**LVMI (g/m^2)**	**p-value**	**MV DT (msec)**	**p-value**	**MVA_{PHT} (cm^2)**	**p-value**	**MV A Vmax (m/s)**	**p-value**
Triacylglycerol	**-0.39**	**0.009**	**-0.36**	**0.014**	0.03	0.866	-0.24	0.111	**-0.30**	**0.047**	-0.13	0.400	**-0.30**	**0.048**	0.23	0.116	-0.22	0.148	**0.30**	**0.045**
Taurine	**-0.36**	**0.015**	**-0.35**	**0.018**	-0.11	0.481	-0.29	0.058	-0.06	0.672	-0.23	0.120	**-0.32**	**0.033**	-0.02	0.888	-0.19	0.194	0.25	0.094
Deoxycholic acid	**-0.36**	**0.016**	**-0.35**	**0.019**	-0.01	0.927	-0.15	0.317	-0.03	0.849	-0.20	0.191	**-0.31**	**0.041**	0.11	0.481	**-0.30**	**0.043**	0.26	0.076
Histidine-betaxanthin	**-0.35**	**0.018**	**-0.33**	**0.028**	-0.14	0.351	-0.26	0.089	-0.10	0.504	-0.23	0.131	-0.28	0.066	-0.04	0.773	-0.14	0.368	0.20	0.185
Propionic acid	0.29	0.051	0.29	0.052	0.06	0.707	**0.30**	**0.046**	0.04	0.779	**0.30**	**0.047**	**0.29**	**0.049**	-0.19	0.205	0.06	0.700	-0.23	0.130
Proline	-0.24	0.106	-0.20	0.180	-0.22	0.149	**-0.30**	**0.043**	-0.12	0.435	**-0.30**	**0.047**	-0.18	0.230	0.18	0.243	-0.02	0.914	**0.32**	**0.033**
Tryptophan	0.24	0.115	0.24	0.106	-0.07	0.665	0.23	0.134	0.19	0.203	0.11	0.487	**0.30**	**0.045**	-0.14	0.353	-0.01	0.928	-0.26	0.076
Choline	-0.22	0.139	-0.21	0.165	-0.06	0.692	**-0.31**	**0.035**	-0.17	0.278	-0.26	0.082	-0.21	0.170	0.00	0.989	-0.10	0.507	0.10	0.527
Guanine	0.21	0.158	0.24	0.105	-0.06	0.682	0.21	0.159	0.07	0.657	0.16	0.294	0.25	0.092	**-0.31**	**0.035**	**0.31**	**0.038**	-0.25	0.091
Histidine	-0.21	0.166	-0.20	0.189	-0.20	0.192	**-0.30**	**0.046**	0.01	0.948	**-0.37**	**0.012**	-0.23	0.131	0.03	0.824	-0.19	0.204	0.22	0.134
alpha-Methylhistidine	0.20	0.184	0.15	0.333	0.24	0.110	-0.02	0.895	-0.17	0.261	-0.29	0.053	-0.05	0.734	0.05	0.731	-0.23	0.131	-0.06	0.694
Isoleucine	-0.20	0.186	-0.17	0.254	-0.12	0.433	-0.29	0.050	-0.22	0.150	-0.18	0.241	-0.20	0.193	0.07	0.626	0.03	0.841	0.17	0.249
Phosphatidylinositol	0.16	0.306	0.19	0.211	-0.08	0.617	-0.03	0.849	0.04	0.775	-0.07	0.647	0.11	0.469	-0.26	0.081	**0.41**	**0.005**	-0.20	0.180
Urea	0.15	0.340	0.06	0.675	0.26	0.087	0.02	0.916	-0.21	0.171	0.19	0.215	-0.09	0.537	0.02	0.920	-0.26	0.077	-0.03	0.822
Methionine	-0.15	0.342	-0.15	0.317	-0.18	0.229	-0.20	0.178	-0.11	0.457	-0.19	0.204	-0.13	0.399	0.23	0.124	-0.04	0.782	**0.34**	**0.020**
Monoglyceride	0.12	0.415	0.13	0.401	0.09	0.567	0.00	0.981	0.26	0.084	-0.11	0.461	0.07	0.664	-0.03	0.821	0.14	0.366	-0.06	0.670
Glutamic acid	0.12	0.449	0.09	0.575	0.18	0.241	0.17	0.266	-0.12	0.424	**0.33**	**0.026**	0.09	0.549	-0.02	0.901	0.01	0.949	-0.07	0.643
Pyruvic acid	-0.06	0.685	-0.04	0.783	-0.23	0.133	-0.22	0.139	-0.03	0.851	-0.29	0.054	-0.04	0.785	0.00	0.989	0.07	0.633	0.12	0.435
3-Methylhistidine	0.01	0.928	-0.03	0.855	**0.31**	**0.040**	-0.15	0.339	-0.19	0.218	0.01	0.929	-0.19	0.202	0.06	0.695	-0.28	0.061	0.09	0.569
Creatine	0.00	0.994	-0.03	0.862	0.15	0.317	-0.06	0.695	-0.09	0.556	0.04	0.774	-0.13	0.377	0.03	0.841	**-0.29**	**0.047**	0.12	0.409

6.3.4 Pathway analysis of the potential tissue specific biomarkers

The potential biomarkers had functional relevance where most of the metabolites were involved in the major energetic pathways and molecular signalling processes. The pathways with high coverage percentage between the total pathway metabolites and the detected metabolites were, arginine biosynthesis and taurine and hypotaurine metabolism (**Table 6.4**). The two pathways are normally involved in the major energetic processes.

Table 6.4. Pathway mapping of the metabolites detected in aortic and mitral valve samples of RHD and degenerative aortic stenosis patients undergoing valve replacement.

	Total pathway metabolites	Metabolites detected	Pathway coverage (%)	Raw p value	FDR adj. p value
Aminoacyl-tRNA biosynthesis	48	9	19	<0.001	<0.001
Arginine biosynthesis	14	5	36	<0.001	<0.001
Arginine and proline metabolism	38	6	16	<0.001	0.001
Glycine, serine and threonine metabolism	33	5	15	<0.001	0.005
Taurine and hypotaurine metabolism	8	3	38	<0.001	0.005
Histidine metabolism	16	3	19	0.002	0.039

Mapping of the detected metabolites onto KEGG database showed that most of the metabolites were associated with amino acids metabolism pathways. The significant and impactful pathways were taurine and hypotaurine metabolism, arginine biosynthesis, and arginine and proline metabolism pathways (**Figure 6.13**). From the pathway mapping, urea, creatine, taurine, histidine, and alpha-methyl histidine were highlighted as some of the potential biomarkers reported in our study.

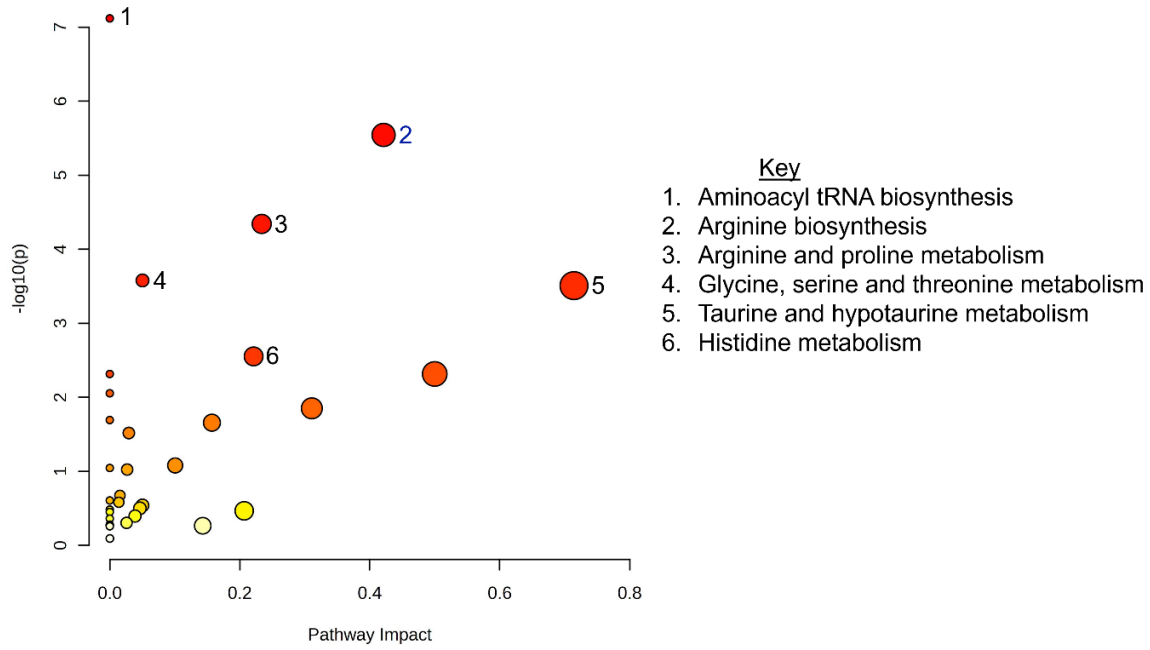


Figure 6.13. Pathway map of the detected metabolites after mapping to human metabolome database in KEGG. The highlighted pathways were significantly mapped after multiple testing adjustment (FDR adj. p value <0.05).

6.4 Discussion

The study suggest that histological findings of calcification moderately associated with lipids, and organonitrogens while myxoid change associated with glutamine on the mitral and aortic valves of RHD and degenerative AS patients. Furthermore, the echocardiography parameters showed weak associations with detected metabolites; specifically, mitral valve area parameters showed association with phosphatidylinositol.

Calcification was associated with decreased levels of choline. Choline is an essential nutrient that is responsible for production of neurotransmitters and glycerophospholipids.[238] Despite the conflicting reports on the risk of choline to cardiovascular diseases most of the studies have not found increased risks of cardiovascular diseases linked to dietary choline intake.[238] Though not detected in this study, trimethylamine N-oxide (TMAO), which is derived from choline, has been shown to promote osteogenic differentiation, which leads to calcification in calcific aortic stenosis.[239] Furthermore, the findings suggests that incidences of calcifications moderately associate with increased levels of fatty acids (propionic acid and hexadecanedioic acid). Propionic acid is a known anti-inflammatory short chain fatty acid, and it has been shown

to confer protection from cardiovascular disease and atherosclerosis.[240] The patients included in this study were at advanced stages of valvular damage due to calcification and therefore the increased levels of anti-inflammatory metabolites could be a biological response for inflammation homeostasis. Hexadecanedioic acid has been associated with beta-oxidation to produce ATP by the mitochondria.[241] Its role in the valve fragments is a concept that was beyond the scope of this study to answer but I could speculate that it was the remnants of the circulatory metabolites. Patients with advanced VHD present with heart failure and failing heart is known to have altered energy metabolism preference from beta-oxidation to glycolysis.[242,243]

Transthoracic echocardiography assessment of the valves is the mainstay for diagnosis of rheumatic heart disease and degenerative aortic stenosis. Some of the parameters used in the diagnosis are the flow dynamics, valve dimensions, and ventricular measurements.[9,244–247] To the best of our knowledge this is one of the few studies that have explored the correlation of the echocardiographic parameters to metabolomics biomarkers. Most of the ventricular function parameters (LA area, LVEDV, LVEF, LVESV and LVMI) and valve morphology parameters (AV Vmean, AV mean PG, MV DT, MVA and MV A Vmax) in this study showed weak to moderate associations with the metabolic biomarkers. It is inherently difficult to determine statistical significance on the metabolic biomarkers detected using untargeted metabolomics, there is therefore a need to perform validation tests on the metabolites that have potential relevance.[234] In addition, changes observed in the potential metabolic biomarkers could be influenced by the physiological changes seen clinically or could be also influenced by the exogenous factors such as environment, diet or exercise.[232–234,248] The samples included in the study were obtained from patients at different stages of valvular pathology and of diverse age groups. The demographic diversity could have contributed to the variations in the detected metabolic biomarkers thus affecting the strengths of the correlations. Furthermore, the functional analysis of the metabolites showed that most of the changed metabolites were involved in the main energetic pathways – a phenomenon that has been reported in other studies.[19,235]

6.5 Limitations

Obtaining healthy control valve biopsies is challenging and is marred by sophisticated ethical issues. Therefore, in this study we compared RHD against degenerative AS valves therefore we could not ascertain if the observed metabolic biomarkers were markers of disease or just markers due to differences of disease physiology. Furthermore, the study could not ascertain the causality of the metabolic biomarkers that was associated with the clinical features of valve pathology. In addition, due to the small sample size owing to the complexity of obtaining the samples, the study could not have enough statistical power to compare the differences between the valve lesions.

6.6 Concluding remarks

In conclusion, in this chapter we have shown the metabolic biomarkers differentially expressed in RHD patients with single or double valve replacement against patients with degenerative AS. Furthermore, we have reported metabolites associating with histopathological and TTE parameters of VHD. As is common to most metabolomics studies this study too could not ascertain the causality of the identified biomarkers. We therefore recommend a follow-up study with a larger sample size and using a different study design comparing specific pathological regions of interest for spatial localisation studies.

6.7 Contributions and acknowledgements

The heart valve samples were collected from patients undergoing valve replacement surgery at the Christiaan Barnard Division of Cardiothoracic Surgery, Groote Schuur Hospital, Cape Town after recruitment and consenting by D Mutithu assisted by Dr R Manganyi, O Aremu, and Dr E Lumngwena. The clinical data was collected and curated by D Mutithu and reviewed by Prof N Ntusi. The samples were processed by D Mutithu. The H&E slides were prepared by D Mutithu assisted by S Govender and reviewed by Prof D Govender and Dr R Roberts. The LC-MS/MS experiments were performed by D Mutithu assisted by A Evans. The LC-

MS/MS data was processed and statistically analysed by D Mutithu and reviewed by Prof J Kirwan.

7 LOCALISATION OF CIRCULATORY AND TISSUE-SPECIFIC BIOMARKERS ON VALVE BIOPSIES USING MALDI-MS IMAGING[6]

7.1 Introduction

Metabolomics reports on the changes in the metabolome within a biological system. It has been shown that metabolomics is valuable in understanding the diseases processes and pathomechanisms due to its close proximity to the organism's physiology.[19,227] Metabolomics techniques are used to perform spatial localisation of metabolic biomarkers which is akin to molecular tissue typing to investigate the microenvironments on the studied tissues.[228–230] RHD and degenerative AS have distinct pathophysiological features with the rheumatic heart dominated by myxoid change, and collagenation while degenerative AS would mostly have features of fibrosis and calcification.[154,231] At the microenvironment level, the physiological changes are preceded by changes in the metabolome of the affected cells. It is therefore expected that different regions of pathology would have different metabolic signatures.[230] MALDI-MSI is used for spatial localisation of ions on tissues by recreating an ions intensity image on the tissues being analysed.[230,235] Few studies have reported spatial localisation of metabolic biomarkers in RHD and degenerative AS. It is however challenging to annotate ions detected using MALDI-MSI. Therefore, MALDI-MSI is suitable for a targeted metabolomics analysis where it is used with another untargeted MS/MS technique where the detected ions are annotated to the respective metabolites.[229,230]

[6]This chapter is based on a manuscript:
Mutithu, D. W., Kirwan, J. A., Adeola, H. A., Aremu, O. O., Lumngwena, E. N., Familusi, M., ... & Ntusi, N. A. (2024). *Major energetic metabolites associated with echocardiographic and histological features of aortic and mitral valve disease* [Unpublished manuscript]. Department of Medicine, University of Cape Town, Cape Town, South Africa.

In this study, we aimed to perform spatial localisation of the important tissue specific biomarkers using targeted MALDI-MSI. The samples were obtained from RHD patients having single (mitral valve only) or double (mitral and aortic valves) valve replacement and degenerative AS patients undergoing AVR (aortic valve only).

7.2 Methods

7.2.1 Spatial localisation of potential biomarkers with MALDI-MSI

MALDI-MSI was used to investigate the localisation of the important metabolites on aortic and mitral valves. Valve samples obtained from the patients described in section 6.3.1 were used for mass spectrometry imaging experiments. The valve samples had been harvested and stored as stated in section 3.2.1. The MALDI-MSI experiments were adapted from Veerasammy et al.[249], Anderson et al.[250], and Zhu et al.[251]. Thin cryosections (5 µm) of frozen valve fragments were thaw-mounted on ITO coated slides (Sigma). The slides were dried in a desiccator for about 15 minutes. Furthermore, 40 mg/ml of 2,5-dihydrobenzoic acid (DHB) matrix was prepared by dissolving 200mg of DHB in 5ml of 50% methanol and mixed with 0.1% formic acid. The solution was sonicated for 15 minutes. Using the HTX TM-Sprayer (HTX Technologies, Chapel Hill, NC, USA), 12 layers of the matrix were applied on the ITO slides. The slides were then dried before the MALDI-MSI analysis using a MALDI- QTOF. The slide images were obtained by scanning and processing the images with CyberViewX and the image uploaded to FlexImaging 5.1 software. The MALDI-MSI data was acquired using rapifleX® MALDI Tissuetyper® with flexControl 4.0 and flexImaging 5.1 (Bruker Scientific LLC, Billerica, MA, USA), in positive reflector mode over an m/z range of 50 - 1,200 Da at 50 µm raster width. To calibrate the MALDI TOF equipment for low molecular weight experiments, a mix of the matrices was used (20 mg of each matrix was dissolved in 1ml of 50% methanol mixed with 0.1% formic acid); DHB (DHB[M+H] *m/z* 155.03444; DHB[2M+H] m/z 309.06105), HCCA (a-Cyano-4-hydroxycinnamic acid) *(*HCCA[M+H] *m/z* 190.04987; HCCA[2M+H] m/z 379.09246), and SA (sinapic acid) (SA [M+H] *m/z* 225.07633; SA[2M+H] m/z 449.14483). The spectra were accumulated from 200 laser shots at 10kHz frequency. All the m/z features detected on tissues

were obtained in positive ionisation mode. The statistical analysis and data visualization were done with SCiLS™ Lab 2020a. During data processing and analysis, the *m/z* values of the metabolites found to be important biomarkers from the tandem LC-MS/MS analysis were localised on the aortic and mitral valves obtained from patients with rheumatic heart disease and degenerative aortic stenosis.

7.2.2 Potential biomarkers statistical analysis

Raw positive mode liquid chromatography tandem mass spectrometry (LC-MS/MS) data were processed with MS-Convert in ProteoWizard 3.0.1908[126] and MS-DIAL 4.38[96] as stated in chapter 6. To identify potentially important features between groups, univariate ANOVA analysis was used, while multidimensional reduction (PCA and PLS-DA) analyses were used to visualise group differences. Furthermore, correlation analyses were performed between perturbed metabolites and; the histopathological features, TTE valve parameters for morphology and flow dynamics, while adjusting for covariates. The important metabolic biomarkers were further localised on the valve fragments using MALDI MSI. The *m/z* features and metabolites with a p-value <0.05 were considered significantly altered.

7.3 Results

7.3.1 MALDI equipment calibration

The equipment was calibrated with a mix of the matrices (DHB, HCCA, and SA) that were targeted at *m/z* DHB (DHB[M+H] *m/z* 155.03444; DHB[2M+H] *m/z* 309.06105), HCCA (HCCA[M+H] *m/z* 190.04987; HCCA[2M+H] *m/z* 379.09246), and SA (SA[M+H] *m/z* 225.07633; SA[2M+H] *m/z* 449.14483) that were all singly charged (**Figure 7.1**). Histidine-betaxanthin (*m/z* 348.991), hexadecadioic acid (*m/z* 309.206), glutamine (*m/z* 147.077), stearamide (*m/z* 284.296), and 3-methylglutaconic acid (*m/z* 145.050) were some of the potential biomarkers identified after ANOVA and correlations analysis. To explore the spatial metabolic profiles, 3 biological repeats were used of mitral valve from RHD undergoing single valve (mitral valve) replacement (RHD 1 MV), mitral valve obtained from RHD patients

undergoing double valve replacement (RHD 2 MV), aortic valve obtained from RHD patients undergoing double valve replacement (RHD 2 AV), and aortic valve from degenerative aortic stenosis patients (AS AV).

7.3.2 MALDI pilot experiments

The aortic and mitral valves that were obtained from RHD patients and degenerative AS patients were histologically assessed by a pathologist using H&E-stained thin sections. The mitral valve obtained from a RHD patient with double valve replacement showed histological evidence of collagenation, fibrosis and focal calcification. There was also evidence of neovascularisation, isolated lymphocytes, and plasma cells but there were no Aschoff bodies. The aortic valve obtained from the same patient showed evidence of calcification, fibrosis due to collagenation, and there was no evidence of inflammation and neovascularisation. The mitral valve obtained from the RHD patient with single valve replacement showed evidence of neovascularisation and stromal fibrosis with collagen deposition. In addition, chronic inflammation due to presence of lymphocytes, plasma cells, Russel bodies, and other immunoglobulins that indicate chronic inflammation. Furthermore, there was evidence of dystrophic calcification. The aortic valve from patients diagnosed with degenerative AS had histological features of dystrophic calcification and fibrosis due to collagenization. In addition, we attempted to determine features that would differentiate between two tissue types using the segmentation algorithm within SCiLS Lab. The m/z features were selected based on their ROC-AUC values of the m/z features differentiating, RHD 1, RHD 2 and degenerative AS.

