UNIVERSITY OF LEICESTER

DISTINGUISHING HEART FAILURE WITH PRESERVED EJECTION FRACTION FROM HEART FAILURE WITH REDUCED EJECTION FRACTION USING PROTEOMICS TECHNIQUES

Thesis submitted for the degree of

Doctor of Philosophy

at the University of Leicester

by

Richard J. Mbasu BSc, MSc

Department of Cardiovascular Sciences

Department of Cancer Studies and Molecular Medicine

University of Leicester

September 2016

Abstract: Distinguishing heart failure with preserved ejection fraction from heart failure with reduced ejection fraction using proteomics techniques.

Heart failure is the second leading cause of morbidity and mortality in the world after cancer. In the UK, over 500,000 people are living with heart failure of which 30-40% die within 1 year of diagnosis. Some biomarkers for diagnosis and prognosis of heart failure have been established. However, they suffer from poor levels of accuracy and efficacy and their roles in clinical use is poorly understood. Thus, new biomarkers are needed.

In this research, mass spectrometry based proteomics was used to profile patients plasma for clinical biomarker discovery due to its ability to perform both quantitative and qualitative protein profiling on clinical samples. Ninety samples from control, heart failure with preserved ejection fraction and heart failure with reduced ejection fraction subjects were used. Plasma protein profiling was performed using an optimised UPLC-IMS-DIA-MS^E label free quantitation method. Bioinformatics analysis was used to analyse the changes observed in the protein profiles to identify potential biomarkers of heart failure.

A novel method, termed mixed mode matrix was used for pilot study prior to main study with lipid removal agent. Samples were analysed using Waters Synapt G2S HDMS QToF mass spectrometer in triplicate on positive mode electrospray ionisation. Statistical comparisons of protein profiles was carried out using Progenesis LC-MS prior to data mining using SPSS, RapidMiner and SIMCA 14 to identify potential biomarkers. Thirty proteins were identified as potential biomarkers and shown to be involved in various pathophysiological processes leading to heart failure. ASL which plays role in nitrogen oxide production in the epithelium was upregulated in heart failure cohort. Conversely, GPX3 which scavenges free radicals in blood preventing apoptosis and necrosis of cells was downregulated in heart failure cohort. These two proteins were proposed as potential biomarkers for heart failure with preserved ejection fraction.

Future studies to validate these biomarkers with the developed targeted LC-MS based MRM assay is needed.

List of publications and presentations in scientific meetings.

Publications:

Richard J. Mbasu, Amirmansoor Hakimi, Jatinderpal K. Sandhu, Liam M. Heaney, Daniel C. Chan, Paulene A. Quinn, Sanjay S. Bhandari, Thong H. Cao, Leong L. Ng & Donald J. L. Jones. A method for plasma protein preparation using a chemical affinity matrix (Manuscript submitted to proteomics). This work was adapted as chapter 3 in this thesis.

Mbasu R.J., Heaney L.M., Molloy B.J., Hughes C.J., Ng L.L., Vissers J.P.C., Langridge J.I., Jones DJL. (**2016**). Advances in quadrupole and time-of-flight mass spectrometry for peptide MRM based translational research analysis. *Proteomics*. 16:2206-2220. This work was adapted as chapter 5 in this thesis.

Mbasu R.J., Hakimi A., Sandhu J.K., Heaney L.M., Quinn P.A., Bhandari S.S., Jones D.J.L., Ng L.L. (**2015**). Proteomics of human plasma in diastolic heart failure (DHF) using novel chemical affinity, mixed mode matrix (M3). *Heart*. 101:A7

Heaney L.M., Jones D.J.L., **Mbasu R.J**., Ng L.L., Suzuki T. High mass accuracy assay for trimethylamine N-oxide using stable-isotope dilution with liquid chromatography coupled to orthogonal acceleration time of flight mass spectrometry with multiple reaction monitoring. Analytical and Bioanalytical Chemistry 2016; 408:797-804

Presentations

Mbasu R.J., Ng L.L., Jones D.J.L. (**2016**). Oral presentation at BMSS conference, Eastbourne, UK

Mbasu R.J., Ng L.L., Jones D.J.L. (**2016**). Poster presentation at ASMS 64th Conference on MS and allied topics, San Antonio, Texas, USA

Mbasu R.J., Ng L.L., Jones D.J.L. (**2016**). Oral presentation at PubhD Leicester, Leicester, UK

Mbasu R.J., Ng L.L., Jones D.J.L. (**2016**). Oral presentation at Thesis forum, University of Leicester, UK

Mbasu R.J., Ng L.L., Jones D.J.L. (**2015**). Oral presentation for biotechnology YES competition, Unilever, UK

Mbasu R.J., Ng L.L., Jones D.J.L. (**2015**). Oral presentation at Café Cardiologique, University of Leicester, UK

Mbasu R.J., Ng L.L., Jones D.J.L. (**2015**). Poster presentation at BMSS conference, Birmingham University, UK

Mbasu R.J., Ng L.L., Jones D.J.L. (**2015**). Poster presentation at BSCR conference, Glasgow University, UK

Mbasu R.J., Ng L.L., Jones D.J.L. (**2015**). Oral presentation at Leicester Adult Education College, Leicester, UK

Mbasu R.J., Ng L.L., Jones D.J.L. (**2015**). Oral presentation at the postgraduate half-day seminar, University of Leicester, UK

Mbasu R.J., Richards C. B. (**2015**). Oral presentation at Café research, University of Leicester, UK

Mbasu R.J., Ng L.L., Jones D.J.L. (**2015**). Poster presentation at the half-day seminar, University of Leicester, UK

Mbasu R.J., Ng L.L., Jones D.J.L. (**2014**). Poster presentation at the half-day seminar, University of Leicester, UK

Acknowledgements

The completion of this work would not have been possible without the assistance and support of a great number of people. Firstly, I would like to thank God for all the wisdom and perseverance he has given me because without his grace this would not have been possible.

Secondly, I would like to thank my primary supervisor Dr Don DL Jones for offering me the opportunity to pursue this PhD. Don has not only been a supervisor to me but also a father figure that I lacked in England. He has not only been there for me in my entire academic journey but also my personal life. I am forever grateful. My secondary supervisor, Professor Leong Ng has been nothing but excellence in my science career. I am better scientist because of your combined efforts.

The entire van Geest group, Liam, Jodie, Paulene, Daniel, Thong, Ben, Zubair, Franscois and Rachel, you have been nothing but happiness in my life. Jodie and Paulene in particular have been the backbone of my mass spectrometry knowledge. Liam, Thong and Daniel have taught me a lot as far as data analysis is concerned. Ben the animation and macro guy, should probably consider another career option. Zubair a.k.a the skysports reporter, Franscois and Rachel my drinking buddies. Could not have asked for a better group.

Thank you to all the members of my office; we have shared very happy moments together. I will really miss you guys especially my desky uncle Hashy. Special thanks to Abi, Meeta and Ashley. Words cannot express how grateful I am, you not only became my good friends but family to me. Meetal, Reshma, Pott and the entire CVS department, I am forever grateful for the moments we have shared together.

Finally, I want to thank my parents and the entire family for their financial and mental support throughout this journey. Without your support I would not have come this far. I am forever grateful.

Contents

List of figures

List of tables

Abbreviations

µg	microgram
µL	microlitre
1-D	one dimension
ESI-MS/MS	electrospray ionisation tandem mass spectrometry
2-D	two dimension
ACN	acetonitrile
ADH	alcohol dehydrogenase
ANOVA	analysis of variance
ATP	adenosine Tri-phosphate
AUC	area under the receiver operating characteristic curve
BCA	bicinchoninic acid
BNP	B-type natriuretic peptide
°C	degree Celsius
CRP	C-reactive protein
CVD	Cardiovascular diseases
Da	Dalton
DHF	diastolic heart failure
DTT	dithiothreitol
ECG	electrocardiogram
EF	ejection fraction

ELISA	enzyme linked immunosorbent assay
ESI	electrospray ionisation
FA	formic acid
FDA	food and drug administration
fmol	femtomole
g	relative centrifuge force
GFP	Glu-fibrinogen peptide
HA	hydroxylapatite
HAPs	high abundant proteins
HDL	high-density lipoprotein
HDMS	high definition mass spectrometry
$HDMS^E$	ion mobility LC-data independent acquisition MS mode
HF	heart failure
HFPEF	heart failure with preserved ejection fraction
HFREF	heart failure with reserved ejection fraction
HPLC	high pressure or high performance liquid chromatography
IAA	iodoacetamide
IMS	ion mobility spectrometry
IQR	interquartile range
kDa	kilo Dalton
LAPs	low abundance proteins

LC	liquid chromatography
LDL	low-density lipoproteins
LRA	lipid removal agent
LVEF	left ventricular ejection fraction
M3	mixed mode matrix
m/z	mass–to–charge ratio
MALDI	matrix assisted laser desorption ionisation
MALDI-MS	matrix assisted laser desorption ionisation mass spectrometry
MARS	Multiple Affinity Removal System
mg	milligram
ml	millilitre
mM	millimolar
MMP	matrix metalloproteinase
MRM	multiple reaction monitoring
MS	mass spectrometry
MS/MS	tandem mass spectrometry
MS^E	non-ion mobility LC-data independent acquisition MS mode
nm	nanometre
PBS	phosphate-buffered saline
PLGS	Protein Lynx Global Server
pmol	picomole

QC	quality control
RP LC	reversed phase liquid chromatography
RP-RP LC	reversed phase – reversed phase liquid chromatography
SD	standard deviation
SHF	systolic heart failure
SNS	sympathetic nervous system
SPE	solid phase extraction
SPSS	Statistical Package for the Social Sciences
SRM	selected reaction monitoring
TFA	trifluoroacetic acid
TOF	time of flight
UK	United Kingdom
UPLC	ultra performance liquid chromatography
vs	versus
USA	United States of America
UV	ultraviolet

Chapter One

INTRODUCTION

1 Introduction

1.1 Cardiovascular diseases

Cardiovascular disease (CVD) also known as heart and circulation disease and is one of the leading causes of death in the world today (Roger *et al.,* 2012). It can be described as a collection of all the diseases that affect the heart and blood vessels that include stroke, atrial fibrillation, diabetes, coronary heart disease and heart failure (HF). Currently, it is reported that more people die from HF than any other condition in the world (Moran *et al.,* 2014). According to Townsend *et al*., 2014, in the UK alone, approximately 160,000 people die of CVD annually, 74,000 of these mortalities resulting from coronary heart disease. They reported that the UK spent approximately 6.8 billion pounds in treating CVD within the National Health Service (NHS) in England in the year 2012/2013. Generally, the risk of cardiovascular disease increases as you get older (75 years and above) and consequently in an aging population cardiovascular disease becomes more prevalent.

Stress, smoking and drinking, high blood pressure and high cholesterol are all common risk factors associated with heart disease. For a long time, heart disease was considered a man's problem but in reality, women suffer more than men do especially as they grow older. This could be attributed to their longer life expectancy compared to that of men (Townsend *et al*., 2014). However, with changes in lifestyle in recent years younger women are increasingly at risk of heart disease. This could be as a result of smoking more than they used to before. Obesity which results to diabetes during pregnancy has also been on the rise (Herring *et al*., 2011). In terms of heart failure, more men than women have been reported to suffer from this condition at younger ages. A new study in the University of Leicester by Bloomer et al. 2012 suggested that men of the same age as women are more prone to this condition due to a sex hormone called testosterone which is linked to increased level of bad cholesterol, the low density lipoprotein (LDL). Yet, after the age of 75 more women suffer from heart failure than men due to their longer life expectancy than men.

In general, the risk in young people is considerably lower; however, with increasing cases of diabetes, high blood pressure and childhood obesity, means that the age at which HF is diagnosed has lowered Townsend *et al*., 2014. The study of the structural functions of the

heart and the molecular pathways involved in case of the disease may also be paramount to maximise intervention strategies.

1.2 Heart failure

1.2.1 Definition of Heart Failure

There is no clear definition for heart failure but clinically, it can be defined as a syndrome where the heart fails to pump adequate amounts of blood to the lungs and the rest of the body. It is the final stage of all the cardiovascular diseases. It affects both sides of the heart and is reported to develop over time as the heart pumping actions grow weaker. There are two types of heart failure; systolic heart failure, also called heart failure with reduced ejection fraction (HFREF) and diastolic heart failure also called heart failure with preserved ejection fraction (HFPEF). Ejection fraction (EF) is expressed as a percentage and is the amount of blood pumped out of the left ventricle (LV) at each heartbeat. Normal ejection fraction may be between 50-70 per cent and is calculated from the left ventricular end diastolic volume (LVEDV) and left ventricular end systolic volume (LVESV) as shown below;

EF = (LVEDV-LVESV)/LVEDV X 100

The European Society of Cardiology (ESC) criteria for normal or mildly abnormal LV function are, left ventricular ejection fraction (LVEF) >50 per cent, LVEDV index <97mL/m2 and LVESV index of <49mL/m2 (Houghton *et al.,* 2013).

HFREF mainly occurs because of systolic dysfunction (section 1.2.3.1) while HFPEF occurs because of diastolic dysfunction (section 1.2.3.2).

1.2.2 Mechanism of heart Failure

When heart failure occurs, cardiac output is reduced significantly. This results in decreased perfusion which stimulates increased sympathetic nervous system activity, renin angiotensin aldosterone system (RAAS) (Demello *et al*., 2009) and inflammation. All these three activities work in a downward spiral to increase heart failure. The sympathetic nervous system is activated due to decreased blood pressure in the aorta that stimulates vasoconstriction. This increases peripheral vascular resistance (PVR) and blood pressure and decreases the cardiac output (Triposkiadis *et al.,* 2009). When RAAS is stimulated, it increases the blood volume and regulates the blood pressure. It is also responsible for the remodelling of the myocardium in the left ventricle making the chamber walls thicker and stiffer thus decreasing the end diastolic volume. However, their long activation causes a lot of damage to the cardiac structure and performance leading to progression of heart failure. All these work in a positive feedback loop and can be reversed by using drugs that can block their mechanisms. For instance, Lim *et al.,* 2001, used angiotensin-converting inhibitors to block the RAAS inhibiting the remodelling of the myocardium in rats according to a study. Beta-blockers on the other hand are used to block the sympathetic receptors preventing vasoconstriction thus maintaining a normal cardiac output. These responses can only support the heart for so long and with time, it fails, a situation referred to as heart failure. Heart failure is divided into two clinical subsets; systolic heart failure (SHF) and diastolic heart failure (DHF).

1.2.3 Types of heart failure

1.2.3.1 Systolic heart failure (SHF)

Systolic heart failure also referred to as heart failure with reduced ejection fraction (HFREF) is a type of heart failure caused by systolic dysfunction. Systolic dysfunction occurs when the left ventricle loses its ability to contract with enough force resulting in inadequate supply of oxygenated blood in the body (EF<40) (Chatterjee, 2007). It represents abnormalities of contraction of the myocardium. HFREF therapy targets the inhibition of the renin-angiotensin-aldosterone and beta-adrenergic systems (neurohormonal blockade). In addition,

since comorbidities like diabetes, dyslipidaemia, hypertension and depression have been common features of HF, treatment of these comorbidities improved the condition substantially (Tschope *et al.,* 2012).

1.2.3.2 Diastolic heart failure (DHF)

Diastolic heart failure also referred to as heart failure with preserved ejection fraction (HFPEF) (Gaasch *et al.,* 2007) is the second most common type of heart failure caused by diastolic dysfunction. It contributes to approximately 50% of heart failure cases and has been known to be associated with the rising cases of morbidity and mortality (Oktay *et al.,* 2013). It represents abnormalities of stretching, filling or relaxation of LV with or without a normal ejection fraction (Table 1.1) and so occurs in conditions where LV becomes less compliant which include ageing, hypertension, LV hypertrophy (Drazner, 2011), myocardial ischaemia and aortic stenosis (Houghton *et al.,* 2013, Martos *et al.,* 2007). HFPEF mostly affects elderly people because of increased collagen cross-linking and increased smooth muscle content (Gutierrez, 2004).

The diagnosis of HFPEF is mainly established with Doppler echocardiography. To diagnose HFPEF, the following criteria must be reached;

i) Patient should show signs or symptoms of heart failure
ii) Systolic LV function should be normal or mildly abnormal
iii) There should be evidence of diastolic LV dysfunction

Treatment of HFPEF mainly focuses on normalising blood pressure, promoting regression of the LV hypertrophy, maintaining normal atrial contraction, treating symptoms of congestion (shortness of breath, swelling of legs etc.) and avoiding tachycardia. Congestion can be prevented by the use of diuretic therapy while beta-blockers are used to prevent tachycardia that prolongs LV diastolic filling time. Angiotensin-converting enzyme inhibitor and angiotensin-receptor blockers have been used to treat diastolic dysfunction (Gutierrez *et al.,* 2004). These two types of heart failure can be detected by blood test (e.g BNP levels). However, they are non-specific and certainly cannot distinguish between the two conditions. An echocardiogram (ECG) plays an important role in distinguishing them by checking the

size, and shape of the heart (Figure 1.1). It uses standard two dimensional, three-dimensional and Doppler ultra sounds to create the sizes and shapes of the heart chambers.

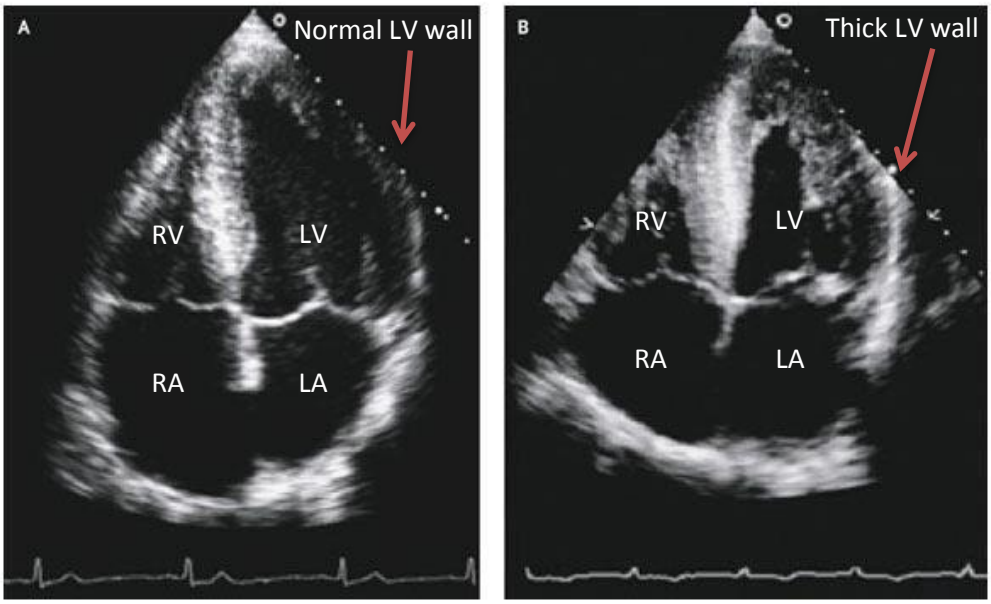


Figure 1.1. *Echocardiographic images showing a person with a normal functioning heart (A) and a patient with diastolic heart failure (B). The patient with heart failure has a thick LV wall (Gaasch et al., 2004). RV=Right Ventricle, RA=Right Atrium, LV=Left Ventricle and LA=Left Atrium.*

1.2.4 Diastolic heart failure vs systolic heart failure

Table 1.1. *Comparison between diastolic and systolic heart failure.*

	DHF	**SHF**
Ejection fraction	>50%	<40%
Pathophysiology	Impaired LV relaxation	Impaired LV contraction
Signs and symptoms	Similar	
Patients characteristics	Older, women, obese, and hypertensive	Prior myocardial infarction
LV stiffness	Increased	Reduced

1.2.5 Plasma and biomarkers

For a long time, scientists have used blood plasma to carry out experiments to ascertain the cause and extent of a disease. This is because blood is rich in proteins (Li *et al.*, 2009) which indicate the biological state of a disease (Anderson *et al*., 2002). Some of these proteins could be potential biomarkers for Heart Failure.

A biomarker is a measurable indicator for the presence of a disease. They play a vital role in understanding the development of chronic diseases like heart failure and can tell you the prognosis and prediction to therapy. Several biomarkers have been discovered and used in the treatment of heart failure including Brain natriuretic peptide (BNP) and C-reactive protein (CRP). When an individual suffers from heart failure, these protein (biomarkers) levels in the blood are measured to check the severity of condition. Blood plasma and serum have long been used in biomarker discovery however, blood plasma is arguably better than serum and most preferred since it has low *ex vivo* protein degradation (Omenn *et al.*, 2005).

A much broader characterisation of the plasma proteome in healthy and disease state could be very useful in identifying biomarkers for early diagnosis and treatment of disease. There are already existing biomarkers (Table 1.3), and chances of finding new ones increase with the number of proteins profiled (Ahmad *et al.*, 2014). However, the most promising source of finding these biomarkers is by investigating the low abundant proteins that probably play a major role in the disease. The low abundant proteins (LAPs) are usually masked by the high abundant proteins (HAPs) (Linke *et al.*, 2007) which constitute the major portion of proteins in the sample (Figure 1.7).

Table 1.2. *A list of the top 14 most abundant proteins in plasma with the top 2 also highlighted. The dominance of these proteins in plasma samples has been highlighted on Figure 1.8 and some of the normal plasma range of these proteins has been included on the table. Accessed on 08/12/2016 from* [*http://scrippslabs.com/protein-levels-in-human-plasma/*](http://scrippslabs.com/protein-levels-in-human-plasma/)

Top 14	**Normal plasma range (mg/dL)**
Albumin	3500-4500
Alpha 1-Acid Glycoprotein	
Alpha 1-Antitrypsin	
Alpha 2-Macroglobulin	400-900
Apolipoprotein A-I	
Apolipoprotein A-II	
Complement C3	
Fibrinogen	
Haptoglobin	100-260
IgA	
IgG	
IgM	
Transferrin	200-320
Transthyretin	10-40

These proteins especially albumin and IgG contribute to more than 70% of total proteins in human plasma. Depleting these proteins is vital for the study of low abundant proteins that could be potential biomarkers for diastolic heart failure.

Table 1.3. *Clinically available biomarkers of heart failure (Adapted from Braunwald, 2008). These biomarkers have been used for diagnosis, prognosis and response to treatment. However, they lack specificity. Thus, more specific biomarkers are needed to treat heart failure.*

Inflammation #*¤	***Oxidative stress #*¶***	***Extracellular-matrix remodelling #*¶***
C-reactive protein	Oxidized low-density lipoproteins	Matrix metalloproteinases
Tumor necrosis factor α	Myeloperoxidase	Tissue inhibitors of metalloproteinases
Fas (APO-1)	Urinary biopyrrins	Collagen propeptides a) Propetide procollagen type 1 b) Plasma procollagen type III
Interleukins 1,6 and 18	Urinary and plasma isoprostanes	Galectin-3
HsCRP soluble ST2	Plasma malondialdehyde	
Neurohormones #*¶	***Myocyte Injury #*¶***	***Myocyte stress *¤¶§***
Norepinephrine	Cardiac specific troponins I and T	Brian natriuretic peptide
Renin	Myosin light-chain kinase I	N-terminal pro-brain natriuretic peptide
Angiotensin II	Heart-type fatty-acid protein	Midregional fragment of proadrenomedullin
Aldosterone	Creatine kinase MB fraction	
Arginine vasopressin		
Endothelin		

#These biomarkers help in explaining the pathogenesis of heart failure

*****These biomarkers provide prognostic information and enhance risk stratification

¤These biomarkers help in identification of subjects at risk of heart failure

¶These biomarkers are potential therapy targets

§ These biomarkers are useful in the diagnosis of heart failure and in monitoring therapy.

1.2.6 Molecular pathways

When the myocardium is damaged, pro inflammatory cytokines (Tumor necrosis factor-α, Interleukin 1, 6 and 18) are produced (Table 1.3). Their production is enhanced by the sympathetic nervous system (SNS). The damaged myocardium together with the hypo-perfused skeletal muscle activates the monocytes (Figure 1.2). After injury these monocytes secrete TNF-α which activates interleukin 1 and 6 (Neumann *et al.,* 1995). When these are activated, some of them are released into the blood stream with the natriuretic peptides (improves the blood circulation) (Anker *et al.,* 2004). These could advance to interfere with the normal function of the ventricles, thus ventricular dysfunction.

In another article Levine *et al.,* 1990 reported high levels of circulating TNF-α and interleukin (1, 6, and 18) in patients with heart failure. They claimed that these were pro inflammatory cytokines and were produced by nucleated cells in the heart where the TNF-α exert its effects. In addition, it mediates cardiac remodelling and ventricular dysfunction after pressure overload according to Sun *et al.,* 2004, and 2007. The cardiac monocytes consist of two TNF receptors, TNFR1 and TNFR2 (Schulz *et al.,* 2009). TNFR-1 is the most expressed and the main receptor since the majority of the deleterious effects by TNF-α seems to be mediated via this receptor. Conversely, TNFR-2 is said to play a more protective role in the heart (Anker *et al.,* 2004). In mice after myocardial infarction, TNFR1 density has been reported to increase significantly for ten days as opposed to TNFR2 that remain unchanged according to a study by Rainer and Gerd 2009. In another study, Sun *et al.,* 2004 investigated the post effects of elevated levels of TNF-α in rat. They discovered that in the wild type rat with acute myocardial infarction (MI) there was free wall cardiac rupture with a frequency of 53.3% that was associated with elevated levels of TNF-α. Conversely, mice without the TNF-α gene had a much-reduced rupture with a frequency of 2.5%. However, Rainer and Gerd 2009 state that both TNFR1 and 2 are down regulated in heart failure while the soluble TNFRs are increased due to proteolytic cleavage and release of TNFR from exosome-like vesicles. An article by Apostolakis *et al.,* 2010 also elaborates that altered cytokine activity in early stages of heart failure especially increased levels of interleukin 6 and TNF-α pose as a risk and have been noted as potential markers for early diagnosis and prognosis. This study confirms that TNF-α plays a major role in myocardial infarction which initiates ventricular remodelling (Sharpe *et al.,* 1992) and heart failure.

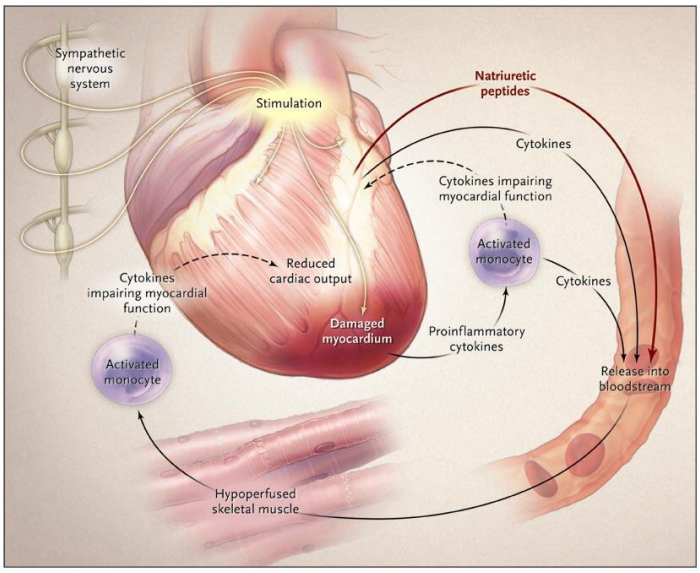


Figure 1.2. *Showing the cytokine hypothesis of heart failure (adapted from Braunwald E., 2008).*

1.2.7 Extra cellular matrix (ECM) remodelling

Remodelling of the ventricles has been reported to play a major role in heart failure. According to Braunwald *et al.*, 1990, this remodelling occurs when there is damage to the heart muscle (Myocardium) due to lack of oxygen. He explains that the ECM plays a major part in this remodelling. It acts as a "skeleton" for myocytes and determine their structure. It consists of various proteins including collagen propeptides (propeptide procollagen type I and plasma procollagen type III), elastin, laminin, fibronectin (Wood *et al.*, 2011), matrix metalloproteinases (MMPs) and the Tissue inhibitors of metalloproteinases (TIMPs). Procollagen type I is cleaved at its C-terminus to produce collagen type I (PIP) which serves as a marker for collagen type I synthesis. In the presence of a failing heart, during degradation, collagen type I is cleaved at specific sites to produce a 36 kDa and 12 kDa carboxy-terminal telopeptide of collagen type I (CITP). This act as a biomarker for collagen type I degradation. The levels of both collagen synthesis and MMPs are enhanced in a failing myocardium in heart failure (Kunishige *et al.*, 2007, Jellis *et al.*, 2011). The MMPs are an endogenous family of Zinc (Zn) enzyme contributing to the remodelling and collagen degradation (Spinale, 2002). They are sub-grouped as shown on Table 1.4.

Table 1.4. *Subgroups of different type of Matrix Metalloproteinase.*

Enzymes	Matrix Metalloproteinase (MMP)
Collagenases	
Collagenase I	MMP 1
Collagenase II	MMP 8
Collagenase III	MMP 13
Collagenase IV	MMP 18
Gelatinases	
Gelatinase A	MMP 2
Gelatinase B	MMP 9
Matrilysins	
Matrilysin I	MMP 3
Matrilysin II	MMP 10
Stromelysins	
Stromelysin I	MMP 7
Stromelysin II	MMP 26
Stromelysin III	MMP 11

They play a key role in other biological processes including, angiogenesis, embryogenesis and wound healing. The MMPs and TIMPs normally exist in equal amounts. However, when the former is more than the latter, left ventricular dilation and remodelling occurs (Ahmed *et al.,* 2006). Due to their involvement in LV remodelling, selective targeting of MMPs can be a potential therapy of heart failure. In 2003, Altieri *et al*. carried out an investigation on increased levels of MMP-2 and MMP-9 in plasma on patients with heart failure. They concluded that increased level of these proteins was associated with persistent extracellular remodelling in HF (Rouet-benzineb *et al.,* 1999) patient and increase in TIMP levels significantly reduced this. In another experiment carried out by King *et al.,* 2003 to inhibit MMP-1 demonstrated that there was a reduced LV wall stress and improved pumping function.

The process of myocardial remodelling not only involves an increase in the myocyte size but also rebuilding the myocardial ECM that consists mainly of collagen fibres as heart failure progresses. Cardiac collagen is synthesised primarily from fibroblasts. Their network is largely influenced by the load with pressure overload reported to increase its content within the LV myocardium. An article by Hamdani and Paulus 2013 emphasise that the rise in ventricular wall stiffness was always accompanied by increased deposition of collagen in patients with HFPEF (Borbely et al., 2005). In addition, Mukherjee *et al*., 1990 explained the significance of collagen in stress-strain relation within the LV myocardium during diastole that is associated with the ventricular wall stiffness (van Heerebeek *et al.,* 2012). There increase in the heart results in stiffening of the LV chamber which in turn alters the pressure/volume relation of diastole in the LV. They form collagen cross-link (Figure 1.3) and shift in isoform transition primarily from collagen I to collagen III. It is also reported that both fibrosis and extracellular matrix turnover have been heavily involved in severity of DHF (Wood *et al.,* 2011). Brilla *et al.,* 2000 also stated that increased myocardial fibrosis might result in deterioration of the LV compliance and diastolic dysfunction in heart failure. Abraham and Krum, 2007 report that in human heart, the ratio of fibroblast to cardiac myocyte is roughly 4:1. They elaborate that when heart failure occurs, fibroblasts are activated by angiotensin II and aldosterone to produce more collagen. The collagen fibres are components of the connective tissue that is responsible for connection, support and separation of different types of tissues in the body. Abraham and colleague emphasise that type III collagen is responsible for the alignment of the myocyte through a weave of struts. These struts can be "pulled apart" when MMPs are activated resulting in cell spillage. However, as much as there is cell spillage and reduction of collagen struts, there might be an increase in the quantity of myocardial interstitial collagen in the heart resulting in diastolic dysfunction (Opie, 2004).

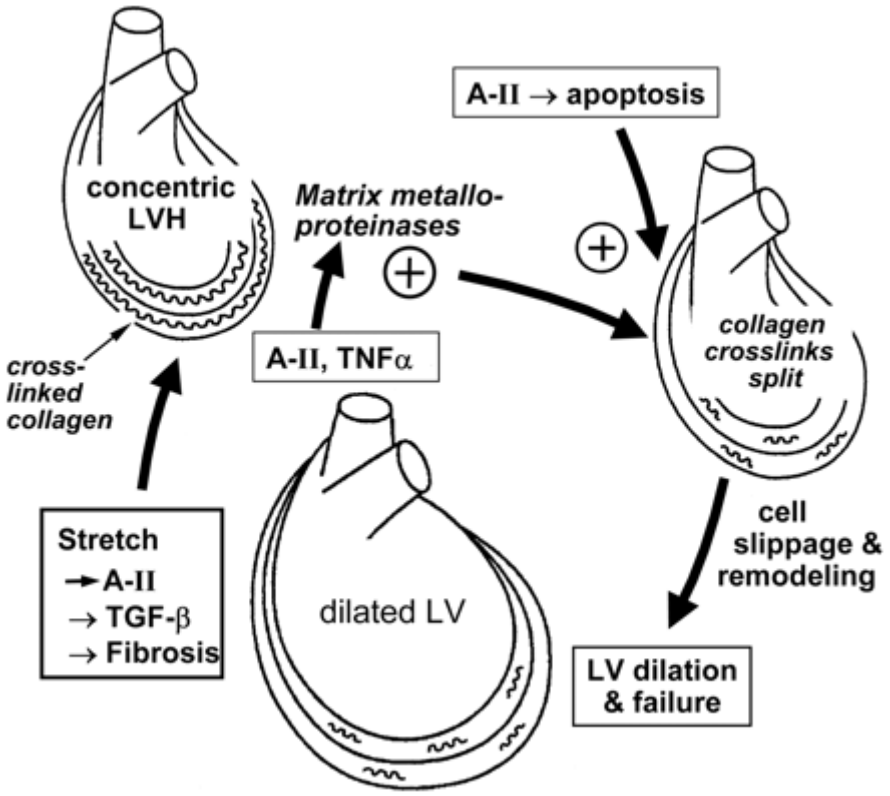


Figure 1.3. *The role of activating Angiotensin II and its downstream effects in the progression of heart failure. LVH=left ventricular hypertrophy, TNF-α-tumor necrosis factor-α, TGF-β=transforming growth factor-β (Adapted from Opie, 2004).*

When blood fills the LV, it stretches releasing angiotensin II (Opie, 2004). This promotes growth and fibrosis by stimulating the transforming growth factor- β (TGF-β). In addition, it stimulates the production of MMPs and tumor necrosis factor-α (TNF-α) which breaks the collagen crosslinks. It also promotes apoptosis that results in the remodelling of the ECM.

MMPs exist in two different types; the membrane bound MMPs and those that are released into the extracellular space. The latter contribute to the majority of MMP species and are reported to be released into the extracellular space as pro enzymes (proMMP) (Spinale, 2002). They are activated in two pathways;

1. Proteinases (proteolytic activation-top pathway)
2. Non proteolytic agent (Chemical activation-bottom pathway)

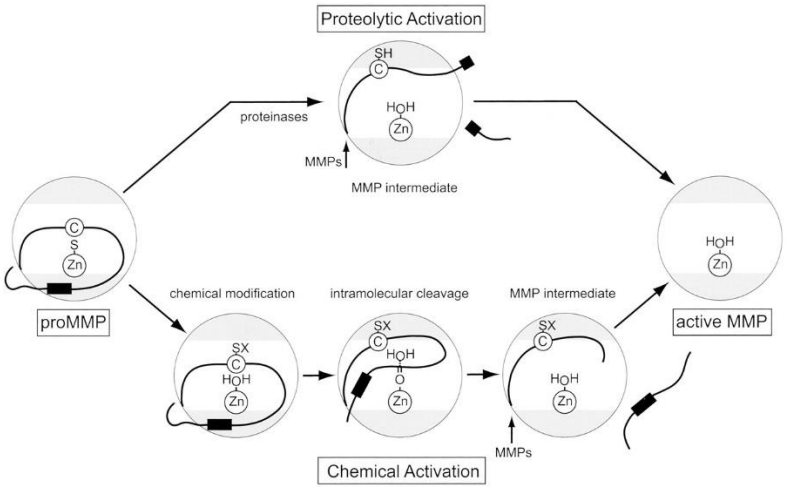


Figure 1.4. *Stepwise activation of the proMMP to active MMP (adapted from Iyer et al., 2006). The circle contains grey catalytic region and an active site in white the region with catalytic site zinc (Zn). The black strip with a rectangle and a cysteine switch (C) represents the propeptide with the rectangle being the bait region (area where cleavage occurs).*

When this bait region is cleaved (dictated by the sequence of each MMP) by cysteine switch sulfhydryl (SX), proMMP is partially activated (Figure 1.4). When the whole propeptide is removed by intermolecular processing, proMMP becomes fully activated (Iyer et al., 2006) and released into the extracellular space where they bind specific ECM proteins (Spinale, 2002).