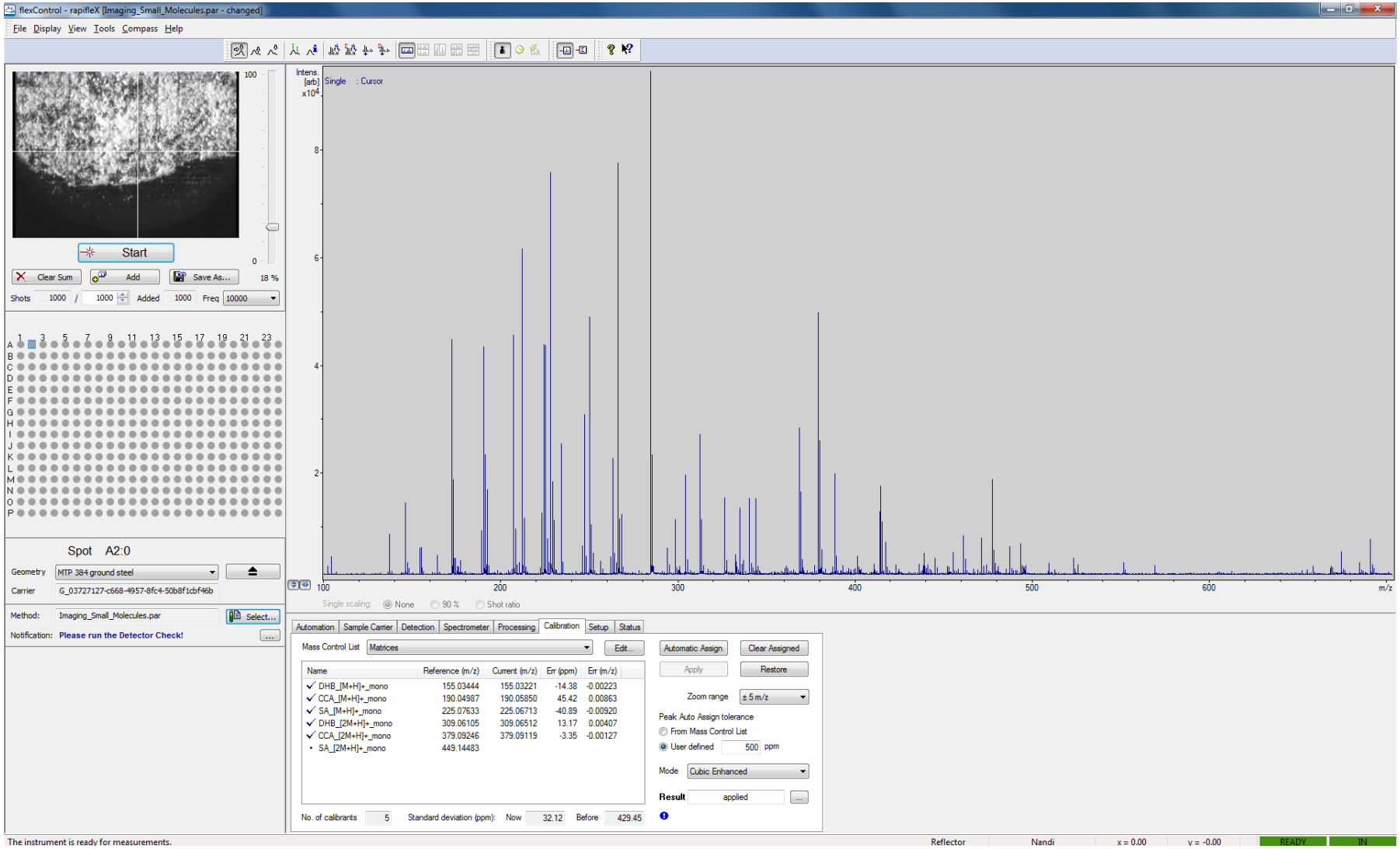


Figure 7.1. Equipment calibration with the matrix mix of 2,5-dihydrobenzoic acid (DHB), a-Cyano-4-hydroxycinnamic acid (HCCA), and sinapinic acid (SA).

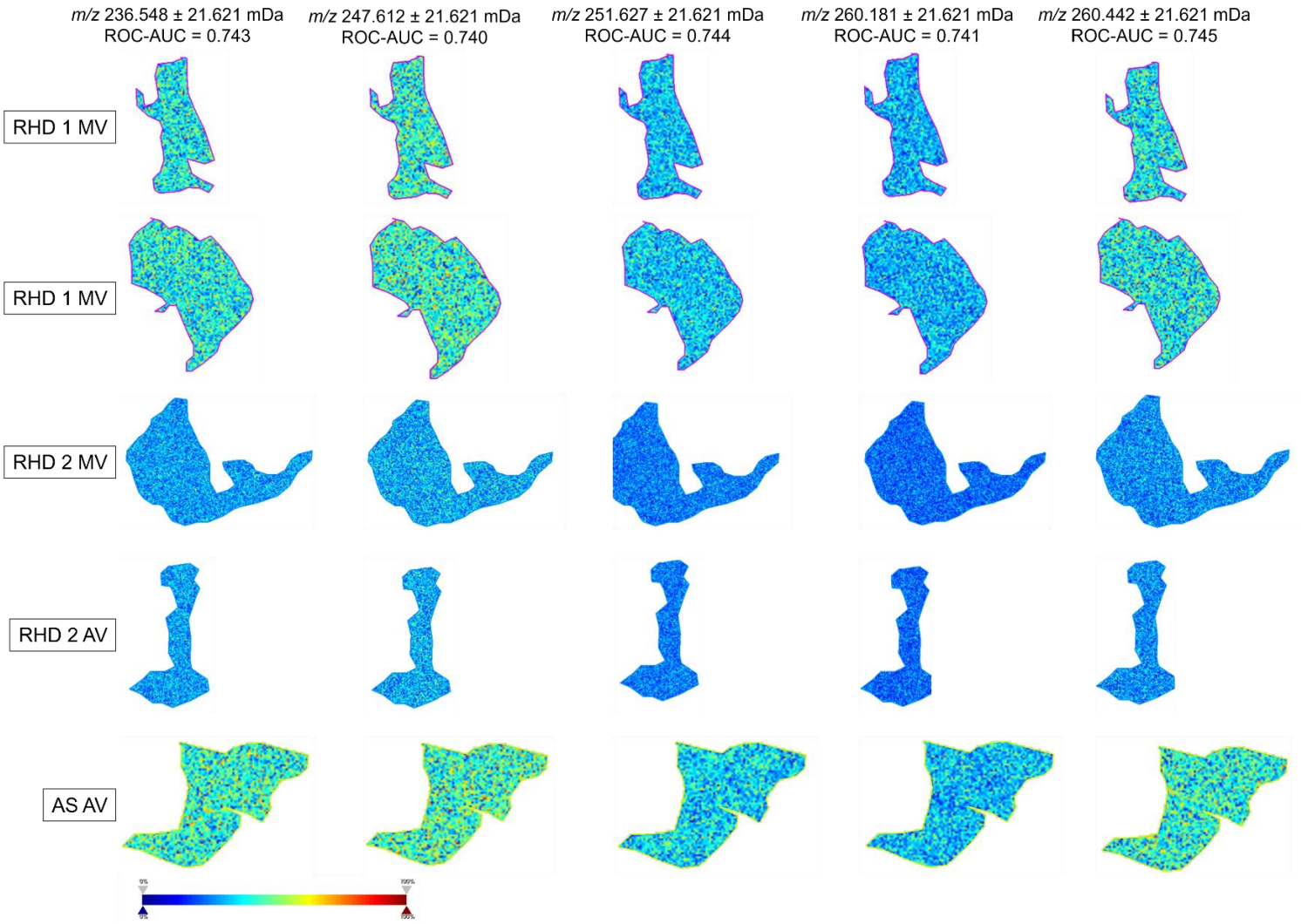


Figure 7.2. The *m/z* images of the ions that differentiated between different valve types while using the SCiLS lab segmentation algorithm. RHD 1 MV; mitral valve obtained from rheumatic heart disease undergoing mitral valve replacement, RHD 2 MV; mitral valve obtained from rheumatic heart disease undergoing double valve replacement, RHD 2 AV; aortic valve obtained from rheumatic heart disease patient undergoing double valve replacement, AS AV; aortic valve obtained from degenerative aortic stenosis patient undergoing aortic valve replacement surgery.

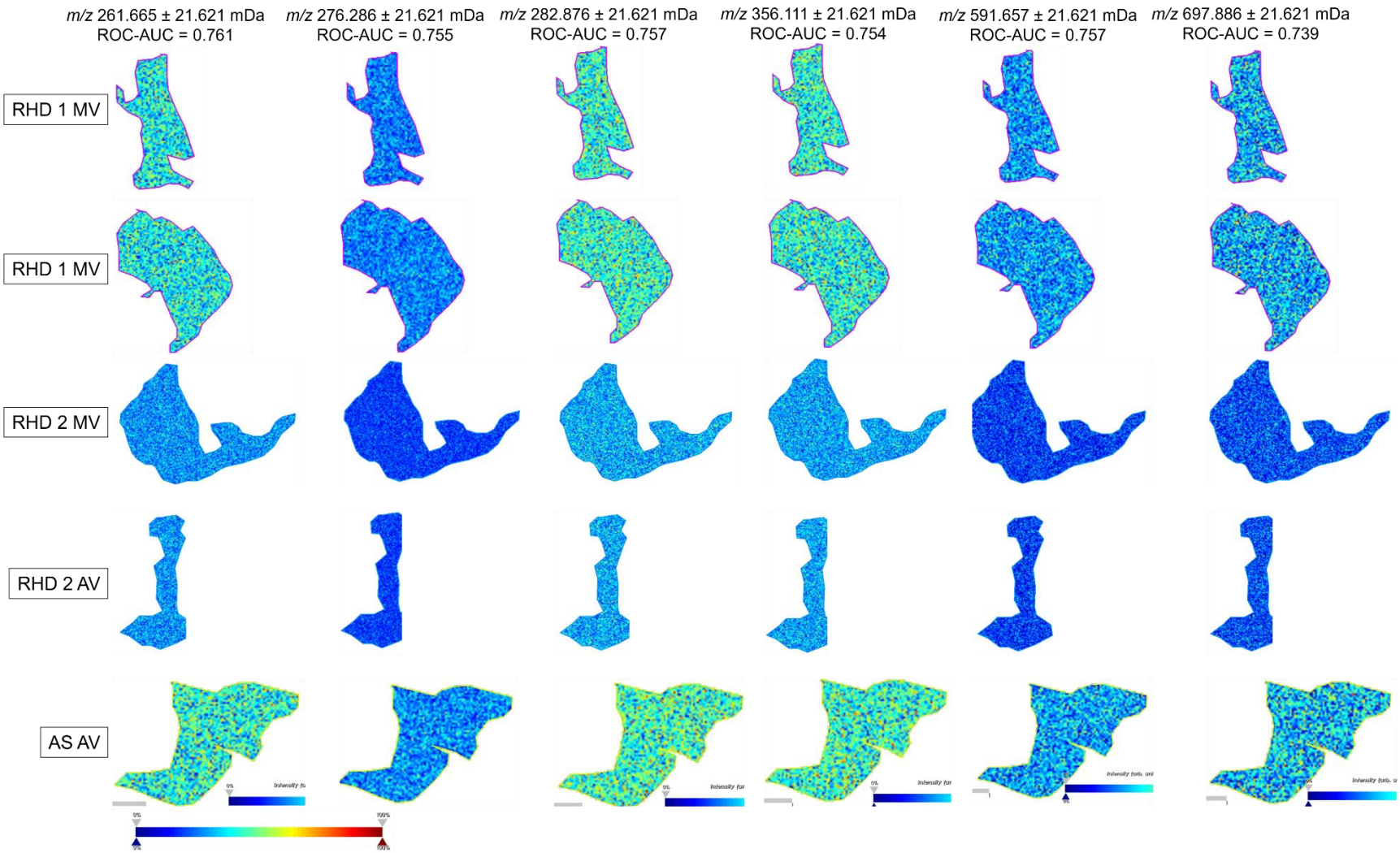


Figure 7.3. Continued m/z images of the ions that differentiated between different valve types while using the SCiLS lab segmentation algorithm. RHD 1 MV; mitral valve obtained from rheumatic heart disease undergoing mitral valve replacement, RHD 2 MV; mitral valve obtained from rheumatic heart disease undergoing double valve replacement, RHD 2 AV; aortic valve obtained from rheumatic heart disease patient undergoing double valve replacement, AS AV; aortic valve obtained from degenerative aortic stenosis patient undergoing aortic valve replacement surgery.

The identified *m/z* features differentiated the valve fragments obtained from RHD patients undergoing single valve or double valve replacement surgery and AS patients having aortic valve replacement surgery (**Figure 7.2 and Figure 7.3**). The identified features were evenly distributed on the tissues and did not show apparent distribution as per the pathology features. In addition, the identified features were difficult to determine which metabolite they represented. Therefore, the findings could not be interpreted into the metabolic biomarkers. To overcome the challenge, I opted to localise the important LC-MS/MS biomarkers discovered in the valve's homogenates.

7.3.3 Spatial localisation of the LC-MS/MS important biomarkers using MALDI-MSI

Histidine-betaxanthin (*m/z* 348.991), hexadecadioic acid (*m/z* 309.206), glutamine (*m/z* 147.077), stearamide (*m/z* 284.296), and 3-methylglutaconic acid (*m/z* 145.050) were some of the potential biomarkers identified after ANOVA and correlations analysis of LC-MS/MS biomarkers. The potential biomarkers also showed patterns of spatial localisation on the aortic and mitral valves after analysis and detection with MALDI-MSI. The different tissue types showed distinct patterns of the detected metabolites. However, the spatial localisation did not correspond to any specific pathology (**Figure 7.4**).

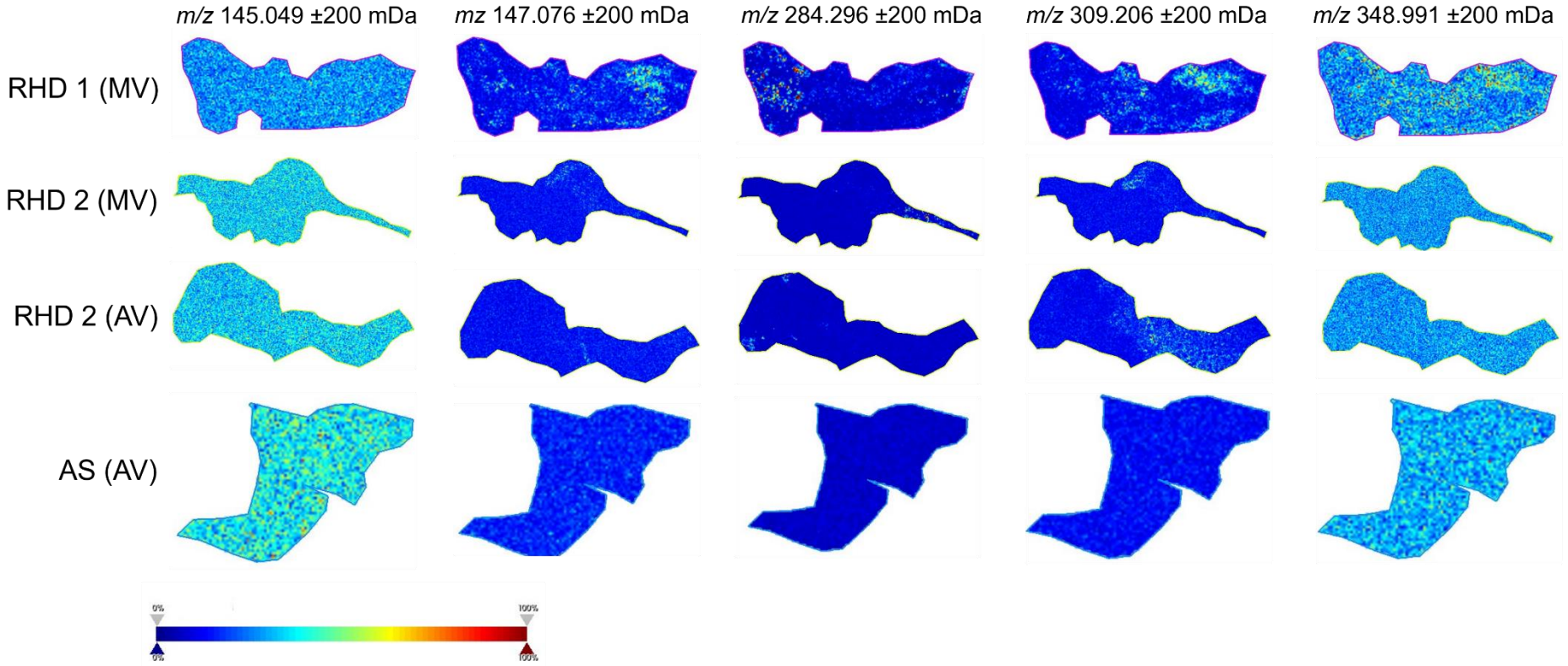


Figure 7.4. MALDI-MSI spatial localisation of the significant metabolites on the mitral (MV) and aortic (AV) valves obtained from rheumatic heart disease patients with single valve replacement (RHD 1) or double valve replacement (RHD 2), and degenerative aortic stenosis (AS). The localised biomarkers were 3-methylglutaconic acid (m/z 145.050), glutamine (m/z 147.077), stearamide (m/z 284.296), hexadecadioic acid (m/z 309.206), and histidine-betaxanthin (m/z 348.991).

7.4 Discussion

Metabolic changes at the microenvironment are associated with physiological pathologies; such findings have extensively been reported in tumour research.[252,253] The advancement of MALDI-MSI technology have enabled close to a single cell analysis of the metabolic changes. In this chapter, I have shown the analysis of *m/z* features that were different among the different valves. I have also shown their potential capabilities to differentiate between the tissue types studied based on their intensities. However, despite their promising applications in the differentiation of the tissues, the *m/z* features would certainly not find application in the biomarker discovery. Annotation of the MALDI-MSI *m/z* features is challenging due to the single dimensionality of the obtained data.[236,251] The observed m/z features with capabilities of differentiating the tissue valves did not show apparent association with the regions of pathology. In any case, the intensities were evenly distributed across the tissue. The matrix used in MALDI-MSI affects the detection of the low molecular weight ions.[236,251] Furthermore, the tissues used in this study were at advanced stages of pathology therefore there were no distinct regions that differentiated between normal and diseased regions. This may perhaps indicate that the metabolic microenvironments are heterogeneous around the tissue. In addition, heart valves are generally considered avascular, nutrition is supplied through diffusion which may mean the metabolites are distributed around the entire sections indiscriminately.[254,255] Due to the challenges in annotating the m/z features identified through MALDI-MSI, the features could not be annotated to the representative tentative metabolites.[229,230,236] We therefore opted to conduct a targeted MALDI-MSI analysis by focusing on the annotated metabolic biomarkers that had been identified from the LC-MS/MS experiments.

The spatial distribution of the potential biomarkers was assessed on thin tissue sections of the fresh frozen valve fragments obtained from RHD and degenerative AS. The LC-MS/MS biomarkers obtained in this study were obtained from the tissue homogenate. From our previous findings,[154] we had shown the histopathological features on the aortic or mitral valves.

Furthermore, there are regions on the excised valves where there are features of disease and there are zones where the tissue is normal. The spatial localisation of the metabolites focuses on the microenvironment metabolome that would show metabolites correlating with zones of pathology on the analysed tissues.[229,230] Some of the limitations of imaging mass spectrometry are the identification of the metabolites, matrix clusters, matrix ion suppression, and resolution of the peaks.[228] Therefore, in this study the spatial localisation of the metabolites that were considered important from the LC-MS/MS analysis of the valve's homogenate were reported. The findings showed feasibility that the 5 important biomarkers from the LC-MS/MS analysis could be traced onto the aortic and mitral valves from RHD patients with single or double valve replacement and aortic valve from degenerative AS patients. To the best of our knowledge, this study is among the few reporting on spatial localisation of metabolites on cardiac valves.[236]

7.5 Limitations

In this study the metabolic signatures were compared between valve biopsies of valvular heart disease and there were no healthy tissues to compare against. The segmentation analysis to determine the features that would differentiate between the tissues from different disease types showed several ions with good ROC-AUC values. However, based on the study design we could not determine how the metabolites and ions vary on normal tissues. In addition, the distribution of the metabolic biomarkers and the discriminant ions did not show associations with specific regions of pathology. This could be due to the heterogeneous distribution of metabolites on the valves since they are supplied through diffusion.

7.6 Concluding remarks

In this chapter we have shown the untargeted MALDI-MSI approach to extract important ions that differentiated between valve tissues obtained from RHD and degenerative AS patients. Furthermore, metabolites that were found significantly different between RHD and degenerative AS were localised on the valve fragments. The study was exploratory, we

therefore recommend further studies using healthy tissues to investigate disease related biomarkers and localise them on the tissues.

7.7 Contributions and acknowledgements

The heart valve samples were collected from patients undergoing valve replacement surgery at the Christiaan Barnard Division of Cardiothoracic Surgery, Groote Schuur Hospital, Cape Town after recruitment and consenting by D Mutithu assisted by Dr R Manganyi, O Aremu, and Dr E Lumngwena. The clinical data was collected and curated by D Mutithu and reviewed by Prof N Ntusi. The samples were processed by D Mutithu. The H&E slides were prepared by D Mutithu assisted by S Govender and reviewed by Prof D Govender and Dr R Roberts. The MALDI-MSI experiments were performed by D Mutithu assisted by N Mehlala and Dr H Adeola. The LC-MS/MS and MALDI-MSI data was processed and statistically analysed by D Mutithu and reviewed by Prof J Kirwan.

8 CONCLUSION

RHD is prevalent in resource limited regions of the world; the prevalence of degenerative AS is also on the rise due to the improving life expectancy, and a rise in lifestyle diseases especially in SSA. There are limited facilities to manage patients with RHD. The increasing prevalence of degenerative AS may worsen the clinical landscape of VHD in the region. There exist diagnostic tools and techniques for VHD, however there are no methods for early detection. Most of the patients presents with advanced VHD and the only feasible treatment is valve replacement or repair. Therefore, the main objective of this doctorate was to explore metabolic biomarkers in patients with RHD and degenerative AS patients. The identified biomarkers are expected to expand our understanding of the pathomechanisms in RHD and degenerative AS and propose biomarkers with diagnostic capabilities.

The histopathological features of degenerative AS and RHD were investigated on the valve biopsies. The most predominant features were calcification in degenerative AS, collagen deposition and inflammation. The findings also confirmed that rheumatic valvulitis could affect the aortic valve even though mitral valve is the most frequently affected. We have also shown a case of acute inflammation on chronic RHD. This supported an existing theory that there could be acute inflammation on chronic inflammation in RHD. The existence of chronic inflammation admixed with acute rheumatic valvulitis support the need for life-long prophylaxis in rheumatic valvulitis patients. Therefore, in ARF endemic regions the presence of ARF should not be ruled out in the differential diagnosis despite not meeting the criteria for diagnosis as per the revised Jones criteria 2015.

The study used untargeted LC-MS/MS metabolomics on serum obtained from patients undergoing valve replacement surgery. The findings suggested that there are metabolites that are differentially expressed in individuals with RHD, degenerative AS, and those without VHD.

The changed metabolites were mainly involved in the major energetic pathways which confirmed that failing hearts has major changes in their energy requirements. Furthermore, I also showed metabolites that had the capability of differentiating RHD and degenerative AS patients from healthy individuals. The gene-metabolite interaction analysis also showed involvement of genes responsible for lipids and fatty acids metabolism. Some of the circulatory metabolites also showed associations with parameters of cardiac remodelling in RHD and degenerative AS patients. The findings showed the involvement of the major energetic pathways possibly because the metabolites were highly abundant and the less abundant were masked by the matrix effect. It could also be that the patients included in the study were mostly at the chronic stages of the disease therefore most of them had heart failure symptoms. Furthermore, the findings found weak associations of the cardiac remodelling parameters with the detected metabolites because the shifts in metabolic signatures detected in the blood samples could have been epiphenomena of pathologies happening elsewhere in the body. Therefore, our study investigated metabolic signatures in the heart valves excised from patients undergoing valve replacement surgery.

As stated above, the changes in circulatory metabolic biomarkers could be because of pathological changes happening elsewhere in the body. Therefore, in this work, tissue specific metabolic biomarkers detected in the heart valves homogenates were explored. The study found metabolites that were differentially expressed in valves obtained from patients diagnosed with RHD and degenerative AS. The findings suggested that the detected choline, carboxylic acids, and fatty acyls associated with the histological and echocardiographic parameters of aortic and mitral valves pathology. The metabolites detected in the mitral and aortic valves and highlighted to be important were associated with the major energetic pathways and molecular signalling processes. The findings described the metabolic biomarkers different between RHD with single or double valve replacement and degenerative AS. However, due to the limitations of obtaining healthy control valve fragments, the study could not compare the differences against healthy samples. Furthermore, the study could not

establish the causal relationship between the changed metabolites and the pathophysiological features observed.

At the microenvironment levels it has been shown that physiological changes are preceded by metabolic changes. Very few studies have explored the metabolic signatures on the valve tissues however very few studies have explored the metabolic signatures at the microenvironment levels on heart valves. The findings using MALDI-MSI reported on the ions that could differentiate the heart valves obtained from RHD or degenerative AS based on the receiver operator area under the curve. Furthermore, I have also shown that metabolites detected in the valve's homogenates could be localized on the valve tissues obtained from RHD and degenerative AS patients. From the histological analysis of the tissues included into the study, the regions of pathology were heterogeneous therefore it was challenging to perform and in situ comparison of regions on the valves. Furthermore, the detected metabolites signals were generally spread across the surface of the valves studied. Since valves are considered avascular, and the nutrition is spread through diffusion which may suggest the difficulties of localising metabolites at specific regions. The study reported the metabolites localisation on valves obtained from patients with VHD, I therefore propose further studies to compare the metabolic signatures between VHD and healthy samples.

While this study has described circulatory and tissue specific metabolic biomarkers in RHD and degenerative AS, it still does not fully explain the metabolic pathomechanisms of VHD. Furthermore, this study had hypothesized that exploring metabolic profile could diagnose RHD and degenerative aortic stenosis. However, due to sample size limitations, the study could not have a validation and a test sample to validate the diagnostic models. In this study the correlation of the metabolic signatures with the clinical parameters was investigated however, the presence and effect of the confounding factors was not explored. The future work is to design and conduct a larger cohort study from different sites to validate the reported biomarkers. In addition, our study proposes future studies to determine the usability of cadaveric tissues as controls in metabolomics studies. The next study could consider

designing RHD and degenerative aortic stenosis animal models to investigate the mechanistic processes leading to pathologies associated with the chronic RHD and degenerative aortic stenosis. We also propose using multi-omics techniques (proteomics, genomics, or transcriptomics) to facilitate a multimodal exploration of the pathomechanisms in RHD and degenerative aortic stenosis. Transportation of samples between study sites is one of the main causes of signal variations in metabolomics studies. Previous studies have shown that the use of dried plasma spots would be a suitable storage method for plasma and other blood samples that does not require specialised handling for shipping. Therefore, we propose a study to investigate the suitability of using dried plasma spots for sample shipping and storage between study sites in SSA.

BIBLIOGRAPHY

1. Nkomo VT. Epidemiology and prevention of valvular heart diseases and infective endocarditis in Africa. *Heart*. 2007;93(12):1510-1519. doi:10.1136/hrt.2007.118810

2. Coffey S, Roberts-Thomson R, Brown A, et al. Global epidemiology of valvular heart disease. *Nat Rev Cardiol*. Published online 2021. doi:10.1038/s41569-021-00570-z

3. Passos LSA, Nunes MCP, Aikawa E. Rheumatic Heart Valve Disease Pathophysiology and Underlying Mechanisms. *Front Cardiovasc Med*. 2021;7:612716. doi:10.3389/fcvm.2020.612716

4. Hurst JR, Kasper KJ, Sule AN, McCormick JK. Streptococcal pharyngitis and rheumatic heart disease: the superantigen hypothesis revisited. *Infect Genet Evol*. 2018;61:160-175. doi:10.1016/j.meegid.2018.03.006

5. Muhamed B, Parks T, Sliwa K. Genetics of rheumatic fever and rheumatic heart disease. *Nat Rev Cardiol*. 2020;17(3):145-154. doi:10.1038/s41569-019-0258-2

6. Antoine C, Mantovani F, Benfari G, et al. Pathophysiology of Degenerative Mitral Regurgitation. *Circ Cardiovasc Imaging*. 2018;11(1):e005971. doi:10.1161/CIRCIMAGING.116.005971

7. Lincoln J, Garg V. Etiology of valvular heart disease-genetic and developmental origins. *Circ J*. 2014;78(8):1801-1807. doi:10.1253/circj.cj-14-0510

8. Goldbarg SH, Elmariah S, Miller MA, Fuster V. Insights Into Degenerative Aortic Valve Disease. *J Am Coll Cardiol*. 2007;50(13):1205-1213. doi:https://doi.org/10.1016/j.jacc.2007.06.024

9. Reményi B, Wilson N, Steer A, et al. World Heart Federation criteria for echocardiographic diagnosis of rheumatic heart disease--an evidence-based guideline. *Nat Rev Cardiol*. 2012;9(5):297-309. doi:10.1038/nrcardio.2012.7

10. Otto CM, Nishimura RA, Bonow RO, et al. 2020 ACC/AHA Guideline for the Management of Patients With Valvular Heart Disease: A Report of the American College of Cardiology/American Heart Association Joint Committee on Clinical Practice Guidelines. *Circulation*. 2021;143(5):e72-e227. doi:10.1161/CIR.0000000000000923

11. Treibel TA, Bennett J, Cavalcante JL. Editorial: Multimodality Imaging in Valvular Heart Disease . *Front Cardiovasc Med* . 2021;8.

12. Bartoli-Leonard F, Aikawa E. Heart Valve Disease: Challenges and New Opportunities . *Front Cardiovasc Med* . 2020;7.

13. Hasin Y, Seldin M, Lusis A. Multi-omics approaches to disease. *Genome Biol*. 2017;18(1):83. doi:10.1186/s13059-017-1215-1

14. Chessa M, Panebianco M, Corbu S, et al. Urinary Metabolomics Study of Patients with Bicuspid Aortic Valve Disease. *Mol* . 2021;26(14). doi:10.3390/molecules26144220

15. Xiong T-Y, Liu C, Liao Y-B, et al. Differences in metabolic profiles between bicuspid and tricuspid aortic stenosis in the setting of transcatheter aortic valve replacement. *BMC Cardiovasc Disord*. 2020;20(1):229. doi:10.1186/s12872-020-01491-4

16. Jiang L, Wang J, Li R, et al. Disturbed energy and amino acid metabolism with their diagnostic potential in mitral valve disease revealed by untargeted plasma metabolic

profiling. *Metabolomics*. Published online 2019:1-12. doi:10.1007/s11306-019-1518-1

17. Disatian S, Lacerda C, Orton EC. Tryptophan hydroxylase 1 expression is increased in phenotype-altered canine and human degenerative myxomatous mitral valves. *J Heart Valve Dis*. 2010;19(1):71-78.

18. Jiang L, Wang J, Li R, et al. Disturbed energy and amino acid metabolism with their diagnostic potential in mitral valve disease revealed by untargeted plasma metabolic profiling. *Metabolomics*. 2019;15(4):57. doi:10.1007/s11306-019-1518-1

19. Mutithu DW, Kirwan JA, Adeola HA, et al. High-Throughput Metabolomics Applications in Pathogenesis and Diagnosis of Valvular Heart Disease. *RCM*. 2023;24(6):169-null. doi:10.31083/j.rcm2406169

20. Naghavi M, Wang H, Lozano R, et al. Global, regional, and national age-sex specific all-cause and cause-specific mortality for 240 causes of death, 1990-2013: A systematic analysis for the Global Burden of Disease Study 2013. *Lancet*. 2015;385(9963):117-171. doi:10.1016/S0140-6736(14)61682-2

21. Watkins DA, Johnson CO, Colquhoun SM, et al. Global, Regional, and National Burden of Rheumatic Heart Disease, 1990-2015. *N Engl J Med*. 2017;377(8):713-722. doi:10.1056/NEJMoa1603693

22. Rothenbühler M, O'Sullivan CJ, Stortecky S, et al. Active surveillance for rheumatic heart disease in endemic regions: a systematic review and meta-analysis of prevalence among children and adolescents. *Lancet Glob Heal*. 2014;2(12):e717-e726. doi:10.1016/S2214-109X(14)70310-9

23. Rwebembera J, Manyilirah W, Zhu ZW, et al. Prevalence and characteristics of primary left-sided valve disease in a cohort of 15,000 patients undergoing echocardiography studies in a tertiary hospital in Uganda. *BMC Cardiovasc Disord*. 2018;18(1). doi:10.1186/s12872-018-0813-5

24. Nkomo VT, Gardin JM, Skelton TN, Gottdiener JS, Scott CG, Enriquez-Sarano M. Burden of valvular heart diseases: a population-based study. *Lancet*. 2006;368(9540):1005-1011. doi:10.1016/S0140-6736(06)69208-8

25. Aikawa E, Schoen FJ. Chapter 9 - Calcific and Degenerative Heart Valve Disease. In: Willis MS, Homeister JW, Stone JRBT-C and MP of CD, eds. Academic Press; 2014:161-180. doi:https://doi.org/10.1016/B978-0-12-405206-2.00009-0

26. Caira FC, Stock SR, Gleason TG, et al. Human degenerative valve disease is associated with up-regulation of low-density lipoprotein receptor-related protein 5 receptor-mediated bone formation. *J Am Coll Cardiol*. 2006;47(8):1707-1712. doi:10.1016/j.jacc.2006.02.040

27. Rostagno C. Heart valve disease in elderly. *World J Cardiol*. 2019;11(2):71.

28. Thanassoulis G, Campbell CY, Owens DS, et al. Genetic Associations with Valvular Calcification and Aortic Stenosis. *N Engl J Med*. 2013;368(6):503-512. doi:10.1056/NEJMoa1109034

29. Lincoln J, Garg V. Etiology of Valvular Heart Disease. *Circ J*. 2014;78(8):1801-1807. doi:10.1253/circj.CJ-14-0510

30. Chen JJ, Manning MA, Frazier AA, Jeudy J, White CS. CT Angiography of the Cardiac Valves: Normal, Diseased, and Postoperative Appearances. *RadioGraphics*. 2009;29(5):1393-1412. doi:10.1148/rg.295095002

31. Mocumbi AO. Rheumatic heart disease in Africa: is there a role for genetic studies? *Cardiovasc J Afr*. 2015;26 (2 H3Af(Supplement 1):21-26. doi:10.5830/CVJA-2015-037

32. Bimerew M, Beletew B, Getie A, Wondmieneh A, Gedefaw G, Demis A. Prevalence of rheumatic heart disease among school children in East Africa: a systematic review and meta-analysis. *Pan Afr Med J*. 2021;38(1).

33. Berhanu H, Mekonnen Y, Workicho A, et al. The prevalence of rheumatic heart disease in Ethiopia: a systematic review and meta-analysis. *Trop Dis Travel Med Vaccines*. 2023;9(1):16. doi:10.1186/s40794-023-00192-y

34. Mebrahtom G, Hailay A, Aberhe W, Zereabruk K, Haile T. Rheumatic Heart Disease in East Africa: A Systematic Review and Meta-Analysis. *Int J Rheumatol*. 2023;2023(1):8834443. doi:https://doi.org/10.1155/2023/8834443

35. Zühlke L, Karthikeyan G, Engel ME, et al. Clinical Outcomes in 3343 Children and Adults With Rheumatic Heart Disease From 14 Low-and Middle-Income CountriesClinical Perspective: Two-Year Follow-Up of the Global Rheumatic Heart Disease Registry (the REMEDY Study). *Circulation*. 2016;134(19):1456-1466. doi:10.1161/CIRCULATIONAHA.116.024769

36. Remenyi B, ElGuindy A, Smith Jr. SC, Yacoub M, Holmes Jr. DR. Valvular aspects of rheumatic heart disease. *Lancet*. 2016;387(10025):1335-1346. doi:10.1016/s0140-6736(16)00547-x

37. Zhang W, Mondo C, Okello E, et al. Presenting features of newly diagnosed rheumatic heart disease patients in Mulago Hospital: A pilot study. *Cardiovasc J Afr*. 2013;24(2):28-33. doi:10.5830/CVJA-2012-076

38. Guilherme L, Kalil J. Rheumatic heart disease: molecules involved in valve tissue inflammation leading to the autoimmune process and anti-S. pyogenes vaccine. *Front Immunol*. 2013;4:352. doi:10.3389/fimmu.2013.00352

39. Engel ME, Stander R, Vogel J, Adeyemo AA, Mayosi BM. Genetic Susceptibility to Acute Rheumatic Fever: A Systematic Review and Meta-Analysis of Twin Studies. *PLoS One*. 2011;6(9). doi:10.1371/journal.pone.0025326

40. Carapetis JR, Beaton A, Cunningham MW, et al. Acute rheumatic fever and rheumatic heart disease. *Nat Rev Dis Prim*. 2016;2:15084. doi:10.1038/nrdp.2015.84

41. Peters F, Karthikeyan G, Abrams J, Muhwava L, Zühlke L. Rheumatic heart disease: current status of diagnosis and therapy. *Cardiovasc Diagnosis Ther Vol 10, No 2 (April 2020) Cardiovasc Diagnosis Ther*. Published online 2019.

42. Martins C de O, Demarchi L, Ferreira FM, et al. Rheumatic Heart Disease and Myxomatous Degeneration: Differences and Similarities of Valve Damage Resulting from Autoimmune Reactions and Matrix Disorganization. *PLoS One*. 2017;12(1):e0170191-e0170191. doi:10.1371/journal.pone.0170191

43. Dass C, Kanmanthareddy A. Rheumatic Heart Disease. In: *StatPearls [Internet]*. StatPearls Publishing; 2019.

44. Veinot JP. Pathology of inflammatory native valvular heart disease. *Cardiovasc Pathol*. 2006;15(5):243-251.

45. Stehbens WE, Zuccollo JM. Anitschkow myocytes or cardiac histiocytes in human hearts. *Pathology*. 1999;31(2):98-101.

46. Fraser WJ, Haffejee Z, Cooper K. Rheumatic Aschoff nodules revisited: an immunohistological reappraisal of the cellular component. *Histopathology*. 1995;27(5):457-461.

47. Love GL, Restrepo C. Aschoff bodies of rheumatic carditis are granulomatous lesions of histiocytic origin. *Mod Pathol an Off J United States Can Acad Pathol Inc*. 1988;1(4):256-261.

48. Spina GS, Sampaio RO, Branco CE, Miranda GB, Rosa VEE, Tarasoutchi F. Incidental histological diagnosis of acute rheumatic myocarditis: case report and review of the literature. *Front Pediatr*. 2014;2:126. doi:10.3389/fped.2014.00126

49. Noubiap JJ, Agbor VN, Bigna JJ, Kaze AD, Nyaga UF, Mayosi BM. Prevalence and progression of rheumatic heart disease: a global systematic review and meta-analysis of population-based echocardiographic studies. *Sci Rep*. 2019;9(1):17022. doi:10.1038/s41598-019-53540-4

50. Muhamed B, Mutithu D, Aremu O, Zühlke L, Sliwa K. Rheumatic fever and rheumatic heart disease : Facts and research progress in Africa. 2019;295:48-55. doi:10.1016/j.ijcard.2019.07.079

51. Mrsic Z, Hopkins SP, Antevil JL, Mullenix PS. Valvular Heart Disease. *Prim Care Clin Off Pract*. 2018;45(1):81-94. doi:https://doi.org/10.1016/j.pop.2017.10.002

52. Lindman BR, Clavel M-A, Mathieu P, et al. Calcific aortic stenosis. *Nat Rev Dis Prim*. 2016;2(1):16006. doi:10.1038/nrdp.2016.6

53. Holland J V, Hardie K, de Dassel J, Ralph AP. Rheumatic Heart Disease Prophylaxis in Older Patients: A Register-Based Audit of Adherence to Guidelines. *Open forum Infect Dis*. 2018;5(6):ofy125-ofy125. doi:10.1093/ofid/ofy125

54. Mayosi BM. Protocols for antibiotic use in primary and secondary prevention of rheumatic fever. *S Afr Med J*. 2006;96(3 Pt 2):240.

55. Turer AT. Using metabolomics to assess myocardial metabolism and energetics in heart failure. *J Mol Cell Cardiol*. 2013;55:12-18. doi:10.1016/j.yjmcc.2012.08.025

56. McGarrah RW, Crown SB, Zhang G-F, Shah SH, Newgard CB. Cardiovascular Metabolomics. *Circ Res*. 2018;122(9):1238-1258. doi:10.1161/CIRCRESAHA.117.311002

57. Bingol K. Recent Advances in Targeted and Untargeted Metabolomics by NMR and MS/NMR Methods. *High-throughput*. 2018;7(2):9. doi:10.3390/ht7020009

58. Villas-Bôas SG, Mas S, Akesson M, Smedsgaard J, Nielsen J. Mass spectrometry in metabolome analysis. *Mass Spectrom Rev*. 2005;24(5):613-646. doi:10.1002/mas.20032

59. Kim HK, Choi YH, Verpoorte R. NMR-based plant metabolomics: where do we stand, where do we go? *Trends Biotechnol*. 2011;29(6):267-275. doi:https://doi.org/10.1016/j.tibtech.2011.02.001

60. Mazumder AG, Banerjee S, Zevictovich F, Ghosh S, Mukherjee A, Chatterjee J. Fourier-transform-infrared-spectroscopy based metabolomic spectral biomarker selection towards optimal diagnostic differentiation of diabetes with and without retinopathy. *Spectrosc Lett*. 2018;51(7):340-349. doi:10.1080/00387010.2018.1471510

61. Silva Elipe MV. Advantages and disadvantages of nuclear magnetic resonance

spectroscopy as a hyphenated technique. *Anal Chim Acta*. 2003;497(1):1-25. doi:https://doi.org/10.1016/j.aca.2003.08.048

62. Pan Z, Raftery D. Comparing and combining NMR spectroscopy and mass spectrometry in metabolomics. *Anal Bioanal Chem*. 2007;387(2):525-527. doi:10.1007/s00216-006-0687-8

63. Forcisi S, Moritz F, Kanawati B, Tziotis D, Lehmann R, Schmitt-Kopplin P. Liquid chromatography-mass spectrometry in metabolomics research: mass analyzers in ultra high pressure liquid chromatography coupling. *J Chromatogr A*. 2013;1292:51-65. doi:10.1016/j.chroma.2013.04.017

64. Nagana Gowda GA, Zhang S, Gu H, Asiago V, Shanaiah N, Raftery D. Metabolomics-Based Methods for Early Disease Diagnostics: A Review. *Expert Rev Mol Diagn*. 2008;8(5):617-633. doi:10.1586/14737159.8.5.617

65. Bupp CR, Wirth MJ. Making Sharper Peaks for Reverse-Phase Liquid Chromatography of Proteins. *Annu Rev Anal Chem*. Published online February 2020. doi:10.1146/annurev-anchem-061318-115009

66. Bhole RP, Jagtap SR, Chadar KB, Zambare YB. Liquid Chromatography-Mass Spectrometry Technique-A Review. *Res J Pharm Technol*. 2020;13(1):505-516.