1.2.8 Advance glycosylation end products (AGEs)

AGEs also referred to as non-enzymatic glycation is the covalent bonding of protein or lipid molecules in the absence of an enzyme. They are reported to cause diastolic dysfunction through the formation of collagen cross-links in the heart (Willemsen *et al.,* 2011) which cause myocardial stiffness (Figure 1.5). This is partly caused by diabetes (Koschinsky *et al*., 1997). Different diabetic complications have been reported to result from protein glycation reactions leading to AGEs suggesting that AGEs are prevalent in diabetes (Singh *et al*., 2014). In another study, Preis *et al.,* 2005 reported that diabetic patients have a high risk of myocardial infarction, vascular dysfunction and diastolic dysfunction. This observation might be because of the uncontrolled amounts of carbohydrates and fats in the blood stream that may undergo reactions resulting in glycation. We could also argue that the high levels of LDL lipoprotein in diabetes bind to the walls of blood vessels blocking them thus causing

heart attacks or stroke leading to heart failure and myocardial remodelling. Hartog *et al.*, 2007 claim that the presence of AGE's is more established in older people and is accelerated by diabetes mellitus that is prevalent in the older generation. AGE's accumulation plays a major role in the pathophysiology of heart failure since it alters the protein structure and properties. They have been reported to cause a significant delay in Calcium ion (Ca^{2+}) uptake that impairs the LV relaxation and contraction causing diastolic dysfunction and systolic dysfunction respectively resulting to HF (Figure 1.5).

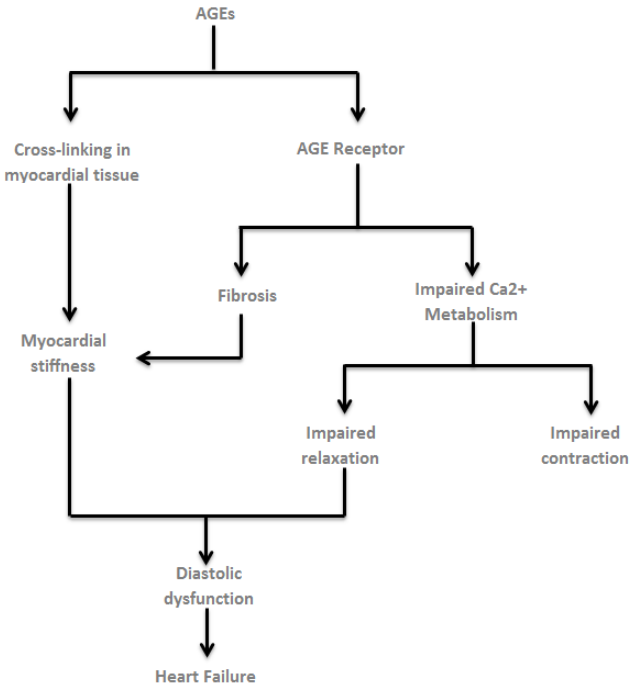


Figure 1.5. *A summary of the pathways which AGE can cause Diastolic heart Failure.*

In another article, Biernacka *et al.*, 2011 suggested that cardiac aging is associated with LV hypertrophy and fibrosis that result in diastolic dysfunction and heart failure with preserved systolic function (HFPEF). They claim that progressive fibrosis (Figure 1.6) is normally an indication of aging in various organs including the heart and kidney. It is a key pathological process in left ventricular hypertrophy (LVH) (Cuspidi *et al.*, 2006). In 1990, Mukherjee *et al.* reported that in an animal model experiment there was increased collagen deposition in aging heart whereas in human subjects there was age-related increase in cardiac fibrosis (Lakatta *et al.*, 2003). This increased fibrosis causes myocardial stiffness and impaired relaxation leading to diastolic dysfunction (Mukherjee *et al.*, 1990).

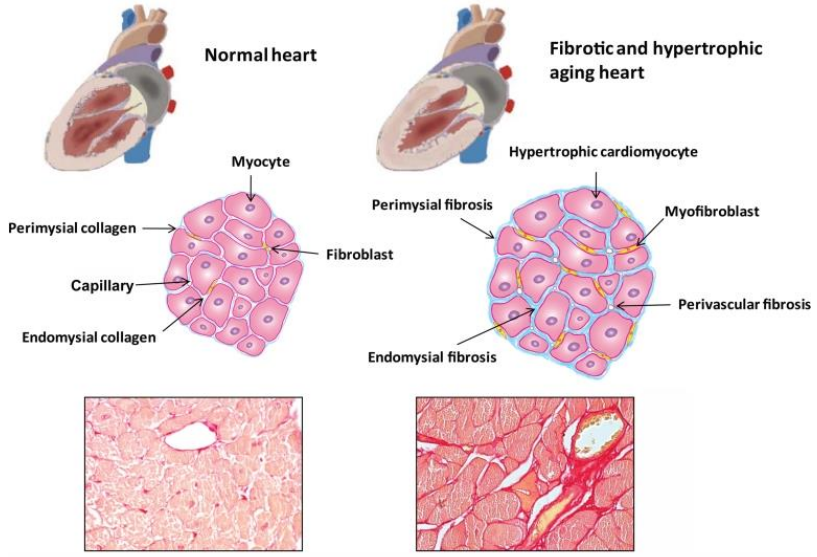


Figure 1.6. *Comparing a normal heart with a fibrotic and hypertrophic aging heart in mice because of AGE and other factors. Adapted from (Biernacka et al., 2011).*

1.3 Proteomics

1.3.1 Background

Proteomics is derived from the word "proteome" which refers to the entire set of proteins expressed by cells, tissues or organisms. Thus, proteomics is the large-scale study of a specific proteome including protein identification, interaction, abundances, their variations/ modifications and cellular processes. Proteins are products of genes that act as machinery of the cells in our bodies. When the genes in our bodies are disrupted, the proteins are also disrupted. Therefore, proteins act as important source of disease/biomarkers discovery (Surinova *et al*., 2011). Proteomics technique varies from top down to bottom up shotgun proteomics. In this research, bottom up shotgun proteomics was used.

1.3.2 Shot gun proteomics

This involves the use of bottom up proteomics techniques for protein identification in complex mixtures with the aid of HPLC and mass spectrometry. Bottom up proteomics is

where by proteins are broken down (digested) into peptides by an enzyme (protease). Their masses/tandem-mass spectra are compared with the masses predicted from a sequence database to identify the peptides. Multiple identifications of these peptides are then assembled into protein identification. This technique provides a better separation and higher sensitivity than top down method where separation is carried out on a protein level.

The main challenge in plasma proteomics is the presence of these HAPs in the sample. The top ten most abundant proteins are reported to occupy up to 90% of the total protein amount while the other proteins exist in a very wide dynamic range, with their concentrations covering more than ten orders of magnitude (Figure 1.7) (Liumbruno *et al.*, 2010). Proteomics technology enables small amounts (femtomole to attomole) of proteins to be detected (Millioni *et al.*, 2011). The LC/MS is inherently limited to only four to five orders of magnitudes and analyses proteins from the most abundant to low abundant. Thus, fractionation or immunodepletion is required to reduce the dynamic range and enable analysis of low abundant proteins.

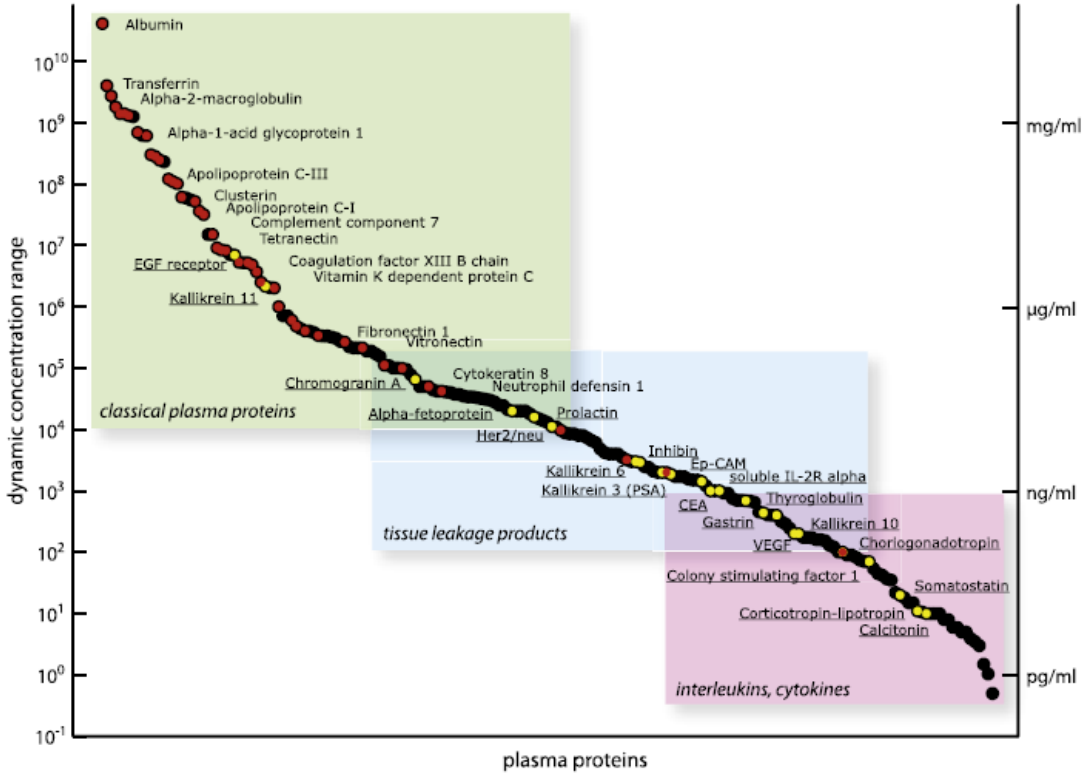


Figure 1.7. *The orders of magnitude of Plasma proteins with Albumin being the most abundant while interleukins and cytokines the low abundant proteins. (Schiess et al., 2009).The red dots represent the proteins identified by HUPO plasma proteome initiative (States et al., 2006) while the yellow dots represent the currently used biomarkers (Anderson et al., 2006).*

Immunodepletion of abundant proteins (Figure 1.8) and fractionation improves the count of low abundance proteins acquired and simplifies analysis with LC/MS. However, this may not always be reliable as proteins of interest may bind to the target proteins (Albumin, IgG) and get depleted. Other variables with depletion include column capacity, buffers and previous use of columns. Omenn *et al.,* 2005 explain that those variations including the use of cheaper dyes like cibaron blue dye may result to protein binding to the dye or albumin, thus limiting the low abundance protein count.

1.3.2.1 The protein dynamic range

The number of proteins occurring in a sample can vary significantly. In some samples, it has been shown that proteins per cell vary from as low as 50 to 1,000,000 for others. In human plasma, the difference between the low abundance proteins (LAPs) and high abundance proteins (HAPs) varies from 10 to 12 orders of magnitude where by, one order of magnitude is 10 to the first. The term dynamic range is used to describe the ratio between the smallest and largest value of proteins of a changeable quantity.

The HAPs are a big challenge in plasma proteomics. They mask the appearance of LAPs due to their dominance in the sample matrix and thus, complicate the process of biomarker discovery that have been reported to be leaked into blood in minute quantities (pg/mL). Thus, these HAPs are normally depleted (section 1.3.3.1) in the sample preparation step prior to mass spectrometry analysis.

1.3.3 Sample preparation

1.3.3.1 MARS HU-14 column depletion

MARS column contains antibodies that specifically bind 14 high abundance proteins (Table 1.2) due to its high affinity nature. When a sample is injected with the mobile phase (buffer A), these high abundance proteins bind to their relative antibodies in the stationary phase which retain them in the column allowing a steady flow of the low abundant proteins through for collection. Buffer B disrupts the affinity of the bound proteins (high abundance) with the antibodies allowing them to flow through the column for the second collection.

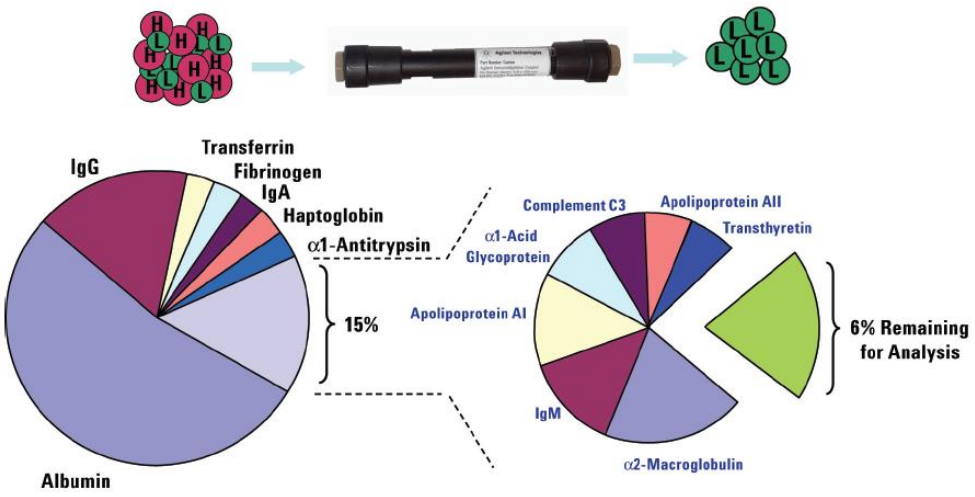


Figure 1.8. *Schematic diagram of the MARS column and the top 14 HAPs it depletes. Figure adapted from Mrozinski et al., 2008.*

1.3.3.2 Reverse phase High Pressure Liquid Chromatography (rpHPLC)

HPLC is one of the most powerful tools used in analytical chemistry today. It is used in separation, identification and quantification of compounds in samples dissolved in liquid based on their polarity. The term "Reverse phase" is used to describe the non-polar stationary phase used in HPLC as opposed to polar. It consists of the mobile phase and the stationary phase. The mobile phase consists of a mixture of solvents (normally 2 polar solvents). In reversed-phase HPLC, A is typically aqueous whilst B is an organic solvent. The proportions of A and B decide the elution of a molecule and can be held static (isocratic) or increasing proportions of B over the time of analysis (gradient). Modifiers such as TFA or formic acid can be added to the mobile phase and can have positive effects on both chromatographic separation and mass spectrometric detection. The mobile phase, stationary phase and gradient will all depend on the separation required and analyte chemistry.

The stationary phase (column) consists of various hydrophobic/ non-polar containing particles (silica). Hydro-carbons chains are attached to the surface of the particle. Most commonly used columns are the C4 (large proteins) and C18 (small peptides) columns (Table 1.5). Column lengths vary from 3-10 cm with an internal diameter ranging from 0.05-4.6 mm. When the mobile phase is pumped into the stationary phase, there is competition between the mobile phase and the silanol groups in the stationary phase (column). Some

compounds in the sample are strongly attracted to the mobile phase thus causing a continuous and steady flow through the column. Other compounds create a strong attraction with the nonpolar molecules causing them to move significantly slower thus leading to a high retention time (Figure 1.9). A detector then detects these compounds.

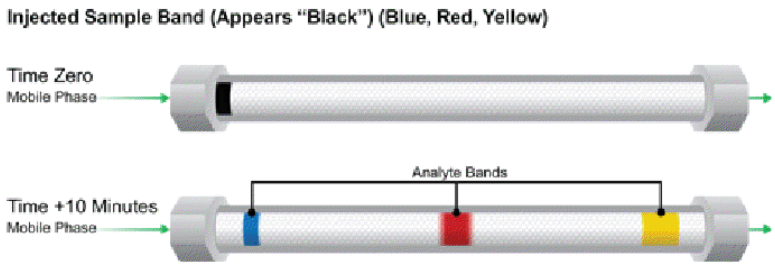


Figure 1.9. *A diagram showing non polar HPLC columns and peptide elution at different intervals. Accessed on 31/10/2013 from* [*http://www.waters.com/waters/en_GB/How-Does-High-Performance-Liquid-Chromatography-Work%3F/nav.htm?cid=10049055&locale=en_GB*](http://www.waters.com/waters/en_GB/How-Does-High-Performance-Liquid-Chromatography-Work%3F/nav.htm?cid=10049055&locale=en_GB)

Table 1.5. *Different HPLC columns used for proteins and peptides. The C18 column is used for small peptides whilst the C4 is typically used for large proteins.*

Stationary phase		
	C 18 columns	
	C 8 columns	
Decreasing hydrophobicity	C 4 columns	Increasing retention of non polar analytes
	Cyano columns	
	Phenyl columns	
	Amino columns	

The detector identifies the compounds eluted from the column. These can be detected in several ways depending on the characteristics and properties of the compound. Some of these ways include the ultra violet (UV) absorbance detector, florescence detector or evaporative light scattering detector (ELSD). Plasma is eluted from the column as an organic compound and absorbed by the UV detector. A chromatogram is a representation of the compounds that have been separated in the column. They are plotted on the computer in form of peaks, each peak forming as each compound is eluted.

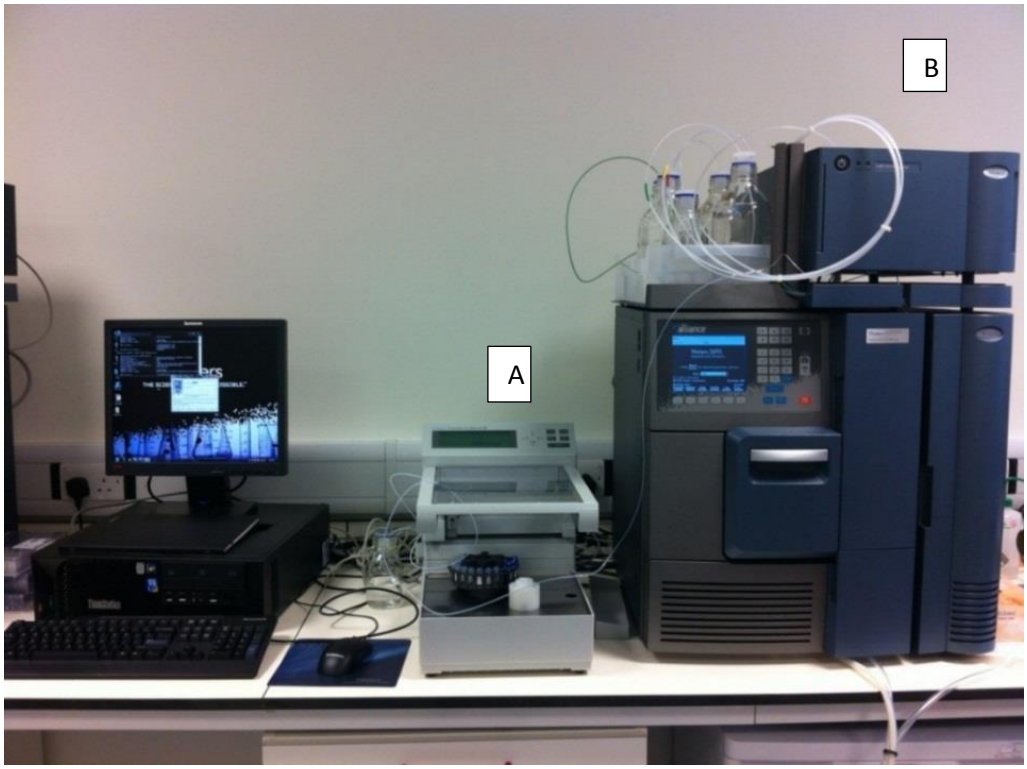


Figure 1.10. *A set-up of the reverse phase HPLC (B) connected with a fractionator (A) for collecting various fractions of samples from the column. The HPLC contains a UV light that detects all the peptides that are being separated in the sample.*

The flow rate, time and amount settings between each collection are set on the computer at the workstation prior to sample injection. The main aim of fractionating the sample was maximise the distribution of abundant proteins in the chromatographic run. According to Millioni *et al.,* 2011, plasma proteome is one of the most complex derived human proteome due to the presence of these high abundant proteins. The current shotgun proteomic technologies have enabled the detection of extremely low amounts of proteins. However, there is a big challenge in detecting and quantifying the proteins present at two to three orders of magnitude lower than the abundant ones. Fractionating the plasma ensures distribution of peptides derived from these high abundant proteins when analysed by mass spectrometry. With complex mixtures, particularly clinical samples that exhibit significant dynamic range, there is insufficient peak capacity to analyse all the analytes in a samples. Consequently, fractionation is employed to effectively increase peak capacity. Thus, maximise the resolution of individual analytes for analysis.

1.3.3.3 Lipid Removal Agent (LRA)

LRA contains calcium silicate hydrate which selectively removes lipids, endotoxins and other bio organic molecules (Gordon *et al.,* 2010). Lipids are hydrocarbons that occur naturally and are poorly soluble in polar solvents like water. They include triglycerides, steroids, phospholipids, glycolipids, lipoproteins and waxes. They serve as structural components in cellular membranes, energy storage and membrane development. However, they interfere with analytical assays and biopharmaceutical production. Phospholipids in particular are reported to cause ion suppression during pre-clinical and clinical studies during pharmaceutical development. Lipoproteins for example apolipoproteins are high abundance proteins which when depleted enable the low abundant proteins to be analysed. The binding affinity residue was used to bind all the HAPs. In this project, LRA was used to deplete high abundance proteins on human plasma samples prior to MARS column depletion and mass spectrometry analysis.

1.3.3.4 Hydroxyapatite (HA)($Ca_5(PO_4)_3OH)_2$

This is a form of calcium phosphate used in chromatographic separation of biomolecules and in bone tissue engineering (Zhou *et al.,* 2011). It has unique separation properties and unparalleled selectivity and resolution. It separates proteins by electrophoretic and other chromatographic techniques. Unlike other chromatography adsorbents, HA is both the ligand and the support matrix. There are two types of HA, type I and type II. Type I is reported to have a higher protein binding capacity and better capacity for acidic proteins while type two has a lower protein binding capacity but with a better resolution of nucleic acid and certain proteins. Type II is also reported to have a very low affinity for albumin and is suitable for purifications of many species and classes of immunoglobulins. In this project, type 1 was used due to its high affinity for proteins and was referred to as M3 for research purposes.

The HA contains positively charged calcium binding groups and negatively charged phosphate groups. Gagnon p *et al*., 2006 reports that when the proteins are bound to HA the amino groups interact with the HA through the cation exchange or calcium affinity. The cation exchange occurs when the amino groups are repelled by the calcium sites but interact ionically with the negatively charged phosphates. These two effects are essential for the

binding of the amino groups. The interactions are later disrupted by the introduction of neutral salts such as NaCl and NaH_2PO_4 (section 3.2.3). In addition, increasing the pH weakens the cation exchange interaction which results in the elution of the proteins. Conversely, the calcium affinity occurs via interaction of the proteins or nucleic acids with the carboxyl or phosphoryl groups. The phosphate groups simultaneously repel these groups. These interactions, also called metal affinity interactions are reported to be 15 and 60 times stronger than ionic interactions (Gagnon P *et al.,* 2006) and unlike cation exchange, they cannot be disrupted by increasing ionic strength with chloride. In fact, increasing the ionic strength increases the species binding through calcium affinity due to ionic shielding of charge repulsion from the phosphate sites. Phosphate salts (NaH2PO4) dissociated these interactions.

1.3.3.5 Ammonium sulphate precipitation

Ammonium sulphate precipitation is a method used in proteomics to purify proteins by changing their solubility. The solubility of proteins is determined by ionic strength of a solution, therefore salt concentration. When the ionic strength is increased, the solubility of proteins is also increased, a process referred to as salting in. However, as the ionic strength increases further, the solubility of proteins begin to decrease. This is because at sufficient high salt concentration the proteins will be almost completely precipitated from the solution, a process referred to as salting out. Since proteins differ significantly in their solubility, salting out is important because it ensures maximum solubility of proteins, thus maximum protein purification. The aim of this experiment was to find ammonium sulphate concentration that would precipitate maximum proportion of desired proteins (LAPs) and leave the rest (HAPs) in solution. The precipitate protein of interest was recovered by centrifugation at 2,500 g.

1.3.4 Sample analysis

Due to the complex nature of plasma samples, they are subjected to depletion and fractionation prior to mass spectrometry analysis. This not only removes the HAPs but also enables us to detect and quantify the LAPs that are secreted in plasma in very small

concentration. Depletion is performed using immunodepletion columns that removes the most abundant proteins from the sample. Fractionation on the other hand reduces the complexity of plasma by distribution of the high abundant proteins in the samples in different fractions

1.3.4.1 NanoACQUITY UPLC

Nanoflow (typically 300 nL/min) UPLC achieves the high level of chromatographic separation found with UPLC but scaled down to the nL/min so that desolvation can occur more efficiently in the electrospray source. Using nanoflow ensures that sensitivity is achieved through maximally obtained ionisation of desolvated molecular species.

1.3.4.2 Mass spectrometry (MS)

Mass spectrometry (Figure 1.11) is an analytical tool used to measure the mass to charge (m/z) of an ion. Molecules within samples can be introduced into a mass spectrometer as a gas, liquid or solid but requires transitioning to the gas phase for ionisation and transmitting within a vacuum to take place. Once molecules are ionised they can be transmitted through a series of analysers which using electromagnetic fields are able to control the transmission of ions through the MS to the detector. When it hits the detection plate, it produces a tiny electrical current that is amplified. The more ions of a specific mass to charge (m/z) ratio, the greater the current and the higher the peak.

The three main parts of a mass spectrometer are; the ionisation source, the analyser(s) and the detector.

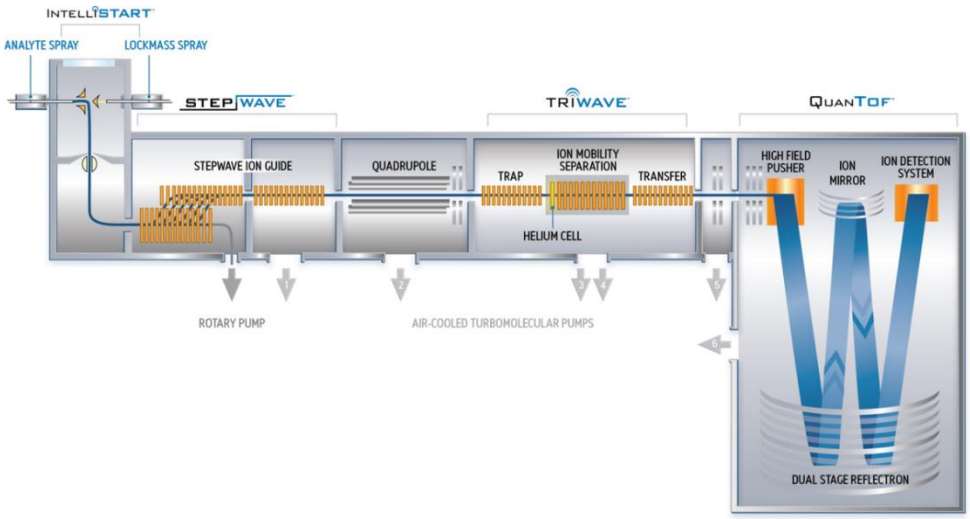


Figure 1.11. *Diagrammatic representation of the G2S Synapt mass spectrometry. It consists of the stepwave which eliminates all the neutral ions from the sample, the triwave with ion mobility separation and quanToF. Adapted from Waters Ltd stock pictures.*

1.3.4.3 The ionisation source

The ionisation source of the MS is where the sample is introduced into the MS and is dependent on the ionisation method used. Several methods of ionisations are available which include atmospheric pressure chemical ionisation (APCI), chemical ionisation (CI) electron impact (EI), electrospray ionisation (ESI), fast atom bombardment (FAB), field desorption (FD/FI), matrix assisted laser desorption ionisation (MALDI) and thermospray ionisation (TSP). However, the most commonly used methods in research are ESI and MALDI because of their soft ionisation technique, which have very little fragmentation of the ions, thus enabling ion stability.

a) Electrospray Ionisation (ESI)

Electrospray ionisation is a technique used in MS to produce ions using an electrospray. The electrospray is made up of stainless steel or quartz silica capillary where the sample is introduced into the MS. It forms ions by using electrical energy to convert solutions into gaseous phase prior to mass spectrometric analysis. This occurs in three steps; (1) distribution of fine spray of charge droplets before (2) solvent evaporation followed by (3) ion injection from the highly charged droplets (Ho *et al.*, 2003) maintained at a voltage of between 2.5-6.0

kV. This creates a mist of highly charged droplets that have the same polarity as the capillary voltage. A nebulising gas (nitrogen) sheaths the emitter and enables enhanced desolvation. The charged droplets at the tip of the electrospray are accelerated down a pressure and potential gradient towards the analyser region of MS. This is aided by the elevated temperature at the ESI source and the stream of nitrogen drying gas that condenses the charged droplets by evaporation (Figure 1.12) of the solvent. This results in an increase in surface charge density and reduction of the droplet radius eventually leading to the ejection of ions from the surface of the droplets into gaseous phase. The emitted ions are then sampled by sampling skimmer cone (Figure 1.13) then accelerated into the mass analyser for molecular mass and ion intensity analysis

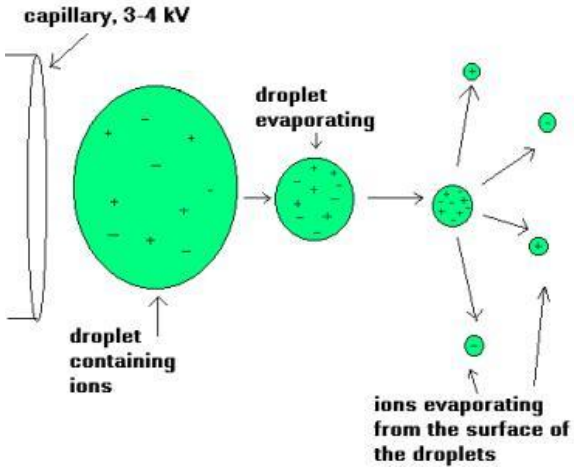


Figure 1.12. *Electrospray ionisation process showing a continuous stream of sample introduction into the MS. The samples are vaporised and ionised as soon as it gets to the tip of the electrospray (Diagram adapted from Ashcroft A., 2016).*

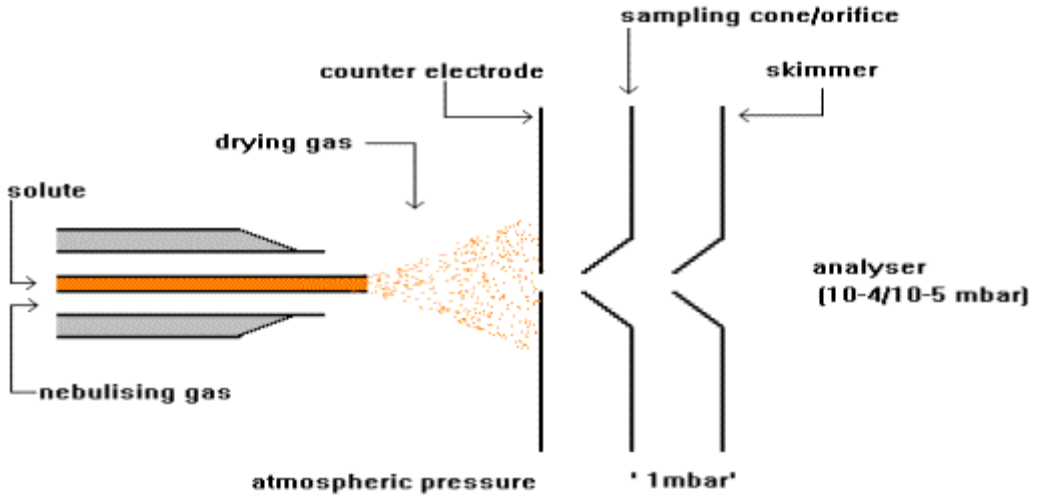


Figure 1.13. *Diagram showing the electrospray ionisation source. Charged ions travel from the metal capillary assisted by the nebulising gas through the counter electrode, orifice and skimmer to the mass spectrometer (Adapted from Ashcroft A., 2016).*

b) Matrix assisted laser desorption ionisation (MALDI)

Matrix assisted laser desorption ionisation is a soft ionisation technique in MS based on the bombardment of sample molecules (DNA, peptides, proteins and sugars) using a laser light resulting to ionisation. During samples preparation, a highly absorbing matrix compound is mixed with the sample. Low concentrations of sample to matrix are normally used for reliable and consistent results. The mixed sample is applied to a metal plate where a pulsed laser irradiates it causing ablation and desorption of the sample and matrix material. The analyte molecules are then ionised by being protonated or deprotonated by the ablated gases prior to analysis in the MS (Ashcroft A., 2016).

1.3.4.4 The analyser

The analyser is used to separate the ions formed in the ionisation source of the MS based on their mass to charge (m/z) ratios. Several types of analysers exist which include quadrupoles, time-of-flight (ToF) analysers, orbitraps and quadrupole ion traps. However, quadrupoles and time-of-flight (ToF) are the most commonly used analysers on MS. In orbitrap, packets of

ions are injected and orbit around a central spindle shaped electrode in harmonic oscillations. These ions are later detected using ion image currents and converted to mass spectra using Fourier Transform (FT) techniques (Makarov, 2000).

a) Quadrupole

A quadrupole mass analyser consist of four rods parallel to each other with opposite electric fields, to either radio frequency (RF) or direct current (DC) applied, a concept introduced in 1953 by Paul and Steinwedel. The quadrupole functions as a mass filter. Under a fixed electric field, the opposing currents separate the masses, restricting only ions of a specific mass to pass through to the detector. To improve sensitivity and specificity, three quadrupoles are assembled one after the other, also called triple quadrupole. In a triple quadrupole mass analyser, the first and last mass analysers are used to filter and transfer ions, while the middle mass analysers which has RF only contains the collision cell which fragments the ions. The middle mass analyser contains an inert gas (typically argon) which causes dissociation of ions. These ions are then taken to the third mass analyser that transfers them to the detector to be measured.

b) Travelling Wave Ion Mobility Separation (TWIMS)

The Triwave consists of the trap, ion mobility separation (IMS) and transfers ion guides. Each ion guide is a series of plates with alternating voltages on them that enable ions to be transmitted and controlled though them. Ion mobility introduces an additional stage of separation with no time penalty. It leads to an increase of 54-69% in peptide identification and 50-60% protein identification. When ions are generated from ESI source, they pass through quadrupole mass analyser/filter where are filtered and guided through to the TWIMS section consisting of the trap, IMS and transfer (Figure 1.14). The trap accumulates the ion in packets with the aid of a potential barrier applied at the gate and releases them to ion mobility for separation. When the previous batch of ions undergo separation, the gates open again for the next set of ions to move in for separation. IMS separates ions based on their size, shape and charge that also determine their drift time before transferring them to Q-ToF for high-resolution mass analysis. IMS uses a series of transient voltages to propel the ions through a

drift tube filled with gas (helium). In the drift tube there is also nitrogen gas that flows from the opposite direction (towards the source) which aids the separation of the ions (Giles *et al.*, 2011). Smaller molecules travel faster than the larger ones; thus, take a shorter time to reach the end of the mobility separator.