67. Silva JC, Denny R, Dorschel CA, et al. Quantitative proteomic analysis by accurate mass retention time pairs. *Anal Chem*. 2005;77(7):2187-2200. doi:10.1021/ac048455k

68. Xinjie Z, Zhou L, Yin P, Xu G. Liquid Chromatography-Mass Spectrometry of Biofluids and Extracts. In: Bjerrum JT, ed. *Metabonomics: Methods and Protocols*. 1st ed. Humana New York; 2015:61-73. doi:doi.org/10.1007/978-1-4939-2377-9

69. Zhao X, Zhou L, Yin P, Xu G. Liquid chromatography-mass spectrometry of biofluids and extracts. *Methods Mol Biol*. 2015;1277:61-73. doi:10.1007/978-1-4939-2377-9_6

70. Furey A, Moriarty M, Bane V, Kinsella B, Lehane M. Ion suppression; A critical review on causes, evaluation, prevention and applications. *Talanta*. 2013;115:104-122. doi:https://doi.org/10.1016/j.talanta.2013.03.048

71. Clark AE, Kaleta EJ, Arora A, Wolk DM. Matrix-Assisted Laser Desorption Ionization–Time of Flight Mass Spectrometry: a Fundamental Shift in the Routine Practice of Clinical Microbiology. *Clin Microbiol Rev*. 2013;26(3):547 LP - 603. doi:10.1128/CMR.00072-12

72. Rockwood AL, Kushnir MM, Clarke NJ. 2 - Mass Spectrometry. In: Rifai N, Horvath AR, Wittwer CTBT-P and A of CMS, eds. Elsevier; 2018:33-65. doi:https://doi.org/10.1016/B978-0-12-816063-3.00002-5

73. Wang R, Yin Y, Zhu Z. Advancing untargeted metabolomics using data-independent acquisition mass spectrometry technology. Published online 2019.

74. Ren J-L, Zhang A-H, Kong L, Wang X-J. Advances in mass spectrometry-based metabolomics for investigation of metabolites. *RSC Adv*. 2018;8(40):22335-22350.

75. Eliuk S, Makarov A. Evolution of Orbitrap Mass Spectrometry Instrumentation. *Annu Rev Anal Chem*. 2015;8(1):61-80. doi:10.1146/annurev-anchem-071114-040325

76. Denisov E, Damoc E, Lange O, Makarov A. Orbitrap mass spectrometry with resolving powers above 1,000,000. *Int J Mass Spectrom*. 2012;325:80-85.

77. Wang H, Zhao Z, Guo Y. Chemical and biochemical applications of MALDI TOF-MS based on analyzing the small organic compounds. *Top Curr Chem*. 2013;331:165-192. doi:10.1007/128_2012_364

78. Mott TM, Everley RA, Wyatt SA, Toney DM, Croley TR. Comparison of MALDI-TOF/MS and LC-QTOF/MS methods for the identification of enteric bacteria. *Int J Mass Spectrom*. 2010;291(1-2):24-32.

79. Wille K, Kiebooms JAL, Claessens M, et al. Development of analytical strategies using U-HPLC-MS/MS and LC-ToF-MS for the quantification of micropollutants in marine organisms. *Anal Bioanal Chem*. 2011;400(5):1459-1472.

80. Allen DR, McWhinney BC. Quadrupole Time-of-Flight Mass Spectrometry: A Paradigm Shift in Toxicology Screening Applications. *Clin Biochem Rev*. 2019;40(3):135.

81. Mellon FA. MASS SPECTROMETRY | Principles and Instrumentation. In: Caballero BBT-E of FS and N (Second E, ed. Academic Press; 2003:3739-3749. doi:https://doi.org/10.1016/B0-12-227055-X/00746-X

82. Cohen LH, Gusev AI. Small molecule analysis by MALDI mass spectrometry. *Anal Bioanal Chem*. 2002;373(7):571-586. doi:10.1007/s00216-002-1321-z

83. Dreisewerd K. Recent methodological advances in MALDI mass spectrometry. *Anal Bioanal Chem*. 2014;406(9):2261-2278. doi:10.1007/s00216-014-7646-6

84. Menger RF, Stutts WL, Anbukumar DS, Bowden JA, Ford DA, Yost RA. MALDI mass spectrometric imaging of cardiac tissue following myocardial infarction in a rat coronary artery ligation model. *Anal Chem*. 2012;84(2):1117-1125. doi:10.1021/ac202779h

85. Lee JH, Kim YH, Kim K-H, et al. Profiling of Serum Metabolites Using MALDI-TOF and Triple-TOF Mass Spectrometry to Develop a Screen for Ovarian Cancer. *Cancer Res Treat*. Published online 2017. doi:10.4143/crt.2017.275

86. Kriegsmann J, Kriegsmann M, Kriegsmann K, Longuespée R, Deininger S-O, Casadonte R. MALDI Imaging for Proteomic Painting of Heterogeneous Tissue Structures. *PROTEOMICS – Clin Appl*. 2019;13(1):1800045. doi:10.1002/prca.201800045

87. Aichler M, Walch A. MALDI Imaging mass spectrometry: current frontiers and perspectives in pathology research and practice. *Lab Investig*. 2015;95(4):422-431. doi:10.1038/labinvest.2014.156

88. Gibb S, Strimmer K. MALDIquant: a versatile R package for the analysis of mass spectrometry data. *Bioinformatics*. 2012;28(17):2270-2271. doi:10.1093/bioinformatics/bts447

89. Xia J, Wishart DS. Using MetaboAnalyst 3.0 for comprehensive metabolomics data analysis. *Curr Protoc Bioinforma*. Published online 2016:10-14. doi:10.1002/cpbi.11

90. Worley B, Powers R. PCA as a practical indicator of OPLS-DA model reliability. *Curr Metabolomics*. 2016;4(2):97-103. doi:10.2174/2213235X04666160613122429

91. Blasco H, Błaszczyński J, Billaut JC, et al. Comparative analysis of targeted metabolomics: Dominance-based rough set approach versus orthogonal partial least square-discriminant analysis. *J Biomed Inform*. 2015;53:291-299. doi:https://doi.org/10.1016/j.jbi.2014.12.001

92. Boiteau RM, Hoyt DW, Nicora CD, Kinmonth-Schultz HA, Ward JK, Bingol K. Structure

elucidation of unknown metabolites in metabolomics by combined NMR and MS/MS prediction. *Metabolites*. 2018;8(1):8. doi:10.3390/metabo8010008

93. Tautenhahn R, Patti GJ, Rinehart D, Siuzdak G. XCMS Online: A Web-Based Platform to Process Untargeted Metabolomic Data. *Anal Chem*. 2012;84(11):5035-5039. doi:10.1021/ac300698c

94. Benton HP, Wong DM, Trauger SA, Siuzdak G. XCMS2: processing tandem mass spectrometry data for metabolite identification and structural characterization. *Anal Chem*. 2008;80(16):6382-6389. doi:10.1021/ac800795f

95. Chong J, Soufan O, Li C, et al. MetaboAnalyst 4.0: towards more transparent and integrative metabolomics analysis. *Nucleic Acids Res*. 2018;46(W1):W486-W494. doi:10.1093/nar/gky310

96. Tsugawa H, Cajka T, Kind T, et al. MS-DIAL: data-independent MS/MS deconvolution for comprehensive metabolome analysis. *Nat Methods*. 2015;12(6):523-526. doi:10.1038/nmeth.3393

97. Wang W, Maimaiti A, Zhao Y, et al. Analysis of Serum Metabolites to Diagnose Bicuspid Aortic Valve. *Sci Rep*. 2016;6:37023. doi:10.1038/srep37023

98. Martínez-Micaelo N, Ligero C, Antequera-González B, Junza A, Yanes O, Alegret JM. Plasma Metabolomic Profiling Associates Bicuspid Aortic Valve Disease and Ascending Aortic Dilation with a Decrease in Antioxidant Capacity. *J Clin Med*. 2020;9(7). doi:10.3390/jcm9072215

99. Xiong T-Y, Liu C, Liao Y-B, et al. Differences in metabolic profiles between bicuspid and tricuspid aortic stenosis in the setting of transcatheter aortic valve replacement. *BMC Cardiovasc Disord*. 2020;20(1):229. doi:10.1186/s12872-020-01491-4

100. Mourino-alvarez L, Gonzalez-calero L, Martinez-laborde C, et al. Patients with calcific aortic stenosis exhibit systemic molecular evidence of ischemia , enhanced coagulation , oxidative stress and impaired cholesterol transport. *Int J Cardiol*. 2016;225:99-106. doi:10.1016/j.ijcard.2016.09.089

101. Olkowicz M, Debski J, Jablonska P, Dadlez M, Smolenski RT. Application of a new procedure for liquid chromatography / mass spectrometry profiling of plasma amino acid-related metabolites and untargeted shotgun proteomics to identify mechanisms and biomarkers of calcific aortic stenosis. *J Chromatogr A*. 2017;1517:66-78. doi:10.1016/j.chroma.2017.08.024

102. Haase D, Bäz L, Bekfani T, et al. Metabolomic profiling of patients with high gradient aortic stenosis undergoing transcatheter aortic valve replacement. *Clin Res Cardiol*. Published online October 2020. doi:10.1007/s00392-020-01754-2

103. Elmariah S, Farrell LA, Furman D, et al. Association of Acylcarnitines With Left Ventricular Remodeling in Patients With Severe Aortic Stenosis Undergoing Transcatheter Aortic Valve Replacement. *JAMA Cardiol*. 2018;3(3):242-246. doi:10.1001/jamacardio.2017.4873

104. Cunningham MW. Molecular Mimicry, Autoimmunity, and Infection: The Cross-Reactive Antigens of Group A Streptococci and their Sequelae. *Microbiol Spectr*. 2019;7(4). doi:10.1128/microbiolspec.GPP3-0045-2018

105. Das S, Kumar Y, Sharma S, et al. An Untargeted LC–MS based approach for identification of altered metabolites in blood plasma of rheumatic heart disease patients. *Sci Rep*. 2022;12(1):5238. doi:10.1038/s41598-022-09191-z

106. Steffens DC, Jiang W, Krishnan KRR, et al. Metabolomic Differences in Heart Failure Patients With and Without Major Depression. *J Geriatr Psychiatry Neurol*. 2010;23(2):138-146. doi:10.1177/0891988709358592

107. Nicoll R, Henein MY. Arterial calcification: Friend or foe? *Int J Cardiol*. 2013;167(2):322-327. doi:10.1016/j.ijcard.2012.06.110

108. Liu C, Li R, Liu Y, et al. Characteristics of Blood Metabolic Profile in Coronary Heart Disease, Dilated Cardiomyopathy and Valvular Heart Disease Induced Heart Failure . *Front Cardiovasc Med* . 2021;7:423.

109. van Driel BO, Schuldt M, Algül S, et al. Metabolomics in Severe Aortic Stenosis Reveals Intermediates of Nitric Oxide Synthesis as Most Distinctive Markers. *Int J Mol Sci*. 2021;22(7):3569. doi:10.3390/ijms22073569

110. Surendran A, Edel A, Chandran M, et al. Metabolomic Signature of Human Aortic Valve Stenosis. *JACC Basic to Transl Sci*. 2020;5(12):1163-1177. doi:https://doi.org/10.1016/j.jacbts.2020.10.001

111. Al Hageh C, Rahy R, Khazen G, et al. Plasma and urine metabolomic analyses in aortic valve stenosis reveal shared and biofluid-specific changes in metabolite levels. *PLoS One*. 2020;15(11):e0242019.

112. Wang W, Maimaiti A, Zhao Y, et al. Analysis of Serum Metabolites to Diagnose Bicuspid Aortic Valve. *Sci Rep*. 2016;6(1):37023. doi:10.1038/srep37023

113. Haase D, Bäz L, Bekfani T, et al. Metabolomic profiling of patients with high gradient aortic stenosis undergoing transcatheter aortic valve replacement. *Clin Res Cardiol*. 2021;110(3):399-410. doi:10.1007/s00392-020-01754-2

114. Naghavi M, Wang H, Lozano R, et al. Global, regional, and national age-sex specific all-cause and cause-specific mortality for 240 causes of death, 1990-2013: A systematic analysis for the Global Burden of Disease Study 2013. *Lancet*. 2015;385(9963):117-171. doi:10.1016/S0140-6736(14)61682-2

115. Drucker E, Krapfenbauer K. Pitfalls and limitations in translation from biomarker discovery to clinical utility in predictive and personalised medicine. *EPMA J*. 2013;4(1):7. doi:10.1186/1878-5085-4-7

116. Kumar RK, Antunes MJ, Beaton A, et al. Contemporary Diagnosis and Management of Rheumatic Heart Disease: Implications for Closing the Gap: A Scientific Statement From the American Heart Association. *Circulation*. 2020;142(20):e337-e357. doi:10.1161/CIR.0000000000000921

117. Vinaixa M, Samino S, Saez I, Duran J, Guinovart JJ, Yanes O. A Guideline to Univariate Statistical Analysis for LC/MS-Based Untargeted Metabolomics-Derived Data. *Metabolites*. 2012;2(4):775-795. doi:10.3390/metabo2040775

118. Faul F, Erdfelder E, Buchner A, Lang A-G. Statistical power analyses using G*Power 3.1: Tests for correlation and regression analyses. *Behav Res Methods*. 2009;41(4):1149-1160. doi:10.3758/BRM.41.4.1149

119. Dong X, Li X, Chang T-W, Scherzer CR, Weiss ST, Qiu W. powerEQTL: an R package and shiny application for sample size and power calculation of bulk tissue and single-cell eQTL analysis. *Bioinformatics*. 2021;37(22):4269-4271.

120. Acharjee A, Larkman J, Cardoso VR, Gkoutos G V. Effect of biomarker identification on power analysis for diagnostics research. *Prepr (Version 1) Res Sq*. Published online

2020. doi:doi.org/10.21203/rs.3.rs-19527/v1

121. Acharjee A, Larkman J, Cardoso VR, Gkoutos G V. PowerTools: A web based user-friendly tool for future translational study design. *Prepr (Version 1) Res Sq*. Published online 2020. doi:doi.org/10.21203/rs.2.23833/v1

122. Zong W, Seney ML, Ketchesin KD, et al. Experimental design and power calculation in omics circadian rhythmicity detection. *bioRxiv*. Published online 2022:2001-2022. doi:doi.org/10.1101/2022.01.19.476930

123. Faul F, Erdfelder E, Lang A-G, Buchner A. G*Power 3: a flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behav Res Methods*. 2007;39(2):175-191. doi:10.3758/bf03193146

124. Xiao F, Zheng R, Yang D, et al. Sex-dependent aortic valve pathology in patients with rheumatic heart disease. *PLoS One*. 2017;12(6):e0180230. doi:10.1371/journal.pone.0180230

125. French WR, Zimmerman LJ, Schilling B, et al. Wavelet-based peak detection and a new charge inference procedure for MS/MS implemented in ProteoWizard's msConvert. *J Proteome Res*. 2015;14(2):1299-1307.

126. Adusumilli R, Mallick P. Data Conversion with ProteoWizard msConvert. *Methods Mol Biol*. 2017;1550:339-368. doi:10.1007/978-1-4939-6747-6_23

127. Tautenhahn R, Böttcher C, Neumann S. Highly sensitive feature detection for high resolution LC/MS. *BMC Bioinformatics*. 2008;9(1):504. doi:10.1186/1471-2105-9-504

128. Hoffmann N, Keck M, Neuweger H, et al. Combining peak- and chromatogram-based retention time alignment algorithms for multiple chromatography-mass spectrometry datasets. *BMC Bioinformatics*. 2012;13(1):214. doi:10.1186/1471-2105-13-214

129. Chong J, Wishart DS, Xia J. Using MetaboAnalyst 4.0 for Comprehensive and Integrative Metabolomics Data Analysis. *Curr Protoc Bioinforma*. 2019;68(1):e86. doi:10.1002/cpbi.86

130. Wulff JE, Mitchell MW. A comparison of various normalization methods for lc/ms metabolomics data. *Adv Biosci Biotechnol*. 2018;9(08):339.

131. Westerhuis JA, Hoefsloot HCJ, Smit S, et al. Assessment of PLSDA cross validation. *Metabolomics*. 2008;4(1):81-89. doi:10.1007/s11306-007-0099-6

132. Liu X, Zhu X-H, Qiu P, Chen W. A correlation-matrix-based hierarchical clustering method for functional connectivity analysis. *J Neurosci Methods*. 2012;211(1):94-102. doi:https://doi.org/10.1016/j.jneumeth.2012.08.016

133. Nothias L-F, Petras D, Schmid R, et al. Feature-based molecular networking in the GNPS analysis environment. *Nat Methods*. 2020;17(9):905-908. doi:10.1038/s41592-020-0933-6

134. Gil-de-la-Fuente A, Godzien J, Saugar S, et al. CEU Mass Mediator 3.0: A Metabolite Annotation Tool. *J Proteome Res*. 2019;18(2):797-802. doi:10.1021/acs.jproteome.8b00720

135. Sumner LW, Amberg A, Barrett D, et al. Proposed minimum reporting standards for chemical analysis. *Metabolomics*. 2007;3(3):211-221. doi:10.1007/s11306-007-0082-2

136. Blaženović I, Kind T, Ji J, Fiehn O. Software Tools and Approaches for Compound

Identification of LC-MS/MS Data in Metabolomics. *Metab* . 2018;8(2). doi:10.3390/metabo8020031

137. Wang M, Carver JJ, Phelan V V, et al. Sharing and community curation of mass spectrometry data with Global Natural Products Social Molecular Networking. *Nat Biotechnol*. 2016;34(8):828-837. doi:10.1038/nbt.3597

138. Bittremieux W, Schmid R, Huber F, van der Hooft JJJ, Wang M, Dorrestein PC. Comparison of Cosine, Modified Cosine, and Neutral Loss Based Spectrum Alignment For Discovery of Structurally Related Molecules. *J Am Soc Mass Spectrom*. 2022;33(9):1733-1744. doi:10.1021/jasms.2c00153

139. Wang S, Kind T, Tantillo DJ, Fiehn O. Predicting in silico electron ionization mass spectra using quantum chemistry. *J Cheminform*. 2020;12(1):63. doi:10.1186/s13321-020-00470-3

140. Aron AT, Gentry EC, McPhail KL, et al. Reproducible molecular networking of untargeted mass spectrometry data using GNPS. *Nat Protoc*. 2020;15(6):1954-1991. doi:10.1038/s41596-020-0317-5

141. Gil de la Fuente A, Godzien J, Fernández López M, Rupérez FJ, Barbas C, Otero A. Knowledge-based metabolite annotation tool: CEU Mass Mediator. *J Pharm Biomed Anal*. 2018;154:138-149. doi:https://doi.org/10.1016/j.jpba.2018.02.046

142. Schiffman C, Petrick L, Perttula K, et al. Filtering procedures for untargeted LC-MS metabolomics data. *BMC Bioinformatics*. 2019;20(1):334. doi:10.1186/s12859-019-2871-9

143. Xia J, Wishart DS. Web-based inference of biological patterns, functions and pathways from metabolomic data using MetaboAnalyst. *Nat Protoc*. 2011;6(6):743-760. doi:10.1038/nprot.2011.319

144. Xia J, Wishart DS. MSEA: a web-based tool to identify biologically meaningful patterns in quantitative metabolomic data. *Nucleic Acids Res*. 2010;38(suppl_2):W71-W77. doi:10.1093/nar/gkq329

145. Hinshaw SJ, H Y Lee A, Gill EE, E W Hancock R. MetaBridge: enabling network-based integrative analysis via direct protein interactors of metabolites. *Bioinformatics*. 2018;34(18):3225-3227. doi:10.1093/bioinformatics/bty331

146. Xia J, Broadhurst DI, Wilson M, Wishart DS. Translational biomarker discovery in clinical metabolomics: an introductory tutorial. *Metabolomics*. 2013;9(2):280-299. doi:10.1007/s11306-012-0482-9

147. Aa N, Lu Y, Yu M, et al. Plasma Metabolites Alert Patients With Chest Pain to Occurrence of Myocardial Infarction. 2021;8(April). doi:10.3389/fcvm.2021.652746

148. Chen J, Li W, Xiang M. Burden of valvular heart disease, 1990-2017: Results from the Global Burden of Disease Study 2017. *J Glob Health*. 2020;10(2):020404. doi:10.7189/jogh.10.020404

149. Lahiri S, Sanyahumbi A. Acute Rheumatic Fever. *Pediatr Rev*. 2021;42(5):221-232. doi:10.1542/pir.2019-0288

150. Olivier C. Rheumatic fever—is it still a problem? *J Antimicrob Chemother*. 2000;45(suppl_1):13-21. doi:10.1093/jac/45.suppl_1.13

151. Culling CFA, Allison RT, Barr WT. *Cellular Pathology Technique*. 4th ed. Elsevier

Science; 2014.