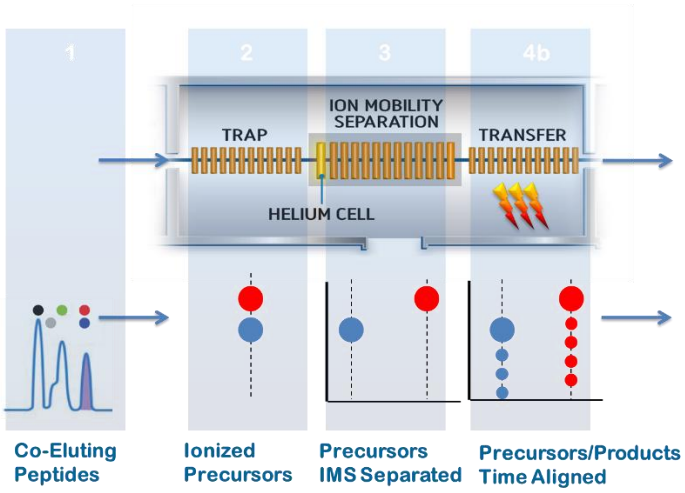


***Figure 1.14**. Diagram showing ion mobility enhanced UPLC/MS^E. The trap holds ions aided by an electric potential gate and releases them to ion mobility separation in packets where they are separated based on their size, shape and charge. This separation is maintained at the transfer region, which aligns the precursors and to the product ion. Adapted from Waters Ltd stock pictures.*

This mobility of ions from the IMS is maintained by the transfer region of TWIMS. At the transfer, the ions are fragmented. Here the product ions are assigned to precursor ions by aligning their retention times (RT) and transferred to ToF.

c) Time-of-flight (ToF)

Stephens first described this method in 1946 before Wiley and McLaren described the linear ToF in 1955 resulting in commercialisation of this instrument (Wiley *et al.*, 1955). Time of flight (ToF) mass analyser is a method of MS that determines or measures the ions m/z ratio via drift time. Synapt G-2S instruments uses orthogonal accelerated TOF (oaToF) (section 5.3.1). This means the line of ions are transmitted and then switched in direction by about 90 degrees. In order to do this the ions are trapped in the ToF region and then accelerated down the flight tube in pulses by an electric field of known strength to the detector. This acceleration results in ions acquiring equal kinetic energy as other ions with the same charge.

When a set of ions of different masses are accelerated and given the same kinetic energy before they are transmitted to the flight tube they all acquire a different velocity. Thus, time is a function of the square root of the m/z value (inversely proportional to the square root of their mass) (de Hoffmann *et al.,* 2007). Therefore, ions with a smaller m/z travel faster than ions with a larger m/z while maintaining a linear relationship over several orders of magnitude (Cotter, 1999).

1.3.4.5 Reflectron

Reflectron is a type of ToF MS consisting of ion source, ion mirror, field free region, ion mirror and ion detector. It uses a static or time dependent electric field in the ion mirror to reverse the direction of an ion in the mass spectrometer (de Hoffmann *et al.,* 2007). Several reflectrons exist but the most common is the ion mirror that provides higher resolution. It consists of a series of evenly spaced electrodes onto which a single, linear, electric field is applied. Reflectrons can be single stage (with one reflectron) or dual stage (Figure 1.15). The dual stage reflectron is used to increase the resolution of your ions before analysis with the detector. The high-field pusher region condenses the ions before releasing them to the reflectron to ensure equal velocity of ions is maintained (de Hoffmann *et al.,* 2007). The higher energy ions travel more and vice versa to the ion mirror where they are reflected to the ion detection system.

Figure 1.15. *Diagram showing a dual stage reflectron consisting of the high field pusher region, ion mirror and ion detector. The pusher region ensures that all the ions have the same velocity before they are transmitted through the reflectron to the detector. Adapted from Waters Ltd stock pictures.*

1.3.5 Tandem mass spectrometry

Tandem mass spectrometry is used to produce structural information of a compound through fragmentation and involves two stages of mass analysis. The first mass analyser (quadrupole) is used to select a precursor peptide for MS/MS that undergoes controlled dissociation while the second mass analyser (ToF) is used to analyse the product ions generated from MS/MS. The product ions can be used to reveal the amino acid sequence of a peptide or structure of the species under investigation.

In tandem MS, ions of interest can be measured or scanned through several ways. This project focused on product ions produced with inert collision gas, thus the following scans were applied:

- **Product or daughter ion scan-**The first mass analyser is static selecting a specific precursor ion for MS/MS while the second mass analyser scans through the product or daughter ions.
- **Precursor ion scan-** The second mass analyser is static to a specific product ion m/z while the first mass analysers scans through all the product ions that generate product or daughter ions.
- **Selected reaction monitoring-**Both the first and second mass analysers are static. The first mass analysers selects the precursor ion m/z while the second analyser selects a specific product or daughter ion characteristic of the analyte of interest.

To acquire product ions from precursor ions, fragmentation has to occur. Several techniques of fragmentation have been reported including collision induced dissociation (CID) (McLafferty *et a*l., 1967), electron capture dissociation (ECD) (Zubarev *et al.,* 1998) and electron transfer dissociation (ETD) (Syka *et al*., 2004). In this project, CID was used where a precursor ion is subjected to multiple collisions with an inert gas molecule (argon). This converts the kinetic energy to translate to an internal vibrational energy which aids the fragmentation of ions (by breaking chemical bonds) at the lowest energy cleavage sites primarily at amide bonds forming b and y ions (Hoffmann *et al.,* 2006).

1.3.5.1.1 Data independent acquisition (DDA)

Data dependant acquisition (DDA) was used to select individual peptides molecular ions and subject them to CID to collect sequence information. In DDA, all the mass analysis is performed on one analyser where the ions are trapped before undergoing dissociation. It is performed on triple quadrupole, orbitrap, Q-ToF (Figure 1.16) and ToF/ToF instruments. Therefore, a scan was undertaken and a number of precursors selected for sequential MS/MS. Each scan can have up to 9 precursors selected. The most intense peptides were switched upon and identified (these may not be the peptides of interest). In addition, other peptides of interest may be eluting in MS/MS mode but are not fragmented or identified. This makes quantification difficult since we cannot use survey data as we do not sample peak effectively and no data is being collected during MS/MS mode.

1.3.5.1.2 Data independent acquisition (DIA)

The data independent acquisition method fragments all ions within a selected m/z range before analysis on the second mass analyser. This process has been explained in section 1.3.5. The two most common types of DIA methods are MS^E (Figure 1.16) and Sequential window acquisition of all theoretical mass spectra (SWATH) (Rosenberger *et al*., 2014). MS^E uses high and low energy scan for fragmentation (Daly *et al*., 2014) while SWATH utilises smaller m/z windows.

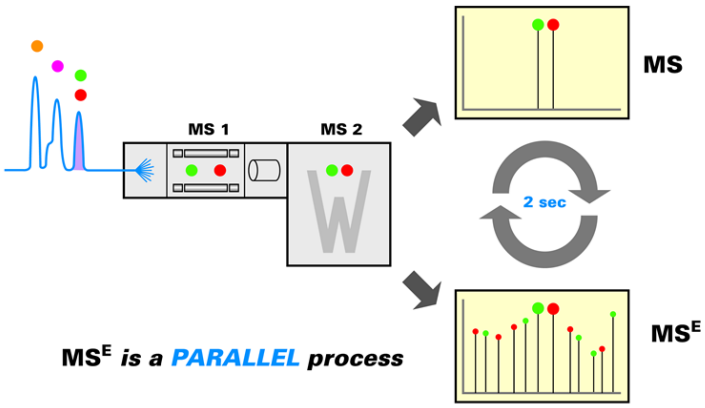


Figure 1.16. *Schematic diagram comparing MS/MS and MS^E. MS^E uses alternating low-energy collision-induced dissociation and high-energy collision-induced dissociation to obtain the precursor and product ion accurate mass. Adapted from Waters Ltd stock pictures.*

1.3.5.2 Quantitative proteomics

Accurate quantification of proteins and peptides in complex biological samples is one of the biggest challenges in proteomics (Wasinger *et al.,* 2013). Quantitation involves determination of the amount of proteins, absolute or relative in grams or moles in a particular sample. Absolute quantification involves determining the exact concentration or amount (ng/mL) of a protein in a sample. Relative quantification on the other hand compares the level of a particular protein in different samples with results being expressed as a relative fold change of protein abundance (Elliot *et al.,* 2009).

1.3.5.3 Labelled approaches

This involves labelling proteins of interest to allow identification of peptide variants in MS spectra or in the MS/MS spectra (Kito *et al*., 2008). This method can analyse more than two or more protein samples at once in the same run. Peptides variants are differentiated by tagged labels. In relative quantification, samples are labelled with different stable isotopes, combined and subjected to MS. Their peak intensity ratio between light and heavy peptides is then measured to determine relative change in protein abundance (Chapter 5). Other techniques of labelling include proteolytic, chemical, isobaric and metabolic.

1.3.5.4 Label free approaches

Unlike the labelled approaches, this method does not use tagged labels, thus label free. Separate quantification runs are performed for each sample which can be either absolute or relative. This method is based on spectral counting or precursor signal intensity. In spectral counting, the numbers of identified spectra for a given peptide in a biological sample are counted and measured based on their mass to charge ratio then the results integrated and quantified. On the other hand, precursor signal intensity is important when applied to high precision mass spectra that utilises the high-resolution power to extract peptide signals on the MS1 level. This as a result separates quantification from the identification process. The high-resolution power facilitates the extraction of peptide signals on the MS1 level and thus uncouples the quantification from the identification process.

1.1.1 Normalisation of data

If sampling of all the samples was identically executed and the instruments used worked exactly the same way for all the samples then the results could be compared directly. However, this is never the case due to variations in sampling and instrument operation resulting in variations in the profiles prompting some form of normalisation. Variations can result from the following;

i) Differences in protein abundances for the groups we are comparing
ii) Differences in protein abundances due to external factors during sample preparation e.g. temperature
iii) Differences due to sample contamination during sample preparation or measurement (weighing scale) or instrument noise
iv) Differences in samples analysis in the Mass spectrometry e.g. the chromatography, Ms spectrum acquisition and MS/MS spectrum acquisition

For normalisation to occur, some criteria have to be met. These include:

i) Samples to be compared have to have the same proteins
ii) Most proteins should have a 1:1 ratio and only a small amount differ in abundance level
iii) The outliers can easily be singled out since the majority of the protein abundances are unchanged

1.3.6 Mass spectrometry data processing

1.3.6.1 ProteinLynx Global Server (PLGS)

The raw data was processed with ProteinLynx Global Server (PLGS) 3.02. PLGS is a fully integrated mass informatics platform for quantitative and qualitative proteomics research. Its primary aim is to identify proteins that are within a sample and its secondary aim is to quantify the identified proteins. In sample preparation, protein mixtures are broken into constituent peptides using proteolytic enzymes (e.g. trypsin). This protease has high cleavage specificity at either C-terminal to Lysine (K) or Arginine (R) except when either is followed by a proline. When searching peptide fragmentation spectra alongside sequence databases,

potentially matching peptide sequences can be required to conform to tryptic specificity (Mann *et al.*, 2004).

PLGS has two main sections, Identity^{E} and Expression^{E}. The former identified and quantified the total amount of proteins loaded onto the column for each protein that was identified while the latter was used to compare the levels of specific proteins between control and disease (heart failure) samples. The user's role is limited to choosing the appropriate database, the right threshold and a suitable false discovery rate (FDR). Identity^{E} uses a unique algorithm to process the raw data files from mass spectrometry. It does this by using different properties of ions that include retention time, precursor/product ions intensity and accurate mass. This is followed by generation of a list of all precursor and product ions by PLGS that are then de-convoluted by a different algorithm called Apex3D. Apex3D is responsible for creating the exact mass and retention time (EMRT) table in low and elevated energy. The table that contains precursor and product ion masses for each peptide is searched against the protein databank (Uniprot) which contains alcohol dehydrogenase for label free quantitation. A decoy database is also created by PLGS where the amino acid sequences are reversed or randomised and linked to the original database. This is used to calculate the FDR. MS^{E} data acquisition was used for protein quantification by the collection of appropriate data points to quantify peak ion intensities. This was implemented by selection of three most intense tryptic peptides (Hi3 method) of a protein as a measure of its abundance (Silva *et al.*, 2006).

PLGS carries out a number of processes (Figure 1.17, Table 1.6) onto the data that mainly include two things;

1.) Data processing (prepares data for database searching)
2.) Workflow (incorporates a series of database searches that integrate a number of rules associated with peptides)

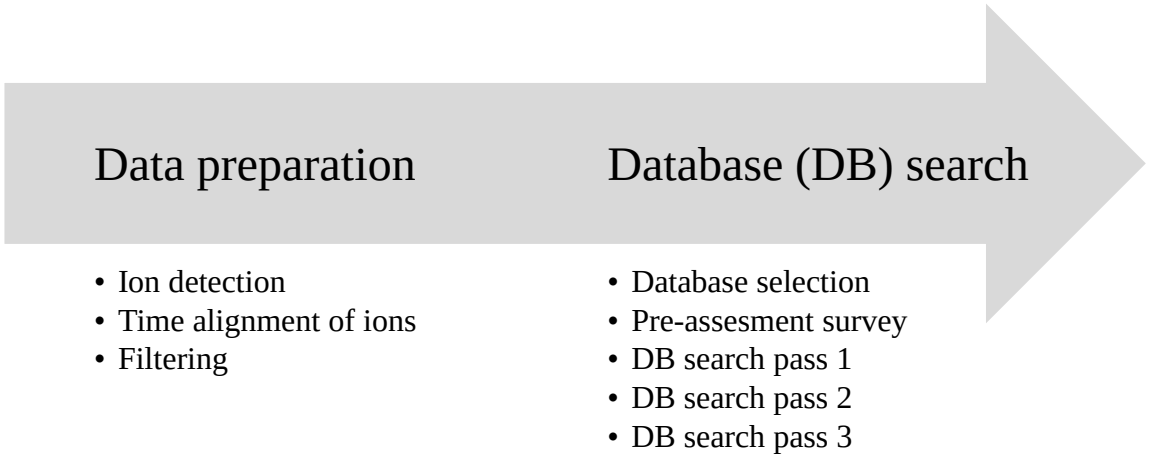


Figure 1.17. *Identity^{E} overview in PLGS. It consists of data preparation phase followed by database search phase that incorporates a number of steps as shown above before generating data.*

Table 1.6. *A detailed summary of identity^{E} overview in PLGS following from Figure 1.17.*

Data preparation	
Ion detection	RT, m/z and intensity are collected
Time alignment	In this phase low and high exact mass and retention time are aligned into product and precursor tables.
Filtering	All low energy precursors <750 Da and elevated product ions below 350 Da are removed
Database (DB) search	
Database selection	A reverse or random DB is created, decoy DB is merged with original DB
Pre-assessment survey	A pre-search of time aligned precursor and product ion tables is conducted.
DB search pass 1	Only the precursor and product ions tables with completely cleaved tryptic peptides are considered
DB search pass 2	PTMs and miss cleavages from pass 1 are considered
DB search pass 3	Searching with no limitations on the intensity of product ions, ranking and scoring until reaching FDR

1.3.7 Multiple Reaction Monitoring (MRM) for validation

Multiple reaction monitoring using mass spectrometry is a very sensitive and selective method for proteins/peptide identification. MRM advantages over the conventional methods (enzyme linked immunosorbent assays-ELISA) of validation have been highlighted in chapter 5.1.

Hypothesis: There is a clinical paucity of biomarkers for heart failure with preserved ejection fraction that provide sufficient specificity and sensitivity. We are testing whether plasma proteomics can reveal novel biomarkers of heart failure with preserved ejection fraction using novel sample preparation technologies.

Aims and objectives

- Create a novel method for discovery proteomics using M3 comparing it with other fractionation methods including ammonium sulphate precipitation, LRA and mRP-C18 column.
- Test M3/ LRA in a clinical cohort for HF.
- Test targeted methods for discovery of biomarkers in HFPEF and HFREF.

Chapter Two

MATERIALS AND METHOD

2 Materials

M3 was obtained from Biorad (Hertfordshire, UK). MARS14, 4.6 x 100 mm column and MARS buffers A and B were obtained from Agilent Technologies (Stockport, UK). Amicon ultracentrifuge devices (0.5 mL 3K MWCO membrane) were obtained from Millipore (Hertfordshire, UK). Spin X centrifuge filters 0.22 µm 2.0 mL spin tubes were obtained from Sigma-Aldrich (Poole, UK). An in-house BCA assay was performed using tartaric acid and cupric sulphate that were obtained from Fisons Scientific Apparatus (Loughborough, UK), while sodium carbonate anhydrous was obtained from Fisher Scientific (Loughborough, UK). Solid phase extraction (SPE) cartridges Empore C18-SD 4 mm/mL were obtained from Bioanalytical Technologies (Naperville, USA). Plasma samples were obtained from healthy donors. The national research ethics committee approved the protocol. The studies complied with the declaration of Helsinki. All other reagents were obtained from Waters (Milford. MA, USA) and Sigma-Aldrich and were of HPLC grade.

2.1 Methods

2.1.1 Sample preparation

2.1.1.1 Plasma samples

Human blood samples were obtained in EDTA blood sample tubes from healthy individuals following informed consent. The plasma was separated from blood via centrifugation at 2,500 g at 4°C for 20 mins using a refrigerated centrifuge and stored at -80°C until use.

2.1.1.2 Hydrophilic Strong Anion Exchange (SAX) chromatography with HPLC

Material and solutions: Trifluoro acetic acid (TFA), Formic acid (FA), water (HPLC grade), Tris hydrochloric acid (HCl) and sodium chloride (NaCl) were used. C18 AS24 column (2 x 250 mm) was obtained from phenomenex (Chesire, UK), was also used. Solvent A contained

20 mM Tris HCl whereas solvent B contained 20 mM Tris HCl in 1 M NaCl both at pH 8. At time zero solvent A was 100% and solvent B 0%. The column was first primed with solvent A for 30 mins. The UV absorbance was set at 280 nm and delay volume at 1.75 mL. A blank run was performed with the gradient for 50 mins at a flow rate of 0.25 mL/min and pressure of about 1800-2000 psi. Pre-digested plasma was then injected onto the column and 12 fractions (1 mL) collected. 0.5 mL of each fraction was taken out and concatenated to make 4 fractions. These fractions were freeze dried overnight, reconstituted in 1 mL 1% TFA and cleaned-up using solid phase extraction (SPE) (section 2.1.1.6). The concatenated fractions (2 mL) from SPE were speed vacuumed; freeze dried and reconstituted in 0.1% FA. ADH (1:1 ratio) was added prior to mass spectrometry analysis.

Solvents A and B (mobile phase) were prepared as below:

Solvent A

20 mM Tris hydrochloric acid (HCl) was prepared by adding 2.4 g of Tris buffer in 1000 mL of water (H_2O) then titrated to pH 8 with HCl.

Solvent B

20 mM Tris HCl at pH 8.0. 1 M NaCl was prepared by adding 58.44 g of NaCl and 2.4 g of Tris and titrated to pH 8 with HCl in 1000 mL of H_2O.

2.1.1.3 Immunodepletion

MARS14 depletion. The MARS14 immunodepletion spin columns contains antibodies that deplete the following 14 human plasma High Abundance Proteins (HAPs): albumin, antitrypsin, alpha2-macroglobulin, alpha1-acid glycoprotein, apolipoprotein AI, apolipoprotein AII, complement C3 fibrinogen, haptoglobin, IgG, IgA, IgM, transferrin, and transthyretin. The samples were prepared and the column was operated in accordance with the manufacturers recommendations. Two hundred µL of plasma were filtered through 0.22 µm spin X centrifuge tube filter and diluted 1 in 4 with buffer A as supplied. The MARS14

depletion column was used on a Waters 1525 binary HPLC instrument with 5 consecutive 40 µL injections of plasma sample (200 µL) to deplete the 14 highest abundant proteins in plasma. The HPLC instrument was equilibrated prior to sample injection. Flow-through fractions (depleted plasma) were collected in clean low bind Eppendorf microcentrifuge tubes and the bound proteins were eluted by buffer B to regenerate the column.

2.1.1.4 Bicinchoninic acid (BCA) protein micro assay

Protein concentration of each sample (aliquot A and the sequential fractions of aliquot B) was determined using an in-house BCA assay measured at 450 nm using a UV spectrophotometer (Thermo Electron Corporation, UK). The BCA assay was prepared by adding 0.8 g sodium carbonate monohydrate/anhydrous and 0.16 g tartic acid to 10 mL of water at pH 11.25 with NaOH for reagent A. Reagent B was prepared by adding 1 g BCA to 25 mL of water while reagent C was 0.4 g $CuSO_4.5H_2O$ in 10 mL of water. The working solution was always made fresh by mixing 50 µL of reagent C with 1.25 mL of reagent B then adding 1.35 mL of reagent A. A 96 well plate was used for sample loading. In each well, 5 µL of sample was added to 95 µL of water (1:20 ratio dilution) before adding 100 µL of the working solution (Table 2.1). This was incubated at 60°C for 45 mins before measuring the absorbance. All standards and samples were analysed in triplicate. FigP was used to generate the standard curve and determine sample concentrations.

Table 2.1. *Standard preparation with water and BSA stock (1 mL ampoule of 1 mg/mL). BSA standard was only added to well A and diluted down to well H with 100 µL of water as shown on the 3rd column.*

Well	µL of 1 mg/mL BSA standard	µL of water per well	Final µg BSA per well
A	80	120	40
B	100	100	20
C	100	100	10
D	100	100	5
E	100	100	2.5
F	100	100	1.25
G	100	100	0.625
H	0	100	0

2.1.1.5 Tryptic digestion

Protein digestion was performed by adding DTT to a final concentration of 15 mM to the samples and incubated at 60°C for 30 mins. Following reduction with DTT, IAA was added to a final concentration of 20 mM and incubated for 30 mins in the dark at room temperature. Trypsin (Promega, Madison, WI) was added to the samples at a ratio of 1:25 (trypsin: sample) and incubated overnight at 37°C. Digestion was stopped by adding TFA to a final 1% concentration to the samples. Tryptic peptides were desalted using Empore C18-SD 4 mm/mL columns. Columns were primed (3 mL of Methanol), conditioned (3 mL of 0.1% formic acid) and 1.5 mL of sample placed in the column. The column was washed with 1 mL of 0.1% FA (x2) and peptides were eluted using 800 mL of 60% followed by 800 mL of 80% acetonitrile. The final eluent was centrifugally evaporated *in vacuo*, freeze dried and reconstituted in 0.1% formic containing *Saccharomyces cerevisiae* alcohol dehydrogenase (ADH) (Waters Ltd., Manchester, UK). One µg of protein as ascertained by peptide assay was injected onto the column.

2.1.1.6 Solid Phase Extraction (SPE)

Materials and solutions: Methanol (MeOH), trifluoroacetic acid (TFA), Water (HPLC grade), Acetonitrile (ACN), Formic acid (FA). Empore C18-SD columns (4 mm/1 mL) were used in this experiment. 0.1% FA was made up in water by adding 100 µL of 100% FA to 99.9 mL of HPLC water. 1% TFA was prepared by adding 1 mL of 100% TFA to 99 mL of HPLC water. 60% of ACN was prepared by adding 2 volumes of HPLC water to 3 volumes of 100% ACN while 80% ACN was made up by adding 1 volume of HPLC water to 4 volumes of 100% ACN. Depleted and concentrated plasma (100 µL) were placed into 1 mL of 1% TFA and left to precipitate for 10 mins then spun hard to remove particulates. The columns were first primed with 3 mL MeOH and washed 3 times with (approx. 1 mL) 0.1% FA prior to sample loading. 6 times 0.1% FA washed was made before bound proteins were eluted with 1 mL 60% and 80% ACN respectively. These elution's were combined, speed vacuumed for 2 h at 12,700 rpm and freeze dried overnight. The dried pellet was reconstituted with 0.1% FA and ADH (50 fmol/µL) for mass spectrometry analysis.

2.1.1.7 Affinity binding

2.1.1.7.1 Protein depletion with Lipid Removal Agent (LRA)

Materials and solution: LRA, Ammonium bicarbonate (ABC), buffer A, buffer B, raw plasma and human 14 MARS column. 50 mg of LRA matrix (Sigma-Aldrich, Germany) was washed five times with 50 mM ABC. The washed matrix was then re-suspended in 480 µL of ABC and 120 µL of neat plasma sample added. Sample was agitated for 2 h at room temperature before the unbound (supernatant) proteins was eluted from the pellet by centrifugation at 2,500 g. The HAPs in the supernatant were further depleted using the MARS14 immunodepletion column (section 2.1.1.3).The protein concentrations of both the depleted sample and the pellet ascertained using the BCA assay (section 2.1.1.4) prior to trypsin digestion (section 2.1.1.5), SPE (section 2.1.1.6) and mass spectrometry analysis.

2.1.1.7.2 CHT Ceramic Hydroxyapatite (HA)

Materials and solution: HA type I 20 µm (Bio-Rad) (also referred to as mixed mode matrix or M3 for the purpose of research presentation in chapters 3 and 4) depleted plasma and MES hydrate solution. Three hundred mg of resin affinity was washed 5 times with MES hydrate solution buffer. This washed resin was then re-suspended with depleted plasma in 5 mM MES hydrate solution buffer and left to agitate overnight at room temperature. The pellet was centrifuged, supernatant collected and washed twice with 5 mM MES hydrate solution buffer prior to elution with different buffers (Table 3.1). The protein concentration of these elutions including the pellet was ascertained using BCA assay (section 2.1.1.4) and digested (section 2.1.1.5) prior to solid phase extraction (section 2.1.1.6) and mass spectrometry analysis.

2.1.1.8 Twenty percent, thirty percent and forty percent ammonium sulphate $(NH_4)_2SO_4$ precipitation

Materials and solutions: $(NH_4)_2SO_4$, PBS, Ammonium Bicarbonate (ABC), amicon filters (3 kDa MWCO), TFA and raw plasma. Plasma samples were centrifuged at 12,000 g, 4° C for 30 mins. Supernatant was collected and diluted with PBS and split into 1 mL aliquots of ~20 mg each. $(NH_4)_2SO_4$ was added as follows; 55, 113, 144, 176, 208, 242, 277, 314 and 351 mg to different aliquots to obtain 10%, 20%, 25%, 30%, 35%, 40%, 45%, 50% and 55% salt concentrations respectively and incubated on ice for 30 mins with occasional mixing. The solution was then centrifuged at 12,000 g, 4°C for 25 mins. The supernatant and precipitate were then concentrated into amicon filters (3 kDa MWCO), buffer exchanged with ammonium bicarbonate (ABC) and split into 2 parts (A and B). Part A was used for western blot (Section 2.1.1.9) while part B was tryptically digested. One percent final concentration TFA was used to stop trypsin digestion (part B) before samples were spun down at 12,000 g, 4°C for 20 mins. Wash through fractions were collected. The pellet in the filter was reconstituted with ABC (50 µL), vortexed, spun down and supernatant collected into different vials for mass spectrometry analysis.

2.1.1.9 Protein visualisation using western blotting

Plasma samples were homogenised and centrifuged to remove debris and kept on ice at all times. The protein concentration of the plasma samples was determined using the BCA assay kit prior to loading in the wells (30 µg/ well). Protein samples were solubilised separately by sodium dodecyl sulfate (SDS) polyacrylamide gel electrophoresis (PAGE). Mini gel apparatus was prepared with 10% polyacrylamide resolving gel overlaid with a 5% polyacrylamide stacking gel. SDS-PAGE loading buffer was made up of 187.5 mM Tris HCl pH 6.8, 6% SDS, 30% glycerol, 150 mM DTT and bromophenol blue. Tris buffer Saline Tween-20 contained 50 mM Tris, 150 Mm NaCl and 0.1% tween 20. Triton-lysis (non-detergent buffer) contained 1% Triton X-100. Running buffer was made up of 25 mM Tris, 192 mM glycine and 0.1% SDS. Transfer buffer contained 25 mM Tris, 192 mM glycine and 0.01% SDS and 10% methanol.Samples were prepared by combing 30 µg of protein in a volume of 10 µL of dH20 with 10 µL of 2x Laemmli buffer and heated at 100°C for 3 min before cooling on ice. Samples were loaded onto the gels with a protein ladder (Thermo Scientific) and electrophoresed for 1 hr at RT. The protein ladder was used to determine the molecular weight of the proteins. Proteins were then transferred from the gels onto nitrocellulose membrane (Amersham) using Towbin transfer buffer at 100 volts for 1 hr at RT. Membranes were stained with Ponceau S and proteins visualised with dH20 before photographing.

2.1.2 Sample analysis

2.1.2.1 Mass spectrometry -Waters Synapt G2S mass spectrometry

A Waters Synapt G2S mass spectrometry (Waters Corp., Herts, UK) coupled to nanoAcquity UPLC instrument was used for separation and analysis of tryptic peptides. Prior to the mass spectrometry analysis, a number of quality control (QC) tests were carried out on the instrument in order the "benchmark" the results obtained. Firstly, the detector voltage was optimised to give a maximum signal. After optimisation, measurements were taken in ADC and edge detection (TDC) mode to ascertain the ion volume correlating to a single ion arrival

at the detector. Secondly, the Time of Flight (ToF) analyser was mass calibrated using a Glu-Fib (GFP) peptide. This peptide is well characterised and its sequence is well established under ESI-MS/MS conditions. The lock spray channel was used to infuse a 100 fmol/µL solution of GFP (m/z 785.84265) (Table 2.2). GFP was then subjected to fragmentation that resulted in a number of ions ranging between m/z 80-1,250. The MS instrument was considered calibrated if a mass accuracy of 2 parts per million (ppm) was attained for 13 of these product ions. Thirdly, a final test of the UPLC instrument was made by injecting a digested protein sample standard (Hela cells). Beyond a threshold of greater than 2,000 proteins, the instrument was considered benchmarked for sample analysis. The source temperature was set at 70ºC, capillary voltage at 3.0 kV, low collision energy at 4V and elevated energy at 15-40 V.

2.1.2.1.1 Data Processing with PLGS

MassLynx v4.1 (Waters Ltd.) software was used to acquire the raw data. Raw data files were then processed against the UNIPROT human database using ProteinLynx Global server (PLGS) 3.02. ADH (50 or 100 fmol/µL) was added as an internal standard to enable label free quantitation of proteins in the sample. The precursor mass tolerance was set at 2 ppm with a maximum of 2 missed cleavages. Carbamidomethyl modification of Cysteine was set as a static modification and oxidation of methionine was set as a dynamic modification. All analysis was filtered using a 1% false discovery rate (FDR).

Table 2.2. *Parameters used for protein identification on PLGS using raw data generated by MS analysis*

Attribute	Value
Chromatographic peak width	Automatic
MS TOF resolution	Automatic
Lock Mass for Charge 1	
Lock Mass for Charge 2	785.84265 Da/e
Lock Mass Window	0.25 Da.
Low Energy Threshold	300.0 Counts
Elevated Energy Threshold	50.0 Counts
Intensity Threshold	750.0 Counts

2.1.2.1.2 Data processing with Progenesis

Raw data files from mass spectrometry were imported into Progenesis LC-MS for expression analysis. The samples were grouped for comparison as shown on Expression^{E} (section 2.1.3.2). Protein identification data files were imported prior to data analysis. A reference run which is a raw data file representing a typical profile in terms of peptide retention time (Rt) and m/z was used to analyse the samples. This data file established a calibration for all the samples to be compared with and all their chromatograms were aligned to this selected run. Multiple filtering criteria were used. Firstly, only features with 2 to 5 charges were included. Secondly, identified peptides with only one hit were excluded and thirdly, the protein lists generated with unacceptable p-values were excluded. The final list were exported into Microsoft excel for further analysis with RapidMiner and SPSS.

2.1.2.2 Mass spectrometry Q-Exactive

LC-MS/MS analyses were performed using a Q-Exactive mass spectrometer (ThermoScientific, Bremen, Germany) coupled to an Ultimate 3,000 RSLC nano HPLC system (Dionex/ThermoFisher Scientific, Bremen, Germany). Samples were digested as

described in section 2.1.1.5. Dried pellets were re-constituted in 10 µL of 0.1% formic acid and 10 µL of 100 fmol/µL alcohol dehydrogenase (ADH) (1:1 ratio) as an internal standard enabling absolute quantitation of the proteins post-analysis. Tryptic peptides were separated on an Ultimate 3000 RSLC nano HPLC system (Dionex/ThermoFisher Scientific, Bremen, Germany). Samples were loaded onto a Cartridge based trap column, using a 300 µm x 5 mm C18 PepMap (5 µm, 100Å) and tryptic peptides were separated using the Easy-Spray pepMap C18 column (75 µm x 50 cm) with a gradient from 3-10% B in 5 mins, 10-50% B in 37 mins, 50-90% in 9 mins and 90-3% in 26 mins, where mobile phase A was 0.1% FA in water and mobile phase B, 80%/20% ACN/Water in 0.1% FA. Total run time of 75 mins. Flow rate was 0.3 µL/min. The column was operated at a constant temperature of 40°C.

The nanoHPLC system was coupled to a Q-Exactive mass spectrometer (ThermoScientific, Bremen, Germany). The Q-Exactive was operated in the data-dependent top 10 mode; full MS scans were acquired at a 70,000 resolution between m/z 200 to 2,000, with an ACG (ion target value) target of 1e6, maximum fill time of 50 ms. MS2 scans were acquired at a resolution of 17,500, with an ACG target of 5e4, maximum fill time of 100 ms. The dynamic exclusion was set at 30 s, to prevent repeat sequencing of peptides

2.1.2.2.1 Data processing with proteome discoverer

The raw data files were processed and peptides were assigned to proteins using Thermo Proteome Discoverer 1.4. All searches were performed against the UniProt human database (Reviewed FASTA format downloaded May 2014, 20,000 entries), with the precursor mass tolerance set at 10 ppm with a maximum of 2 missed cleavages and fragment tolerance setting at 0.6 Da. Sequest HT search engine was used to process all the data. Carbamidomethyl modification of cysteine was set as a static modification and oxidation of methionine was set as a dynamic modification. A decoy database was used where protein amino acid sequences were reversed or randomised and concatenated to the original database. This was used to calculate the peptide false discovery rate (FDR) of 1%.

2.1.3 Data analysis

2.1.3.1 Protein centre v3.14

Protein centre is a bioinformatics tool used to compare and interpret proteomics data sets. It enables single and combined data sets to be filtered, clustered, compared and analysed statistically. All the processed data (protein lists) from PLGS were exported to Microsoft Excel file which were converted into csv format prior to exporting to protein centre. Each cohort had 10 samples with 5 fractions each analysed in triplicate which generated 15 lists of protein per sample. These 15 lists were then merged and the protein quantities summed to produce one final list per sample. Unique proteins of each of the 3 cohorts (Control, HFPEF and HFREF) were then acquired by removing the duplicates and compared against each other (Figure 4.6). The Gene Ontology (GO) (section 2) slim molecular functions in the control, HFPEF and HFREF group were then obtained (Figure 4.7).

2.1.3.1.1 Gene Ontology (GO)

Gene ontology is used to perform enrichment analysis of gene sets. It does this by finding which GO terms are over/under represented using annotations for that gene set. GO can be performed using a number of tools including AmiGO, OBO-Edit and Protein centre.

To obtain an overall view of the molecular functions in the three groups, a gene ontology analysis was performed using protein centre. When comparing the total protein profiles, the gene ontology summaries were very similar in all the groups (Figure 4.7) with the highest represented categories being protein binding, catalytic activity and metal ion binding regions. The KEGG pathways repository speeds up and facilitates the extraction of biologically meaningful information and statistical analysis of MS data. Data analysis with SPSS

2.1.3.2 Expression^{E} with PLGS

Protein expression was carried out using **identity^{E}** (section 1.3.6.1). The processed raw data files (Protein lists) were placed in 3 groups (Control, HFPEF and HFREF) for comparison.

The protein lists were normalised using ADH that was the internal standard added to the samples for quantification purposes. All proteins with a p-value of less than 0.05 were considered significant. These were put on the final list and exported to Microsoft excel for further analysis with SIMCA, RapidMiner and SPSS.

2.1.3.3 Statistical Package for the Social Sciences (SPSS)

Statistical Package for the Social Sciences is a software package used for statistical analysis. This software is used for a variety of things including data management, and data documentation. Some of the statistics that could be performed with SPSS are descriptive analysis, bivariate analysis (analysis of variance-ANOVA, t-test, means, correlation and nonparametric tests), prediction for numeral outcomes (linear regression) and prediction for identifying groups (cluster and discriminant analysis).

2.1.3.4 Soft Independent Modelling of Class Analogies (SIMCA)

SIMCA is software used for methods of principle component analysis (PCA) and partial least square (PLS) regression. It helps in the analysis of process variations, identification of critical parameters and prediction of the final product quality. With SIMCA we can structure information to find hidden details in the data and also extract true predictive information using PLS. Its graphical interface enables us to easily interpret data and draw conclusions.