152. Taqi SA, Sami SA, Sami LB, Zaki SA. A review of artifacts in histopathology. *J Oral Maxillofac Pathol*. 2018;22(2):279. doi:10.4103/jomfp.JOMFP_125_15

153. Bindhu P, Krishnapillai R, Thomas P, Jayanthi P. Facts in artifacts. *J Oral Maxillofac Pathol*. 2013;17(3):397-401. doi:10.4103/0973-029X.125206

154. Mutithu DW, Roberts R, Manganyi R, Ntusi NAB. Chronic rheumatic heart disease with recrudescence of acute rheumatic fever on histology: a case report. *Eur Hear J - Case Reports*. 2022;6(7):ytac278. doi:10.1093/ehjcr/ytac278

155. Gjertsson P, Caidahl K, Odén A, Bech-Hanssen O. Diagnostic and referral delay in patients with aortic stenosis is common and negatively affects outcome. *Scand Cardiovasc J*. 2007;41(1):12-18. doi:10.1080/14017430601115935

156. Rosenhek R, Maurer G, Baumgartner H. Should early elective surgery be performed in patients with severe but asymptomatic aortic stenosis? *Eur Heart J*. 2002;23(18):1417-1421. doi:10.1053/euhj.2002.3163

157. Antunes MJ. The Global Burden of Rheumatic Heart Disease: Population-Related Differences (It is Not All the Same!). *Brazilian J Cardiovasc Surg*. 2020;35(6):958-963. doi:10.21470/1678-9741-2020-0514

158. Negi PC, Kandoria A, Asotra S, et al. Gender differences in the epidemiology of Rheumatic Fever/Rheumatic heart disease (RF/RHD) patient population of hill state of northern India; 9 years prospective hospital based, HP-RHD registry. *Indian Heart J*. 2020;72(6):552-556. doi:https://doi.org/10.1016/j.ihj.2020.09.011

159. Carabello BA, Paulus WJ. Aortic stenosis. *Lancet*. 2009;373(9667):956-966. doi:https://doi.org/10.1016/S0140-6736(09)60211-7

160. Fukui M, Gupta A, Abdelkarim I, et al. Association of Structural and Functional Cardiac Changes With Transcatheter Aortic Valve Replacement Outcomes in Patients With Aortic Stenosis. *JAMA Cardiol*. 2019;4(3):215-222. doi:10.1001/jamacardio.2018.4830

161. Marijon E, Mirabel M, Celermajer DS, Jouven X. Rheumatic heart disease. *Lancet*. 2012;379(9819):953-964. doi:10.1016/s0140-6736(11)61171-9

162. Mutagaywa RK, Mwakigonja A, Chillo P, et al. Histopathological evaluation of chronic rheumatic mitral valve stenosis: the association with clinical presentation, pathogenesis, and management at a National Cardiac Institute, Tanzania. *Cardiovasc Pathol*. 2022;60:107434. doi:https://doi.org/10.1016/j.carpath.2022.107434

163. Wallby L, Steffensen T, Jonasson L, Broqvist M. Inflammatory Characteristics of Stenotic Aortic Valves: A Comparison between Rheumatic and Nonrheumatic Aortic Stenosis. Cicoira M, ed. *Cardiol Res Pract*. 2013;2013:895215. doi:10.1155/2013/895215

164. Chandrasoma P, Taylor CR. Chapter 5. Chronic Inflammation. In: *Concise Pathology, 3e*. The McGraw-Hill Companies; 1998.

165. Sika-Paotonu D, Beaton A, Raghu A, Steer A, Carapetis J. Acute rheumatic fever and rheumatic heart disease. *Streptococcus pyogenes Basic Biol to Clin Manifestations [Internet]*. Published online 2017.

166. Dunn WB, Broadhurst DI, Deepak SM, et al. Serum metabolomics reveals many novel metabolic markers of heart failure, including pseudouridine and 2-oxoglutarate.

Metabolomics. 2007;3(4):413-426. doi:10.1007/s11306-007-0063-5

167. Li Q, Laflamme DP, Bauer JE. Serum untargeted metabolomic changes in response to diet intervention in dogs with preclinical myxomatous mitral valve disease. *PLoS One*. 2020;15(6):e0234404. doi:10.1371/journal.pone.0234404

168. Li Q, Freeman LM, Rush JE, et al. Veterinary Medicine and Multi-Omics Research for Future Nutrition Targets: Metabolomics and Transcriptomics of the Common Degenerative Mitral Valve Disease in Dogs. *OMICS*. 2015;19(8):461-470. doi:10.1089/omi.2015.0057

169. Schwaiger M, Schoeny H, El Abiead Y, Hermann G, Rampler E, Koellensperger G. Merging metabolomics and lipidomics into one analytical run. *Analyst*. 2019;144(1):220-229.

170. Adusumilli R, Mallick P. Data Conversion with ProteoWizard msConvert BT - Proteomics: Methods and Protocols. In: Comai L, Katz JE, Mallick P, eds. Springer New York; 2017:339-368. doi:10.1007/978-1-4939-6747-6_23

171. Pang Z, Chong J, Zhou G, et al. MetaboAnalyst 5.0: narrowing the gap between raw spectra and functional insights. *Nucleic Acids Res*. 2021;49(W1):W388-W396. doi:10.1093/nar/gkab382

172. Dührkop K, Fleischauer M, Ludwig M, et al. SIRIUS 4: a rapid tool for turning tandem mass spectra into metabolite structure information. *Nat Methods*. 2019;16(4):299-302. doi:10.1038/s41592-019-0344-8

173. Salek RM, Steinbeck C, Viant MR, Goodacre R, Dunn WB. The role of reporting standards for metabolite annotation and identification in metabolomic studies. *Gigascience*. 2013;2(1):2013-2047. doi:10.1186/2047-217X-2-13

174. Lepoittevin M, Blancart-Remaury Q, Kerforne T, Pellerin L, Hauet T, Thuillier R. Comparison between 5 extractions methods in either plasma or serum to determine the optimal extraction and matrix combination for human metabolomics. *Cell Mol Biol Lett*. 2023;28(1):43. doi:10.1186/s11658-023-00452-x

175. Limian Z, Megan J. *Protein Precipitation for Biological Fluid Samples Using Agilent Captiva EMR — Lipid 96-Well Plates*. Agilent technologies; 2018.

176. Barberini L, Noto A, Saba L, et al. Multivariate data validation for investigating primary HCMV infection in pregnancy. *Data Br*. 2016;9:220-230. doi:https://doi.org/10.1016/j.dib.2016.08.050

177. Bijlsma S, Bobeldijk I, Verheij ER, et al. Large-Scale Human Metabolomics Studies: A Strategy for Data (Pre-) Processing and Validation. *Anal Chem*. 2006;78(2):567-574. doi:10.1021/ac051495j

178. Nitz K, Lacy M, Atzler D. Amino Acids and Their Metabolism in Atherosclerosis. *Arterioscler Thromb Vasc Biol*. 2019;39(3):319-330. doi:10.1161/ATVBAHA.118.311572

179. Neumann J, Hofmann B, Dhein S, Gergs U. Cardiac Roles of Serotonin (5-HT) and 5-HT-Receptors in Health and Disease. *Int J Mol Sci*. 2023;24(5). doi:10.3390/ijms24054765

180. Zilla P, Bolman RM, Yacoub MH, et al. The Cape Town Declaration on Access to Cardiac Surgery in the Developing World. *Cardiovasc J Afr*. 2018;29(4):256-259. doi:10.5830/CVJA-2018-046

181. Das UN. Essential Fatty Acids and Their Metabolites in the Pathobiology of Inflammation and Its Resolution. *Biomolecules*. 2021;11(12). doi:10.3390/biom11121873

182. Guilherme L, Kalil J, Cunningham M. Molecular mimicry in the autoimmune pathogenesis of rheumatic heart disease. *Autoimmunity*. 2006;39(1):31-39. doi:10.1080/08916930500484674

183. Chopra P, Gulwani H. Pathology and pathogenesis of rheumatic heart disease. *Indian J Pathol Microbiol*. 2007;50(4):685-697.

184. Habeeb NM, Al Hadidi IS. Ongoing inflammation in children with rheumatic heart disease. *Cardiol Young*. 2011;21(3):334-339. doi:10.1017/s1047951111000047

185. Huang Y, Zhou M, Sun H, Wang Y. Branched-chain amino acid metabolism in heart disease: an epiphenomenon or a real culprit? *Cardiovasc Res*. 2011;90(2):220-223. doi:10.1093/cvr/cvr070

186. Sun H, Olson KC, Gao C, et al. Catabolic Defect of Branched-Chain Amino Acids Promotes Heart Failure. *Circulation*. 2016;133(21):2038-2049. doi:10.1161/CIRCULATIONAHA.115.020226

187. Patel DA, Lavie CJ, Milani R V, Shah S, Gilliland Y. Clinical implications of left atrial enlargement: a review. *Ochsner J*. 2009;9(4):191-196.

188. Guan C, Niu Y, Chen S-C, et al. Structural insights into the inhibition mechanism of human sterol O-acyltransferase 1 by a competitive inhibitor. *Nat Commun*. 2020;11(1):2478. doi:10.1038/s41467-020-16288-4

189. Ikonen E, Zhou X. Cholesterol transport between cellular membranes: A balancing act between interconnected lipid fluxes. *Dev Cell*. 2021;56(10):1430-1436. doi:https://doi.org/10.1016/j.devcel.2021.04.025

190. Hofmann AF, Hagey LR. Bile Acids: Chemistry, Pathochemistry, Biology, Pathobiology, and Therapeutics. *Cell Mol Life Sci*. 2008;65(16):2461-2483. doi:10.1007/s00018-008-7568-6

191. Park T, Oh J, Lee K. Dietary taurine or glycine supplementation reduces plasma and liver cholesterol and triglyceride concentrations in rats fed a cholesterol-free diet. *Nutr Res*. 1999;19(12):1777-1789. doi:https://doi.org/10.1016/S0271-5317(99)00118-9

192. Ploier B, Korber M, Schmidt C, Koch B, Leitner E, Daum G. Regulatory link between steryl ester formation and hydrolysis in the yeast Saccharomyces cerevisiae. *Biochim Biophys Acta - Mol Cell Biol Lipids*. 2015;1851(7):977-986. doi:https://doi.org/10.1016/j.bbalip.2015.02.011

193. Ramaraj R, Sorrell VL. Degenerative aortic stenosis. *BMJ*. 2008;336(7643):550 LP - 555. doi:10.1136/bmj.39478.498819.AD

194. Kamstrup PR, Hung M-Y, Witztum JL, Tsimikas S, Nordestgaard BG. Oxidized Phospholipids and Risk of Calcific Aortic Valve Disease: The Copenhagen General Population Study. *Arterioscler Thromb Vasc Biol*. 2017;37(8):1570-1578. doi:10.1161/ATVBAHA.116.308761

195. Capoulade R, Chan KL, Yeang C, et al. Oxidized Phospholipids, Lipoprotein(a), and Progression of Calcific Aortic Valve Stenosis. *J Am Coll Cardiol*. 2015;66(11):1236-1246. doi:10.1016/j.jacc.2015.07.020

196. Bouchareb R, Mahmut A, Nsaibia MJ, et al. Autotaxin Derived From Lipoprotein(a) and

Valve Interstitial Cells Promotes Inflammation and Mineralization of the Aortic Valve. *Circulation*. 2015;132(8):677-690. doi:10.1161/CIRCULATIONAHA.115.016757

197. Torzewski M, Ravandi A, Yeang C, et al. Lipoprotein(a)-Associated Molecules Are Prominent Components in Plasma and Valve Leaflets in Calcific Aortic Valve Stenosis. *JACC Basic to Transl Sci*. 2017;2(3):229-240. doi:https://doi.org/10.1016/j.jacbts.2017.02.004

198. Wu J, Lu J, Huang J, et al. Variations in Energy Metabolism Precede Alterations in Cardiac Structure and Function in Hypertrophic Preconditioning. *Front Cardiovasc Med*. 2020;7:602100. doi:10.3389/fcvm.2020.602100

199. Kundu BK, Zhong M, Sen S, Davogustto G, Keller SR, Taegtmeyer H. Remodeling of glucose metabolism precedes pressure overload-induced left ventricular hypertrophy: review of a hypothesis. *Cardiology*. 2015;130(4):211-220.

200. Li J, Kemp BA, Howell NL, et al. Metabolic Changes in Spontaneously Hypertensive Rat Hearts Precede Cardiac Dysfunction and Left Ventricular Hypertrophy. *J Am Heart Assoc*. 2019;8(4):e010926. doi:10.1161/JAHA.118.010926

201. Jameel MN, Zhang J. Myocardial energetics in left ventricular hypertrophy. *Curr Cardiol Rev*. 2009;5(3):243-250. doi:10.2174/157340309788970379

202. Rajamannan NM, Nealis TB, Subramaniam M, et al. Calcified Rheumatic Valve Neoangiogenesis Is Associated With Vascular Endothelial Growth Factor Expression and Osteoblast-Like Bone Formation. *Circulation*. 2005;111(24):3296-3301. doi:10.1161/CIRCULATIONAHA.104.473165

203. Rajamannan NM, Antonini-Canterin F, Moura L, et al. Medical therapy for rheumatic heart disease: is it time to be proactive rather than reactive? *Indian Heart J*. 2009;61(1):14-23.

204. Kontush A, Lhomme M, Chapman MJ. Unraveling the complexities of the HDL lipidome. *J Lipid Res*. 2013;54(11):2950-2963. doi:10.1194/jlr.R036095

205. Chen H, Wang Z, Qin M, et al. Comprehensive Metabolomics Identified the Prominent Role of Glycerophospholipid Metabolism in Coronary Artery Disease Progression . *Front Mol Biosci* . 2021;8.

206. C. NP, David G, Donna R, K. GC, A. DE. Phospholipase A2 regulates eicosanoid class switching during inflammasome activation. *Proc Natl Acad Sci*. 2014;111(35):12746-12751. doi:10.1073/pnas.1404372111

207. Vorkas PA, Isaac G, Holmgren A, et al. Perturbations in fatty acid metabolism and apoptosis are manifested in calcific coronary artery disease: An exploratory lipidomic study. *Int J Cardiol*. 2015;197:192-199. doi:https://doi.org/10.1016/j.ijcard.2015.06.048

208. Tintut Y, Hsu JJ, Demer LL. Lipoproteins in Cardiovascular Calcification: Potential Targets and Challenges. *Front Cardiovasc Med*. 2018;5:172. doi:10.3389/fcvm.2018.00172

209. Akram M. Citric Acid Cycle and Role of its Intermediates in Metabolism. *Cell Biochem Biophys*. 2014;68(3):475-478. doi:10.1007/s12013-013-9750-1

210. Yuyun MF, Sliwa K, Kengne AP, Mocumbi AO, Bukhman G. Cardiovascular Diseases in Sub-Saharan Africa Compared to High-Income Countries: An Epidemiological Perspective. *Glob Heart*. 2020;15(1):15. doi:10.5334/gh.403

211. Wang W, Maimaiti A, Zhao Y, Zhang L, Tao H, Nian H. Analysis of Serum Metabolites to Diagnose Bicuspid Aortic Valve. *Nat Publ Gr*. 2016;(November):1-8. doi:10.1038/srep37023

212. Li S-S, Liu Y, Li H, et al. Identification of psoriasis vulgaris biomarkers in human plasma by non-targeted metabolomics based on UPLC-Q-TOF/MS. *Eur Rev Med Pharmacol Sci*. 2019;23(9):3940-3950. doi:10.26355/eurrev_201905_17823

213. Chong Nguyen C, Duboc D, Rainteau D, et al. Circulating bile acids concentration is predictive of coronary artery disease in human. *Sci Rep*. 2021;11(1):22661. doi:10.1038/s41598-021-02144-y

214. Charach G, Argov O, Geiger K, Charach L, Rogowski O, Grosskopf I. Diminished bile acids excretion is a risk factor for coronary artery disease: 20-year follow up and long-term outcome. *Therap Adv Gastroenterol*. 2017;11:1756283X17743420-1756283X17743420. doi:10.1177/1756283X17743420

215. Essop MR, Nkomo VT. Rheumatic and Nonrheumatic Valvular Heart Disease. *Circulation*. 2005;112(23):3584-3591. doi:10.1161/CIRCULATIONAHA.105.539775

216. Zuehlke L, Engel ME, Karthikeyan G, et al. Characteristics, complications, and gaps in evidence-based interventions in rheumatic heart disease: the Global Rheumatic Heart Disease Registry (the REMEDY study). *Eur Heart J*. 2015;36(18):1115-U29. doi:10.1093/eurheartj/ehu449

217. Gaasch WH, Meyer TE. Left Ventricular Response to Mitral Regurgitation. *Circulation*. 2008;118(22):2298-2303. doi:10.1161/CIRCULATIONAHA.107.755942

218. Rusinaru D, Bohbot Y, Kowalski C, Ringle A, Maréchaux S, Tribouilloy C. Left Atrial Volume and Mortality in Patients With Aortic Stenosis. *J Am Heart Assoc*. 2022;6(11):e006615. doi:10.1161/JAHA.117.006615

219. Balestrino M. Role of Creatine in the Heart: Health and Disease. *Nutrients*. 2021;13(4):1215. doi:10.3390/nu13041215

220. Guo W, Wang Z, Jue H, Dong C, Yang C. Cardioprotective Effect of (Z)-2-Acetoxy-3-(3,4-Dihydroxyphenyl) Acrylic Acid: Inhibition of Apoptosis in Cardiomyocytes. *Cardiovasc Ther*. 2020;2020.