The analysis cycle

A new project is created by importing the primary data file (1) which is modified by generating new variables as functions of existing ones (2). The default workset is the whole data set with variables as X while the default model (unfitted) is a principle component model of X (3). The role of the variables is transformed to fit the model (4). An estimation is then done at this stage to fit the model (5). A score scatter plot is used to detect any presence of outliers and other patterns in the data that are normally excluded from the work set at this

stage and step 4 repeated to fit a new model. Once the fit has been adjusted, the whole spectrum of plots can then be used for model interpretation (6). If the effect of the fitted model is satisfied a prediction set is built from primary or secondary data sets to do prediction (7) (Figure 2.1).

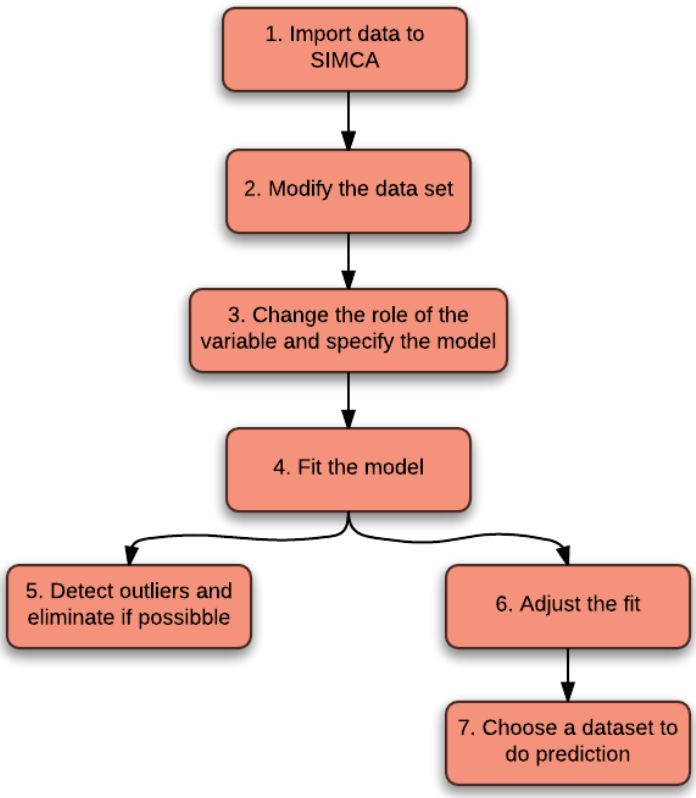


Figure 2.1. *A road map of data processing with SIMCA. All the boxes have been numbered to show the order of data processing.*

2.1.3.5 RapidMiner

RapidMiner, formerly known as YALE (yet another learning environment) is software used for data analysis. This software is used to carry out several data mining procedures including data loading and transformation, predictive analysis and statistical modelling as well as data pre-processing and visualisation.

Chapter Three

METHOD DEVELOPMENT

3 Mixed Mode Matrix (M3) - A method for plasma protein preparation using novel chemical affinity matrix

3.1 Introduction

Blood plasma is a complex mixture containing over 5,000 proteins (Anderson *et al.,* 2004, Wu *et al.,* 2011, Farrah *et al.,* 2011) that are vital for the smooth running of the body. These proteins provide important biological information that could be used in the diagnosis and treatment of diseases (Xie *et al.,* 2014), (Hu *et al.,* 2006). Thus, due to its composition, it is a primary clinical specimen for biomarker discovery.

However, the use of plasma for biomarker discovery has several challenges including the huge dynamic range of proteins that are dominated by high abundance proteins (HAPs) (Tu *et al.,* 2010). These HAPs act as a barrier in the detection of medium and low abundance proteins in proteomic analyses (Hakimi *et al.,* 2014, Shi *et al.,* 2012). As a result, many strategies have been attempted to overcome these challenges including protein precipitation (Kay *et al.,* 2008) and immunodepletion (Tu *et al.,* 2010), both of which have been successfully utilised to deplete HAPs. Protein precipitation is utilised due to its low cost and minimal method development requirements and ammonium sulphate has been used extensivelyd to precipitate the low abundance proteins prior to mass spectrometry analysis. Protein precipitation with ammonium sulphate (Mahn *et al.,* 2011) occurs due to the high salt concentration causing neutralisation of charges on the protein surface and consequently proteins inherently aggregate and thus precipitate (Englard *et al.,* 1990). Immunodepletion, which uses antibody columns, was first introduced in 2003 (Pieper *et al.,* 2003). The columns contain antibodies which target 7, 14 (Tu *et al.,* 2010) or 20 abundant plasma proteins (Smith *et al.,* 2011, Yadav *et al.,* 2011). More so, a new set of commercially available columns have medium abundance proteins antibodies that capture the next 50 most abundant proteins. (Shi *et al.,* 2012, Carr *et al.,* 2015). These columns were employed to improve the depth of protein identification and sensitivity for targeted analysis of proteins (Tu *et al.,* 2010).

3.1.1 Limitations of precipitation and immunodepletion

Despite the success of depletion columns and precipitation, there are still some limitations in the detection of LAPS. One of the main limitations is the high costs of reagent/depletion kits (MARS7, MARS14, Seppro IgY-14/Supermix and ProteoPrep20) and despite removal of almost 99% of protein content, there are still significant amounts of HAPs (serum albumin) in the sample implying that the large dynamic range problem is not fully addressed (Farrah *et al.,* 2014, Shi *et al.,* 2012). There are also concerns that the bound proteins may interact with potentially unique proteins (LAPs) that are then overlooked from subsequent proteomic analysis. Like immunodepletion, ammonium sulphate precipitation lacks specificity. HAPs are precipitated alongside the LAPs defeating the purpose of HAPs depletion. In addition, it requires intensive cleaning process to remove salts from the sample prior to mass spectrometry analysis.

3.1.2 Protein enrichment

Alternatively, enrichment of proteins could be attempted. The use of combinatorial peptide ligand libraries has been successfully applied to clinical samples (Hakimi *et al.,* 2014, Righetti *et al.,* 2010, Boschetti *et al.,* 2007, Craig *et al.,* 2006). This strategy has also been commonly employed in analysing post-translational modification (Moreno-Gonzalo *et al.,* 2014) particularly with peptide enrichment (Gu *et al.,* 2015).

Mixed mode matrix (M3) for instance could be used for plasma protein enrichment. It is a form of calcium phosphate containing positively charged calcium binding groups (C-sites) and negatively charged phosphate group (P-sites) (Shepard *et al.,* 2000, Cummings *et al.,* 2009) that interact with amino acids (Lee *et al.,* 2013) through cation exchange or calcium affinity (Gagnon *et al.,* 2006). These interactions are later disrupted by the introduction of neutral salts such as NaCl and NaH_2PO_4.

In addition to the depletion or enrichment strategies, extensive fractionation (Such-Sanmartín *et al.,* 2014, Greening *et al.,* 2011) has to take place in order to penetrate the plasma proteome to the extent of ng/mL. Fractionation on a peptide level will improve the number of identified proteins substantially (Anderson *et al.,* 2004, Ly *et al.,* 2011). However, despite the

significant increase in protein numbers with fractionation the extension of dynamic range is limited by the fact that peptides from the dominant proteins still dominate most of the fractions. In addition, LC fractionation has a negative impact on throughput. Therefore, either depletion or enrichment coupled with fractionation leads to significant increases in the cost of analysing a sample (capital costs of HPLC, columns, consumables as well as the increased MS time with incurred operators time etc.).

In addition to M3, lipid removal agent (LRA) has also been used to enhance protein detection of lipoproteins (Heink *et al.*, 2015). LRA uses its high affinity for lipids and lipoproteins particles to reduce the complexity of plasma.

Thus, the aim of this chapter is to analyse and compare various plasma proteomics methods with and without this developed novel protocol called Mixed Mode Matrix (M3), which utilises ammonium sulphate precipitation in conjunction with M3 for plasma protein enrichment.

3.2 Methods

Several methods were utilised in this chapter including Hydrophilic Strong Anion Exchange (SAX) (Section 2.1.1.2), LRA binding affinity (Section 2.1.1.7.1) and ammonium sulphate precipitation (20%, 30% and 40% concentration) (Section 2.1.1.8). The protocol for ammonium sulphate was adjusted slightly (Section 3.2.1) and a new method called Mixed Mode Matrix (M3) (Section 3.2.2) used in conjunction with ammonium sulphate precipitation (20%) was included in the analysis.

3.2.1 Twenty percent ammonium sulphate (AS) precipitation

Five hundred µL of plasma were centrifuged at 12,000 g at 4°C for 30 mins using a Centrifuge 5418 R (Eppendorf) to remove any particulates. Supernatant was collected, diluted with 400 µL of PBS and split into two equal aliquots to a protein amount of ~20 mg each. To each aliquot, 113 mg of AS was added to obtain 20% salt concentrations before incubating on ice for 30 mins with occasional mixing. The solution was centrifuged at 12,000 g, 4°C for 25 mins. The precipitate was separated and concentrated using an Amicon filter

(3K MWCO) and buffer exchanged with 5 mM 2-(N-morpholino)ethanesulfonic acid (MES) buffer pH 6.5.

3.2.2 LAPs enrichment with M3

An experiment was established to assess M3 treatment on depleted plasma (section 2.1.1.3), crude plasma and 20% AS precipitated plasma (section 3.2.1). The experiment was done in the presence and absence of M3 to generate 6 samples *(*Figure 3.1) in total for comparison. The depleted and undepleted fractions were concentrated on an Amicon ultracentrifuge device (3K MWCO) and subsequently buffer exchanged with 5 mM MES buffer pH 6.5. Twenty percent AS protein-precipitated plasma was treated as described in section 3.2.1.

Three tubes of M3 (300 mg) were pre washed (5x) with 5 mM MES buffer pH 6.5. To each tube of M3, either depleted plasma, crude plasma or 20% AS precipitated plasma were added. A further three tubes with no M3 either had depleted plasma, crude plasma or 20% AS precipitated plasma added. One mL of 5 mM MES buffer pH 6.5 was added to each tube and agitated for 10-14 h at room temperature. Unbound plasma proteins were removed by centrifugation at 2,500 g and bound proteins were eluted using four consecutive, increasing concentrations of NaH_2PO_4 solutions (1.5 mL) (Table 3.1).

After each addition of 1.5 mL NaH_2PO_4 solution, the samples were placed on a rotating agitator for 10 mins before centrifuging at 15,000 g for 2 mins and the eluent removed. Two washes were made with 5 mM MES buffer pH 6.5 after each elution and added to their respective fractions. Twenty percent ammonium hydroxide was added to adjust to pH 8.0. The protein concentration was determined (section 2.1.1.4) and the six samples were subject to reduction, alkylation and digestion (Section 2.1.1.5) with trypsin overnight (12-14 h) prior to SPE (section 2.1.1.6) and LC-MS/MS analysis.

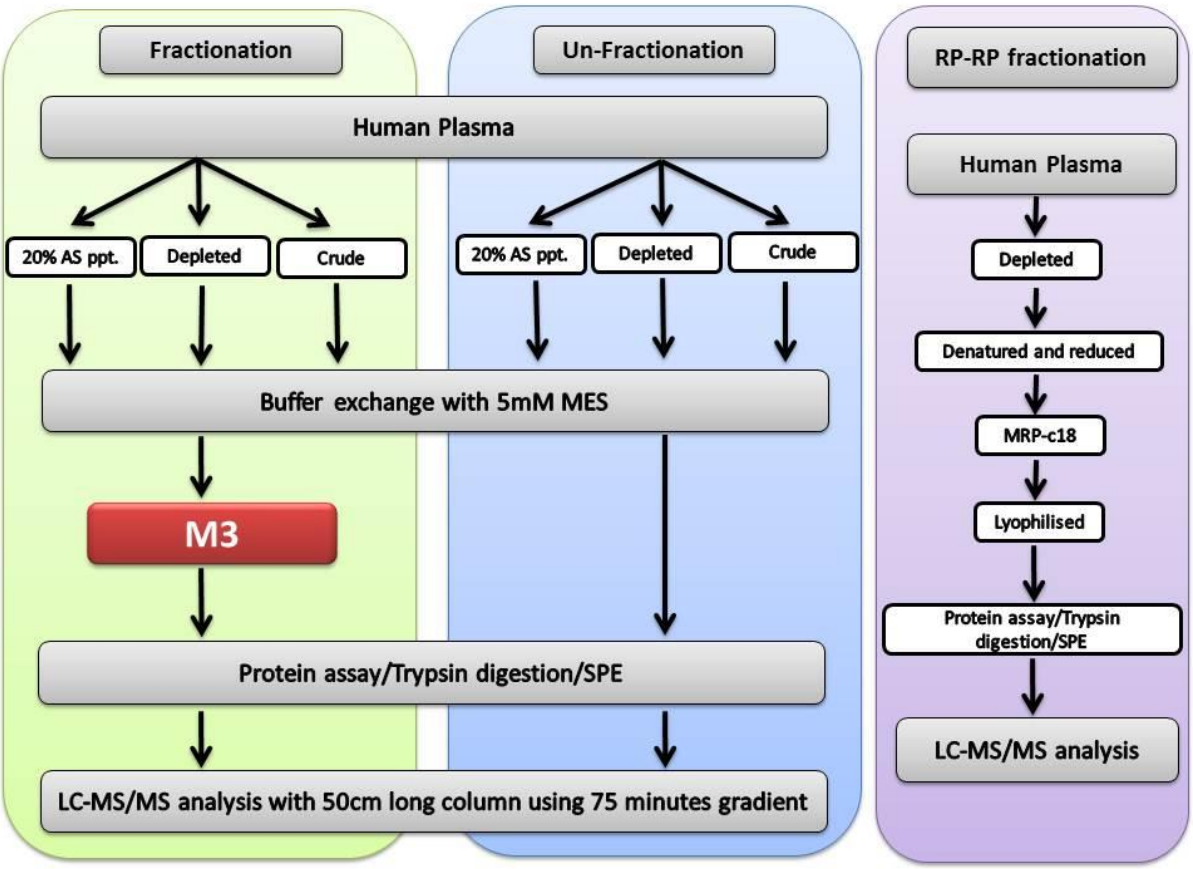


Figure 3.1. *General workflow used for the proteomic analyses of human plasma samples.*

Table 3.1. *Five different fractions from M3 particles with different salt concentrations.*

Fraction	Elution buffer	Volume
1	200 mM NaCl in MES	1.5 mL
2	120 mM NaH_2PO_4 pH 6.5	1.5 mL
3	200 mM NaH_2PO_4 pH 6.5	1.5 mL
4	300 mM NaH_2PO_4 pH 6.5	1.5 mL
5	Unbound supernatant	1.5 mL

All LC-MS/MS analyses were performed using a Q-Exactive mass spectrometer (ThermoScientific, Bremen, Germany) coupled to an Ultimate 3000 RSLC nano HPLC system (Dionex/ThermoFisher Scientific, Bremen, Germany) (Section 2.1.2.2). The raw data files were processed and peptides were assigned to proteins using Thermo Proteome Discoverer 1.4 (Section 2.1.2.2.1)

3.2.3 NaH_2PO_4 vs. NaCl

A test experiment was carried out to ascertain the number of proteins attained when eluted with NaCl and NaH_2PO_4 in two separate experiments (Table 3.2). The same conditions were applied in both experiments. It was evident that more proteins were eluted with NaH_2PO_4. In addition, all the proteins eluted with NaCl were also present in the NaH_2PO_4 fractions. Further analysis was done to narrow down the number of fractions with NaH_2PO_4. All the fractions had unique proteins but the ones eluted by 400 mM NaH_2PO_4 all overlapped with the other fractions. Therefore, 100 mM, 200 mM, 300 mM and 500 mM elutions and the supernatant were used in the clinical study.

Table 3.2. *Total number of proteins attained when 200 µl of plasma was bound on HA and eluted with increasing concentration of NaCl and NaH_2PO_4 in separate experiments. More proteins were eluted with NaH_2PO_4 which was used in the study.*

Concentration (mM)	NaH_2PO_4	NaCl
100	143	50
200	143	46
300	156	43
400	150	67
500	146	24
Supernatant	85	92
Total unique proteins	**245**	**111**

3.2.4 Reversed-Phase Fractionation of Proteins using macroporous Reverse-Phase C18 (mRP-C18) column

The mRP-C18 column (Agilent) is a macroporous reversed-phase chromatographic column with octadecylsilane functionality. Fractionation with this column is based on previously described methods (Martosella *et al.*, 2008). The column was conditioned according to the manufacturer's instructions. The LAP flow-through fraction from the MARS14 column was denatured in 6 M urea at room temperature for 30 mins. The sample was reduced with 10 mM tris-2-carboxyethyl-phosphine (TCEP) for 15 mins. IAA was added to a final concentration of 20 mM and allowed to alkylate the sample for 30 mins in the dark (section 2.1.1.5). The

sample was then fractionated on a binary gradient, using a Waters Alliance 2765 HPLC, into 5 fractions using the mRP-C18 column at a column temperature of 80 °C with a flow rate of 0.75 mL/min, using an elution gradient of 0-3 mins 5% B, 3-3.5 mins 5-30% B, 3.5-6.5 mins 30-45% B, 6.5-7.5 mins 45-100% B, 7.5-12.5 mins 100% B, 12.5-18 mins 100-5% B. Buffer A was water with 0.1% TFA (v/v) and buffer B was acetonitrile with 0.08% TFA (v/v). Five consecutive 1 min fractions (0.75 mL each) were collected between 4.8-9.8 mins using a Waters Fraction Collector III. Collected fractions were centrifugally evaporated in vacuum and freeze-dried overnight. Lyophilised proteins were resuspended in water. A protein assay was performed (section 2.1.1.4). Ammonium bicarbonate was added to a concentration of 10 mM with a resultant pH of 7.8. Trypsin digestion was performed as described.

3.3 Results

A comparison between fractionated (Hydrophilic Strong Anion Exchange-SAX) and unfractionated samples was carried out. SAX did not yield many proteins, perhaps due to fractionation on a peptide level that capitulated a number of challenges including sensitivity levels, system volumes flow rates and fraction collection. Due to these low protein numbers obtained with SAX, another experiment with LRA binding affinity was carried out. We hypothesised that the LRA might have a higher binding affinity for proteins. Thus, we investigated, to observe if elution could be improved by the addition of an MS compatible detergent in the protocol that could aid protein elution from LRA. However, some proteins could not be eluted from the matrix. This resulted in substituting the LRA affinity matrix with M3. Prior to the M3 experiment, ammonium sulphate concentrations (20%, 30% and 40%) were used to precipitate LAPs from crude plasma. This experiment was carried out to ascertain which concentration would give the highest number of LAPs. The precipitate with the highest LAPs was then used in conjunction with M3 for protein analysis. The ability of M3 to retain proteins was tested on three preparations of plasma namely (i) LAPs precipitated with 20% AS; (ii) depleted plasma (MARS 14); and (iii) crude plasma without any abundant protein depletion. (NB: Ammonium sulphate was used as a comparison for a standard protein precipitation method that is commonly employed in biochemistry experiments).

3.3.1 Hydrophilic Strong Anion Exchange (SAX) chromatography with HPLC

A 50 mins gradient was used in this fractionation. Twelve fractions were collected. These fractions were concatenated as shown in Table 3.3.

Table 3.3. *Twelve peptide fractions were concatenated into 4 concatenated fractions* ***(CF).***

CF 1	CF 2	CF 3	CF 4
1	2	3	4
5	6	7	8
9	10	11	12

An unfractionated peptide sample was collected prior to fractionation to make a comparison with the fractionated samples. These fractions were analysed on the Q-Exactive using the orbitrap technology (Section 2.1.2.2).

Table 3.4. *Protein amounts loaded onto the column (Q-Exactive system) and protein hits (1%FDR) obtained from all the fractions by proteome discoverer 1.4 database. CF= concatenated fraction.*

CF	Proteins amounts on column (ng)	Protein Hits
1	1225	89
2	68.9	86
3	160	91
4	123	43
Unfractionated sample	415	45

The unfractionated sample had the lowest number of proteins as compared to the fractionated sample. However, the total protein hits obtained after fractionation was comparatively low to the proteins obtained in literature (Millioni *et al.*, 2011). As a result, a second experiment with LRA binding affinity was carried out (Section 2.1.1.7.1).

3.3.2 Lipid Removal Agent (LRA) binding affinity

We used LRA to isolate lipids and lipoprotein particles to reduce the complexity of plasma. Both the supernatant containing lipid free plasma and the bound proteins on the LRA were analysed. When the matrix was digested, 224 proteins were identified from 1875 peptides as compared to the supernatant, which identified 191 proteins from 1280 peptides (Table 3.6). This shows the high affinity nature of the LRA matrix (Bhandari, manuscript in prep). In addition, some low abundant proteins were also identified in the matrix (Table A- 6). This suggests a possible sample/protein loss with this workflow.

Therefore, it was suspected that due to LAPs being retained on the matrix potential biomarkers were being lost. Minimising the loss onto the matrix would be crucial to the success of the method. This prompted the use of the mixed mode matrix (M3) which not only had little retention but also enriched the plasma proteins. Prior to the M3 experiment, ammonium sulphate concentrations (20%, 30% and 40%) were used to precipitate LAPs from crude plasma.

3.3.3 Ammonium sulphate precipitation

An evaluation of ammonium sulphate precipitation (Section 2.1.1.8) was carried out. A protein assay was used to ascertain the amount of proteins in the ammonium sulphate ppt. The aim of this experiment was to precipitate only the low abundance proteins from the sample. Figure 3.2 shows that when ammonium sulphate concentration was increased, more proteins some of which, high abundance proteins were precipitated. This was visualised on the 1D gel electrophoresis (Figure 3.4) where the albumin band could been seen getting bigger with increasing concentration of ammonium sulphate.

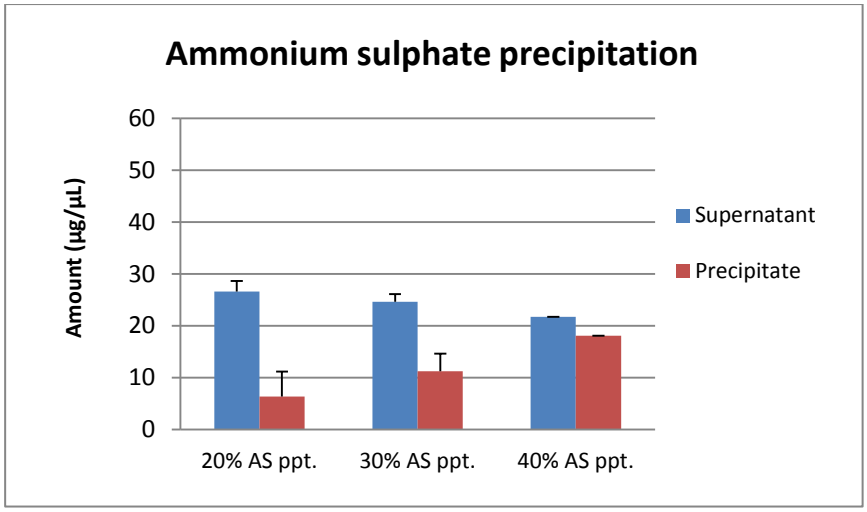


Figure 3.2. *Bar Graphs of human plasma precipitation with 20%, 30% and 40% ammonium sulphate. The supernatant and precipitate were collected and run separately on the gel (Figure 3.4). More proteins were precipitated out of the supernatant as ammonium sulphate concentration was increased.*

1-D-SDS-PAGE analysis of pre-fractionated raw plasma with ammonium sulphate $(NH_4)_2SO_4$) showed increased numbers of proteins bands in the precipitate and respective supernatant with 20% and 30% $(NH_4)_2SO_4$ than the 40% $(NH_4)_2SO_4$ (Figure 3.4).

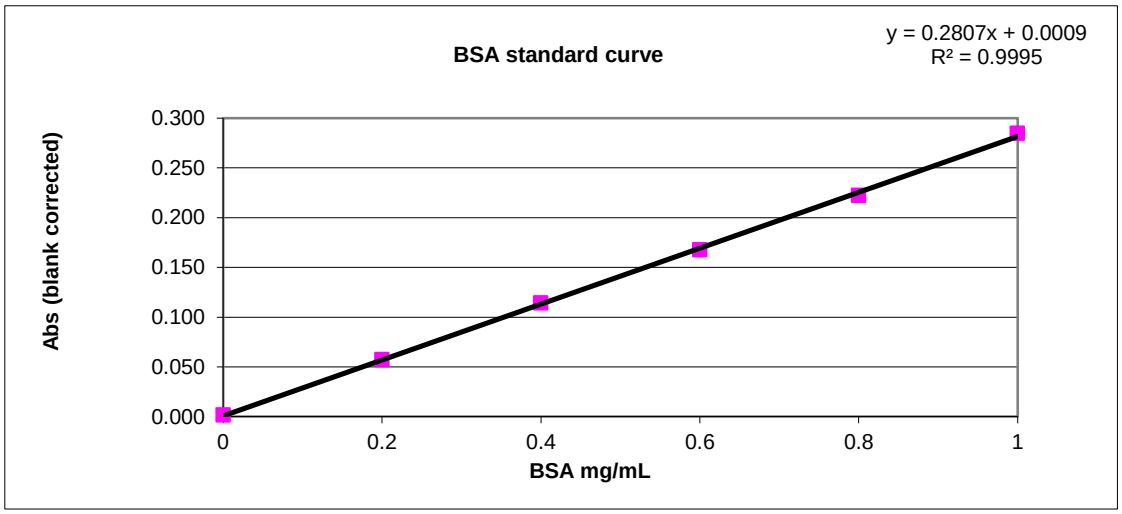


Figure 3.3. *BSA standard curve of concentration vs absorbance. R^2 value is close to 1 suggesting that the trend line is very close approximation to the actual values.*

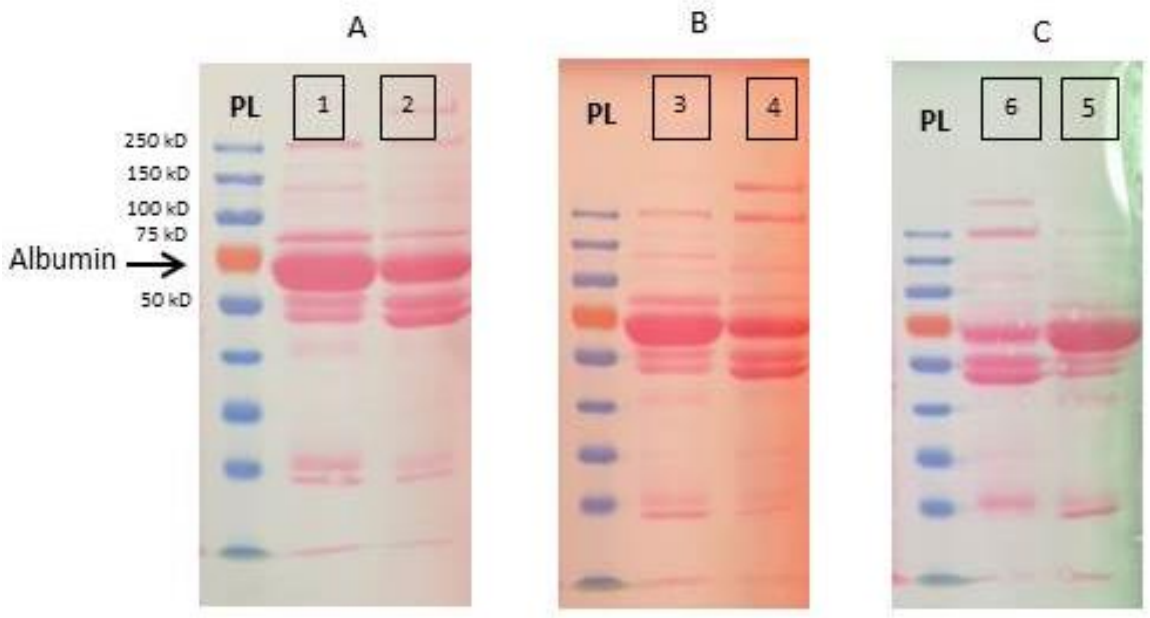


Figure 3.4. *One dimensional gel electrophoresis of 20%, 30% and 40% ammonium sulphate $(NH_4)_2SO_4$) precipitate and supernatant showing protein bands. The molecular weight of albumin is 66.5 kDa which could potentially be the orange band indicated by the arrow between 50kDa and 75kDa. A=20% $(NH_4)_2SO_4$ supernatant (1) and precipitate (2), B=30% $(NH_4)_2SO_4$ supernatant (3) and precipitate (4) and C=40% $(NH_4)_2SO_4$ supernatant (6) and precipitate (5) from normal plasma. PL= protein ladder*

The $(NH_4)_2SO_4$ precipitated the majority of the low abundance proteins leaving the high abundance proteins in the supernatant which included albumin, alpha 2 macroglobulin, transferrin, fibronectin and haptoglobin. This was evident on the 1-D gel electrophoresis analysis shown in Figure 3.4. which showed the high abundance proteins in much higher intensities in the supernatant as opposed to the precipitate. Albumin (most abundant plasma protein) bands in particular were much bigger on the supernatant as opposed to the precipitate. Transferrin and alpha 2 macroglobulin bands are also dominant on the supernatant as opposed to the precipitate. However, Fibronectin- beta and APO-E bands are much bigger in the precipitate than the supernatant. As a result, 20% AS precipitation (Section 3.2.1) was chosen to be used in conjuction with the M3 in future experiments.

3.3.4 Unfractionated samples analysis

To establish a baseline for protein numbers, a comparison of the number of proteins and peptides attained in the three different experiments was carried out in the absence of M3 matrix. The depleted plasma provided the highest proteome coverage identifying 281 proteins from 2,725 peptides. Crude plasma identified 204 proteins from 2,020 peptides which was 15% more than the proteins identified (174) in 20% AS precipitate (Table 3.6). In addition, there was a significant overlap in the three experiments with 117 unique proteins identified with depleted plasma as compared to 30 and 26 for precipitated plasma and crude plasma respectively (Figure 3.8). One hundred and fourteen proteins were common to all three preparations (between 20% AS precipitate, crude plasma and depleted plasma). The numbers described here are comparable to previous reports (Hakimi *et al.*, 2014, Kussmann *et al.*, 2013). Both 1% and 5% FDR are shown in Table 3.6 so that comparisons with other published results can be appropriately and correctly made.

Table 3.5. *Number of proteins identified in each triplicate run of the unfractionated plasma samples. Mean refers to observed number of proteins in triplicate. CoV=Coefficient of variance.*

Triplicate	Depleted Plasma	Crude plasma	20% AS ppt. plasma	LRA matrix	LRA supernatant
1	228	166	148	176	152
2	226	170	143	192	151
3	212	172	138	185	139
Mean (CoV)	222 (3%)	169 (1%)	143 (2%)	184 (3%)	147 (4%)

The triplicate analysis demonstrated a consistency in the number of proteins identified in each experiment with standard deviations of <7 which yields CoVs of less than 4% (Table 3.5) (Figure 3.5). This shows that there was minimal variation in the sample analysis.

Table 3.6. *Summary of the total number of proteins identified in each experiment of the unfractionated plasma samples. FDR = False Discovery Rate, Protocol descriptions:*

Experiment	Sample Volume (µL)	Sample Prep. Time	Identified Proteins 1% FDR	Unique Peptides 1% FDR	Identified Proteins 5%FDR	Unique Peptides 5% FDR
Depleted plasma[a]	200	48 h	281	2725	555	2930
Crude plasma[b]	200	45 h	204	2020	376	2069
20% AS precipitated plasma[c]	200	48 h	174	1419	299	1494
LRA matrix[b]	200	48 h	224	1875	451	2262
LRA supernatant[b]	800	46 h	191	1280	314	1481

a *Depletion, buffer exchange, Tryptic digestion, SPE, LC-MS/MS,*
b *Buffer exchange, Tryptic digestion, SPE, LC-MS/MS,*
c *Precipitation, Buffer exchange, Tryptic digestion, SPE, LC-MS/MS*

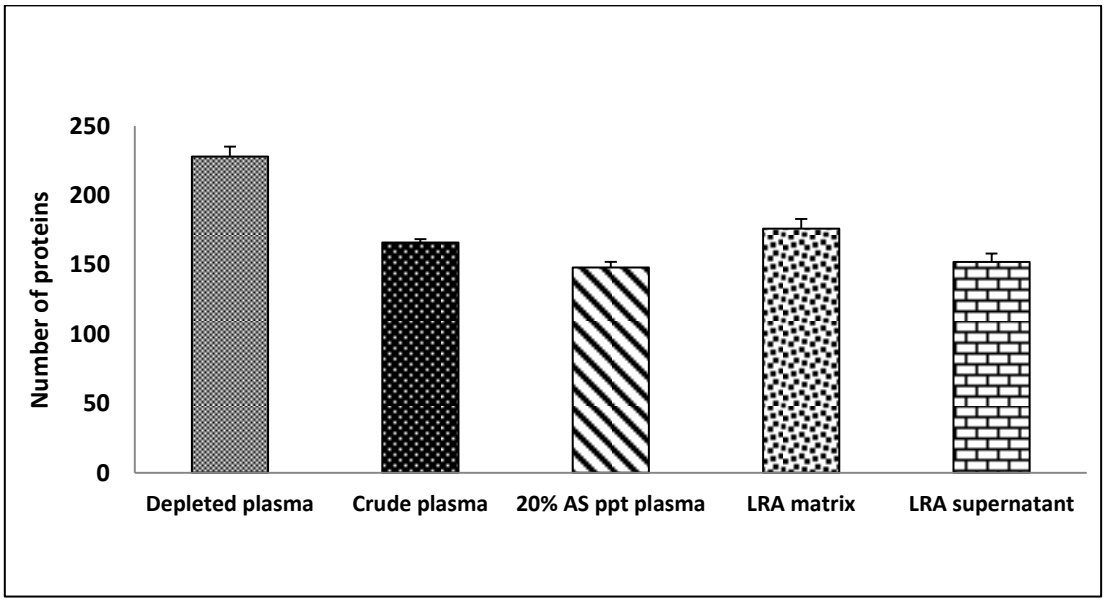


Figure 3.5*. Mean proteins identified in each triplicate run of the unfractionated plasma samples (n=1) with CoV of <4. In this experiment, when samples are depleted and not fractionated, more proteins are achieved as compared to other methods as shown on the figure.*

3.3.5 M3 analysis

M3 enhanced protein numbers in each of the three experiments (Figure 3.8) namely; Depleted plasma identified 330 proteins (up from 281) crude plasma 311 proteins (up from 204) and 20% AS precipitated plasma 245 proteins (up from 174) (1% FDR). One hundred and twenty-eight proteins were commonly identified in all three experiments. The depleted plasma provided the highest number of unique proteins, which was 21% and 45% more than the unique proteins in crude plasma and 20% AS precipitate respectively (Figure 3.9). The increase in protein identifications was achieved despite significant reductions in peptide identifications for depleted and crude plasma when both were treated with M3 (Table 3.7).

Table 3.7. *Summary of the total number of proteins identified in each experiment of the fractionated plasma samples using M3.*

Experiment	Sample Volume (µL)	Sample Prep. Time	Protein Identifications 1% FDR	Unique Peptides 1% FDR	Protein Identifications 5% FDR	Unique Peptides 5% FDR
Depleted plasma[a]	200	72 h	330	2018	642	2473
Crude plasma[b]	200	68 h	311	1654	635	2055
20% AS precipitated plasma[c]	200	72 h	245	1426	474	1653

a Depletion, M3, fractionation, Tryptic digestion, SPE, LC-MS/MS,
b M3, Fractionation, Tryptic digestion, SPE, LC-MS/MS,
c Precipitation, M3, Fractionation, Tryptic digestion, SPE, LC-MS/MS

The triplicate analysis of the five fractions per experiment showed consistency in protein hits with standard deviations of < 11 (Table 3.9, Figure 3.7). This demonstrates acceptable variation in the sample analysis. In addition, it is evident that there is a trend in improving reproducibility when plasma is depleted and incorporated with M3 (Figure 3.6) as compared to AS precipitation with M3 and crude plasma with M3.