221. Iravani S, Varma RS. Advanced Drug Delivery Micro- and Nanosystems for Cardiovascular Diseases. *Molecules*. 2022;27(18):5843. doi:10.3390/molecules27185843

222. Li Z, Vance DE. Thematic Review Series: Glycerolipids. Phosphatidylcholine and choline homeostasis. *J Lipid Res*. 2008;49(6):1187-1194. doi:https://doi.org/10.1194/jlr.R700019-JLR200

223. Wilson TWH, Fredrik B, Ulf L, L. HS. Intestinal Microbiota in Cardiovascular Health and Disease. *J Am Coll Cardiol*. 2019;73(16):2089-2105. doi:10.1016/j.jacc.2019.03.024

224. Amrein M, Li XS, Walter J, et al. Gut microbiota-dependent metabolite trimethylamine N-oxide (TMAO) and cardiovascular risk in patients with suspected functionally relevant coronary artery disease (fCAD). *Clin Res Cardiol*. 2022;111(6):692-704. doi:10.1007/s00392-022-01992-6

225. Zilla P, Bolman III RM, Boateng P, Sliwa K. A glimpse of hope: cardiac surgery in low- and middle-income countries (LMICs). *Cardiovasc Diagn Ther*. 2019;10(2):336–349. doi:10.21037/cdt.2019.11.03

226. Raphael DM, Roos L, Myovela V, et al. Heart diseases and echocardiography in rural Tanzania: Occurrence, characteristics, and etiologies of underappreciated cardiac pathologies. *PLoS One*. 2018;13(12):e0208931. doi:10.1371/journal.pone.0208931

227. Shah NJ, Sureshkumar S, Shewade DG. Metabolomics: A Tool Ahead for Understanding Molecular Mechanisms of Drugs and Diseases. *Indian J Clin Biochem*. 2015;30(3):247-254. doi:10.1007/s12291-014-0455-z

228. Chughtai K, Heeren RMA. Mass Spectrometric Imaging for Biomedical Tissue Analysis. *Chem Rev*. 2010;110(5):3237-3277. doi:10.1021/cr100012c

229. Ščupáková K, Balluff B, Tressler C, et al. Cellular resolution in clinical MALDI mass spectrometry imaging: the latest advancements and current challenges. *Clin Chem Lab Med*. 2020;58(6):914-929. doi:doi:10.1515/cclm-2019-0858

230. Schulz S, Becker M, Groseclose MR, Schadt S, Hopf C. Advanced MALDI mass spectrometry imaging in pharmaceutical research and drug development. *Curr Opin Biotechnol*. 2019;55:51-59. doi:https://doi.org/10.1016/j.copbio.2018.08.003

231. Blaser MC, Kraler S, Lüscher TF, Aikawa E. Multi-Omics Approaches to Define Calcific Aortic Valve Disease Pathogenesis. *Circ Res*. 2021;128(9):1371-1397. doi:10.1161/CIRCRESAHA.120.317979

232. Ulmer CZ, Maus A, Hines J, Singh R. Challenges in Translating Clinical Metabolomics Data Sets from the Bench to the Bedside. *Clin Chem*. 2021;67(12):1581-1583. doi:10.1093/clinchem/hvab210

233. Cheng S, Shah SH, Corwin EJ, et al. Potential impact and study considerations of metabolomics in cardiovascular health and disease: a scientific statement from the American Heart Association. *Circ Genomic Precis Med*. 2017;10(2):e000032. doi:10.1161/HCG.0000000000000032

234. Di Minno A, Gelzo M, Caterino M, Costanzo M, Ruoppolo M, Castaldo G. Challenges in Metabolomics-Based Tests, Biomarkers Revealed by Metabolomic Analysis, and the Promise of the Application of Metabolomics in Precision Medicine. *Int J Mol Sci*. 2022;23(9). doi:10.3390/ijms23095213

235. Bengel P, Elkenani M, Beuthner BE, et al. Metabolomic Profiling in Patients with Different Hemodynamic Subtypes of Severe Aortic Valve Stenosis. *Biomolecules*. 2023;13(1). doi:10.3390/biom13010095

236. Mezger STP, Mingels AMA, Bekers O, Cillero-Pastor B, Heeren RMA. Trends in mass spectrometry imaging for cardiovascular diseases. *Anal Bioanal Chem*. 2019;411(17):3709-3720. doi:10.1007/s00216-019-01780-8

237. Pang Z, Zhou G, Ewald J, et al. Using MetaboAnalyst 5.0 for LC–HRMS spectra processing, multi-omics integration and covariate adjustment of global metabolomics data. *Nat Protoc*. 2022;17(8):1735-1761. doi:10.1038/s41596-022-00710-w

238. Meyer KA, Shea JW. Dietary Choline and Betaine and Risk of CVD: A Systematic Review and Meta-Analysis of Prospective Studies. *Nutrients*. 2017;9(7). doi:10.3390/nu9070711

239. Li J, Zeng Q, Xiong Z, et al. Trimethylamine N-oxide induces osteogenic responses in human aortic valve interstitial cells in vitro and aggravates aortic valve lesions in mice. *Cardiovasc Res*. 2022;118(8):2018-2030. doi:10.1093/cvr/cvab243

240. Bartolomaeus H, Balogh A, Yakoub M, et al. Short-Chain Fatty Acid Propionate Protects

From Hypertensive Cardiovascular Damage. *Circulation*. 2019;139(11):1407-1421. doi:10.1161/CIRCULATIONAHA.118.036652

241. Zhang C, Zhou S, Chang H, et al. Metabolomic Profiling Identified Serum Metabolite Biomarkers and Related Metabolic Pathways of Colorectal Cancer. Belardinilli F, ed. *Dis Markers*. 2021;2021:6858809. doi:10.1155/2021/6858809

242. Kankuri E, Finckenberg P, Leinonen J, et al. Altered acylcarnitine metabolism and inflexible mitochondrial fuel utilization characterize the loss of neonatal myocardial regeneration capacity. *Exp Mol Med*. 2023;55(4):806-817. doi:10.1038/s12276-023-00967-5

243. Karwi QG, Uddin GM, Ho KL, Lopaschuk GD. Loss of Metabolic Flexibility in the Failing Heart. *Front Cardiovasc Med*. 2018;5. doi:https://doi.org/10.3389/fcvm.2018.00068

244. Rwebembera J, Marangou J, Mwita JC, et al. 2023 World Heart Federation guidelines for the echocardiographic diagnosis of rheumatic heart disease. *Nat Rev Cardiol*. Published online 2023. doi:10.1038/s41569-023-00940-9

245. K. TV, Vijeta B, H. PM, Katrina A, H. WJ. Echocardiographic Surveillance of Valvular Heart Disease in Different Sociodemographic Groups. *JACC Cardiovasc Imaging*. 2019;12(4):751-752. doi:10.1016/j.jcmg.2018.05.025

246. Pandian NG, Kim JK, Arias-Godinez JA, et al. Recommendations for the Use of Echocardiography in the Evaluation of Rheumatic Heart Disease: A Report from the American Society of Echocardiography. *J Am Soc Echocardiogr*. 2023;36(1):3-28. doi:10.1016/j.echo.2022.10.009

247. Otto CM, Nishimura RA, Bonow RO, et al. 2020 ACC/AHA Guideline for the Management of Patients With Valvular Heart Disease: Executive Summary: A Report of the American College of Cardiology/American Heart Association Joint Committee on Clinical Practice Guidelines. *Circulation*. 2021;143(5):e35-e71. doi:10.1161/CIR.0000000000000932

248. Qiu S, Cai Y, Yao H, et al. Small molecule metabolites: discovery of biomarkers and therapeutic targets. *Signal Transduct Target Ther*. 2023;8(1):132. doi:10.1038/s41392-023-01399-3

249. Veerasammy K, Chen YX, Sauma S, et al. Sample preparation for metabolic profiling using MALDI mass spectrometry imaging. *JoVE (Journal Vis Exp*. 2020;(166):e62008. doi:10.3791/62008-v

250. Andersen MK, Høiem TS, Claes BSR, et al. Spatial differentiation of metabolism in prostate cancer tissue by MALDI-TOF MSI. *Cancer Metab*. 2021;9(1):9. doi:10.1186/s40170-021-00242-z

251. Zhu X, Xu T, Peng C, Wu S. Advances in MALDI mass spectrometry imaging single cell and tissues. *Front Chem*. 2022;9:782432. doi:10.3389/fchem.2021.782432

252. Ma B, Zhang Y, Ma J, Chen X, Sun C, Qin C. Spatially resolved visualization of reprogrammed metabolism in hepatocellular carcinoma by mass spectrometry imaging. *Cancer Cell Int*. 2023;23(1):177. doi:10.1186/s12935-023-03027-0

253. Reina-Campos M, Moscat J, Diaz-Meco M. Metabolism shapes the tumor microenvironment. *Curr Opin Cell Biol*. 2017;48:47-53. doi:10.1016/j.ceb.2017.05.006

254. Mendelson K, Schoen FJ. Heart valve tissue engineering: concepts, approaches, progress, and challenges. *Ann Biomed Eng*. 2006;34(12):1799-1819.

doi:10.1007/s10439-006-9163-z

255. Saini N, Saikia UN, Sahni D, Singh RS. Vascularity of human atrioventricular valves: A myth or fact? *J Thorac Cardiovasc Surg*. 2014;147(1):517-521. doi:https://doi.org/10.1016/j.jtcvs.2013.03.031

Appendix

Appendix 1. Ethics and protocol amendment approval

UNIVERSITY OF CAPE TOWN
Faculty of Health Sciences
Human Research Ethics Committee

Room E53-46 Old Main Building
Groote Schuur Hospital
Observatory 7925
Telephone [021] 406 6626
Email: shuretta.thomas@uct.ac.za
Website: www.health.uct.ac.za/fhs/research/humanethics/forms

26 September 2018

HREC REF: 574/2018

Prof Ntobeko Ntusi
Medicine
J-Floor, OMB

Dear Prof Ntusi

PROJECT TITLE: UNDERSTANDING CHRONOLOGY OF INFALAMMATORY HEART DISEASE AND STRESS INDUCED PROLIFERATIVE MECHANISMS USING METABOLIC BIOMARKERS (PHD Candidate - Mr D Mutithu) SUB-STUDY LINKED TO 554/2017

Thank you for submitting your response to the Faculty of Health Sciences Human Research Ethics Committee.

It is a pleasure to inform you that the HREC has **formally approved** the above-mentioned study.

Approval is granted for one year until the 30 September 2019.

Please submit a progress form, using the standardised Annual Report Form if the study continues beyond the approval period. Please submit a Standard Closure form if the study is completed within the approval period.
(Forms can be found on our website: www.health.uct.ac.za/fhs/research/humanethics/forms)

Please quote the HREC REF in all your correspondence.

Please note that the ongoing ethical conduct of the study remains the responsibility of the principal investigator.

Please note that for all studies approved by the HREC, the principal investigator **must** obtain appropriate institutional approval, where necessary, before the research may occur.

The HREC acknowledge that the student, Daniel Mutithu will also be involved in this study.

Yours sincerely

PROFESSOR M BLOCKMAN
CHAIRPERSON, FHS HUMAN RESEARCH ETHICS COMMITTEE
Federal Wide Assurance Number: FWA00001637.
Institutional Review Board (IRB) number: IRB00001938

This serves to confirm that the University of Cape Town Human Research Ethics Committee complies to the Ethics Standards for Clinical Research with a new drug in patients, based on the Medical Research Council (MRC-SA), Food and Drug Administration (FDA-USA), International Convention on Harmonisation Good Clinical Practice (ICH GCP), South African Good Clinical Practice Guidelines (DoH 2006), based on the Association of the British Pharmaceutical Industry Guidelines (ABPI), and Declaration of Helsinki (2013) guidelines.
The Human Research Ethics Committee granting this approval is in compliance with the ICH Harmonised Tripartite Guidelines E6: Note for Guidance on Good Clinical Practice (CPMP/ICH/135/95) and FDA Code Federal Regulation Part 50, 56 and 312.

UNIVERSITY OF CAPE TOWN
IYUNIVESITHI YASEKAPA • UNIVERSITEIT VAN KAAPSTAD

FACULTY OF HEALTH SCIENCES
Human Research Ethics Committee

Form FHS006: Protocol Amendment

HREC office use only (FWA00001637; IRB00001938)			
☒ Approved	☒ Type of review: Expedited	☐ Full committee	
This serves as notification that all changes and documentation described below are approved.			
Signature HREC Chairperson / Designee	AJM	Date	1/12/22

Note: All **Major** amendments must include a **Cover Letter** and a local **PI Synopsis** justifying the changes for the amendment. Please note that incomplete amendment submissions will not be reviewed.

Please email this form and supporting documents (if applicable) in a combined pdf-file to hrec-enquiries@uct.ac.za with subject line: FHS006 + (HREC Reference number).

The latest forms are found on our website.

http://www.health.uct.ac.za/fhs/research/humanethics/forms

Please also clarify your plan for research-related activities during COVID-19 lockdown.

Comments from the HREC to the Principal Investigator:
Note: The approval of this protocol amendment does not grant annual approval. Please complete the FHS016 / FHS017 form for annual approval at least one month before study expiration.

Principal Investigator to complete the following:

1. Protocol information

Date (when submitting this form)	11/8/2022		
HREC REF Number	574/2018		
Protocol Title	Understanding Chronology of Inflammatory Heart Disease and Stress Induced Proliferative Mechanisms using Metabolic Biomarkers		
Protocol Number (if applicable)	Version 1.1		
Principal Investigator	Prof Ntobeko Ntusi		
Department / Office Internal Mail Address	Medicine, J floor Room 52.23, Old Main Building		
1.1 Is this a major or a minor amendment? (see FHS006hlp) Major (tick box) Minor (tick box)		☐ Major	✓ ☐ Minor
1.2 Does this protocol receive US Federal funding?		☐ Yes	✓ ☐ No

1.3 If the amendment is a major amendment and receives US Federal Funding, does the amendment require full committee approval? **Note:** Any protocol amendments for **Full Committee Review** MUST be submitted on the monthly HREC submission dates. (Please email an electronic copy to hrec-enquiries@uct.ac.za)	☐ Yes	✓ ☐ No
1.4 Did the initial study require UCT No-Fault Insurance	☐ Yes	✓ ☐ No

2. List of Proposed Amendments with Revised Version Numbers and Dates

Please itemise on the page below, all amendments with revised version numbers and dates, which need approval.
This page will be detached, signed and returned to the PI as notification of approval. Please add extra pages if necessary.

1. We have amended the original protocol (version 1.0). In the amended protocol (version 1.1, date 14/7/2022) we have added a pilot study for assessing the suitability of using dried plasma spot as a technique for sample handling for metabolomics in resource limited regions. At the abstract of the amended protocol (version 1.1, date 14/7/2022) we have added **lines 26-28** stating **"Further, dried plasma spots will be investigated as an alternative method of sample collection and storage regarding their future use for research and diagnostics in resource limited environments."**
2. In the amended protocol we have omitted a section that explained into details the mass spectrometry techniques as used in metabolomics. **Lines 206-226 and 241-144** as was in version 1.0 have been deleted in amended protocol (version 1.1, date 14/7/2022).
3. In the amended protocol (version 1.1, date 14/7/2022) under the introduction section, we have added some background information on the use of dried blood spots and dried plasma spots in metabolomics studies **(lines 256-304)**
4. In the amended protocol (version 1.1, date 14/7/2022) at the rationale section, we have added information detailing the rationale of using dried plasma spot as an alternative method for sample handing for metabolomics studies especially in sub-Saharan Africa **(lines 341-350)**.
5. In the amended protocol (version 1.1, date 14/7/2022) we have edited the aims and objectives and re-organized them sequentially as per the proposed experiments **(lines 306-349)**.
6. In the amended protocol (version 1.1, date 14/7/2022) we have added a new aim and objectives regarding the pilot experiment on assessing the suitability of using dried plasma spot as a method of sample handling in metabolomics studies **(lines 341-349)**.
7. At the study design section of the previous protocol version 1.0 we had stated we will compare metabolic dysregulation patterns in tissues. In the amended protocol (version 1.1, date 14/7/2022) we have stated specifically that we will compare metabolic patterns in tissues obtained from RHD, degenerative aortic valve stenosis patients against cadaveric tissue donors as controls **(lines 353-356) "Tissue specific biomarkers analysis will be conducted using valve biopsies from RHD and degenerative AS patients obtained during valve replacement surgery and cadaveric tissue donors obtained during scheduled autopsies."**
8. In the amended protocol (version 1.1, date 14/7/2022) under the study design section we have added how the pilot experiment samples will be obtained and the groups to be compared **(lines 362-364) "To investigate the suitability of DPS as a method for sample collection, peripheral blood will be collected from the consenting cardiovascular disease patients to make DBS and DPS using card-based filtration systems."**
9. In the amended protocol (version 1.1, date 14/7/2022) under the study setting section, we have stated how the cadaveric tissues will be obtained **(lines 374-376) "The control valve biopsies in this study will be obtained from the stored biopsies collected by the HREC approved main study (HREC REF 061/2018)."**
10. In the amended protocol (version 1.1, date 14/7/2022) under the study population and sampling section we have added **lines 383-398**. Explaining the sampling population for negative control

tissues **(lines 383-385)**, and the population and sampling for dried plasma spot pilot studies **(lines 386-398)**.

11. In the amended protocol (version 1.1, date 14/7/2022) at the collection of samples section we have added the collection of samples for dried plasma spot experiments **(lines 416-417) "Blood samples will be collected from cardiovascular disease patients for dried blood spot (DBS), dried plasma spot (DPS), and EDTA-liquid plasma."**
12. In the amended protocol (version 1.1, date 14/7/2022) under the measurements sub-section, we have added explanations for the experimental design for exploring the metabolic profile of liquid plasma, dried blood spot and the dried plasma spot pilot experiments **(lines 443-458)**.
13. In the amended protocol (version 1.1, date 14/7/2022) under the benefits sub-section, we have added a statement explaining the potential benefits of dried plasma spot experiments findings to metabolomics researchers in resource limited regions of Africa **(lines 501-503). "Further, metabolomics researchers will benefit from the assessment of DPS as an alternative sample collection method suitable for sample collection and shipping from remote and resource limited regions of Africa."**

3. Protocol status (tick ✓)

✓☐	Open to enrolment
☐	No participants have been enrolled
	Closed to enrolment (tick ✓)
✓☐	Research-related activities are ongoing
☐	Research-related activities are complete, long-term follow-up only
☐	Research-related activities are complete, data analysis only

4. Proposed changes will affect: (tick ✓ all the categories that apply)

Protocol	
✓☐	Study objectives, design (including investigator's brochure, clinical activities, study length)
☐	Study instruments, questionnaires, interview schedules
✓☐	Sample size
☐	Recruitment methods
☐	Eligibility criteria (inclusion and exclusion criteria)
☐	Drug/device (composition, amount, schedule, route of administration, combination with other drugs/devices, safety information)
✓☐	Data collection/ analysis
☐	Principal Investigator. (Please attach revised conflict of interest and PI declaration statements. Refer: sections 7 and 8.4 in the New Protocol Application Form FHS013)
☐	Consent form and information sheet
☐	Recruitment materials (e.g. advertisements)

☐	Administrative (e.g. change in sponsor's name, change in contact information)
☐	Other. Please specify:

**Note: Amendment changes involving study length, sample size, additional sites and eligibility criteria (i.e. inclusion of minors and /or pregnant woman) need to be declared to the Insurance office. Please liaise via fhs.sponsorship@uct.ac.za regarding the required documentation and information to be submitted to obtain an updated UCT No-fault Insurance Certificate- it should be included herewith*

4.1 In your opinion, will there be any **increase** in risk, discomfort or inconvenience to participants?	☐ Yes	✓☐ No
If yes, please provide a detailed justification/explanation:		

4.2 What follow-up action do you propose for participants who are already enrolled in the study?	
☐	Inform current participants as soon as possible
☐	Re-consent current participants with revised consent/assent forms (append)
✓☐	No action required
☐	Other. Please describe:

5. Detailed description of the change(s)

Please attach, for each amendment, a summary of all changes which clearly indicates:

i. Old wording (e.g. ~~strikethrough~~ text, CHANGED FROM and CHANGED TO)

ii. New wording (e.g. *italicized*, **bold**, tracked)

iii. Detailed rationale/ justification/ explanation for each change

6. Ethics Review for Amendment Levy – cost including vat

Amendment Review Costs including VAT

Please tick amount to be billed:

Submission Type	*Description*	*New fee (Vat Incl.)*	*tick ✓*
Research funded solely from UCT departmental/ divisional/group budget	Major/ Minor Amendments	**R0,00**	✓☐
Non-sponsored student research for degree purposes at UCT/Other Universities & Colleges	Major/ Minor Amendments	**R0,00**	☐
Protocol amendment - Major (FHS006 Form)	Clinical Trial & International Grant Funded Research - Any changes to the protocol that requires Full Committee review	**R8 000.00**	☐
Protocol amendment - Major (FHS006 Form)	Clinical Trial & International Grant Funded Research - Any change to the protocol that requires Expedited review that does not require Full Committee Review	**R5 000,00**	☐
Protocol amendment - Minor (FHS006 Form)	Clinical Trial & International Grant Funded Research - Minor amendments, administrative changes that do not affect study design e.g. changes to informed consent form, changes in study staff, etc.	**R2 250,00**	☐
Protocol amendment - Major (FHS006 Form)	National grant funded research - Any change to the protocol that requires Full Committee review	**R7 000,00**	☐
Protocol amendment - Major (FHS006 Form)	National grant funded research - Any change to the protocol that requires Expedited review that does not require Full Committee review	**R2 500,00**	☐
Protocol amendment - Minor (FHS006 Form)	National grant funded research - Minor amendments, administrative changes that do not affect study design e.g. changes to informed consent form, changes in study staff, etc.	**R1 000,00**	☐

NB: Protocols funded by UCT (e.g. departmental funding / student research) and by certain grant funding organizations (e.g. MRC, NRF, CANSA,) are exempt from these charges.

Please provide details for Invoicing, either complete section 1 or 2 :

1. Invoice billing – Directly to Sponsor

Sponsor's name	
Billing Address of Sponsor:	
Vat Number:	
Contact person:	
Telephone number:	
Email Address:	

2. Internal Journal Billing:

Fund Number:	
Cost Centre Number:	
Account Holder Name:	Prof Ntobeko Ntusi
Division of Account Holder:	Cardiology

7. Amendment Submission checklist (tick ✓)

7.1 Please tick that all the documents are attached before submitting to the HREC. NB: Incomplete submissions will not be processed	
✓☐	Latest FHS006 form completed with all sections completed as per our website
✓☐	Cover Letter
✓☐	PI Justification/ Summary for the reasons for the amendment
✓☐	Protocol - Track changes & Clean Copy (where necessary)
☐	Informed Consent Forms (ICF), if applicable (Any changes made to ICF tracked & clean copy)
✓☐	Any other additional documentation in support of amendment
☐	Updated no fault insurance certificate (if applicable)

Please email this form and supporting documents (if applicable) in a combined pdf-file to hrec-enquiries@uct.ac.za with subject line: FHS006 + (HREC Reference number). The latest forms are found on our website.

8. Signature

My signature certifies that I will maintain the anonymity and/ or confidentiality of information collected in this research. If at any time I want to share or re-use the information for purposes other than those disclosed in the original approval, I will seek further approval from the HREC.			
Signature of PI		Date	12/08/2022

Appendix 2. Participant information sheet and consenting form

PARTICIPANT INFORMATION

Title

Investigating metabolic change as factors causing Rheumatic Heart Diseases (RHD) development

Introduction

You are invited to take part in a study on RHD. We will give a brief description and what you are required to do as a participant.

Please read the information leaflet carefully and you can discuss with your family, friends or doctor.

If there is anything you don't understand or need further explanation, please feel free to ask.

Purpose of study

Patients with rheumatic heart disease (RHD) have soreness of the heart valves. These causes some of the valves to narrow and thicken which may cause disease of the heart muscle. If allowed to continue for a long time it may lead to enlarged heart chambers. We think changes in small molecules called metabolites causes development of the disease. These molecules produced at the heart valve and heart muscles could help us understand RHD development. Changes in these molecules can then be used to identify patients before serious signs appear.

In addition, we want to investigate changes of these molecules in blood. We want to understand how presence of specific molecules could show damages of heart muscles and valves.

Additionally, we want to study RHD using CMR and Echocardiography scans. We will compare both results with changes in molecules at several stages of continuing RHD.

If you accept to take part in this study, we will request to collect some blood before your heart operation. If you have not had a recent ultrasound scan, you will be requested to participate in one. We will then examine the structure and function of your heart with CMR scans.

Why have I been invited?

You have been invited because you have been identified with RHD and require a valve change. This study will help to understand your improvement after operation of the damaged valve(s).

Is it a must I take part in the study?

Taking part in this study is on free will and you are not forced to participate. If you decide to take part, you are free to withdraw your consent at any time without giving a reason. Your decision will not affect the quality of care you receive at Groote Schuur hospital. If you decide to no longer continue with the study after agreeing, we will fully remove your information from our study records.

What happens to me if I take part?

The doctor will do a full clinical examination that includes measurement of your pulse, blood pressure, weight and height. He will also ask some questions about your health and regular treatments. An ultrasound and CMR scan of your heart will be performed before your operation. Both exams are not painful and will have no negative effect on your body. The scans help to properly understand level of RHD and to guide the surgeons and treating doctors.

After operation, the cardiologist will place you on some medication that will help reduce the advancement of RHD.

What do I benefit from taking part in the study?

You will get regular meeting with cardiologist, free ultrasound and CMR scans. This will help to closely observe your healing after operation.

Below, all the tests mentioned above are discussed in a little bit more detail:

Clinical observation

The observation will be about your health and past medical conditions, using a planned survey. The physical exam will include your pulse, blood pressure, weight and height, as well as an examination of your circulatory system.

The heart CMR scan (approximately 60 minutes)

The CMR scan of your heart will be the most important part of this study. It is painless and involve use of a strong magnetic device. Please let the researcher and your doctor know if you have the following conditions before the scans:

- a permanent pacemaker
- metal pins in blood vessels or head
- metal specks in the eye
- insulin pump
- Injuries from bomb or bullet
- automatic implants in the head
- implants in the ear

The CMR machine is shaped like a large doughnut, the opening measures about 1 meter wide, with a table that slides in and out. You will be asked to change into comfortable hospital robe and to lie still on your back on the table. During the scan you will also be asked to breath in and out and hold your breath for several seconds.

Pictures of the heart are created using a magnetic field, radio waves and computers. When images are being taken, the CMR machine makes some sound. These sounds cause no harm to you. You will however be provided with earphones to protect your ears from the sounds and for communicating with the operator. It is important that you stay motionless during the scan.

To evaluate the heart and the blood circulation some contrast dye, called gadolinium will be injected through a drip in your arm. The dye has been tested and approved by cardiologist and have no harm.

What about travel expenses?

We will refund travel costs to and from the hospital when you come for repeat scans after your operation.

What will I have to do, if I agree to take part in this study?

Agree to take part in this study by signing a consent form.

Your heart will be scanned as described above before your operation. Some 20ml of blood will be taken from you by a nurse while you are in the hospital ward waiting your operation. Remains of the heart valves removed during your valve change operation will be taken. They will be used to study changes in molecules due to RHD.

You will have repeat visit to Groote Schuur Hospital for another scan 6 months after operation. During your repeat visit, some blood will also be collected for study.

It is important that you do not eat 4-6 hours before scans, so you might have to come early in the morning.

Are there any possible risks from taking part?

CMR scanner is normally used in hospitals to get images of various body parts.
CMR scans are safe and do not involve any radiations.
However, it is good to note: -

- **Narrowed space**: Some people with fear of narrowed spaces might find it uncomfortable.
- **Noise**: The scanner will produce some mild noisy sound. We provide earphones to protect your ears.

- **Magnetic field**: The scanner consists of a powerful magnet, it may attract certain metallic objects.
- **Pregnancy**: as a precaution we advise you to tell us if there is any chance you might be pregnant.
- **Needle injection:** Some people find having a drip in their arm uncomfortable.
- **Gadolinium Injection**: In rare cases some people have reported warm feeling at injection site, slight nausea, a metallic taste or a rash. In case of discomfort talk to the technician and the injection will be stopped immediately.

What will happen to the samples?

We will collect and store blood, heart valves, heart muscle biopsies and health information of participants. We will prepare your blood samples and tissues right away and store them for analysis at a future date. When this project is complete, we would like to store your blood and heart tissues and information. We will store it together with other samples that people have given. The materials will be stored in University of Cape Town protected by material transfer agreements approved by University of Cape Town Human Research Ethics Committee. Samples might be stored for some time by the investigators. No samples will be sent out of University of Cape Town or South Africa before a material transfer agreement has been permitted. The samples will not be sold, but investigators may develop products based on studying your samples. If this happens, you will not be able to share in any profits.

We would like to use your leftover blood samples and tissues for future research on heart diseases. If we use your blood samples for future heart disease research, we will label them with a code instead of your name. No information obtained from this research will point back to you.

Do you agree to let us store your samples for future research on heart diseases? Below you will find the question clearly asking for your permission to store and use the specimens for future work on heart diseases.

You may still take part in this study if you don't allow samples to be used in the future. This type of research is being done to answer research questions, not to provide you with care. You and your doctor will not be given the results of these tests.

What happens when the research study stops?

The end of the study will not affect the treatment you receive from your doctors. The end of the study will mark the official end of your participation in this project. If you still require repeat appointments, you will be referred to a cardiologist who will see you at your ease. Copies of any publications obtained from this study will be available on request from Dr Ntusi.

Will my taking part in the study be kept private?

Yes. We will follow ethical and legal procedures and all information about you will be handled in confidence. We will follow information management and protection policies as stated by the University of Cape Town. If you take part in the study, some of the information collected from the study would be looked at by authorised persons from the University of Cape Town. All investigators have a duty of privacy to you as a research participant, and nothing that can reveal your identity would be disclosed to third parties. Information collected from the study will be given unidentified code that you would not be identifiable.

What if relevant new information becomes available?

Sometimes, the investigators might get new information about the process of development of the condition being studied. If this happens, one of us will tell you and discuss whether you should continue in the study. If there is enough evidence to suggest you may be harmed from participating in this study, the study could be stopped.

Unexpected findings on your scan

In the unlikely event of us seeing any structural abnormalities on your MRI scan, a clinical specialist will discuss the suggestions with you. However, the scans are not for investigative purposes, and therefore these scans are not a substitute for your hospital appointments. So, if we find anything unusual it would be appropriate for us to contact your doctor. But we would only do this after we and the specialist have discussed your options and gained your permission.

What will happen if I don't want to carry on with the study?

You are free to withdraw from the study at any time. However, it will be a nice thing to participate till the end as this will help in the future treatment of the condition. If you decide to no longer continue with the study after consenting, we will fully remove your collected information from our study records.

What will happen to the results of the research study?

The results of this research are planned to improve on the understanding of RHD and treatment options. We plan to publish the results in scientific papers for the benefit of the wider medical community. However, individual patients will not be identified in any publication and your personal and clinical details will remain strictly private.

Any scientific publications coming out of this study will be available on request to all participants. You would have no legal right to a share of any profits that may arise from the research.

Will your test results be shared with you?

We will show you the images we get from the ultrasound and MRI scans. The results of the other tests will only be available on publication of the results. If results of any of the tests are completely abnormal, we will contact you to discuss before contacting your doctor.

Who is organising and funding the research?

The study is organised and done by researchers from the University of Cape Town and Groote Schuur Hospital. The studies are funded, in part, by National Research Foundation of South Africa.

Who has reviewed the study?

The University of Cape Town Human Research Ethics Committee has reviewed and permitted the study.

Further information and contact details

Should you wish to know more about any this study, please contact **Dr Ntusi** on **(021) 406 6200**.

The **UCT's Faculty of Health Sciences Human Research Ethics Committee** can be contacted on **(021) 406 6626** in case you have any ethical concerns or questions about your rights or welfare as a participant on this research study.

CONSENT FORM

Study Full Title	Understanding Chronology of Inflammatory Heart Disease and Stress Induced Proliferative Mechanisms using Metabolic Biomarkers
Patient ID	
Principal investigator	Prof. Ntobeko Ntusi
Researcher	Daniel Mutithu (MSc)

I confirm that I have read and understood the information sheet for the above study. I have had the opportunity to consider the information, ask questions and have had been answered fully.	YES	NO
I understand that my participation is by free will and that I am free to withdraw at any time without giving any reason and without my medical care or legal rights being affected.	YES	NO
I understand that applicable sections of my medical history and biochemical data collected during the study may be looked at by the researchers on this project at the University of Cape Town, during my taking part in this research.	YES	NO
I understand that my doctor will, with my permission, be informed of the results of medical tests performed as part of the research, which are important for my health care.	YES	NO
I understand that I may be invited for a repeat blood collection, echo and MRI scan after the operation.	YES	NO
There may be future studies that you may qualify to take part in. Can we contact you in the future about future research studies?	YES	NO
There could be some leftovers of some of the tissues collected from you at the end of this study. Do you agree to let us store your samples for future heart diseases research?	YES	NO

Name of participant	Signature	Date
Name of person taking consent	Signature	Date

Appendix 3. Recruitment and sampling checklist

RHD – Proteomics/Molecular study – Sampling check list

Name of Patient			
Study No.			
Date			
GSH File No.			
Date of Birth		Sex:	
Height		Weight:	
Ethnic Group			
Pulse:	BP:		

Consent Signed			
History Obtained		ARF:	Other:
Echo (Confirmed)			
CMR		Date:	
Date of Surgery:			

Specimen Collected

1. Blood	**Date and Time:**		**Processed:**
		Serum	Plasma
	Proteomics		
	Metabolomics		
	Antibody Array		
	MiRNA		
	Cytokines		
	Exosomes		
	Histones		
		PBMC total Cell count	
		Count per Aliquot	
2. Right Atrium	**Date and Time:**		**Processed:**
	Proteomics		
	Metabolomics		
	Right Atrium 3		
	Right Atrium 4		
3. Left Ventricle	**Date and time:**		**Processed:**
	Proteomics		
	MTT		
	Metabolomics		
4. Oral Swabs	**Date and Time:**		**Processed:**
5. Pericardial fluid	**Date and Time:**		**Processed:**
6. Valve Tissue:	**Date and Time:**		**Processed:**
		Mitral	Aortic
	Proteomics		
	Metabolomics		
	Autophagy		
	Histology		
	MiRNA		
	IPMS		
	Histones		

RHD Study
Sample collection log
Version 1.0
26 Sept 2018

Materials checklist

	Item	Qty	Checked
1	Tissue vials	25	
2	Vials labels	25	
3	Biopsy needles	2	
4	scalpels and blades	2	
5	Forceps	2	
6	Liquid Nitrogen	1	
7	10% NB formalin	1	
8	Phosphate buffered saline	1	
9	Marker pens	1	
10	Dissecting board	1	
11	Tubes rack	1	
12	Sample check list	1	
13	Signed consent forms	1	
14	15ml blue cap tubes	5	

RHD Study **Healthy Participant Sampling Check list** Ntusi group

Name of Patient			
Study No.			
Date			
GSH File No.			
Date of Birth		Sex:	
Height		Weight:	
Race			
Pulse:	BP:		

Consent Signed			
History Obtained			
Echo (Confirmed)			
CMR		Date:	

History

Chronic illness	
Arthritis	
Diabetes	
CVD	
Hypertension	
Liver disease	
Kidney disease	
Dyslipidemia	
Asthma or pulmonary disease	
SLE, RA, IBS, Psoriasis	
Crohn's disease	
Inflammatory bowel disease	
On medication (specify)	
Smoker	
Exercise	
Other medical history	
Do you have metal implants?	

RHD Study **Healthy Participant Sampling Check list** Ntusi group

Specimen Collected

1. Blood	Date and Time:			Processed:
		Serum	Plasma	
	Proteomics			
	Metabolomics			
	Antibody Array			
	MiRNA			
	Cytokines			
	Exosomes			
	Histones			
		PBMC total Cell count		
		Count per Aliquot		

Oral pharyngeal swabs	

Materials checklist

	Item	Qty	Checked
1	Tissue vials	25	
2	Vials labels	25	
3	Biopsy needles	2	
4	scalpels and blades	2	
5	Forceps	2	
6	Liquid Nitrogen	1	
7	10% NB formalin	1	
8	Phosphate buffered saline	1	
9	Marker pens	1	
10	Dissecting board	1	
11	Tubes rack	1	
12	Sample check list	1	
13	Signed consent forms	1	
14	15ml blue cap tubes	5	

Appendix 4. Protocol for FFPE tissue processing and H&E staining

FFPE tissue processing and H&E staining

Materials

- Tissue processor
- Tissue cassettes
- Scalpel blades
- Microtome blades
- Dissecting board
- Tissue Embedder
- Embedding moulds
- Forceps
- Painters brush
- Water bath
- Ice Tray
- Microscope glass slides
- Glass coverslips
- Filter paper
- Microscope
- Microtome
- Hot plate

Reagents

- 10% buffered formalin
- Paraffin wax
- 70%, 80%, 95%, 96% and 100% ethanol
- Xylene
- Mayers Haematoxylin
- Phloxine Eosin
- Ammoniated water
- Entellan

Method

The tissues for histological examination were fixed in formalin immediately after excision, processed in a tissue processor overnight, and followed by embedding in paraffin wax the following day. The blocks were stored in appropriate conditions until required for further study. Below follows the detailed procedure for tissue processing in the processor:

Table 1: Tissue Processing Schedule (Leica Tissue Processor)

REAGENT	TIME
10% Formalin	Optional delay start
70% Ethanol	2 hours
96% Ethanol	2 hours
96% Ethanol	2 hours
100% Ethanol	2 hours

100% Ethanol	2 hours
100% Ethanol	2 hours
100% Ethanol	2 hours
Xylene	2 hours
Xylene	2 hours
Paraffin wax (55-56°C)	2 hours
Paraffin wax (55-56°C)	2 hours

The above first step in formalin did not involve a standard time as the tissues were already collected and properly fixed in formalin. This step was only used for the tissues to sit in for a delay start up. Tissues were embedded using the Leica EG1140H embedder and the Leica EG1140C chiller plate (Leica microsystems, Nussloch GmbH, Heidelberger Str., Nussloch, Germany) was used to cool and harden the wax tissue blocks.

Methods on how to make Haematoxylin and Eosin

To make Mayer's haematoxylin mix the following reagents

- Hematoxylin 1g
- Potassium 50g
- Sodium iodate 0.2g
- Citric acid 1g
- Chloral hydrate 50g
- Distilled water 1000ml

Dissolve hematoxylin and sodium iodate in water. Add chloral hydrate and citric acid. Boil for 5 minutes, cool and then filter.

To make eosin mix the following reagents

- 1% Phloxine B 10ml
- 1% eosin y 100ml
- 95% alcohol 780ml
- Glacial acetic acid 4ml

Appendix 5. Protocol for metabolites extraction and mass spectrometry analyses

Metabolite extraction

Materials

- Serum and valve biopsies
- Reverse phase chromatography column
- HPLC pump system and accessories
- QTOF mass spectrometer
- Nitrogen gas sample dryer
- BSL II safety cabinet
- High speed (>13,000 g) and low speed cooled centrifuge

Reagents

- 100% acetonitrile, purity (GC)≥99.9%
- 100% methanol
- Ammonium formate
- Formic Acid
- HPLC grade deionized water
- NIST 1950 SRM – Human plasma metabolites standard

Metabolites extraction from serum

Prepare a batch with randomly selected 3 Cases, 3 Negative Controls, 3 Positive controls, 3 pooled QCs, 3 standard QCs and blanks.

1. Create pooled sample by collecting 50uL of all the samples into 1 aliquot. This forms the pooled-QC sample.
2. Prepare the standard sample QC by reconstituting the provided external standard (NIST SRM 1950) add 50uL of NIST 1950 to 50uL of H2O and process for metabolites extraction as samples and pooled samples.
3. For protein precipitation, add 400uL of ice-cold extraction solvent to 100uL plasma/serum samples/pooled QCs/external standard QCs into an Eppendorf tube.
4. Vortex the mixture for 2 minutes at high speed.
5. Allow the mixture to sit overnight in −20°C.
6. Centrifuge the mixture at 14,000xg for 15 minutes in a cooled centrifuge (4-8°C).
7. Decant the supernatant into a 3ml glass test tubes and dry the supernatant by blowing nitrogen gas at a block heated to 40°C.
8. Store the pellet at −80°C until ready for analysis or reconstitute immediately.
9. Reconstitute the pellet with 150uL of ACN: H2O (98:2) v/v with 0.1FA and 2mM Ammonium Formate and add it into Eppendorf tube.
10. Conduct a quick centrifugation of the extract at about 14,000xg for 5 minutes to ensure there are no precipitates in the sample.

11. After centrifugation add the samples into micro-inserts in the LC vials ready for loading into the autosampler.

Metabolites extraction from valve biopsies

To prepare a batch prepare case tissues, negative control, positive control, and pooled QC samples.

To obtain pooled QC samples obtain 20-50mg of all the tissues which will later be aliquoted to control for all the batches designed for all the analysis tissues. Homogenize the pooled QCs by crushing them first into a powder. The crushed powder is then aliquoted and stored in -80℃ until analysis. Process the pooled QC samples same as test samples.

1. Obtain 40-100mg of fresh frozen tissues and should be mixed with the extraction solvent in the ratio 1mg/3uL of solvent i.e. ACN:MeOH (90:10) v/v in 1% formic acid (FA).
2. Homogenize the tissue in homogenization tubes with 1.4mm ceramic beads and homogenize for three times over 20s at 5500 rpm with 30s pause intervals.
3. Allow to sit overnight in −20°C.
4. Centrifuge at 14,000xg for 10 minutes at (4-8°C)
5. Add the supernatant into 3ml glass test tubes and lyophilize the supernatant with nitrogen gas at 30°C.
6. Store the pellet at −80°C until ready for analysis.
7. Reconstitute the pellet with 100uL of ACN: H2O (98:2) v/v in 0.1FA and 2mM Ammonium Formate and add it into GC vials with micro-inserts.

Chromatography

1. Reverse phase liquid chromatogram was used in this study.
2. Stationary phase was Omega Polar C18 (particle size 1.6um, pore size 100, length 10cm, internal diameter 3mm) (Phenomenex Inc., Torrance, CA, USA) operated at pH 8.5., temperatures 30°C.
3. Mobile phase (A) HPLC grade water with 2mM ammonium formate in 0.1% FA. (B) 100% ACN:MeOH (1:1) v/v with 2mM ammonium formate in 0.1% FA.
4. Total length of the gradient was 60 minutes at a flow rate of 0.4ml/min. The gradient is set as 0.5mins 5%B, 45mins 95% B, 47min 95%B, 47.10min 2%B, 60min 2%B.
5. The flow rate set at 210uL/Min.
6. 0.1% Formic acid is used to modify for positive ion mode and 0.1% ammonium hydroxide is used to modify for negative ion mode.
7. The autosampler injection volume is 5uL/s set at 15°C with rinsing set at 2s at 35uL/s.
8. Column oven set at 40°C.

MS Method

- Data acquisition is performed on +ve and -ve modes using data dependent methods (IDA) with the x500R q-TOF with dual ESI ion sources
- Duration 60minutes, total scan time 0.713 seconds.
- TOF mass range 50-1200 Da, accumulation time 0.1s with a declustering potential 70V and collision energy at 10V. time bins to sum was set at 4. CAD gas at 7.
- Maximum candidate ions was set at 7, intensity threshold 50 counts/second and was set to exclude former candidate ions. Mass tolerance was set at +/- 50mDa

- TOFMSMS IDA ionspray voltage at 5500V, with an accumulation time of 0.08s, declustering potential 70V, with collision energy at 35V, collision energy spread 15V, Q1 unit resolution, time bins to sum 8

Appendix 6. Data processing parameters with XCMSOnline

XCMSOnline version 2.7.2
XCMS version 1.47.3
CAMERA version 1.34.0

---

1. General parameters
    Polarity positive
    Retention time format minutes
2. Feature detection
    method : centWave
    ppm 10
    snthr 6
    peakwidth 5 60
    mzdiff 0.01
    prefilter peaks 3
    prefilter intensity 100
    noise 0
    Feature detection results :

1T36.mzML	[RHD_4]	-->	12356 Features.
1T43.mzML	[RHD_4]	-->	10464 Features.
2T49.mzML	[RHD_4]	-->	10742 Features.
4T51.mzML	[RHD_4]	-->	12356 Features.
1N41.mzML	[NRHD_5]	-->	4342 Features.
2N33.mzML	[NRHD_5]	-->	9572 Features.
3N35.mzML	[NRHD_5]	-->	16336 Features.
3N42.mzML	[NRHD_5]	-->	13612 Features.
4N45.mzML	[NRHD_5]	-->	12421 Features.
1C24.mzML	[CTL_4]	-->	10461 Features.
2C28.mzML	[CTL_4]	-->	7578 Features.
3C22.mzML	[CTL_4]	-->	8942 Features.
3C9.mzML	[CTL_4]	-->	10444 Features.
4C16.mzML	[CTL_4]	-->	14111 Features.
1QC-P1.mzML	[QC_P4]	-->	10717 Features.
1QC-P2.mzML	[QC_P4]	-->	4104 Features.
2QC-P2.mzML	[QC_P4]	-->	9314 Features.
3QC-P2.mzML	[QC_P4]	-->	12425 Features.
4QC_P1.mzML	[QC_P4]	-->	13709 Features.
1QC-S1.mzML	[QC_S4]	-->	6510 Features.
2QC-S1.mzML	[QC_S4]	-->	6718 Features.
2QC-S2.mzML	[QC_S4]	-->	5420 Features.
3QC-S1.mzML	[QC_S4]	-->	9811 Features.
4QC_S1.mzML	[QC_S4]	-->	8895 Features.
1B1.mzML	[BLK_4]	-->	7400 Features.
2B1.mzML	[BLK_4]	-->	6602 Features.
2B2.mzML	[BLK_4]	-->	4478 Features.