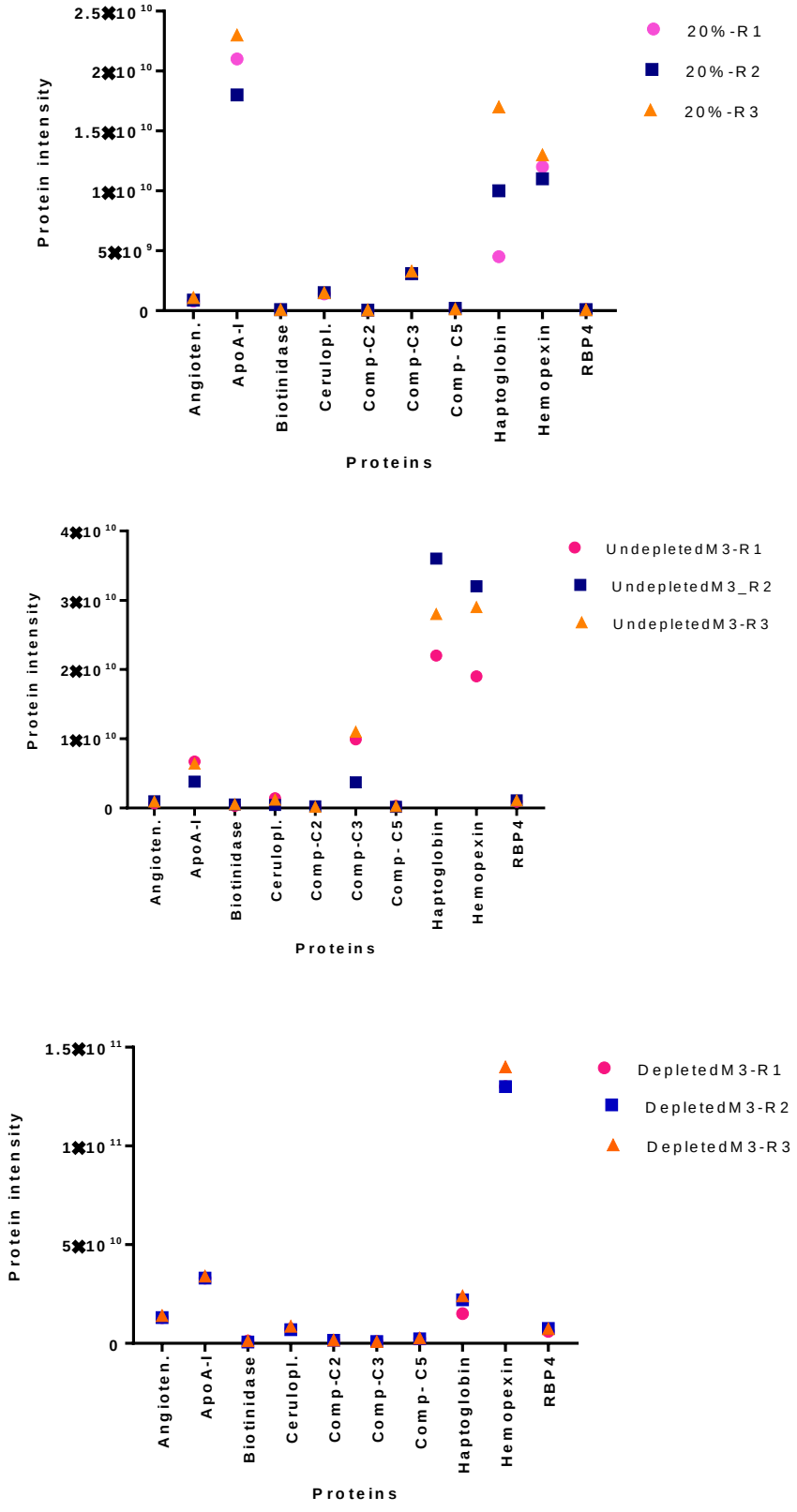


Figure 3.6. *Diagram showing a triplicate analysis of Ten FDA markers identified with Fractionated-M3 plasma samples. Y axis shows overall protein intensity normalised to alcohol dehydrogenase (ADH). The trend shows a good reproducibility between the triplicate analysis in all the groups with depleted plasma + M3 having the best outcome. Angioten=Angiotensinogen, ApoA-1=Apolipoprotein A-1, cerulopl. =Ceruloplasmin, Comp=Complement, RBP4=retinol-binding protein 4.*

Table 3.8. *A triplicate analysis of the ion intensities of Ten FDA markers identified with Fractionated-M3 plasma samples. These data was normalised relative to ADH (50 fmols). All the triplicate analysis for all the proteins showed good reproducibility.*

Description	AS 20% ppt.-R1	AS 20% ppt.-R2	AS 20%-ppt.-R3	Undepleted M3-R1	Undepleted M3_R2	UndepletedM3-R3	DepletedM3-R1	DepletedM3-R2	DepletedM3-R3
Angiotensinogen	8.00E+08	8.80E+08	1.10E+09	6.70E+08	9.70E+08	9.60E+08	1.30E+10	1.30E+10	1.40E+10
Apolipoprotein A-I	2.10E+10	1.80E+10	2.30E+10	6.70E+09	3.80E+09	6.40E+09	3.30E+10	3.30E+10	3.40E+10
Biotinidase	9.90E+07	8.80E+07	9.60E+07	3.20E+08	5.10E+08	5.50E+08	1.20E+09	6.60E+08	1.30E+09
Ceruloplasmin	1.40E+09	1.50E+09	1.50E+09	1.40E+09	4.50E+08	1.20E+09	7.20E+09	6.90E+09	8.70E+09
Complement C2	5.30E+07	5.00E+07	7.90E+07	2.30E+08	2.60E+08	1.90E+08	1.60E+09	1.60E+09	1.70E+09
Complement C3	3.10E+09	3.10E+09	3.30E+09	9.90E+09	3.70E+09	1.10E+10	9.80E+08	9.90E+08	9.50E+08
Complement C5	1.60E+08	1.90E+08	1.60E+08	3.00E+08	1.80E+08	2.80E+08	2.30E+09	2.30E+09	2.70E+09
Haptoglobin	4.50E+09	1.00E+10	1.70E+10	2.20E+10	3.60E+10	2.80E+10	1.50E+10	2.20E+10	2.40E+10
Hemopexin	1.20E+10	1.10E+10	1.30E+10	1.90E+10	3.20E+10	2.90E+10	1.30E+11	1.30E+11	1.40E+11
Retinol-binding protein 4	1.00E+08	8.70E+07	9.00E+07	9.10E+08	1.10E+09	1.10E+09	6.10E+09	7.50E+09	7.40E+09

Table 3.9. *Number of proteins identified in each triplicate run per fraction (refer to Table 3.1) with M3 on plasma samples (Coefficient of Variance (CoV) of the triplicate runs).*

Fraction	Depleted plasma + M3 Fractionation	Crude plasma + M3 Fractionation	20% AS ppt. plasma + M3 Fractionation	Depleted plasma + mRP-C18 Fractionation
1	90 102 (6%) 103	113 132 (8%) 136	80 78 (4%) 86	113 115 (1%) 114
2	174 169 (1%) 171	101 107 (3.5%) 110	105 106 (0.5%) 106	133 135 (1%) 133
3	135 147 (5%) 151	115 100 (7%) 99	94 99 (5%) 106	127 119 (4%) 117
4	118 123 (5%) 110	126 117 (5%) 111	115 108 (3%) 114	148 143 (4%) 137
5	89 64 (14%) 88	113 98 (6%) 106	49 51 (2%) 52	140 137 (1%) 137

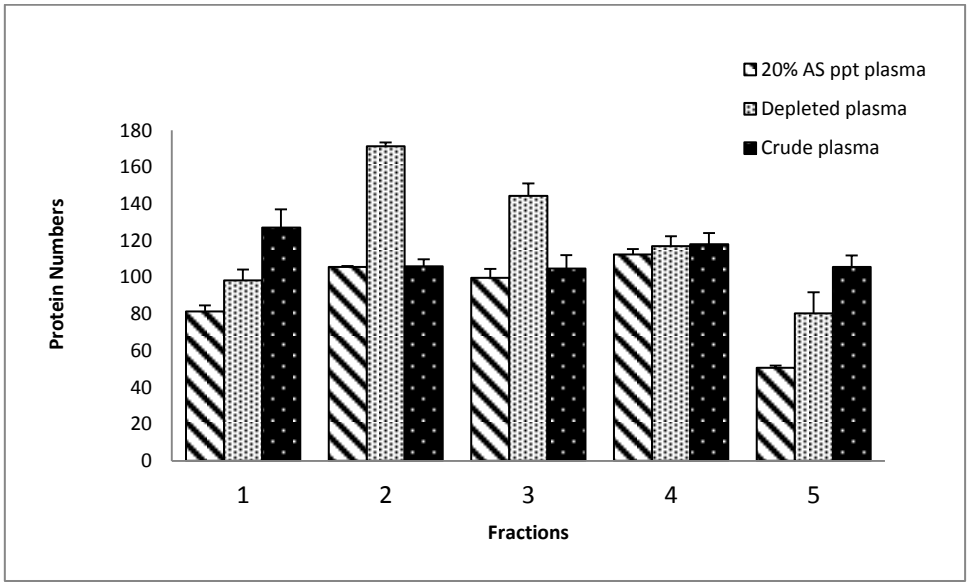


Figure 3.7*. Mean average of proteins identified in each triplicate run of the Fractionated-M3 plasma samples. The figure shows that the triplicate analysis was reproducible across all the 5 fractions with CoV of <10.1=200 nM NaCl, 2=120 nM* NaH_2PO_4*, 3=200* NaH_2PO_4*, 4=300* NaH_2PO_4 *and 5=supernatant.*

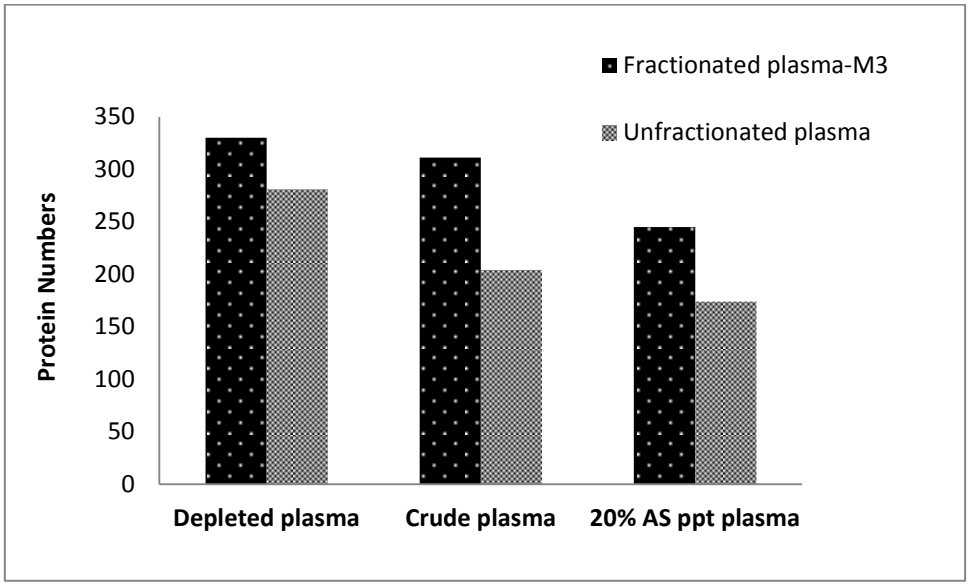


Figure 3.8. *Total number of novel proteins identified in all the six experiments. Depleted plasma, crude plasma and 20% AS precipitated plasma were all analysed with and without M3 treatment. The presence of M3 showed an increased number of proteins in all the 3 experiments.*

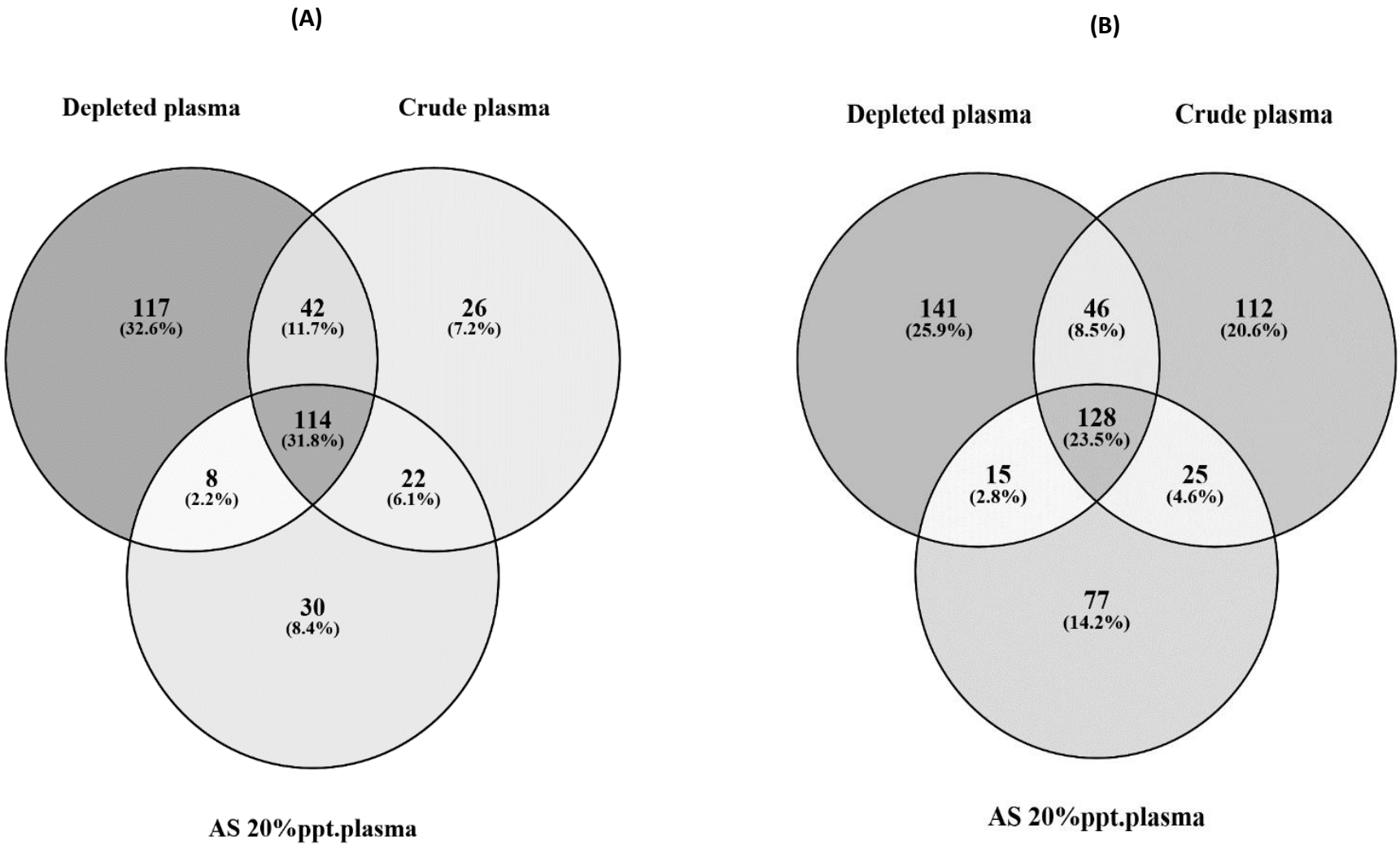


Figure 3.9. *Comparison of the number of unique proteins identified in unfractionated (A) and Fractionated-M3 plasma (B). The venn diagrams shows a significant increase in unique proteins in the presence of M3. See Table A-5 for complete list of identified proteins*

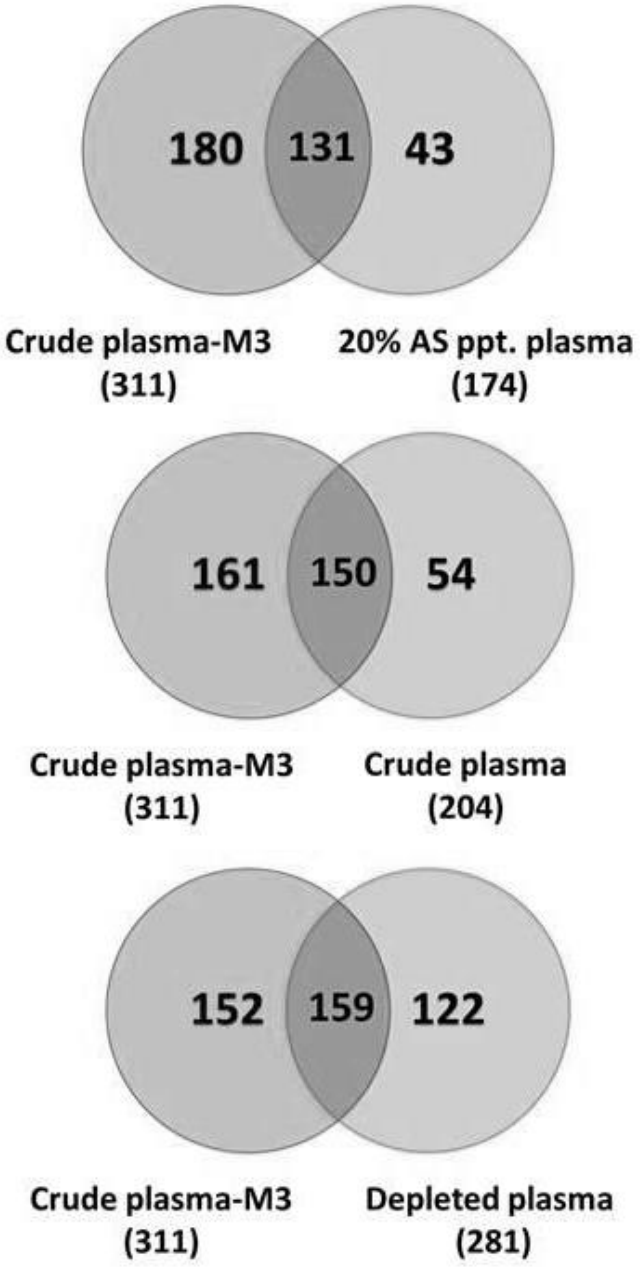


Figure 3.10. *Comparison of the number of unique proteins identified in crude plasma when M3 was added against 20% AS ppt., crude plasma and depleted plasma. The overlap showed a significant increase in unique proteins in the presence of M3 on crude plasma. The fact that unique proteins are described in all the experiments indicates that the mechanism of M3 is not only specific but also provides a different profile of proteins that can be vital in biomarker discovery.*

3.3.6 Comparison with Protein RP Fractionation on mRP-C18 column

The M3 method was compared directly to protein RP fractionation on the mRP-C18 column as described. Depleted plasma samples were fractionated on the mRP-C18 column with similar downstream processing as M3 samples. Two hundred and ten unique proteins were attained. When compared with other methods (Table 3.7), it was evident that the lowest number of proteins was acquired when samples were analysed with this method. In addition, only 9.3% of unique proteins were attained as compared to 21.3%, 17.3%, and 12.8% with depleted, crude and AS 20% ppt. plasma with M3, respectively (Figure 3.11).

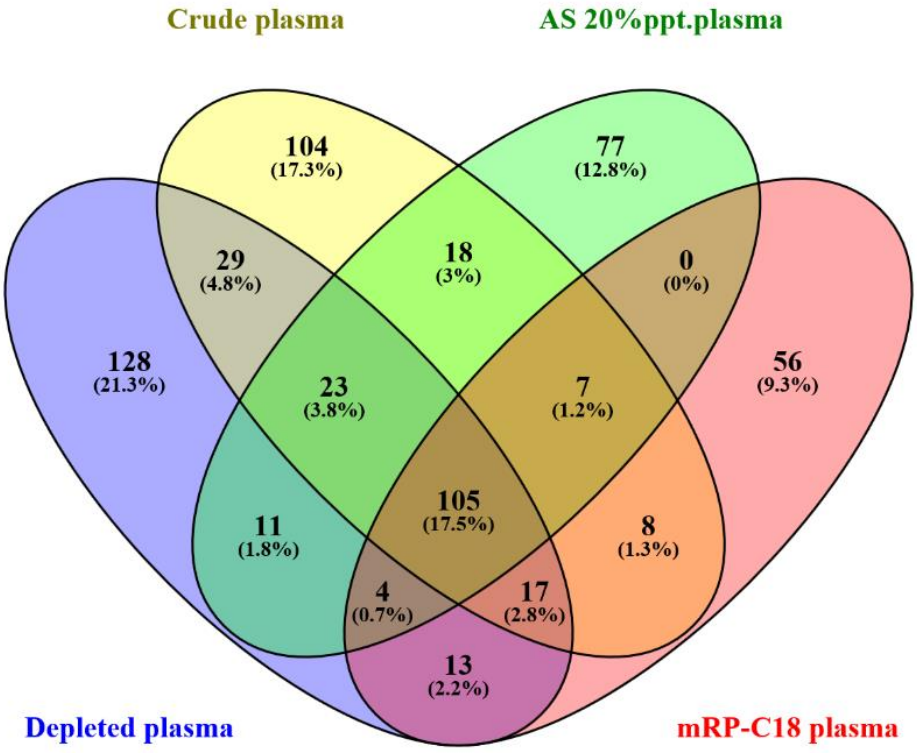


Figure 3.11. *A comparison of M3 method (depleted, crude and AS 20% ppt. plasma) with protein RP fractionation on mRP-C18 column.*

3.3.7 Detection of Food and Drug Administration (FDA) approved markers

In order to provide evidence that this method could provide utility in the discovery of biomarkers we then compared the protein list obtained in each of experiments with an accepted list of FDA approved markers (Anderson *et al.,* 2010). Plasma treated with M3 contained 17 FDA approved proteins, whilst the precipitated supernatant treated with M3 revealed 19 FDA approved markers. The depleted plasma treated with M3 contained 31 approved markers (Table 3.10). These numbers demonstrate that the methods can pick up proteins at concentrations that can yield clinically relevant markers for disease management.

Table 3.10. *Detected FDA approved markers. X indicates that the proteins were present in the sample. Most of these proteins have >6 orders of magnitude (Figure 1.7) yet mass spectrometry is limited to only 4-5 orders of magnitude. Thus, it is challenging detecting these proteins at very low concentration.*

FDA approved markers	***Depleted + M3***	***Undepleted + M3***	***20%ppt. + M3***
Adiponectin	x		
Angiotensinogen	x	x	x
Apolipoprotein A-I	x	x	x
Biotinidase	x	x	x
Ceruloplasmin	x	x	x
Cholinesterase	x		x
Complement C2	x	x	x
Complement C3	x	x	x
Complement C5	x	x	x
Cystatin-C	x	x	
Fibronectin	x	x	x
Glutathione peroxidase 3	x		x
Haptoglobin	x	x	x
Hemopexin	x	x	x
Ig gamma-1 chain C region	x		x
Ig gamma-2 chain C region	x		x
Ig gamma-3 chain C region	x		x
Ig gamma-4 chain C region	x		x
Insulin-like growth factor I	x		
Insulin-like growth factor II	x		
Intercellular adhesion molecule 1	x		
Interleukin-1 receptor accessory protein	x		
Mannan-binding lectin serine protease 1	x		
Mannan-binding lectin serine protease 2	x	x	
Plasma kallikrein	x	x	
Plasminogen		x	x
Protein-S-isoprenylcysteine O-methyltransferase			x
Retinol-binding protein 4	x	x	x
Thyroxine-binding globulin	x		
Vitamin D-binding protein	x	x	
von Willebrand factor	x	x	x

3.4 Discussion

Proteomic strategies for biomarker discovery offer great potential in identifying functional molecules that could be beneficial to the clinical management of disease. However, biomarker discovery using MS-based proteomics has proved challenging due to the dynamic range and complex nature of plasma samples. In addition, cost, lack of throughput, detection sensitivity and quantitative precision (Keshishian *et al.*, 2015) has made it even more difficult to acquire reliable results. In this project, a novel method to address these limitations has been developed. This involved improving the dynamic range through intensive depletions and reducing the complexity of plasma through LAPs enrichment with M3 and ammonium sulphate precipitation.

When a comparison between fractionated (RP-RP) and unfractionated samples was carried out, the fractionated samples had twice as many protein hits compared to the unfractionated sample (Table 3.4). Fractionation has a dual role in separation. It enables greater peak capacity so that ions of low abundance can be resolved and thus observed. It also reduces the dominance of high abundance peptide ions within the peak capacity space. Higher peak capacity provides the opportunity to expand dynamic range. This RP-RP method only yielded 91 proteins (Table 3.4) which is three times less than the proteins numbers obtained by Millioni *et al.*, 2011. The low protein numbers suggest that fractionation on a peptide level (rp-rp) does not address the abundant proteins issue. As a result, a second experiment was introduced which involved depletion with the LRA binding residue which has high affinity for lipoproteins. In this experiment, both the supernatant and the LRA matrix were digested and analysed. This did not yield many proteins either due to LRA's high protein retention. The LRA pellet had the highest number (224) of proteins as compared to the supernatant (191) (Table 3.6). This explains LRA's high binding and retention nature. The majority of the proteins (possibly potential biomarkers) bound on the pellet could not be eluted and there were still substantial amounts of HAPs in the supernatant. Hence, this method of analysis was not viable for reducing the dynamic range of protein concentration. Possibly, the raw plasma could be depleted with MARS column and analysed with MS to curb the protein loss through the LRA affinity binding residue pellet. Therefore, a third experiment with ammonium sulphate was introduced.

According to Mahn *et al.*, 2011 ammonium sulphate precipitation has been used by scientists to deplete high abundant proteins in plasma. The aim of this experiment was to precipitate LAPs with ammonium sulphate of various concentrations (20, 30 and 40%) and discard the HAPs that remained in the solution (supernatant). The concentration with the highest number of LAPs was used in conjunction with this novel method Mixed Mode Matrix (M3) for protein enrichment. Both the precipitate and the supernatant were analysed by 1D-SDS-PAGE. The 1D-SDS-PAGE was only used to visualise the protein bands.

1D-SDS-PAGE analysis (Figure 3.4) showed most of the HAPs dominant on the supernatant as opposed to the precipitate. However, HAPs such as Fibronectin-beta and APO-E bands were much bigger in the precipitate than the supernatant. This indicated that some of the HAPs were precipitated with the LAPs. When 30 µg of proteins loaded on the wells, the protein concentration of the precipitate increased with increasing concentration of $(NH_4)_2SO_4$. Conversely, the protein concentration of the supernatant decreased with increasing $(NH_4)_2SO_4$ concentration. From this, we can deduce that increasing $(NH_4)_2SO_4$ concentration resulted in further precipitation of low abundant proteins and depletion of the high abundant proteins. However, when these proteins were analysed with 1D-SDS-PAGE analysis (Figure 3.4) we confirmed that the 30 and 40% precipitates had increased levels of HAPs as compared to 20%. Thus, 20% AS ppt. was used in conjunction with M3.

Incorporation of M3 provides significant improvements in protein numbers whilst reducing the typical impact of throughput that the standard number of fractionations would necessitate for SCX or RP-RP fractionation. Additionally, the cost is significantly less as the M3 material is inexpensive (£100) or 5p per sample, especially in comparison to immunodepletion columns (£3,000), macroporous MRP-c18 column (£600) or equalizer beads (£550).

An assessment of combining immunodepletion with M3 was also made and compared with just M3. In addition, the utilisation of a well-established protein precipitation method commonly used in biochemistry was examined in order to provide evidence of specificity. The fact that unique proteins are described in both experiments indicates that the mechanism of M3 is not only specific but also provides a different profile of proteins and could help to identify novel biomarkers. In this study, we demonstrated that when M3 is used in conjunction with immunodepletion, more proteins are attained than just immunodepletion alone (Table 3.6). Additionally, when these samples were depleted (LAPs), enriched with M3 and fractionated, the protein numbers increased by over 15% (Table 3.7; Figure 3.8).

However, it was evident that we also acquired more proteins (311) when crude plasma was enriched with M3 as demonstrated by a 34% increase in protein numbers compared to crude plasma without M3 (Table 3.6). A marginal increase of 5.8% was observed with depleted plasma with M3 (330) as compared with crude plasma prepared with M3 (311). Despite the increase in numbers when crude plasma was enriched with M3, the profile of proteins attained were mostly IgGs, which could not help to identify novel biomarkers. Nonetheless, it showed the importance of enrichment with M3.

There was a significant overlap in proteins acquired in the M3 experiment between the three sample preparation groups. Depleted samples had more unique proteins (141) as compared to crude (112) and 20% AS precipitated (77) samples (Figure 3.9). Notably, the number of unique proteins in crude plasma was enhanced from 26 to 112 proteins, a 76% increase when M3 was used. In addition, the overlap analysis showed a significant number of proteins identified with crude plasma with M3 as compared to other methods (Figure 3.10). The increase in unique proteins observed in depleted samples could be because of other proteins depleted alongside the top 14 abundant proteins, possibly due to non-specific interactions in the MARS 14 column. Some of the proteins we identified with depletion include adiponectin, interleukin-1 receptor accessory protein and titin, while enrichment showed the presence of cardiomyopathy associated protein 5, metalloproteinase inhibitor 1, extracellular matrix protein FRAS 1 and cytochrome p450 4F12. These are relevant in disease mechanisms (Ahmed *et al.*, 2006, Gaggin *et al.*, 2013, Sente *et al*., 2016). Using this methodology, we demonstrated that we can achieve impressive coverage of low abundance proteins using a cost effective method even without the use of immunodepletion columns and other variables.

Recently, there have been significant advances in the numbers of proteins observed in human plasma (Keshishian *et al.*, 2015, Kussmann *et al.*, 2013, Shi *et al.*, 2012). Isotopic labelling of proteins, extensive fractionation (thus huge reduction in throughput) and the use of multiple costly immunodepletion columns, has brought about this increase. The discovery of biomarkers requires careful consideration of experimental design and a greater number of clinical samples utilised in the discovery phase. M3 is more applicable to these kinds of experiments. M3 enables separation at the protein level that promotes the monitoring of individual proteins in each fraction following elution. This provides an advantage when considering possible future validation strategies for candidate proteins from the discovery phase. Potentially, proteins (e.g cystatin C, Plasma serine protease inhibitor and Macrophage colony-stimulating factor 1 receptor) could be resolved and analysed from a single fraction

increasing throughput considerably (Table A-7). This approach could also be employed as a fractionation method prior to top down analysis.

Conclusion:

We have demonstrated the use of M3 to provide deeper coverage of the plasma proteome than other comparable technologies. Moreover, we have shown it to be robust, reproducible and depending on the application could be used to improve throughput. Finally, the M3 is particularly inexpensive, especially when compared to mRP-C18 columns and equalizer beads and thus should be considered in biomarker research.

Chapter Four

CLINICAL STUDY

4 Clinical study phase

In the world, 23 million people suffer from HF (Braunwald, 2013) and the global cost of HF treatment estimated at 108 billion US dollars (Cook *et al.,* 2014). These statistics signify a major public health problem and urgent need for new treatment. Several biomarkers for diagnosing and managing HF have been proposed including cardiac troponins and CRP, which reflect intricate pathophysiology of HF. This resulted in improved survival rate among HFREF patients. On the other hand, patients with HFPEF have seen little improvement in survival rate. Several models have been suggested to explain these differences in HFREF and HFPEF, which include hemodynamic, neurohumoral and cardiorenal models. Signs and symptoms of risk factors in patients with HF have been evaluated through physical examinations, laboratory tests and echocardiography (Eckstein *et al.,* 2016). Urine (Rossing *et al.,* 2016) and plasma proteomics have been explored in research to find new biomarkers of heart failure. However, these methods require reproducibility, high-throughput, sensitivity and a wide coverage of the proteome in study.

Hypothesis: Using plasma proteomics techniques with the state of the art mass spectrometry we can reveal novel biomarkers of HFPEF.

Aims and objectives

- Carry out a pilot study using M3
- Carry out a clinical study using M3/LRA

4.1 Discovery proteomics

The examination of the protein content of plasma derived from three cohorts using proteomics methods (Figure 4.1) developed in chapter 3 (section 3.2.2 and section 3.3.2) will be described. Protocols for M3 and LRA were adjusted as shown on section 4.2.1.2 and 4.2.1.3.

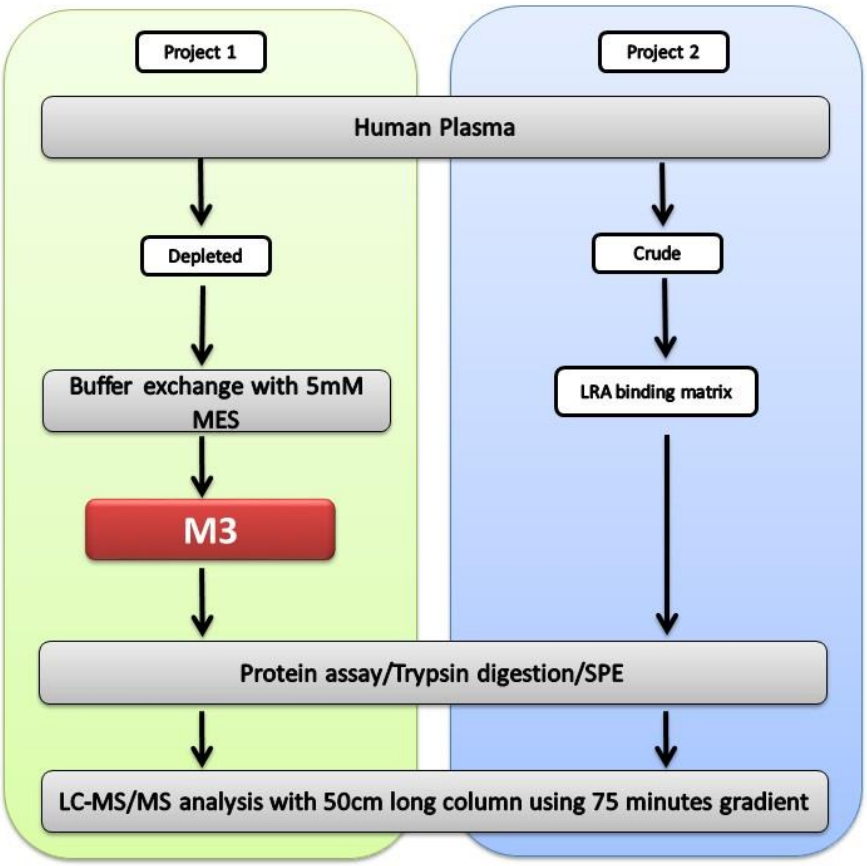


Figure 4.1. *Overview of the proteomics workflow used for proteomics analysis of the healthy volunteers and the disease samples*

4.2 Methods

4.2.1 Sample preparation

4.2.1.1 Plasma samples

Human blood samples were obtained in EDTA blood sample tubes from 10 healthy individuals, 10 HFPEF and 10 HFREF patients for project 1 and 30 healthy individuals, 30 HFPEF and 30 HFREF patients for project 2 (Figure 4.1). The protocol was approved by the national research ethics. The studies complied with the declaration of Helsinki, which is a set of ethical principles regarding research using human experimentation developed for the medical community by the world medical association (WMA) (WMA, 2013). The NRES Committee East Midlands granted approval for this study (REC reference number: 12/EM/0222) (Appendix section 7.5). The blood was centrifuged at 15,000 g at 4°C for 30 mins using a refrigerated centrifuge. The plasma was separated from blood cells and stored at -80°C until use.

4.2.1.2 Protocol for low abundance proteins with M3

Protocol for low abundance proteins binding with M3 was slightly altered due to the low numbers of proteins eluted with NaCl, (section 3.2.3) and its overlap of proteins eluted with NaH_2PO_4. Elution was carried out with 1.5 mL of 100, 200, 400 and 500 mM of NaH_2PO_4 pH 6.5. The supernatant was also retained for MS analysis.

4.2.1.3 Protocol for Lipid Removal Agent (LRA)

Materials and solution: LRA, Ammonium bicarbonate (ABC), Tris(2-carboxyethyl)phosphine hydrochloride (TCEP), IAA, ammonium salt of deoxycholic acid, (ADC), trypsin, buffer A, buffer B and raw. Fifty microlitres of plasma was bound for 1 hr to 100 µL of pre washed LRA and washed 5 times with 50 mM ABC to a final volume of 500 µL. TCEP (10 mM) was added and incubated for 5 mins prior to adding 15 mM of IAA and

incubating at room temperature in the dark for 30 mins. Five hundred µL of 1% ADC was added to give final concentration 0.5% before incubating at 60 degrees. Protein concentration was assumed at 4 mg per tube. Trypsin was added at 1:50 per tube and incubated at 37 °C for 12 h and stopped by adding 1% final concentration FA. Samples was spun hard using centrifuge at 12,000 × g for 5 mins, peptides collected in the supernatant and desalted using empore columns. Bound peptides on empore columns were eluted sequentially with 60% and 80% ACN, speed vacuumed for 2 hr and freeze dried overnight. The dried pellet was reconstituted with 0.1% FA and ADH (50 fmol/µL) for mass spectrometry analysis.

4.2.1.4 Fluorescence peptide assay

The peptide assay was carried out in a 96-well plate format. Fifty microliters of water was placed in the wells for the standards and the unknown samples. Eight milligrams of Phthaldialdehyde (OPA) was added to 100 µL of Dimethylformamide (DMF) and 10 µL of this solutions added to 1 mL of 100 mM Boric acid + 1 mg/mL Brij. Two µL of mercaptoethanol (which made the solution unstable) was added to this solution (pH 9-10). For the top standard, 90 µL of water was added to well A and 50 µL of water from well B-H. Rock peptide (1 mg/mL) was then made up in water and 10 µL added to well A. This top standard was mixed and 50 µL serial dilutions made in wells B, C, D, E and F. Wells G and H were left with only water, these were calculated as zeros. The top well contained 5 µg of peptide. Samples volume were made up to 50 µL with water and added to their respective wells. A hundred µL of Boric mixture was added to each well and read at excitation 340 nm and emission 490 nm after 2 mins incubation at room temperature. The net abundance was then acquired by subtracting the zero from all the values to get net abundance. A standard curve was constructed in FigP to determine the unknown peptide concentrations.

4.2.2 Sample analysis

All samples were analysed on a Synapt G2S coupled to a nanoacquity nanoUPLC (section 2.1.2.1) in triplicate. The protein concentration of each sample was determined prior to sample loading. These protein amounts were used to estimate how much proteins were loaded

on the column to avoid over saturation of the LC-MS system. The same volume (0.5 µL) of sample was then loaded on column to avoid any bias.

4.2.3 Data processing and analysis

MS raw data files were processed using PLGS (section 2.1.2.1.1) and Progenesis QI (Section 2.1.2.1.2). One percent FDR was used for all the samples except for HELA cells (QC) where 4% FDR was used. Data analysis was carried out using protein centre (section 2.1.3.1), SPSS (section 2.1.3.3), SIMCA (section 2.1.3.4) and RapidMiner (section 2.1.3.5).

4.3 Results

Project 1 using mixed mode matrix (M3)

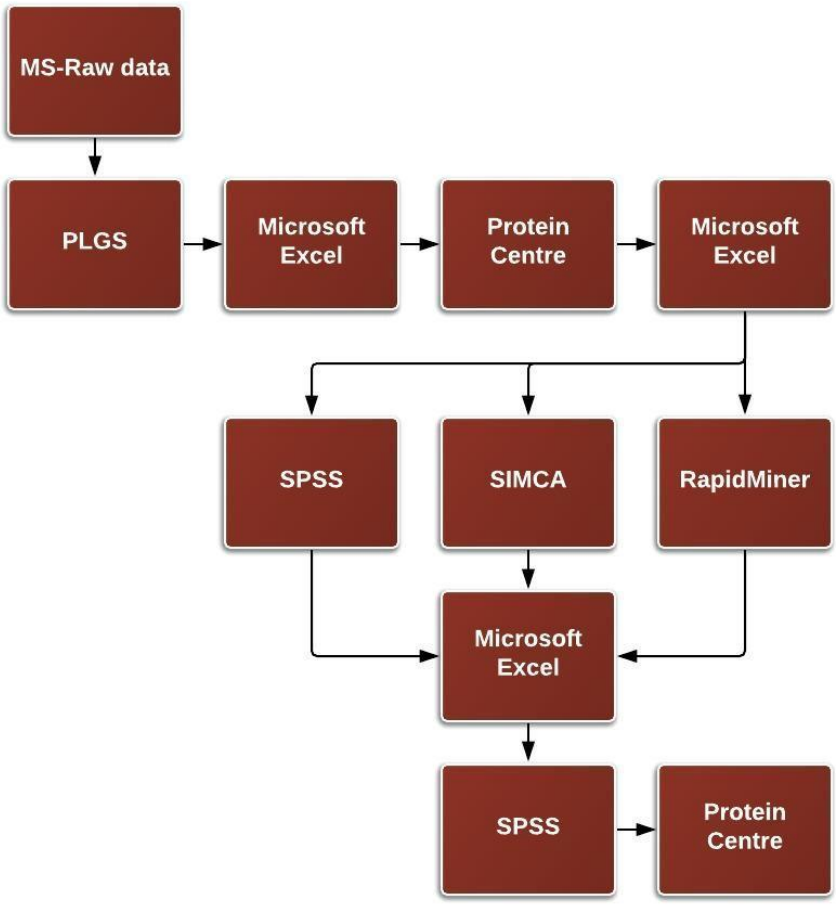


***Figure 4.2**. General data analysis work flow used in the clinical study with M3.*

4.3.1.1.1 Healthy volunteers and disease samples demographics

Table 4.1. *A summary of the 10 healthy donors, 10 heart failure with preserved ejection fraction (HFPEF) patients, and 10 heart failure with reduced ejection fraction (HFREF) patients based on their age, sex and ethnicity. eGFR readings for 7 subjects were not recorded in this data.*

Subject	Age	Sex		Creatinine	EGFR	CKD
		Male	Female			
Healthy	73.8 ± 1.6	5	5	73.1 ± 4.3	71.0 ± 1.9	1.5 ± 0.2
HFPEF	74.9 ± 2.2	5	5	102.8 ± 8.8	57.3 ± 5.7	2.4 ± 0.2
HFREF	72.4 ± 4.2	5	5	107.7 ± 8.9	53.7 ± 5.6	2.4 ± 0.3

Ethnicity

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Caucasian	26	86.7	86.7	86.7
	South Asian	3	10.0	10.0	96.7
	Black	1	3.3	3.3	100.0
	Total	30	100.0	100.0	

The levels of estimated glomerular filtration rate (eGFR) in both heart failure groups were lower than the control group (Figure 4.4). According to Bowling *et al*., 2008, low levels of eGFR are associated with poor prognosis in heart failure. Figure 4.3 shows the normal, mildly abnormal and extreme levels of eGFR resulting to kidney failure which is associated with heart failure.

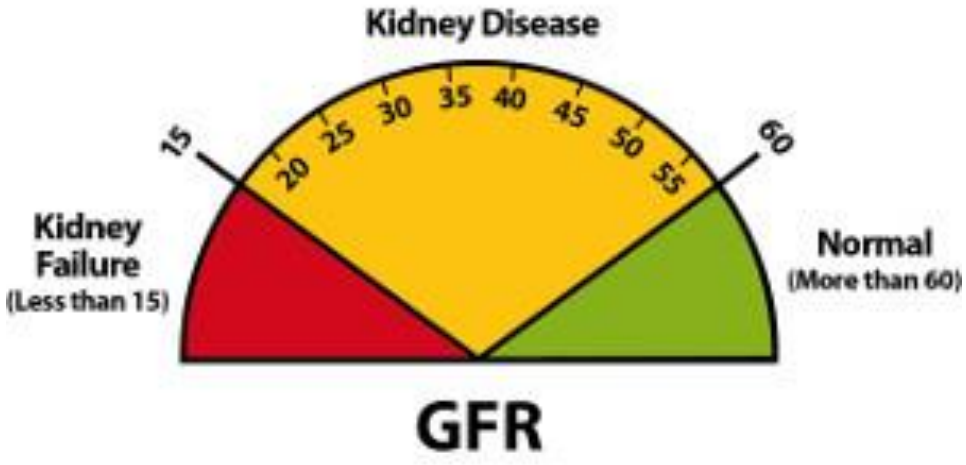


Figure 4.3. *A summary of the estimated glomerular filtration rate (eGFR) showing the normal range, kidney disease and kidney failure which are associated with heart failure. Accessed on 08/11/2016 from [https://www.niddk.nih.gov/health-information/health-communication-programs/nkdep/a-z/explaining-kidney-test-results/Pages/explaining-kidney-test-results.aspx.](https://www.niddk.nih.gov/health-information/health-communication-programs/nkdep/a-z/explaining-kidney-test-results/Pages/explaining-kidney-test-results.aspx)*

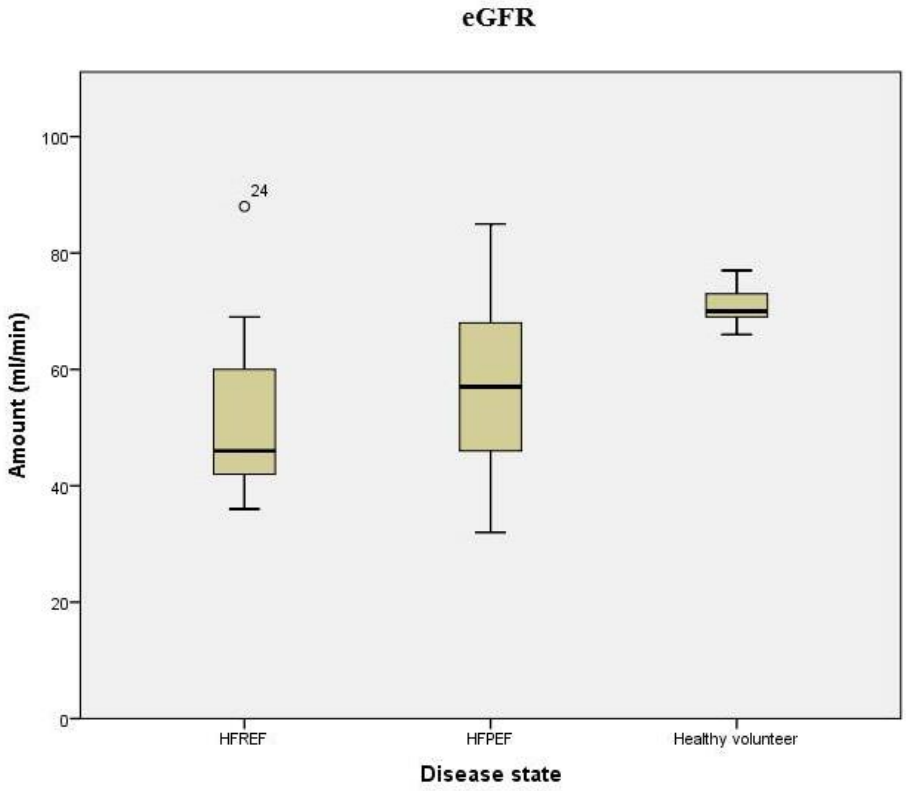


Figure 4.4. *Box plots showing the differences of the glomerular filtration rate **(**eGFR) levels between HFREF, HFPEF and healthy group. There was a significant change ($p<0.05$) between the healthy volunteers and the heart failure cohorts but there was no significant change between the heart failure groups. The mean eGFR levels for control, HFPEF and HFREF was 71.00 ml/min, 57.33 ml/min and 53.67 ml/min respectively (Table 4.1).*

4.3.2 Mixed Mode Matrix (M3) HFPEF project

In this project, 10 samples from control, HFPEF and HFREF cohorts were used. Sample preparation with M3 was carried out as shown on section 3.2.2. Elutions were adjusted as shown on section 4.2.1.2. Sample analysis was carried out as shown on Figure 4.2. PLGS was used for data processing.

All the processed data files were exported to protein centre to merge and sum the protein amounts of each sample.

4.3.2.1 Protein centre analysis

4.3.2.2 Total unique proteins obtained from protein centre

All the protein lists were imported from PLGS into protein centre in Microsoft Excel (csv format) to merge all the fractions per sample into individual list, filtering all the duplicates. One thousand four hundred and fifty unique proteins were identified when the individual protein lists obtained were merged and duplicates removed. In HFREF group, 926 proteins were identified as compared to 791 and 667 proteins from control and HFPEF groups respectively (1% FDR). Three hundred and seventy proteins were commonly identified in all three groups. The HFREF plasma had the highest number of unique proteins that was 25.7% and 53.5% more than the unique proteins in control plasma and HFPEF respectively. This increase in protein numbers with HFREF patients could be explained by the fact that these patients' heart failure was more severe than HFPEF. It has been reported that HFPEF patients' EF reduced by 5.8% over 5 years with prevalence in older patients (Sanderson, 2014). In addition, Sanderson reported that 39% of HFPEF patients had an EF of <50% after 5 years suggesting that they had developed HFREF (Figure 4.5). Therefore, it is possible that the increase in the number of proteins in HFPEF is because of HFPEF transforming into HFREF. A table (Table 4.2) and a pie chart (Figure 4.7) showing a summary of the GO molecular functions was generated on protein centre. Further data analysis was carried out of SPSS (section4.3.2.3), SIMCA (section 4.3.2.4) and RapidMiner (section 4.3.2.5).

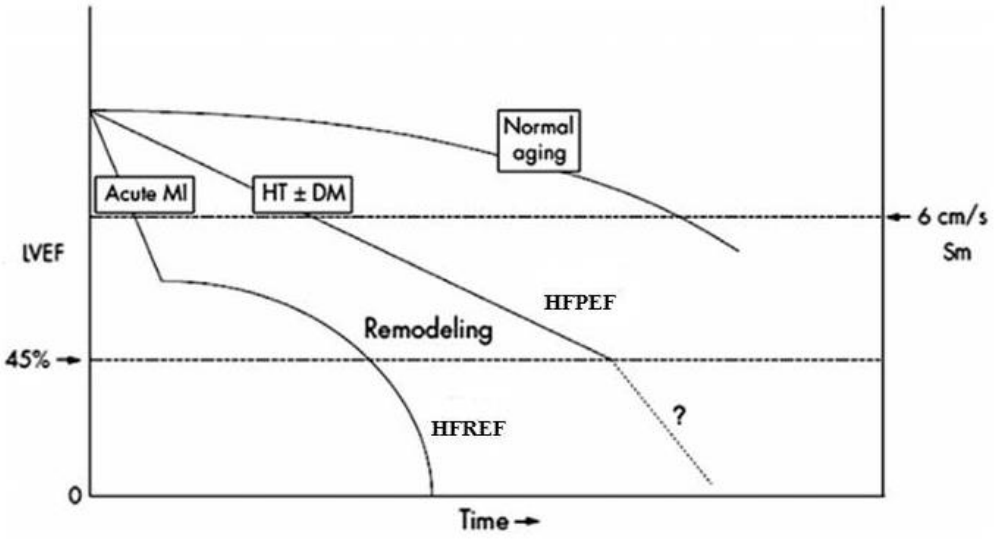


Figure 4.5. *Time course and development pattern of reduction in Left Ventricular Ejection Fraction (LVEF) in heart failure showing normal, HFPEF and HFREF. MI=myocardial infarction, HT=hypertension, DM=diabetes mellitus, HFREF=heart failure with reduced ejection fraction, HFPEF=heart failure with preserved ejection fraction (figure adapted from Sanderson, 2014).*

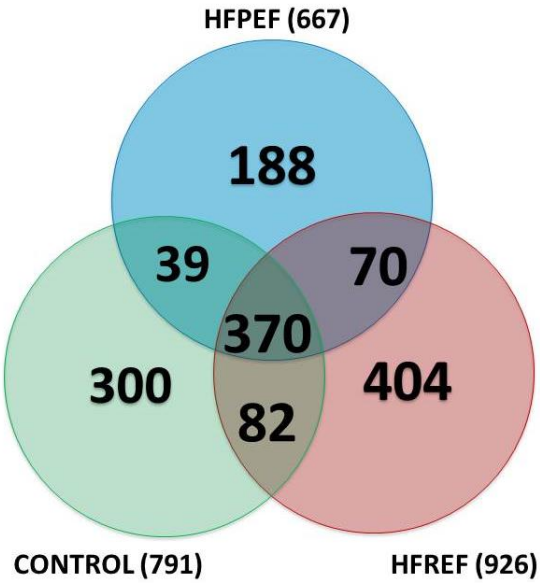


Figure 4.6. *Venn diagram showing unique and common proteins in the three different plasma sample groups. Each sample contained 5 fractions which were all analysed in triplicate. One thousand four hundred and fifty three unique proteins were identified overall.*

Table 4.2 summarises the molecular functions of the proteins identified in the analysis with M3 between control, HFPEF and HFREF groups. Most proteins were involved in protein binding with the majority observed in HFREF (83.8%), control (73.8%) and HFPEF (61.9%). Catalytic activity had the second highest protein involvement followed by metal binding which showed a similar trend in dominance within the cohort as observed in protein binding. This trend could be influenced by the total number of proteins observed per cohort as highlighted on Figure 4.6. The visual representation of these molecular functions has been shown on Figure 4.7.

Table 4.2. *Table showing a summary of the GO Slim molecular functions in the control, HFPEF and HFREF group. Most of the proteins observed in the 3 cohorts were involved in protein binding*

GO Slim Molecular Functions	Total Control	%	Total HFPEF	%	Total HFREF	%
transporter activity	49	6.1	41	5.1	58	7.3
translation regulator activity	0	0	0	0	2	0.2
structural molecule activity	73	9.2	71	8.9	79	9.9
signal transducer activity	45	5.6	27	3.4	48	6
RNA binding	54	6.8	38	4.8	49	6.1
receptor activity	61	7.7	43	5.4	59	7.4
protein binding	584	73.8	490	61.9	663	83.8
nucleotide binding	97	12.2	66	8.3	80	10.1
motor activity	37	4.6	26	3.2	26	3.2
metal ion binding	172	21.7	148	18.7	184	23.2
enzyme regulator activity	88	11.1	80	10.1	91	11.5
DNA binding	45	5.6	41	5.1	50	6.3
catalytic activity	283	35.7	219	27.6	287	36.2
antioxidant activity	13	1.6	15	1.8	15	1.8
Unannotated	71	8.9	60	7.5	85	10.7

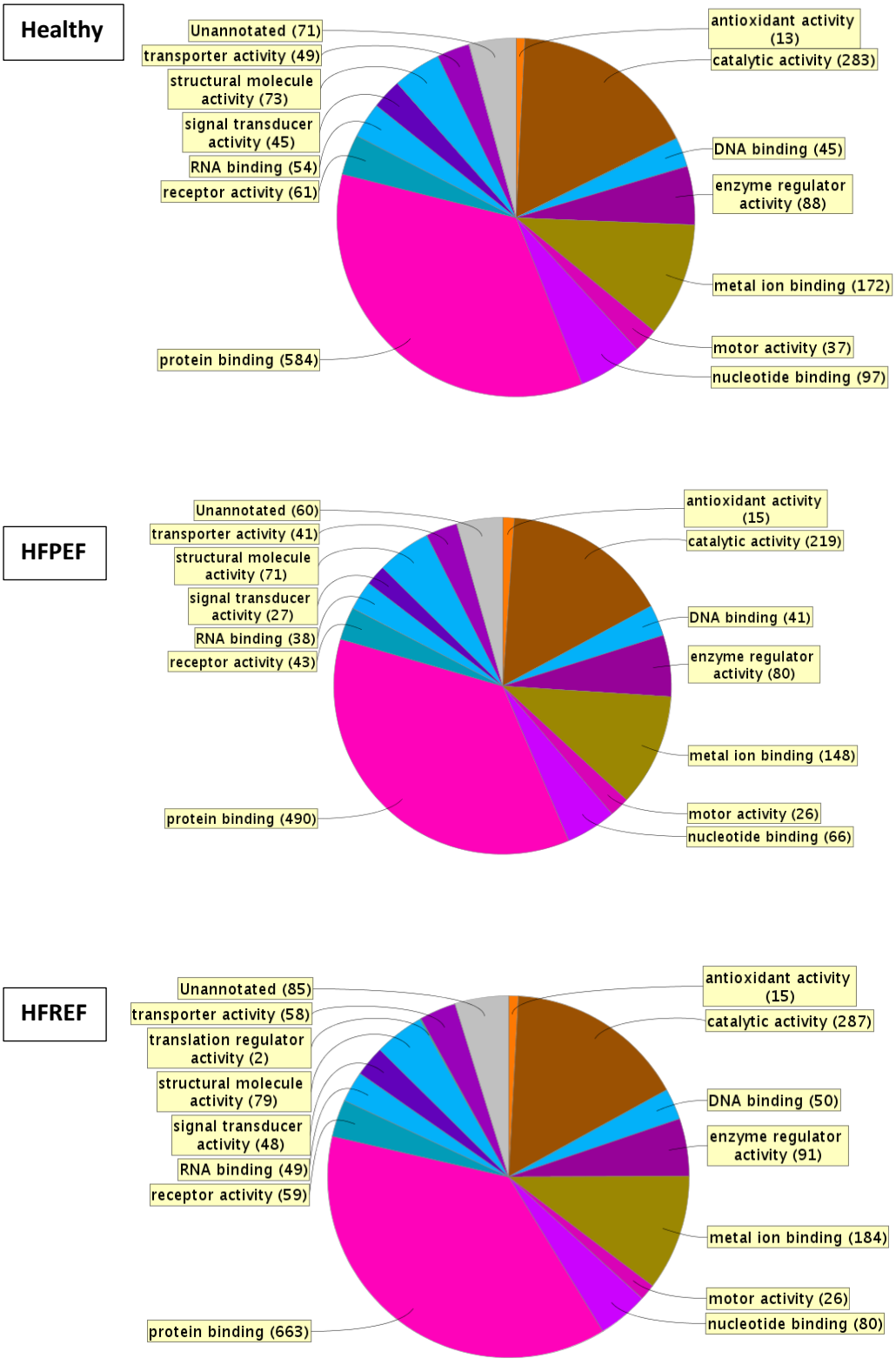


Figure 4.7. *Pie charts created in Gene Ontology term molecular function category showing distribution of human plasma proteins in the three cohorts.*

4.3.2.3 M3 HFPEF data analysis with SPSS

Prior to importing data into RapidMiner, all the HAPs and proteins not quantified were filtered to a final list of 820 proteins from the initial 1450 proteins. These 820 proteins (supplementary Table B-1) were normalised by interquartile range (IQ-range) and ANOVA (Kruskal-Wallis test) carried out on SPSS. All proteins with a p-value greater than 0.05 were filtered to a final list of 75 proteins. These 75 proteins were then imported into SIMCA (section 4.3.2.4) and RapidMiner (section 4.3.2.5) for classification analysis.

4.3.2.4 M3 HFPEF data analysis with SIMCA

In SIMCA, principle component analysis (PCA) (section 4.3.2.4.1) and partial least square (PLS) (section 4.3.2.4.2) regression analysis were performed. A model was created and data analysis carried out.

4.3.2.4.1 PCA

The PCA was used as a descriptive method to reduce the original variables into a small number, visualise correlation among the original variables and visualise proximities among statistical units (Figure 4.8) Orthogonal partial least squares-discriminant analyses (OPLS-DA) (Figure 4.9) for control subjects vs. those with HFPEF and HFREF was performed and sample runs that deviated significantly from the Hotelling's T^2 95% confidence interval were excluded and the model refitted. Peptides considered as contributing to the supervised separation of groupings were identified by consultation of the accompanying S-plot.

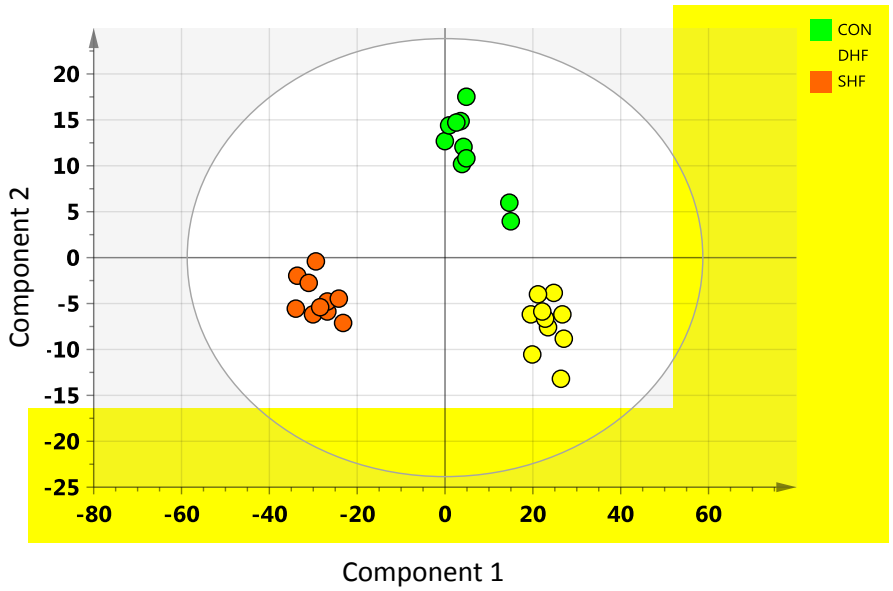


Figure 4.8. *The scores Control, DHF (HFPEF) and SHF (HFREF), are new variables summarising the X-variables (Figure 4.9). The scores are orthogonal, i.e., completely independent of each other.*

4.3.2.4.2 Partial Least Square regression (PLS) with SIMCA

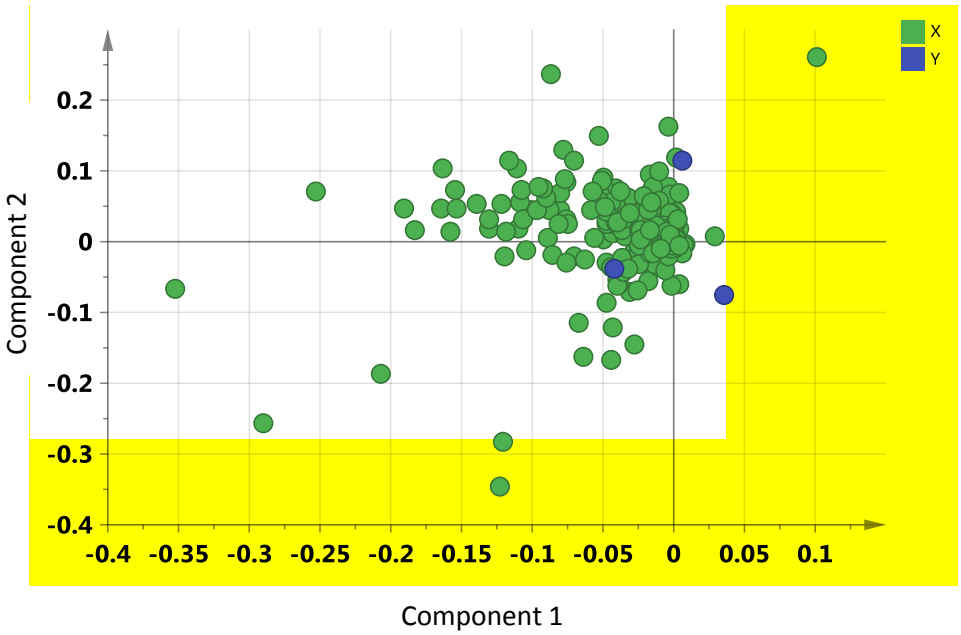


Figure 4.9. *An OPLS-DA of Control, HFPEF and HFREF showing distribution of proteins. The horizontal axis displays the X-loadings p and Y-loadings q of the predictive component. X-variables situated near the dummy Y-variables have the highest discriminatory power between the classes.*

4.3.2.5 M3 HFPEF data analysis with RapidMiner

The 75 proteins obtained from SPSS (section 4.3.2.3) were imported into RapidMiner for classification analysis. Super-operators were used for cross validation.

Support vector machine (SVM) was used for training data and apply model for testing the data. Dot kernel 1 and C parameter 0 were used. This SVM method was in Table 4.3 and Table 4.4.

Firstly, the data set was tested using genetic algorithm (GA) with voted SVM (Table 4.3) before testing the same data set with forward selection cross validation 10 times (Table 4.4)

When SVM with GA was used to classify control vs. HFPEF vs. HFREF groups, a predictive accuracy of 93.33% was achieved with false positives in HFPEF (Table 4.3). Conversely, when SVM with polynomial kernel was used, a predictive accuracy of 90% was achieved with false positives in control and HFPEF groups (Table 4.4).

***Table 4.3**. Classification analysis of the Con vs. HFPEF vs. HFREF using genetic algorithm (GA) with voted support vector machine (SVM).*

Con vs. HFPEF vs. HFREF				
Using SVM dot kernel C=1				
		GA p initialise 0.9		
Accuracy 93.33%				
	True Control	True HFPEF	True HFREF	Class precision
Predictive Control	10	0	1	90.91%
Predictive HFPEF	0	10	1	90.91%
Predictive HFREF	0	0	8	100.00%
Class recall	100.00%	100.00%	80.00%	

Proteins identified using SVM with GA.

Plasma protease C1 inhibitor
Afamin
Complement component C9
Alpha2 anti-plasmin
Keratin typeII cytoskeletal 1
Hepatocyte growth factor activator
Apolipoprotein F
Complement factor I
Complement factor D
Histidine rich glycoprotein
Actin cytoplasmic 1
Serum amyloid P component
Complement factor H
Plasma kallikrein

Table 4.4. *Classification analysis of the Con vs. HFPEF vs. HFREF using polynomial kernel with support vector machine (SVM). Cross validation was used in this analysis. C is the parameter for nonlinear SVM.*

Con vs. HFPEF vs. HFREF				
Using SVM dot kernel C=1				
	Cross validated 10x			
Accuracy 90%				
	True Control	True HFPEF	True HFREF	Class precision
Predictive Control	9	0	1	90.00%
Predictive HFPEF	1	10	1	83.33%
Predictive HFREF	0	0	8	100.00%
Class recall	90.00%	100.00%	80.00%	

Proteins identified using SVM with polynomial kernel.

Actin, cytoplasmic 1
Alpha-1B-glycoprotein
Extracellular matrix protein 1
Isoform 2 Alpha-1B glycoprotein
Isoform2Vitamin D binding protein
Isoform 4 CDK5 regulatory subunit associated protein 2
Kallistatin
Protein Z dependent protease inhibitor
Retinol-binding protein 4

Using M3 not only improved the dynamic range but also increased the number of LAPs in plasma. However, this method had a number of challenges. Firstly, plasma had to be depleted (immunodepletion), bound on the M3 matrix for 12- 14 h, eluted and cleaned up using SPE. These suggested that there were too many steps involved, which could have led to potential sample loss. Secondly, this fractionation method introduced reproducibility issues since proteins eluted were smeared across all the fractions. The fact that plasma was subjected to immunodepletion could mean that some low abundant proteins could have been lost through non-specific binding on the depletion column. Nonetheless, the LRA matrix that had previously been used in method development (chapter 3.3.2) had been subsequently optimised and proved a much better matrix for plasma samples analysis. This LRA method not only improved throughput but also was very reproducible which alleviated fractionation and normalisation issues. The addition of the 1% ADC in the LRA protocol alongside TCEP aided the unfolding of the protein structures prior to tryptic digestion thus improving the protein numbers.

4.3.3 LRA HFPEF project

Project 2 using the LRA matrix.

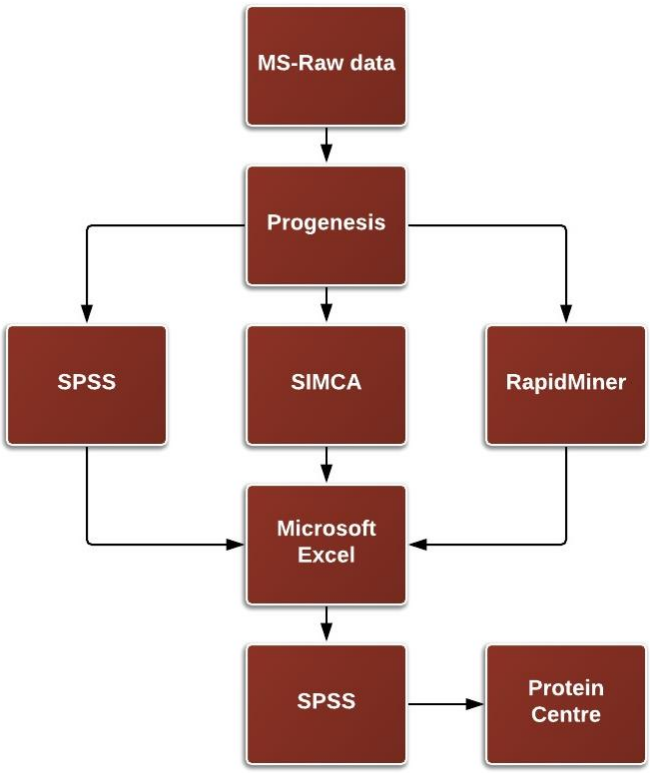


Figure 4.10. *General data analysis work flow used in the clinical study with LRA.*

Raw data files were processed on Progenesis (section 2.1.2.1.2). About 800 unique proteins were identified from the 3 cohorts (Control n=30, HFPEF n=30 and HFREF n=30). All the proteins that were not quantified and the top 20 most abundant proteins were filtered to a final list of 664 proteins (supplementary Table B-2). SPSS, SIMCA and RapidMiner were then used for data mining and protein centre for identified the molecular pathways of the significant proteins identified (Figure 4.10).

Table 4.5. *A summary of the 30 healthy donors, 30 heart failure with preserved ejection fraction (HFPEF) patients, and 30 heart failure with reduced ejection fraction (HFREF) patients based on their age, sex and ethnicity.*

Subject	Age	Sex		Creatinine	EGFR	CKD
		Male	Female			
Healthy	72.5 ±0.9	30	30	74.1 ± 3.7	74.4 ± 2.5	1.5 ± 0.1
HFPEF	70.8 ± 1.5	30	30	94.5 ± 6.3	64.8 ± 3.8	2.3 ± 0.1
HFREF	72.5 ± 1.5	30	30	101.5 ± 5.4	71.2 ± 3.2	2.5 ± 0.1

Ethnicity

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Black	1	1.1	1.1	1.1
	Caucasian	83	92.2	92.2	93.3
	South Asian	6	6.7	6.7	100.0
	Total	90	100.0	100.0	

Data analysis was carried out as shown with SPSS (section 4.3.3.1), SIMCA (section 4.3.3.2) and RapidMiner (section 4.3.3.3).