3B1.mzML [BLK_4] --> 9686 Features.
4B1.mzML [BLK_4] --> 5670 Features.
3. Retention time correction
method : obiwarp
profStep 0.1
4. Grouping
method : density
bw 30
mzwid 0.01
minfrac 0.5
minsamp 3
5. FillPeaks
6. Diffreport
classes RHD_4
classes NRHD_5
classes CTL_4
classes QC_P4
classes QC_S4
classes BLK_4
statistical test Kruskal Wallis
statistics.threshold.pvalue 0.01
statistics.diffReport.value into
statistics.normalization Median Fold Change
Removing QC_P4 from the statistical analysis
Making Coeff. Var. from QC sample: QC_P4
Finished Running Statistical tests
7. Additional Plots & Statistics
Running mummichog
Printing MDS plot
Printing static PCA and Select Scaling plot
8. Annotation (isotopes & adducts)
featureAnnotation.CAMERA.annotate isotopes
featureAnnotation.CAMERA.mzabs 0.015
featureAnnotation.CAMERA.ppm 10
featureAnnotation.CAMERA.sigma 6
featureAnnotation.CAMERA.perfwhm 0.6
featureAnnotation.CAMERA.maxcharge 3
featureAnnotation.CAMERA.maxiso 4
featureAnnotation.CAMERA.intensity into
9. Putative ID's (METLIN)
identification.METLIN.ppm 10
identification.METLIN.adducts M+H
Found 1519 total, MS^2 spectra to match

Appendix 7. Data processing parameters with MS-DIAL

MS-DIAL Data processing positive mode

MS-DIAL ver. 4.48	
#Project	
MS1 Data type	Centroid
MS2 Data type	Centroid
Ion mode	Positive
Target	Metabolomics
Mode	ddMSMS
#Data collection parameters	
Retention time begin	0
Retention time end	60
Mass range begin	50
Mass range end	1200
MS2 mass range begin	50
MS2 mass range end	1200
#Centroid parameters	
MS1 tolerance	0.01
MS2 tolerance	0.025
#Isotope recognition	
Maximum charged number	2
#Data processing	
Number of threads	1
#Peak detection parameters	
Smoothing method	LinearWeightedMovingAverage
Smoothing level	3
Minimum peak width	5
Minimum peak height	1
#Peak spotting parameters	
Mass slice width	0.05
Exclusion mass list (mass & tolerance)	
#Deconvolution parameters	
Sigma window value	0.5
MS2Dec amplitude cut off	0
Exclude after precursor	TRUE
Keep isotope until	3
Keep original precursor isotopes	TRUE
#MSP file and MS/MS identification setting	
MSP file	None
Retention time tolerance	
Accurate mass tolerance (MS1)	
Accurate mass tolerance (MS2)	
Identification score cut off	
Using retention time for scoring	
Using retention time for filtering	
#Text file and post identification (retention time and accurate mass based) setting	
Text file	
Retention time tolerance	0.1
Accurate mass tolerance	0.01
Identification score cut off	85
#Advanced setting for identification	
Relative abundance cut off	0

Top candidate report	TRUE
#Adduct ion setting	
#Alignment parameters setting	
Reference file	(File directory of QC samples)
Retention time tolerance	0.05
MS1 tolerance	0.01
Retention time factor	0.5
MS1 factor	0.5
Peak count filter	0
N% detected in at least one group	0
Remove feature based on peak height fold-change	TRUE
Sample max / blank average	5
Sample average / blank average	5
Keep identified and annotated metabolites	FALSE
Keep removable features and assign the tag for checking	FALSE
Gap filling by compulsion	TRUE
#Tracking of isotope labels	
Tracking of isotopic labels	FALSE
#Ion mobility	
Ion mobility data	FALSE

Data processing negative mode

MS-DIAL ver. 4.70	
#Project	
MS1 Data type	Centroid
MS2 Data type	Centroid
Ion mode	Negative
Target	Metabolomics
Mode	ddMSMS
#Data collection parameters	
Retention time begin	0
Retention time end	100
Mass range begin	50
Mass range end	1200
MS2 mass range begin	50
MS2 mass range end	1200
#Centroid parameters	
MS1 tolerance	0.01
MS2 tolerance	0.025
#Isotope recognition	
Maximum charged number	2
#Data processing	
Number of threads	1
#Peak detection parameters	
Smoothing method	LinearWeightedMovingAverage
Smoothing level	3
Minimum peak width	5
Minimum peak height	1
#Peak spotting parameters	
Mass slice width	0.05
Exclusion mass list (mass & tolerance)	

#Deconvolution parameters	
Sigma window value	0.1
MS2Dec amplitude cut off	0
Exclude after precursor	TRUE
Keep isotope until	0.5
Keep original precursor isotopes	TRUE
#MSP file and MS/MS identification setting	
MSP file	None
Retention time tolerance	
Accurate mass tolerance (MS1)	
Accurate mass tolerance (MS2)	
Identification score cut off	
Using retention time for scoring	
Using retention time for filtering	
#Text file and post identification (retention time and accurate mass based) setting	
Text file	
Retention time tolerance	0.1
Accurate mass tolerance	0.01
Identification score cut off	85
#Advanced setting for identification	
Relative abundance cut off	0
Top candidate report	TRUE
#Adduct ion setting	
Reference file	(file directory of QC samples)
Retention time tolerance	0.05
MS1 tolerance	0.01
Retention time factor	0.5
MS1 factor	0.5
Peak count filter	0
N% detected in at least one group	30
Remove feature based on peak height fold-change	TRUE
Sample max / blank average	5
Sample average / blank average	5
Keep identified and annotated metabolites	FALSE
Keep removable features and assign the tag for checking	FALSE
Gap filling by compulsion	TRUE
#Tracking of isotope labels	
Tracking of isotopic labels	FALSE
#Ion mobility	
Ion mobility data	FALSE

Appendix 8. Protocol for MALDI MSI of fresh frozen tissues

Cryosectioning of tissues

1. Immediately after excision the tissues are diced and placed in labelled cryovials then snap frozen in liquid nitrogen.
2. The tissues are then stored in -80°C until further analysis
3. Prepare the cryostat microtome to be chilled to -20°C. prechill the sample stage and the ITO slides to -20°C.
4. Also allow the tissue to acclimatize with the temperatures of the cryostat microtome.
5. Mount the tissue with HPLC grade water by placing a drop on the prechilled stage and immediately mounting the tissue on the stage.
6. While mounting the tissue ensure it is positioned in the desired orientation of the sections.
7. Trim the tissue to obtain a flat surface.
8. Obtain sections of 5µM and immediately thaw-mount the sections on the prechilled ITO slides by placing a warm finger behind the slide.
9. Dry the slides by placing them in a desiccator for 15min. Ensure the desiccator is wrapped with aluminum or placed in 50ml blue cap tubes wrapped in aluminum. The sections should be protected from direct sunlight to avoid photodegradation.

Tissue preparation for MALDI TOF

1. The tissue sections should not be washed at all. Store it in beakers wrapped in aluminum foil.
2. Prepare the 2,5-dihydroxybenzoic acid (DHB) matrix by dissolving 40mg/mL of DHB in 0.1 FA and 50% methanol.
3. To prepare 5ml of 40mg/mL DHB add 200mg of DHB in 50% Methanol in 0.1% FA.
4. Dissolve the matrix in a sonicator for 15 minutes.
5. For matrix spraying set up the fume hood, the HTX TM-Sprayer and the computer loaded with the HTX control software.
6. Use 50% Methanol as the push solution in the HPLC pump.
7. While preparing the equipment and loading the matrix make sure the sprayer knob is on LOAD.
8. Open the software and set up the required protocol for your experiment. Choose appropriate method. Confirm the settings as (velocity = 1250 mm/min, flow rate = 0.05mL/min, passes = 12, line spacing = 3mm, matrix density = 0.0128 mg/mm^{2,} nozzle temperature = 80°C).
9. On the software adjust the spraying according to the number of slides you need to spray.
10. Set the temperature of the sprayer according to the values indicated in the protocol.
11. Turn the knob behind the sprayer for nitrogen gas to the pressure recommended by the protocol.
12. Wait for the temperatures to rise and click start on the TMX- control software.
13. Push 5ml of the matrix though the syringe corresponding to the DHB matrix to be used for this experiment when the temperatures have risen above 30°C.
14. Switch the sprayer knob from LOAD to SPRAY.
15. Prepare the standard mix for calibration. Use the standards within the range of the targeted metabolites less than 1Da either on positive or negative mode. Mix the prepared group of standards with 40mg/mL DHB matrix this will be applied on the slides after spraying with DHB.

16. After the sprayed slides are airdry, apply a spot on a cleaned area on the slide (can be wiped with methanol) using 1mL of the prepared standard mixture. Apply several layers by allowing it to dry before applying the next layer. Apply at least 4 layers of the standard.
17. Draw orientation marks (teaching marks) using Tipp-Ex pen at the corners of the slide ensuring the tissue is within the markings.
18. Scan the tissue with the scanner using the CyberviewX application, scan settings at 3200dpi and 8-bit colour.
19. Open scan>prescan>prescan current frame. Adjust the view area on the pre-scanned image. Click scan> scan current frame.
20. Open the image that is automatically saved in the files, adjust the orientation according to the way the tissue will be loaded in into MALDI scanner.

Equipment calibration

Using Matrix peaks

1. Prepare a matrix mix (2,5-dihydroxy-benzoic acid (DHB), α-cyano-4-hydroxycinnamic acid (HCCA) and 3,5-dimethoxy-4-hydroxycinnamic acid (SA)).
2. Dissolve the matrices with the solvent used to apply the matrix on the tissue to make 10mg/mL of each of the matrices.
3. Make a spot on the ITO slide with 1uL of the calibrant added with multiple layers.

MALDI MSI scanning of the tissue

1. Insert slide into Rapilflex slide holder.
2. Use a 96well plate lid to provide the orientation within the flex control environment.
3. Mark the tissue orientation where the X marks locations are on the 96 well plate lid.
4. Also highlight the location where the calibration spot is.
5. Place the cover lid on top of MTP384 ground steel plate and use it for orientation.
6. Open the flex control.
7. Insert the slide carrier into rapiflex.
8. Select the correct geometry (MTP slide adaptor II).
9. Select method (check method for small molecules).
10. Press the insert button.
11. Navigate to the calibration spot and press start to shoot.
12. If the spectra is of good quality “add” repeat 3 times
13. Display sum spectra.
14. At the calibration tab select the relevant mass control list of calibrations. Select the suitable calibration method.
15. Click on automatic assign and click on apply when at least 3 mass standards are assigned.
16. Open flex imaging.
17. Set up new imaging run.
18. Assign a name to your image run.
19. Acquisition settings raster 50µm, method (suitable for small molecules).
20. Processing options select “perform smoothing” and perform base line subtraction.
21. Select the image that was scanned.
22. Create 3 teaching points on image at the fleximaging guided by the X marks that had been made.
23. On the flex imaging create ROIs around the tissue edges using polygon function.

24. Draw the ROIs with the mouse then click “start automatic run” navigate to flex control and wait until the image run is complete.
25. Navigate back to flex imaging and click load results. The data will be saved in D:\data drive.
26. Inspect the loaded spectra.
27. Conduct another scan as stated above however only using the areas with only the matrix as an ROI. This will be used to remove matrix effect.

MALDI data analysis with ScilsLab

1. Plug in scilslab software dongle.
2. Open the scils lab from the icon.
3. To make independent data sets Load the dataset into the scils lab by clicking on File>New>select TOF> Next.
4. Use the red+ to add the data sets you need to analyze.
5. To combine two or more data sets into one .sl file click on New (combine data sets into a new .sl file.
6. Select TOF as the instrument type> next>click on the red+ to navigate to the data sets stored in D:data file keep adding as many as the data sets you want to compare.
7. When all the data sets have been imported then you can conduct downstream analysis.
8. Name the new data set formed based on the datasets merged.

Statistical analysis in scils lab to extract significant features

1. Follow the scils lab resources and webinars to extract/discover *m/z* that correlate with the annotated regions.
2. Show the ROC data to show the discriminatory power of the idenfied marker *m/z*.
3. Use find colocalized *m/z* values tools to determine features that are exclusive for an ROI by selecting a correlation threshold of 0.5.

ScilsLab statistics analysis to extract significant features

1. Import the images into one datafile as previously explained by importing raw spectra data.
2. The pipelines for data pre-processing. The pipelines help with peak picking, alignment, and look for similar or different *m/z* features between the cases. There are two pipelines. Classification pipeline and the segmentation pipeline.
3. The classification pipeline looks for features distinguishing the ROIs. When using classification, provide the training, validating and classification dataset. One needs to divide the datasets to help with the classification. Use the obtained *m/z* intervals and generate *m/z* images and investigate their distribution on the tissues. Use the intensity boxplots and the ROC curves to determine how well they differ between the ROIs of interest. Outputs from Classification pipeline is different thus conducting a discriminant analysis is invalid.
4. While using the segmentation pipeline, it does not consider the differences in the tissues. Conducts spectra alignment and peak picking. To determine discriminating *m/z* features, use component analysis.
5. PCA analysis. Uses the aligned *m/z* intervals. After the PCA outputs investigate the important *m/z* features from the loadings. Draw the 95 CI ellipse and anything outside the ellipse is the important features. Explains the most to the differences in the ROIs of interest. Also check the output in the *m/z* intervals with the most features. Check the

segmentation pipeline peaks output. Feed the results into the viewer and also use it to obtain the PCA analysis.

6. To obtain the features, ensure that the identified features generate or forms clear peaks after alignment and peak picking. The formed peak can be used to obtain the *m/z* image that can be assessed on its distribution on the tissues, check the AUC, check the intensity box plot.
7. After loading the PCA component that explains the most variation, choose to see the spectra as gel. The brighter bands show *m/z* features with high loading values. Generate the *m/z* images and inspect their distribution around the tissue using the intensity boxplots and AUC
8. After the analysis export the data as table include the peaks output and then include other columns necessary for further analysis and presentation of the data
9. If one has a set of m/z features of interest, use the peaks results and manually generate the *m/z* image and inspect the ion's distribution around the tissue.

Appendix 9. Data management plan

Understanding Chronology of Inflammatory Heart Disease and Stress Induced Proliferative Mechanisms using Metabolic Biomarkers

A Data Management Plan created using DMPRoadmap

Creator: Daniel Mutithu

Affiliation: University of Cape Town (UCT-Generic)

Template: National Research Foundation - South Africa

ORCID iD: https://orcid.org/0000-0001-8587-0416

Grant number: 104839 and 105858

Project abstract:

The case-control study intends to understand the pathogenesis of rheumatic heart diseases (RHD) using metabolomics techniques. Metabolomics studies will provide robust data to identify pathways and biomarkers that signal RHD etiology. Targeted and untargeted mass spectroscopy experiments will be conducted to determine the quantitative concentrations of the identified metabolites in serum/plasma samples. Matrix Assisted Laser Desorption/Ionization Mass Spectroscopy Imaging (MALDI MSI) technique will be used for spatial localization of metabolites on tissues. Results from this study will also contribute towards creating a mathematical model for computational simulation of left ventricular remodeling during myocardial inflammation/fibrosis. From the study we intent to obtain cardiac medical history of the participants. Further we will obtain metabolomics data from liquid chromatography mass spectrometry (LC-MS) experiments using plasma/serum or cardiac tissues. We will further obtain transcription analysis data to understand control of pathways affecting the detected pathways.

Last modified: 21-12-2020

Understanding Chronology of Inflammatory Heart Disease and Stress Induced Proliferative Mechanisms using Metabolic Biomarkers - CHED DMP

Other DMP requirements

Have you already written a DMP for this project?

- No

Data collection

What are your data sources?

- I am generating primary research data

What data will you be collecting?

New MS data will be collected for this study from cases and controls. The LCMS and MALDI-MSI data will be obtained directly from the mass spectrometry. The clinical data retrospectively collected from the patient's file will be used to help in stratification of participants into the study groups. Histology data collected as H&E images from the tissues will be used for histopathology assessment. Transcription data will be obtained from an RT-PCR equipment.

Where will you be collecting your data?

University of Cape Town and Groote Schuur Hospital, Cape Town, Western Cape, South Africa

What format(s) will your data use?

- Tables and spreadsheets .xls
- Mass spectrometry data .mzML
- Image formats .tiff, .jpg
- Text formats .txt, .doc

How much data will you be collecting?

Data collected from about 150 participants

Storage

Indicate where your data be stored during the research. Indicate what security measures will be in place, and your plans for backing-up your data.

Raw data will be stored on a secure external harddrive accessible only by the data collector and the PI. Processed data, with identifiers removed, will be placed on the UCT's G-drive account where it can be shared with the full team, including external collaborators. We have adopted the Purdue University guide for file-naming for all the raw files obtained from LCMS data analysis batches. All data will be backed-up on monthly intervals to the secure external harddrive.

Metadata: Indicate how the data will be documented and described (metadata is descriptive information that helps others understand or categorise your data).

The metadata collected is guided by Metabolomics Standards Initiative (MSI)
Data preservation will be guided by the MSI standards. Pre-processed LCMS raw data from the experiments will be archived for future studies. The raw files will be stored as .mzML formats for easy sharing within the metabolomics community. All the raw data files and the metadata will be anonymized. All data sharing and presentations will be guided by UCT Research Data Management Policy.

Data permissions, data sharing and ethical considerations

If you are using secondary data, indicate how will you obtain permission to reuse the secondary data?

N/A

If you are collecting primary data, how will you obtain informed consent from participantsfor participation in your study?

N/A

If you are collecting either primary or secondary data, how will you gain permission to disseminate the results of the study? What steps do you plan to take to ensure that the identity of participants is protected when you publish results?

Participants will be recruited at the Groote Schuur Hospital. They will be taken through the study information materials. During the consenting the participants will indicate that they have given us permission to publish the results of the study. They will be informed of their rights and how their privacy will be preserved. They will then sign their informed consent that will be stored in a secure location at UCT.
Clinical data about the participants will be anonymized with unique identifiers. Source files with personal information will be stored in a secure location with controlled access only to the researchers. The LCMS, Histology, RT-PCR and MALDI data will be anonymized with unique identifiers that cannot be used to identify the research participants.
The clinical data collect, and the signed informed consents contains information that can be used to directly lead to the participant. Using the participants names in the data collected puts their privacy at risk. To mitigate the risk, Helsinki declaration and good clinical practices will be strictly adhered to while handling participants' data. We had to apply for HREC clearance to allowed to commence the study and data handling plans was stated in the approved protocol.

If you are collecting either primary or secondary data, do you plan to share the dataset(s)? How will you gain

permission to share this data? What steps do you plan to take to ensure that the identity of participants is protected when you publish the data?

During consenting, the participants will be asked to allow us to share the datasets coming from the study. To share data with collaborators, data and matterial trasfer agreements will be signed as guided by UCT and HREC guidelines. The processed and analyzed data will be suitable for open access sharing. UCT's Intellectual Property Policy and Research Data Management policies will guide the data sharing.
The data can be used by metabolomics researchers, bioinformatics researchers, clinicians and students. The data will be stored in public repositories for easy access by the different disciplines. The data will be shared under Creative commons licenses.

If you are collecting either primary or secondary data, how will you ensure that sensitive data is stored securely?

Clinical data about the participants will be anonymized with unique identifiers. Source files with personal information will be stored in a secure location with controlled access only to the researchers. The LCMS, Histology, RT-PCR and MALDI data will be anonymized with unique identifiers that cannot be used to identify the research participants. Data with sensitive information will be stored in a password controlled computer accessible to only the research team.

Data retention and preservation

Indicate which data (if any) should be retained once your research is complete.

Mass spectrometry will be retained for resuse in the future for bioinformatics studies and projects focusing on metabolomics. RT-PCR data will be retained to coroborate MS data to understand metabolic dysregulation patterns.

Indicate for what period of time your data should be retained.

The data will be converted to open source formats and deposited in secure repositories for future use by the metabolomics community.

Indicate where your data will be stored in the relative long-term (while it is being retained).

The retained data will be deposited in UCT controlled repository (ZivaHub) for longterm access by the research community.

Implementing and resourcing your data management process

Indicate what resources or additional guidance you will need to implement the above plan, if any.

Publication and data curatuion charges will be covered by the project funding.