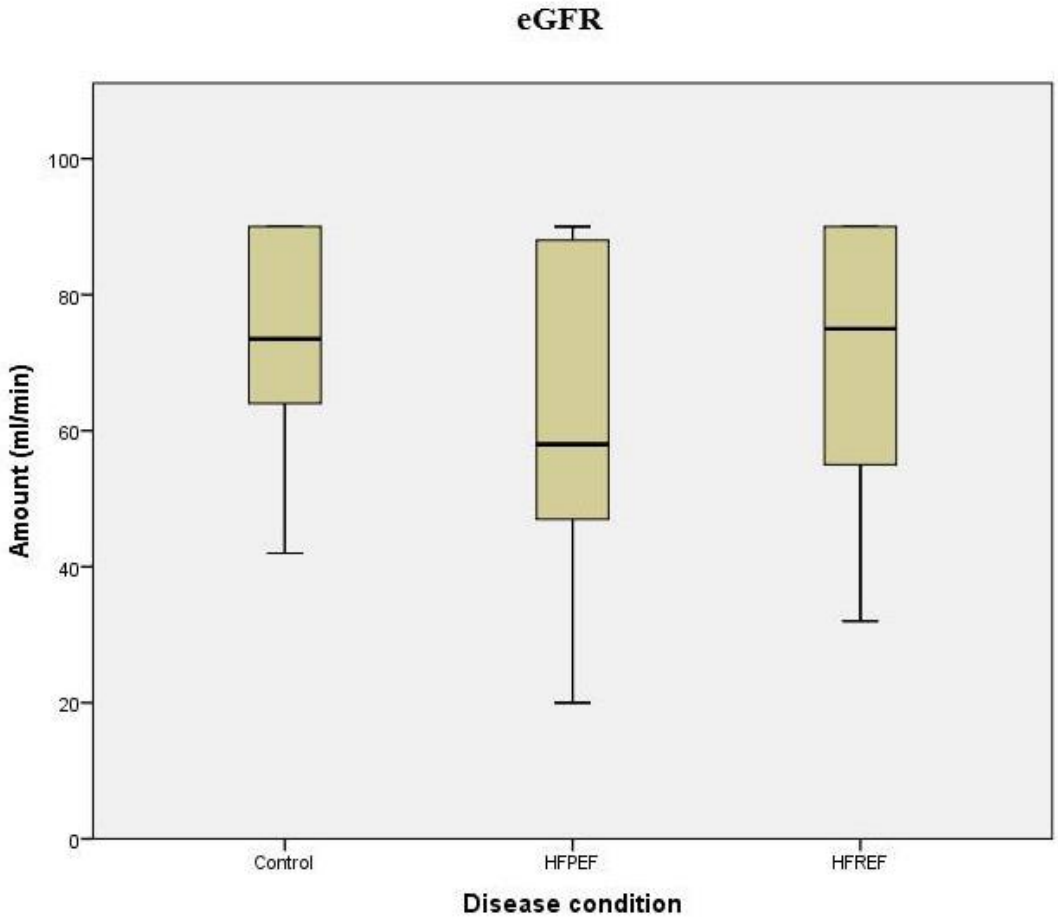
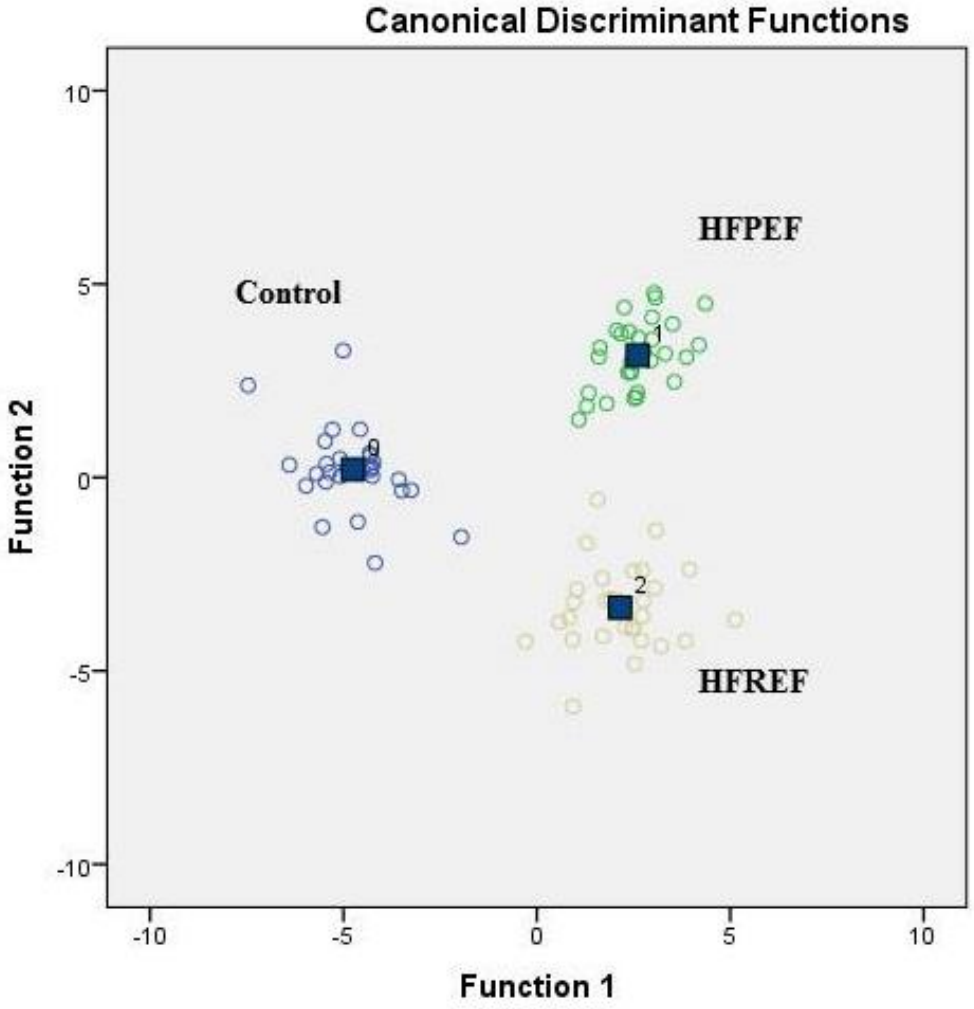


Figure 4.11. *Box plots showing the differences of the glomerular filtration rate* ***(****eGFR) levels between HFREF, HFPEF and healthy group. There was a significant change between control and HFPEF ($p<0.05$) but there was no significant change between control and HFREF. The mean eGFR levels for control, HFPEF and HFREF was74.43 ml/min, 64.77 ml/min and 71.17 ml/min respectively (Table 4.5).*

Discriminant analysis (Figure 4.12) between the three cohorts showed that three groups could be classified using statistical models. Molecular features that differed between the three groups were responsible for this discrimination. To try to elucidate the features responsible for the separation a variety for statistical approaches were applied.



	Function	
	1	2
Control	-4.758	.207
HFPEF	2.606	3.157
HFREF	2.152	-3.365

Unstandardized canonical discriminant functions evaluated at group means.

Figure 4.12. *Disciminant analysis of control, HFPEF and HFREF cohorts showing a good separation. Irrespective of the method used, we could still discriminate between the 3 cohorts.*

4.3.3.1 SPSS data analysis for LRA HFPEF project

Eight hundred unique proteins were identified from the 3 cohorts (Control n=30, HFPEF n=30 and HFREF n=30) with progenesis. All the proteins that were not quantified and the top 20 most abundant proteins were filtered to a final list of 664 proteins (supplementary Table B-2). These 664 proteins (supplementary Table B-1) were normalised by interquartile range (IQ-range) and ANOVA (Kruskal-Wallis test) carried out. All proteins with a p-value greater than 0.05 were filtered to a final list of 80 proteins (supplementary Table C-2). The following analysis were carried out with the 80 proteins:

- Logistic regression between control and HF groups (Table 4.6 and Figure 4.13).

- Logistic regression between HFPEF and HFREF (Table 4.7 and Figure 4.14).

Control vs. Heart Failure (HFPEF and HFREF)

Logistic regressions were performed to calculate the probabilities of heart failure prediction for each individual peptide of interest, and as a combination of all these proteins. Receiver operator characteristic curves (ROCs) were produced using these probabilities and the areas under the curve (AUC) were calculated. All tests with a two-tailed p value of <0.05 were deemed as statistically significant.

Firstly, logistic regression analysis for control vs. HF group using 80 proteins (section 4.3.3.1) was done using SPSS. Ten significant proteins that predicted control from heart failure samples were identified (Table 4.6 and Figure 4.13). NCBP2L,HIST1H2AH, DMRT2, COL18A1, TUBAC3, ACSM5 and PDCD6 were found to be statistically significant. Thus, predicted healthy samples from disease. Another analysis was carried out to predict HFPEF group from HFREF group using the 80 proteins. Among the proteins measured AKR1D1, SERPINA3 and SPATA5L1 (Table 4.8) were significant predictors of HFPEF and HFREF. The risk prediction was improved by combining the three proteins as shown on Table 4.8 and Figure 4.13.

***Table 4.6**. Logistic regression for control vs. heart failure groups. The test result variables: HIST1H2AH and COL18A1 have at least one tie between the positive actual state group and the negative actual state group. a. under the nonparametric assumption. b. Null hypothesis: true area =0.5. Roc curve is shown on the figure below. ASL=Isoform 2 of Argininosuccinate lyase, NCBP2L=Nuclear cap-binding protein subunit 2-like, HIST1H2AH=Histone H2A type 1-H, CBX3=Chromobox protein homolog 3, DMRT2=Doublesex- and mab-3-related transcription factor 2, COL18A1=Isoform 3 of Collagen alpha-1(XVIII) chain, TUBAC3=Tubulin alpha-3C/D chain, LUC7L3=Luc7-like protein 3, ACSM5=Acyl-coenzyme A synthetase ACSM5, mitochondrial and PDCD6=Isoform 2 of Programmed cell death protein 6.Std=standard, sig.=significance.*

Test Result variable	Area	Std Error	Significance	95% Confidence Interval	
				Lower bound	Upper Bound
ASL	0.183	0.048	0.000	0.088	0.278
NCBP2L	0.687	0.062	0.004	0.566	0.808
HISTH2AH	0.633	0.066	0.040	0.503	0.763
CBX3	0.274	0.054	0.001	0.169	0.380
DMRT2	0.667	0.060	0.010	0.549	0.786
COL18A1	0.578	0.066	0.231	0.449	0.706
TUBA3C	0.648	0.063	0.023	0.525	0.771
LUC7L3	0.329	0.058	0.009	0.215	0.444
ACSM5	0.589	0.064	0.168	0.464	0.715
PDCD6	0.669	0.060	0.009	0.552	0.786
Probability (HFvsCON)	0.000	0.000	0.000	0.000	0.000

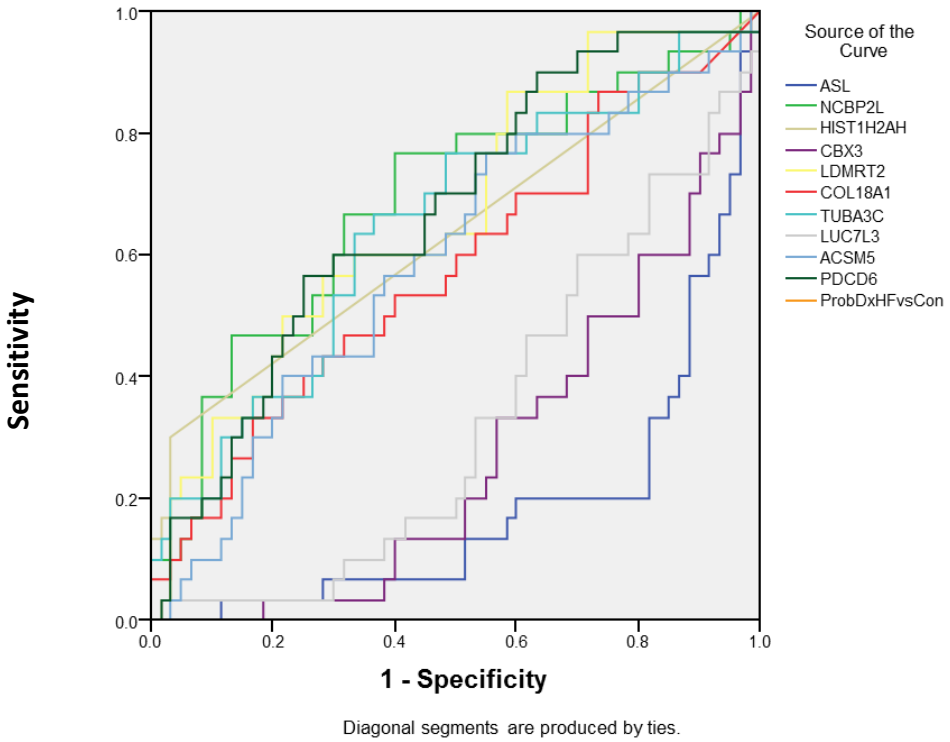


***Figure 4.13**. Roc curve showing significant peaks (10 proteins) after Kruskal-Wallis (KW) test on 664 proteins and regression analysis using the 80 significant proteins identified.*

HFPEF vs. HFREF

Table 4.7. *Logistic regression analysis for detecting HFPEF vs. HFREF groups. The test result variables: AKR1D1and SPATA5L1 have at least one tie between the positive actual state group and the negative actual state group a. under the nonparametric assumption. b. Null hypothesis: true area =0.5. Roc curve is shown on the figure below. AKR1D1=3-oxo-5-beta-steroid 4-dehydrogenase, SERPINA3=Alpha-1-antichymotrysin and SPATA5L1= Spermatogenesis-associated protein 5-like protein 1. Std=standard, sig. =significance.*

Test Result variable	Area	Std Error	Significance	95% Confidence Interval	
				Lower bound	Upper Bound
AKR1D1	0.746	0.066	0.001	0.617	0.874
SERPINA3	0.703	0.069	0.007	0.569	0.838
SPATA5LT	0.740	0.067	0.001	0.609	0.871
Probability (HFPEFvsHFREF)	0.832	0.054	0.000	0.727	0.938

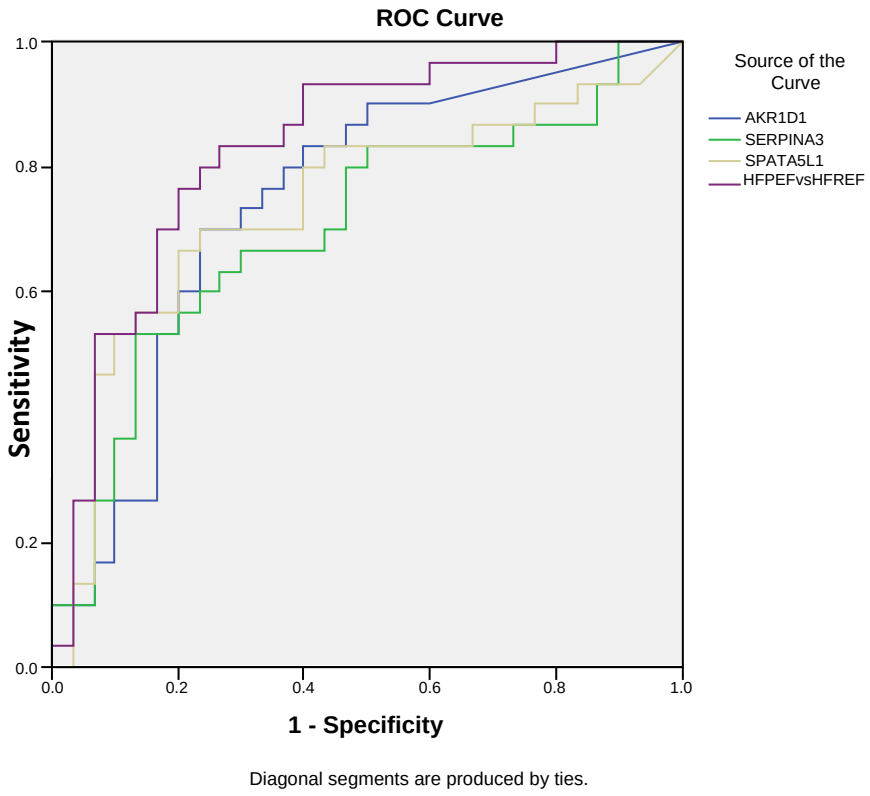


Figure 4.14*. Three potential biomarkers combined and ROC curve generated*

The proteins observed in the regression analysis between control and HF groups and HFPEF vs. HFREF groups were added to a final list of proteins obtained from SIMCA and

RapidMiner for pathway analysis study. An expression analysis was performed to ascertain their levels in control and disease groups (Figure 4.19-Figure 4.22).

4.3.3.2 SIMCA data analysis for LRA HFPEF project

In SIMCA, a multivariate analysis (PCA and OPLS-DA) (Figure 4.15-Figure 4.18) was carried out using the 664 proteins identified (section 4.3.3.1). A near complete separation was observed when the three cohorts were analysed together. The proteins that influenced the separation of the three cohorts are shown on Figure 4.16.

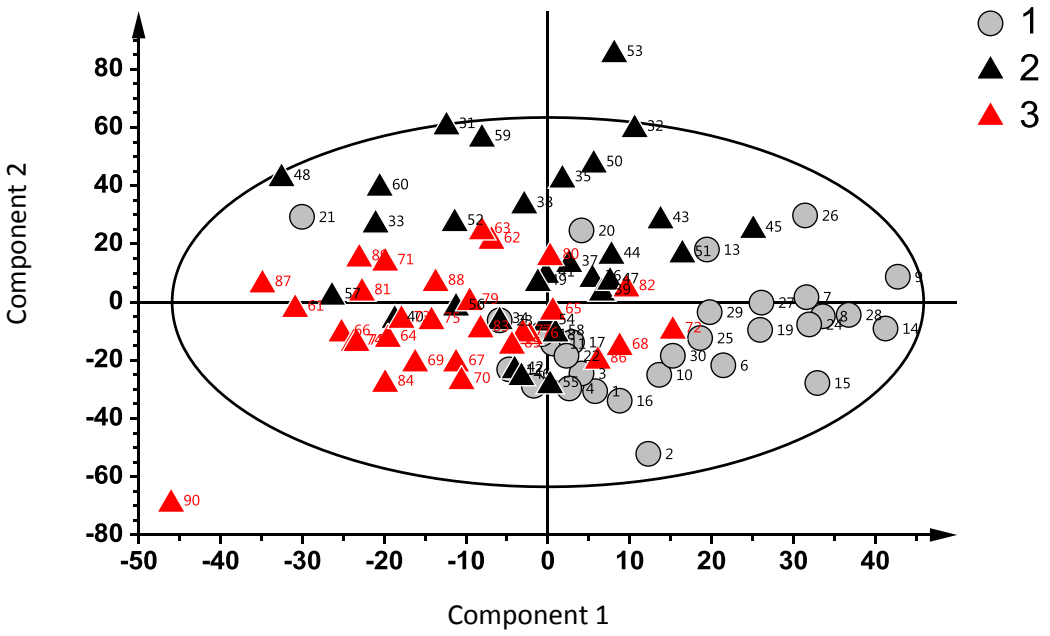


Figure 4.15. *A score plot of control (1), HFPEF (2) and HFREF (3) groups showing distribution of 30 patients per group. SIMCA drew a confidence ellipse based on Hotelling's T2, by default at significance level 0.05. Observations situated outside the ellipse are considered outliers.*

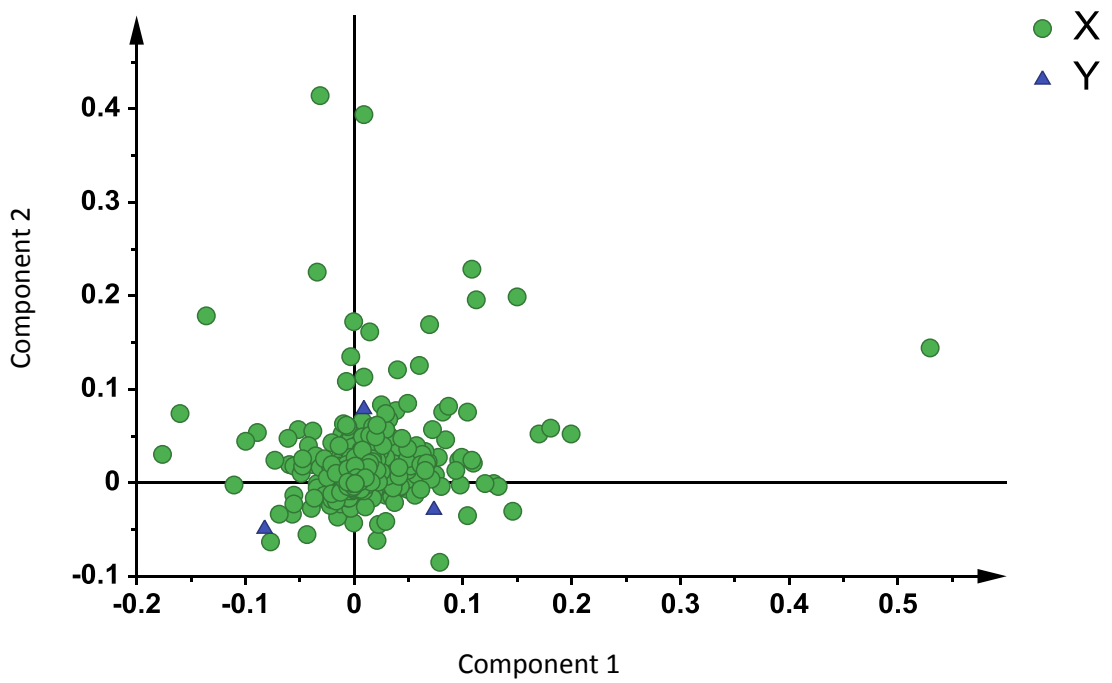


Figure 4.16. *An OPLS-DA of Control, HFPEF and HFREF showing distribution of 664 unique proteins. The horizontal axis displays the X-loadings p and Y-loadings q of the predictive component. X-variables situated near the dummy Y-variables have the highest discriminatory power between the classes.*

The classification between control vs. HF groups and HFREF vs. HFPEF group is shown on (PCA) and (OPLS-DA). All the observations that were outside the ellipse for example sample 32 and 90 on were considered outliers and were eliminated in the data set. Significant proteins contributing mostly to the separation between control vs. HF group and HFPEF vs. HFREF were identified (Table 4.8) by analysing the loadings distribution, covariance verses correlation. These proteins were added to the proteins obtained from RapidMiner and SPSS (supplementary Table C-1) for further investigation of their roles in HF. The molecular functions of these proteins (Table 4.11) were obtained from protein centre GO.

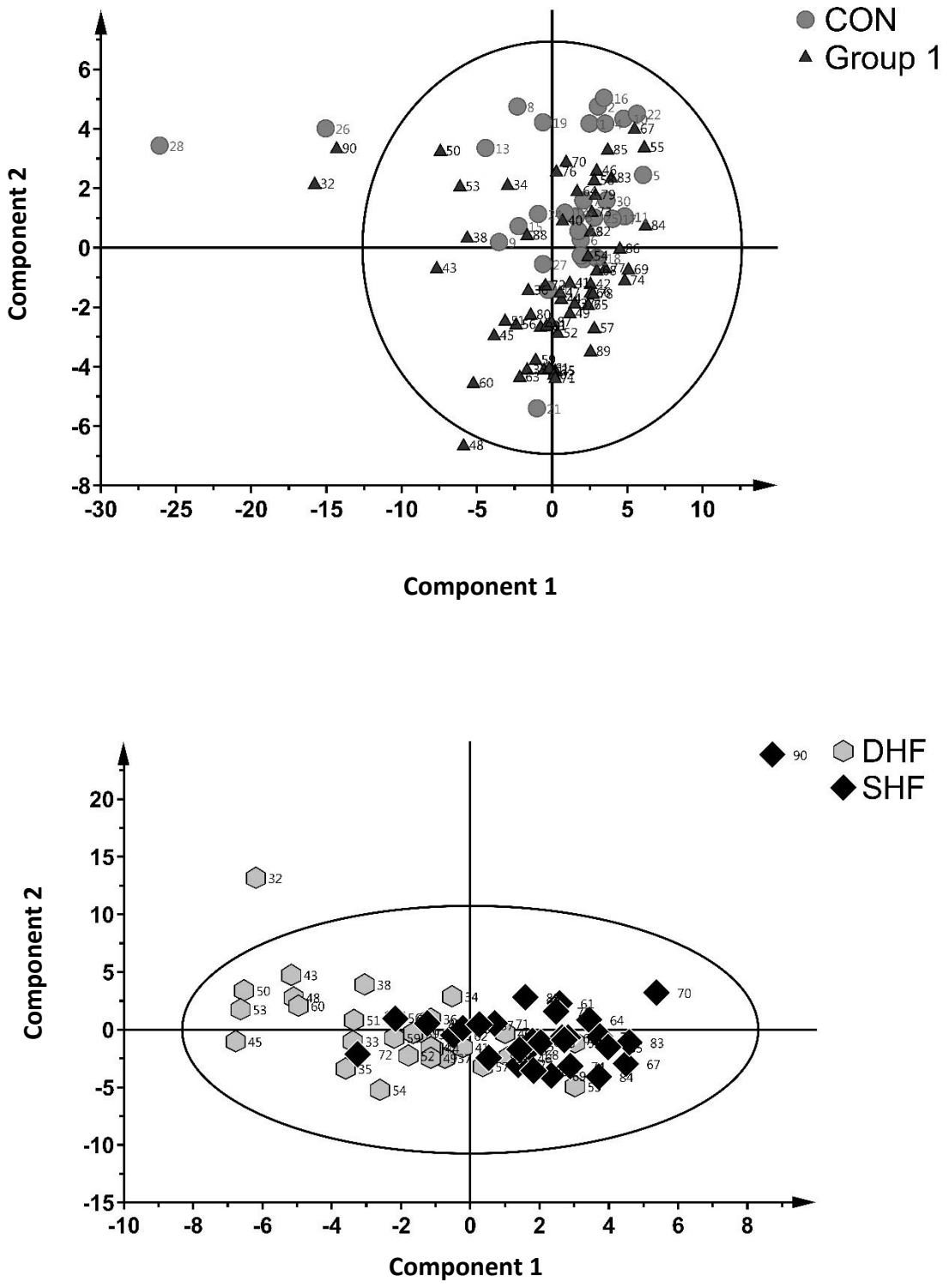


Figure 4.17*. PCA Score plots of control vs HF (group 1) n=90, control vs. DHF (HFPEF) n=60 and DHF vs. SHF (HFREF) n=60 for 74 variables. SIMCA drew a confidence ellipse based on Hotelling's T2, by default at significance level 0.05. Observations situated outside the ellipse are considered outliers.*

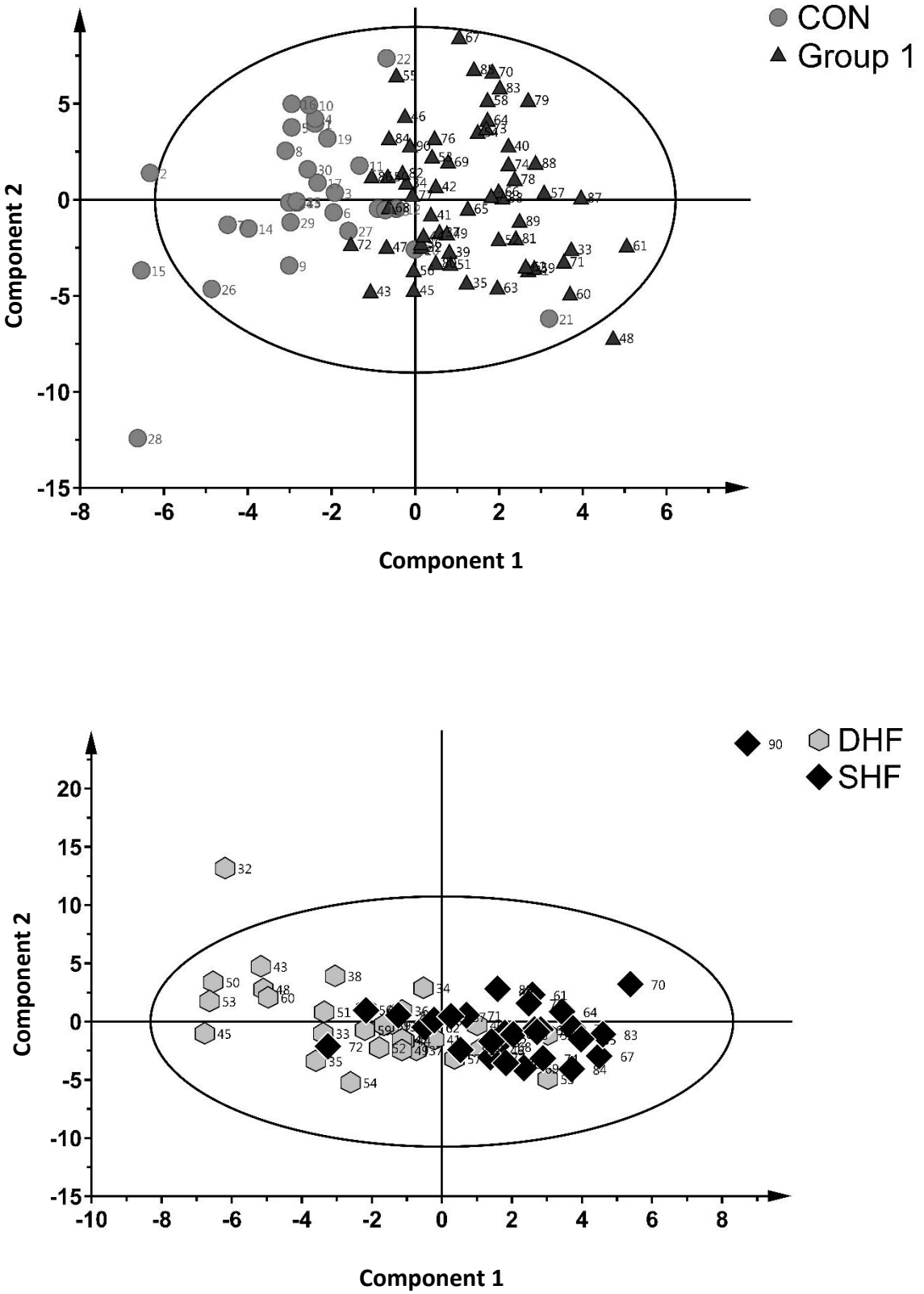


Figure 4.18. *OPLS-DA score plots of control vs. HF (group 1) n=90, control vs. DHF (HFPEF) n=60 and DHF vs. SHF (HFREF n=60 for 74 variables. SIMCA drew a confidence ellipse based on Hotelling's T2, by default at significance level 0.05. Observations situated outside the ellipse are considered outliers.*

***Table 4.8**. A list of the significant proteins picked from SIMCA using non-parametric test.*

CON VS. HF	
ACCESION NO.	**Protein description**
P04424-2	Isoform 2 of Argininosuccinate lyase (ASL)
Q52LW3	Rho GTPase-activating protein 29 (ARHGAP29)
Q6ZN30	Zinc finger protein basonuclin-2 (BNC2)
Q92538	Golgi-specific brefeldin A-resistance guanine nucleotide exchange factor 1 (GBF1)

HFPEF VS. HFREF	
ACCESSION NO.	**Protein description**
Q9Y5X5	Neuropeptide FF receptor 2 (NPFFR2)
Q9H2M9	Rab3 GTPase-activating protein non-catalytic subunit (RAB3GAP2)
A6NK21	Putative zinc finger protein LOC730110
Q8IWJ2	GRIP and coiled-coil domain-containing protein 2 (GCC2)
O00391	Sulfhydryl oxidase 1 (QSOX1)
P31942	Heterogeneous nuclear ribonucleoprotein H3 (HNRNPH3)

4.3.3.3 RapidMiner analysis for the LRA HFPEF project

The 75 proteins obtained from SPSS (section 4.3.2.3) were imported into RapidMiner for classification analysis. Super-operators were used for cross validation. Support vector machine (SVM) was used for training data and apply model for testing the data. Dot kernel 1 and C parameter 0 were used. This SVM method was in Table 4.9 and Table 4.10.

When control vs. heart failure groups were classified (Table 4.9), there was an 80% chance of predicting control from heart failure group. This gave an overall predictive accuracy of 91%. When HFPEF vs. HFREF groups were classified (Table 4.10), there was an 86.67% chance of predicting HFPEF from HFREF group. This gave an overall predictive accuracy of 78.33%. The proteins identified in these classifications were shortlisted in the final analysis as potential biomarkers.

Table 4.9. *Classification analysis of the control vs. heart failure groups using polynomial kernel with support vector machine (SVM). C is the parameter for nonlinear SVM.*

CON VS. ALL HF

	SVM C=1 dot kernel		
	Cross validated 10x		
ACCURACY 91%	True CON	True HF	Class precision
PRED. CON	24	2	92.31%
PRED. HF	6	58	90.62%
CLASS RECALL	80.00%	96.67%	
ACCESSION NO.	**Protein description**		
Q6NUN0	Acyl-coenzyme A synthetase ACSM5, mitochondrial		
P05413	Fatty acid-binding protein, heart,FAB3 (FABP)		
P22352	Glutathione peroxidase 3 (GPX3)		
P31942	Heterogeneous nuclear ribonucleoprotein (HNRNPH3)		
P04424-2	Isoform 2 of Argininosuccinate lyase (ASL)		
P49796-2	Isoform 2 of Regulator of G-protein signaling 3		
Q8TCZ2-4	Isoform 4 of CD99 antigen-like protein 2		
A6NK21	Putative zinc finger protein LOC730110 (LOC730110)		

Table 4.10. *Classification analysis of the HFPEF vs. HFREF groups using logistic on RapidMiner to forward select.*

HFPEF VS. HFREF

	SVM C=1 dot kernel		
ACCURACY 78.33%	True HFPEF	True HFREF	Class precision
PREDICTIVE DHF	26	9	74.29%
PRED. SHF	4	21	84.00%
CLASS RECALL	86.67%	70.00%	
ACCESION NO.	**Protein description**		
P51857	3-oxo-5-beta-steroid 4-dehydrogenase (AKR1D1)		
P23528	Cofilin-1 (CFL1)		
Q6X9E4	F-box/WD repeat-containing protein 12 (FBXW12)		
Q15485	Ficolin-2 (FCN2)		
P39060-2	Isoform 3 of Collagen alpha-1(XVIII) chain (COL18A1		
Q86UD1	Out at first protein homolog (OAF)		
P80108	Phosphatidylinositol-glycan-specific phospholipase D (GPLD1)		
Q9BVQ7	Spermatogenesis-associated protein 5-like protein 1 (SPATA5L1)		

The proteins observed in the classification analysis between control and HF groups and HFPEF vs. HFREF groups were added to a final list of proteins obtained from SPSS and SIMCA. Only the top 30 significant proteins were selected from the list for further studies. An expression analysis was performed to ascertain their levels in control and disease groups. Molecular functions of these proteins were also ascertained and from Table 4.11 it is evident that most proteins were involved in protein binding, catalytic and metal binding. The least involvement in the molecular functions was seen in antioxidant, enzyme regulator, transport and structural molecule. In the expression analysis (Figure 4.22), 5 of the 30 proteins shortlisted for biomarkers investigation had very low protein levels hence could not be analysed on box plots. These include AKR1D1, MUCL1, ZNF201, HIST1H2AH and COL18A1. The remaining 25 proteins had either been up or down regulated in the control or HF (HFPEF/HFREF) groups.

Table 4.11. *List of molecular functions of the 30 potential biomarkers that could discriminate between the 3 cohorts. AIG1, CBX3, GBF1 and NPFFR2 were also only involved in protein binding. LSMD1 and MUC1 were not categorised in this analysis.*

	Antioxidant	Catalytic	DNA binding	Metal binding	Nucleotide binding	Protein binding	RNA binding	Enzyme regulator	Transporter	Structural molecule
ASL	✓	✓				✓				
GPX-3	✓	✓				✓				
ACSM5		✓		✓	✓					
PDC6		✓		✓		✓				
MYH9		✓			✓	✓	✓			
TUBA3C		✓			✓	✓				✓
AKR1D1		✓								
CFD		✓								
QSOX1		✓								
DMRT2			✓	✓		✓				
BNC2			✓	✓						
ZNF701			✓	✓						
Luc7			✓			✓	✓			
HIST1H2AH			✓			✓				
ARHGAP29				✓		✓		✓		
KCN1P2				✓		✓			✓	
LOC730110				✓						
HNRNPH3					✓	✓	✓			
NCBP2L					✓		✓			
C4BPA						✓	✓			
Col18a1						✓	✓			
RAB						✓		✓		
FABP						✓			✓	
GCC2						✓				✓

An area under the curve analysis was done for the 30 proteins idenfied on Table 4.11. This analysis was implemented to predict control from heart failure group (HFPEF/HFREF) and HFPEF from HFREF group as shown Table 4.12 and Table 4.13 respectively.

Eight proteins (AKR1D1, KCNIP2, MUCL1, ZNF701, HIST1H2AH, CBX3, COL18A1 and GBF1) were found to be significant in predicting control patients from heart failure patients (Table 4.12).

On the other hand, 5 proteins (AKR1D1, KCNIP2, MUCL1, HIST1H2AH and COL18A1) were found to be significant in predicting HFPEF from HFREF group (Table 4.13). These 5 proteins were combined and compared with 30 proteins combined in predicting HFPEF. There was a 22% (Table 4.14) increase in prediction of HFPEF when 30 proteins were used together as compared to 5 proteins.

Table 4.12 *A list of proteins with their area under the curve (AUC). The ROC curve analysis for prediction of control from heart failure group is also shown. The test result variable(s): AKR1D1, KCNIP2, MUCL1, ZNF701, HIST1H2AH, CBX3, COL18A1 and GBF1 have at least one tie between the positive actual state group and the negative actual state group.*

Area Under the Curve

Test Result Variable(s)	Area
ASL	.818
DMRT2	.333
NCBP2L	.313
HNRNPH3	.256
AKR1D1	.583
C4BPA	.687
CFD	.618
FABP3	.268
GPX3	.350
AIG1	.331
KCNIP2	.317
LSMD1	.374
MUCL1	.339
MYH9	.324
ZNF701	.654
HIST1H2AH	.367
CBX3	.725
COL18A1	.422
TUBA3C	.352
LUC7L3	.671
ACSM5	.411
PDCD6	.331
ARHGAP29	.235
BNC2	.711
GBF1	.744
NPFFR2	.342
RAB3GAP2	.547
LOC730110	.546
GCC2	.529
QSOX1	.616

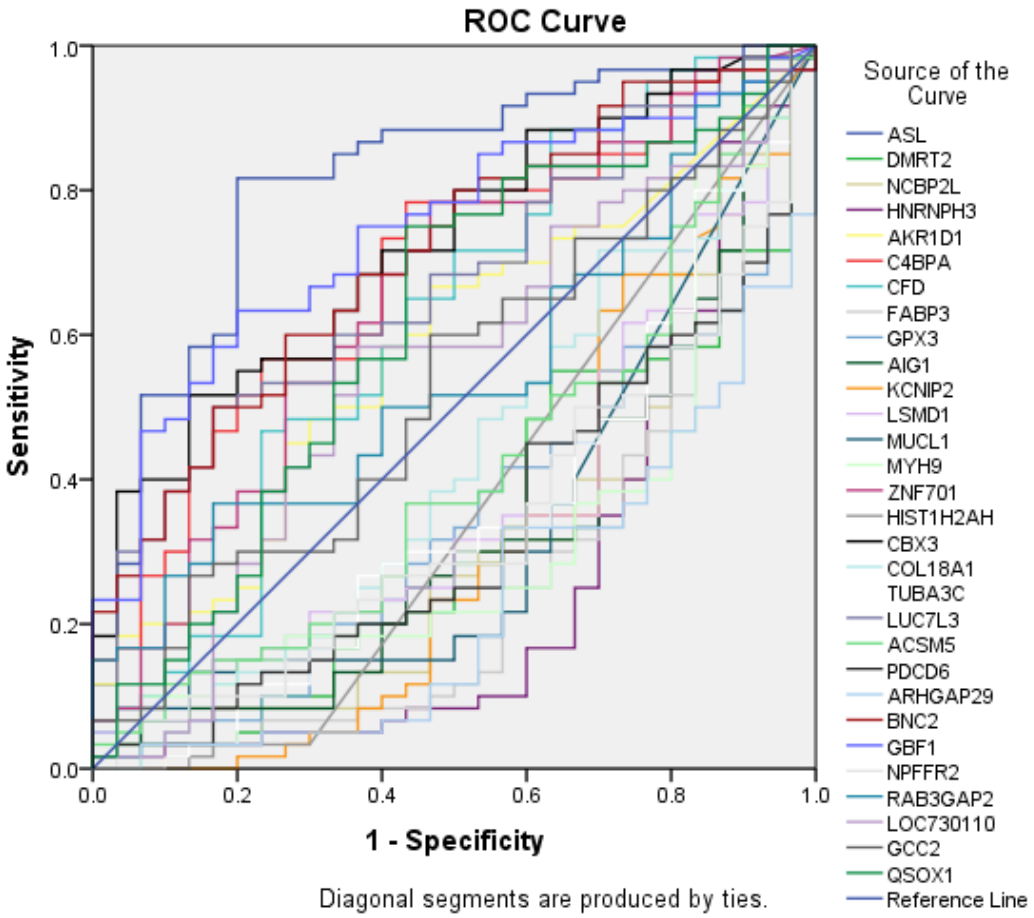


Table 4.13. *A list of proteins with their area under the curve (AUC). The ROC curve analysis for prediction of HFPEF from HFREF is also shown. The test result variable(s): AKR1D1, KCNIP2, MUCL1, HIST1H2AH and COL18A1 has at least one tie between the positive actual state group and the negative actual state group.*

Area Under the Curve

Test Result Variable(s)	Area
ASL	.439
DMRT2	.329
NCBP2L	.323
HNRNPH3	.350
AKR1D1	.254
C4BPA	.478
CFD	.361
FABP3	.339
GPX3	.401
AIG1	.313
KCNIP2	.536
LSMD1	.298
MUCL1	.552
MYH9	.431
ZNF701	.281
HIST1H2AH	.501
CBX3	.458
COL18A1	.277
TUBA3C	.366
LUC7L3	.429
ACSM5	.312
PDCD6	.476
ARHGAP29	.434
BNC2	.488
GBF1	.457
NPFFR2	.559
RAB3GAP2	.308
LOC730110	.299
GCC2	.277
QSOX1	.364

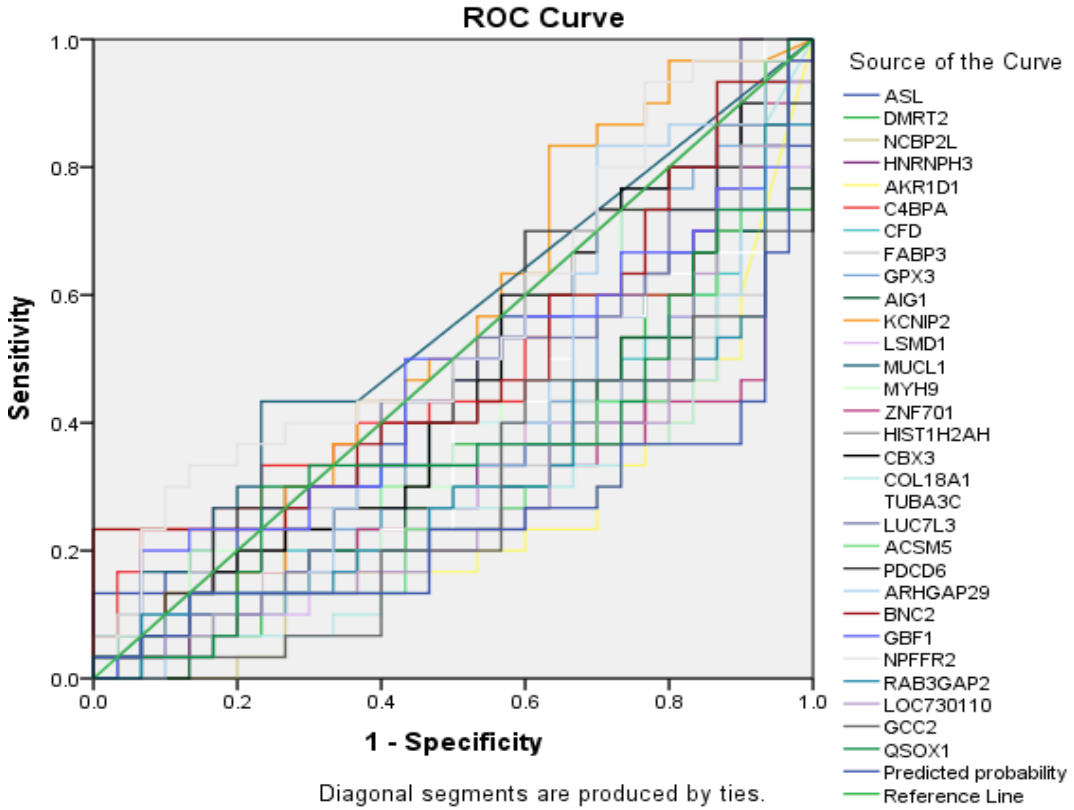
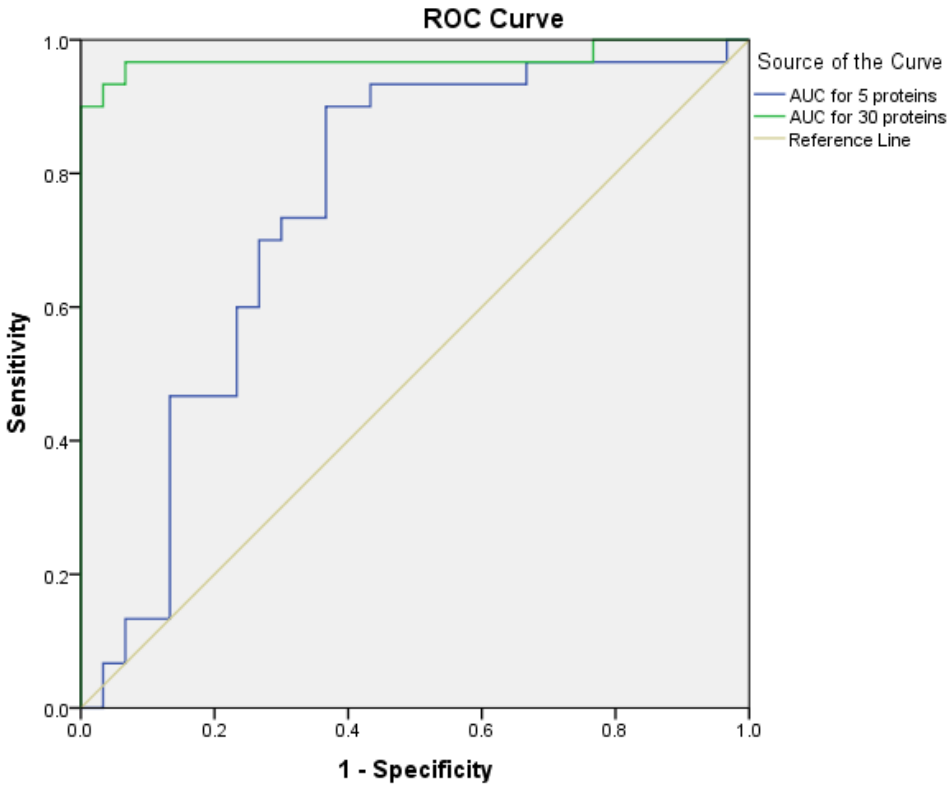


Table 4.14. *Area under the curve and ROC curve for 5 proteins (Table 4.13) and 30 proteins (Table 4.11) predicting biomarkers of HFPEF from HFREF.*

Area Under the Curve

Test Result Variable(s)	Area
AUC for 5 proteins	.751
AUC for 30 proteins	.971



Box plots of 25/30 proteins identified on Table 4.11 were generated (Figure 4.19-Figure 4.21). The data for the remaining 5 proteins with low protein amounts has been shown on Figure 4.22. The affiliations/associations of these proteins with disease has been summarised on section 6.2. Prior to mass spectrometry analysis, the dried pellets (samples) were reconstituted with 0.1% FA and ADH (50 fmol/µL). The ADH was later used to normalise the data in Progenesis for Proteomics via Log10 normalised intensity and sample groups compared.

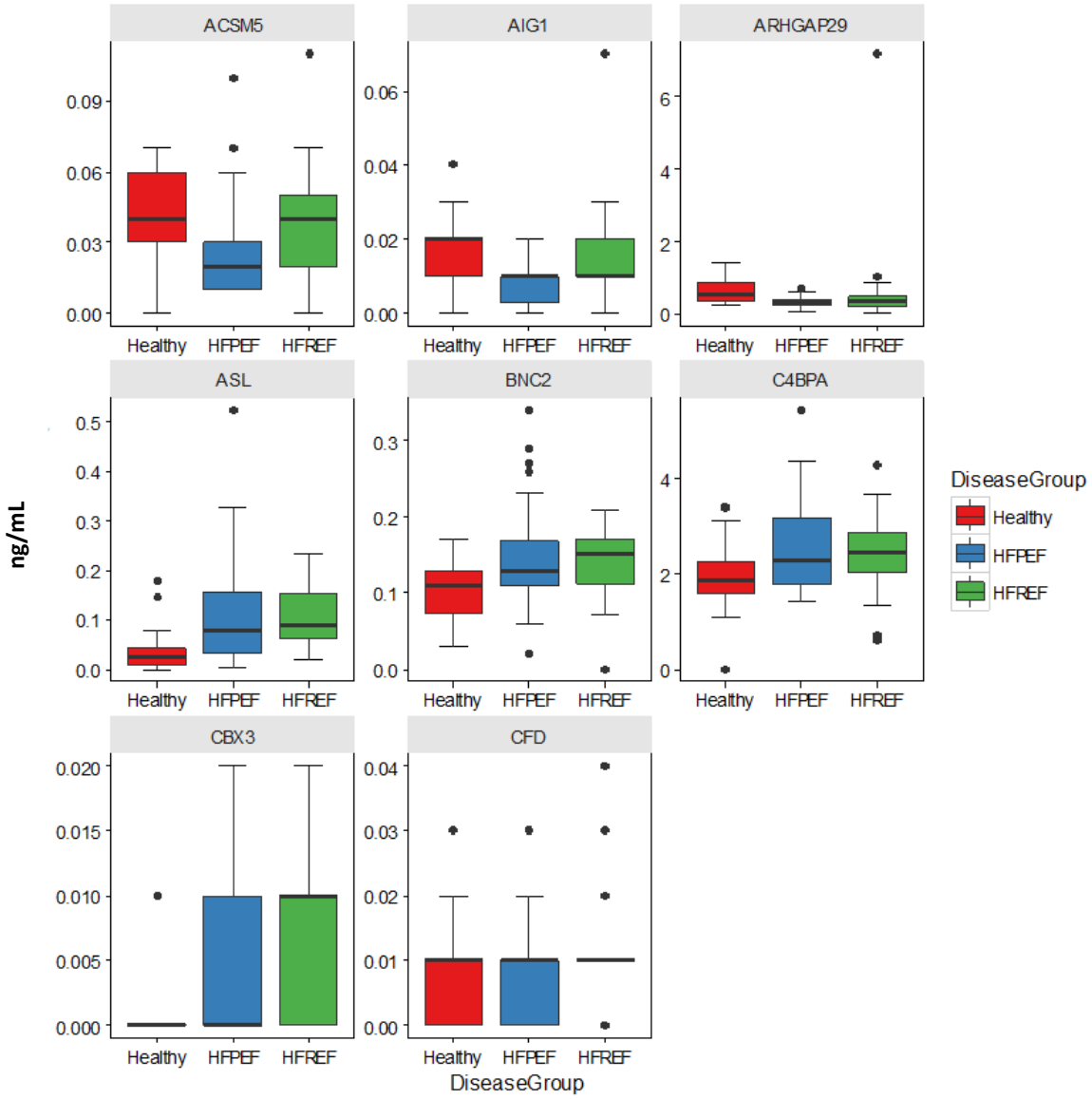


Figure 4.19. *Box plots comparing protein regulation in HFREF, HFPEF and healthy cohorts. The box plots show the up and down regulated proteins in HFREF, HFPEF and the healthy group. The dots represent those who had more extreme cases than the other subjects had in the group and are considered as outliers. All the proteins selected from the samples (n=90) were all brought forward from multivariate analysis as they had a p<0.05, but some (CFD) failed to have significance in univariate analysis.*

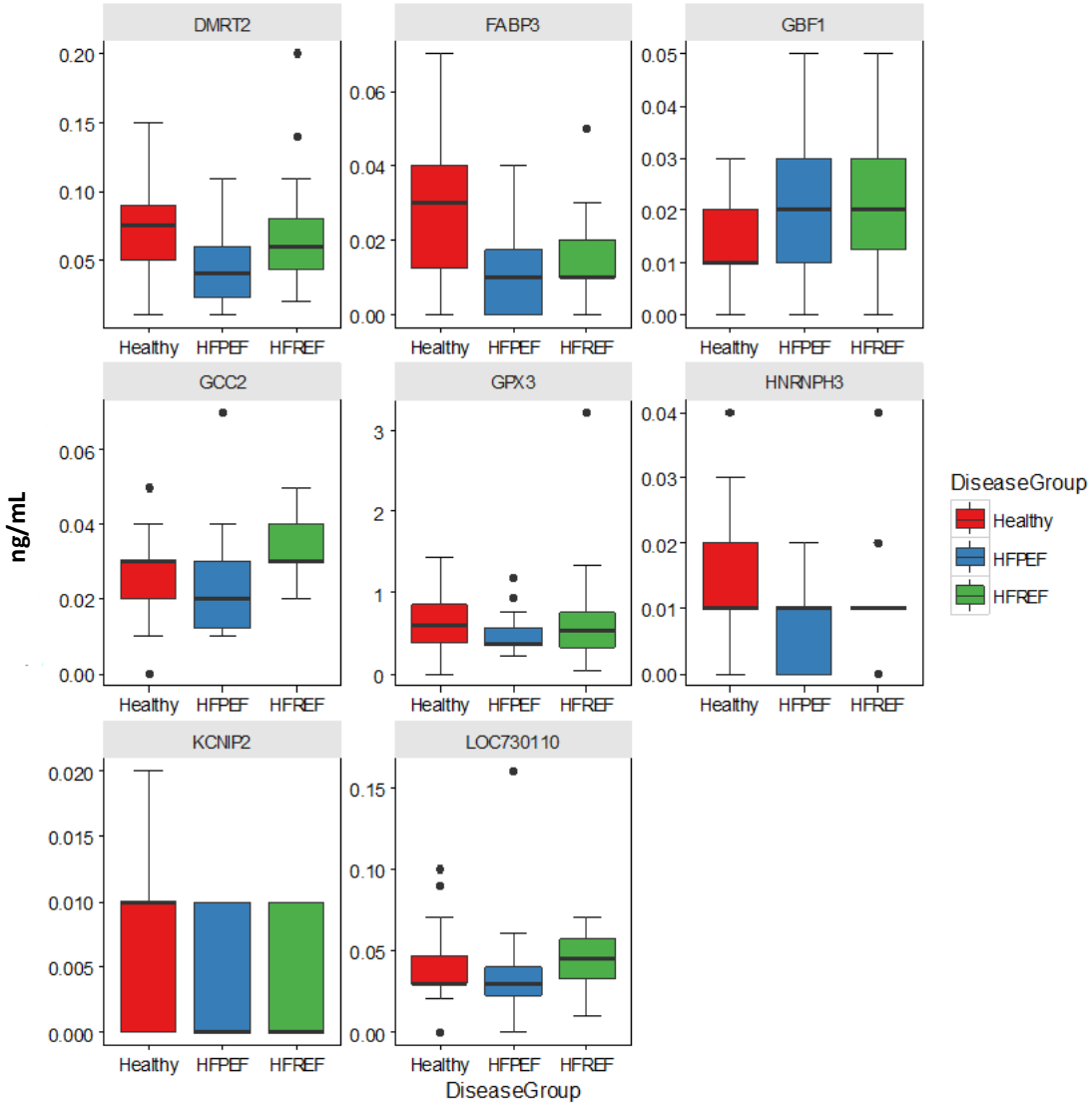


Figure 4.20. *Box plots comparing protein regulation in HFREF, HFPEF and healthy cohorts. The box plots show the up and down regulated proteins in HFREF, HFPEF and the healthy group. The dots represent those who had more extreme cases than the other subjects had in the group and are considered as outliers. All the proteins selected from the samples (n=90) were all brought forward from multivariate analysis as they had a $p<0.05$, but some (HNRNPH3) failed to have significance in univariate analysis.*

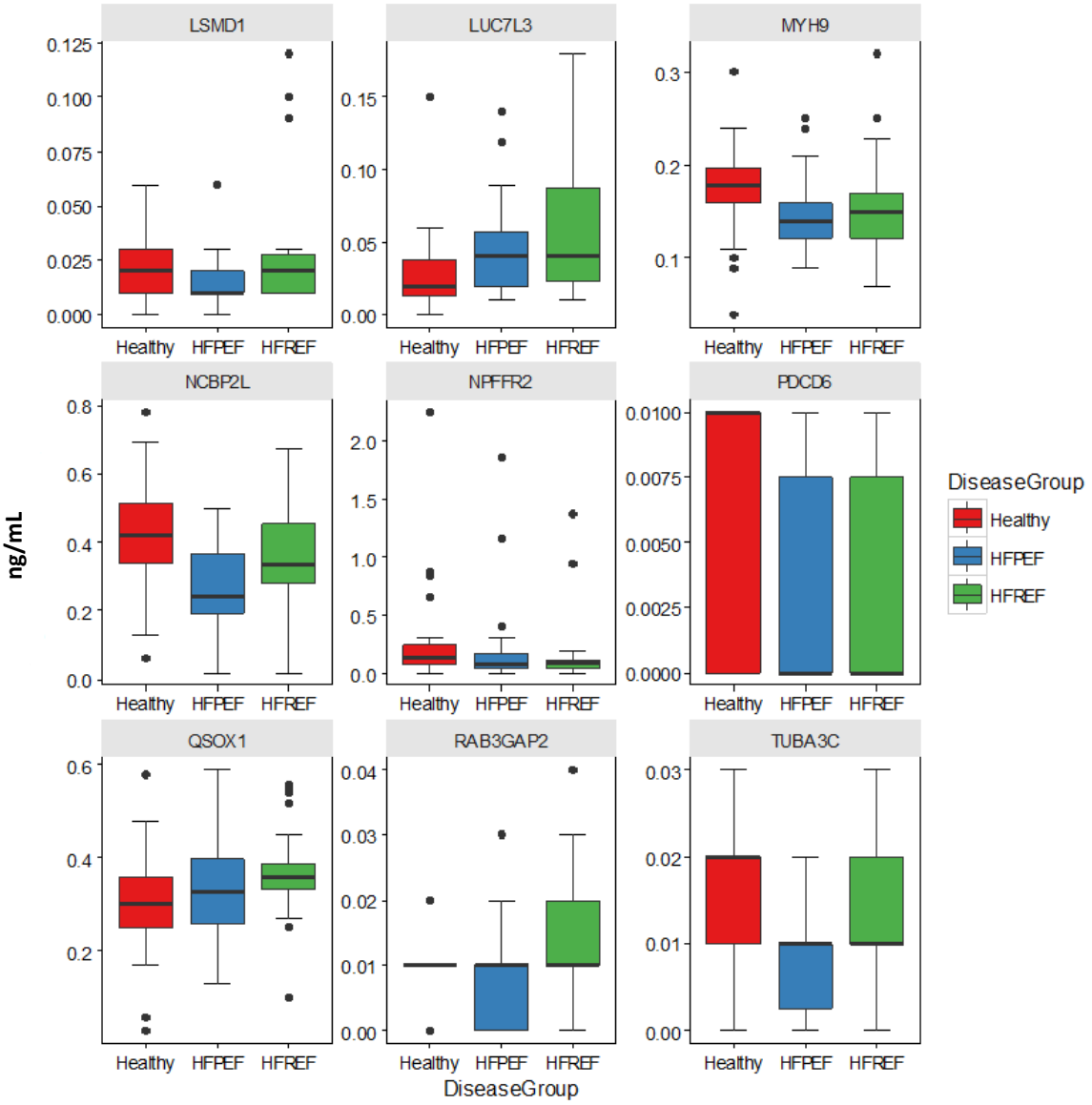


Figure 4.21. *Box plots comparing protein regulation in HFREF, HFPEF and healthy cohorts. The box plots show the up and down regulated proteins in HFREF, HFPEF and the healthy group. The dots represent those who had more extreme cases than the other subjects had in the group and are considered as outliers. All the proteins selected from the samples (n=90) were all brought forward from multivariate analysis as they had a p<0.05, but some (RAB3GAP2) failed to have significance in univariate analysis.*

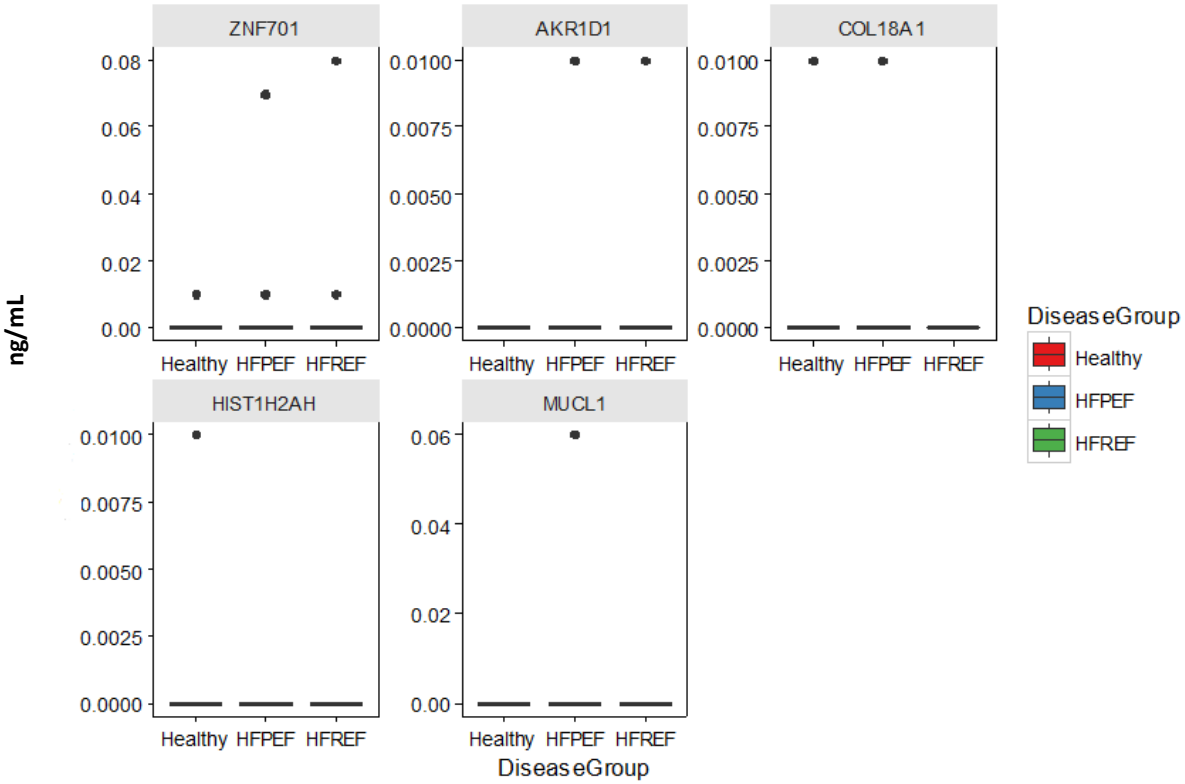


Figure 4.22. *Box plots comparing protein regulation in HFREF, HFPEF and healthy cohorts. Box plots for these proteins were below detection due to low protein amounts. All the proteins selected from the samples (n=90) were all brought forward from multivariate analysis as they had a p<0.05, but failed to have significance in univariate analysis.*

The mean and fold change between the three cohorts of all the 30 potential biomarkers have been shown on Table 4.15 and Table 4.16. The peptide count and unique proteins for these 30 proteins have been highlighted on the tables.

Table 4.15. *Raw data showing the mean and fold change of 25 potential biomarkers that could discriminate between the 3 cohorts. The 2 chosen biomarkers have been highlighted in yellow. A detailed analysis of the peptide analysis has been shown on Table 4.17-Table 4.19. The expression analysis of these proteins has been shown in Figure 4.19-Figure 4.21.*

Description	Peptide count	Unique peptides	Confidence score	Mean CON	Mean HFPEF	Mean HFREF	Mean HF	Fold CON vs. HFPEF	Fold CON vs. HFREF	Fold HFPEF vs. HFREF	Fold CON vs. HF
ASL	8	2	39.50	0.04	0.10	0.11	0.11	0.35	0.33	0.92	0.34
GPX-3	26	13	171.90	0.68	0.45	0.61	0.55	1.53	1.12	0.73	1.23
ACSM5	6	2	23.62	0.04	0.02	0.04	0.03	1.57	1.01	0.64	1.14
PDC6	4	2	34.58	0.01	0.00	0.00	0.00	1.37	1.45	1.06	1.42
MYH9	229	56	1402.23	0.17	0.14	0.16	0.15	1.21	1.11	0.92	1.14
TUBA3C	28	1	181.37	0.02	0.01	0.01	0.01	1.64	1.29	0.79	1.39
AKR1D1	3	2	16.16	0.00	0.00	0.00	0.00	1.14	0.42	0.36	0.53
CFD	7	4	47.77	0.01	0.01	0.01	0.01	1.01	0.73	0.73	0.81
QSOX1	107	61	683.91	0.31	0.32	0.36	0.35	0.97	0.85	0.87	0.88
DMRT2	8	4	56.85	0.07	0.05	0.06	0.06	1.63	1.20	0.74	1.32
BNC2	8	4	50.49	0.10	0.14	0.15	0.14	0.75	0.70	0.94	0.72
ZNF701	3	2	13.31	0.00	0.00	0.01	0.01	1.05	0.29	0.28	0.38
Luc7	9	4	67.43	0.03	0.04	0.05	0.05	0.65	0.54	0.83	0.57
HIST1H2AH	2	1	12.45	0.00	0.00	0.00	0.00	396.33	21064.36	53.15	1145.86
ARHGAP29	3	1	23.59	0.01	0.01	0.01	0.01	1.04	0.86	0.83	0.91
KCN1P2	3	1	18.00	0.04	0.03	0.03	0.03	1.59	1.26	0.79	1.35
LOC730110	2	1	16.96	0.04	0.03	0.04	0.04	1.25	0.87	0.69	0.97
HNRNPH3	3	3	12.30	0.01	0.01	0.01	0.01	2.09	1.49	0.71	1.65
NCBP2L	2	1	11.27	0.42	0.27	0.35	0.32	1.56	1.21	0.78	1.31
C4BPA	99	61	608.75	1.96	2.37	2.58	2.51	0.83	0.76	0.92	0.78
Col18a1	7	2	28.08	0.00	0.00	0.00	0.00	3.09	1.27	0.41	1.58
RAB	5	2	24.74	0.01	0.01	0.01	0.01	1.19	0.69	0.58	0.80
FABP	2	1	11.95	0.03	0.01	0.01	0.01	2.15	1.76	0.82	1.87
GCC2	24	6	130.45	0.03	0.02	0.03	0.03	1.14	0.87	0.76	0.95
AIG1	2	2	17.36	0.02	0.01	0.01	0.01	1.57	1.14	0.73	1.26

Table 4.16. *Raw data showing the mean and fold change of 5 potential biomarkers that could discriminate between the 3 cohorts. The expression analysis of these proteins has been shown in Figure 4.22. The box plots were not generated possibly due to low protein amounts in the 3 cohorts.*

Description	Peptide count	Unique peptides	Confidence score	Mean CON	Mean HFPEF	Mean HFREF	Mean HF	FoldCON vs. HFPEF	Fold CON vs. HFREF	Fold HFPEF vs. HFREF	Fold CON vs. HF
CBX3	3	2	16.13	0.00	0.01	0.01	0.01	0.60	0.48	0.80	0.51
GBF1	5	4	27.45	0.01	0.02	0.02	0.02	0.60	0.54	0.90	0.56
NPFFR2	2	1	9.75	0.27	0.16	0.19	0.18	1.62	1.44	0.89	1.50
LSMD1	4	3	34.73	0.02	0.01	0.02	0.02	1.63	0.96	0.59	1.11
MUCL1	1	1	6.87	0.00	0.00	0.00	0.00	0.03	0.89	35.57	0.07

The absolute abundances of the peptides for the 2 proposed potential biomarkers (ASL and GPX-3) have been shown in details on Table 4.17-Table 4.19. Post-translational modification can also be seen in some of the peptides in these proteins that could have affected the quantitation of these two proteins (ASL and GPX-3). GPX-3 had three peptides used for quantitation as highlighted on table Table 4.17. All these peptides were higher in control compared to HF group as shown on the protein level. Two of the peptides are essentially the same peptide with and without neutral loss. Taking this forward, and judging by the intensities identified, WNFEK without the neutral loss, would be used. In addition, they were multiply charged except two, had no conflict (did not belonged to another protein) and were not modified. Conversely, ASL had 2 peptides used for quantitation as shown on Table 4.17. Both peptides were multiply charged and had no conflict. However, the expression levels between control and HF group were different. This is because one peptide (AELNFGAITLNSMDATSER) had 3 modification and their expression levels did not match that on the protein level. Thus, could not have been used for quantitation.

Table 4.17. *A summary of the peptides used for quantitation of ASL and GPX-3 showing average abundances between control and heart failure samples.*

Sequence	Peptide ion	Score	Hits	Mass	Charge	Conflicts	Modifications	Drift time	Average abundance Control	Average abundance HF
GPX-3										
MDILSYMRR	3800	4.48	1	1183.6241	2	0		4.09	5293.67	4768.51
WNFEK	124	6.55	4	722.3270	1	0		7.62	4.52e+004	3.44e+004
WNFEK	27233	0.00	1	704.3336	1	0	[C-term] neutral loss	7.69	1939.47	1305.14
ASL										
AELNFGAITLNSMDATSER	78769	5.71	1	2199.8766	3	0	[4] Deamination N [12] Phosphoryl STY [16] Phosphoryl STY	4.22	27.36	21.75
EFSFVQLSDAYSTGSSLMPQK	8234	5.60	1	2321.0732	3	0		4.43	1766.22	4954.10

Table 4.18. *A summary showing comparison of 26 GPX3 peptides absolute abundances between control and heart failure samples. HF=heart failure*

Sequence	Peptide Ion	Score	Hits	Mass	Charge	Tags	Conflicts	Modifications	In quantitation	Drift time (ms)	Average Normalised Abundances	
											Control	HF
EQKFYTFLK	1512	5.33	1	1202.6322	2		5		no	4.29	9666.79	1.12e+004
EQKFYTFLK	36186	5.86	5	1202.6528	3		1		no	2.84	103.48	155.81
EQKFYTFLK	29941	5.67	1	1202.6149	2		2		no	4.16	86.23	81.38
EQKFYTFLK	23649	---	---	1202.6417	3		5		no	2.91	311.95	431.24
FLVGPDGIPIMR	10859	7.08	57	1313.7247	2		0		no	4.92	725.47	624.91
FYTFLK	51355	6.55	8	817.4377	1		0		no	9.63	27.65	24.51
FYTFLK	27127	6.55	7	817.4594	2		0		no	3.12	155.02	133.19
FYTFLK	20471	5.02	1	817.4223	1		0		no	9.07	133.90	125.34
GDVN̄GEKEQK	936	5.16	1	1103.5608	2		2	[4] Deamidation N	no	3.95	1.19e+004	1.54e+004
GDVN̄GEKEQK	11755	---	---	1103.5311	1		2	[4] Deamidation N	no	13.02	510.98	603.85
LFWEPMK	49184	5.35	1	949.4501	1		1		no	10.80	45.32	32.99
LFWEPMK	20127	4.91	3	949.4655	1		1		no	10.32	222.89	142.44
LFWEPMK	569	5.62	1	949.4740	1		2		no	10.53	1.04e+004	7451.23
LFWEPMK	10728	6.83	21	949.4770	2		2		no	3.53	1802.85	1067.39
LFWEPMK	6576	6.95	5	949.4724	1		1		no	11.08	1357.07	775.21
LFWEPMK	12701	5.76	1	949.4738	1		0		no	11.08	322.69	257.16
LFWEPMK	18687	4.79	1	949.4709	2		1		no	3.60	39.91	200.98
LFWEPM̄K	9024	---	---	965.4822	2		2	[6] Oxidation M	no	3.53	2368.98	1747.22
LFWEPM̄K	9431	5.61	1	965.4678	1		2	[6] Oxidation M	no	10.94	1114.76	818.80
MDILSYMR	92	6.22	1	1027.4926	1		3		no	11.63	1.05e+005	1.02e+005
MDILSYMR	23154	6.10	9	1027.4774	1		1		no	11.43	104.30	124.33
MDILSYMR	235	6.22	1	1027.5200	2		3		no	3.74	6.84e+004	7.42e+004
MDILSYMRR	3800	4.48	1	1183.6241	2		0		yes	4.09	5293.67	4768.51
MDILSYMRR	5337	5.38	2	1183.6000	2		1		no	4.29	1668.06	1952.32
MDILSYMRR	20177	5.46	1	1183.6019	2		0		no	4.02	342.67	363.40
MDILSYMRR	48486	5.17	1	1183.5506	1		2		no	13.50	92.59	51.50

MDILSYMRR	73238	6.64	1	1183.5740	1		0		no	13.16	0.08	0.00
NFEK	6320	0.00	2	536.2588	1		2		no	3.81	752.80	584.43
NSC̄PPTSELLGTSDR	23562	6.52	14	1632.7300	2		0	[3] Carbamidomethyl C	no	5.54	449.16	397.41
NSC̄PPTSELLGTSDR	22122	6.90	39	1632.7431	2		0	[3] Carbamidomethyl C	no	5.54	457.17	389.33
QAALGVK	61948	5.22	1	685.4146	1		0		no	3.19	24.03	20.96
QEPGENSEILPTLK	15942	7.13	55	1553.7927	2		1		no	5.33	760.39	604.07
QEPGENSEILPTLK	45199	5.91	1	1553.7630	2		1		no	5.68	54.84	50.83
QEPGENSEILPTLK	63824	5.35	2	1553.7614	2		1		no	5.89	20.86	37.46
TFLK	36775	0.00	1	507.3142	1		0		no	5.82	53.50	48.60
TT̄VSNVK	48202	5.10	1	827.3995	2		0	[2] Phosphoryl STY	no	2.98	82.97	101.70
TTVS̄NVK	9752	5.73	1	827.3773	1		2	[4] Phosphoryl STY	no	8.59	743.21	985.33
TTVSN̄VK	53334	4.24	1	748.4084	1		0	[5] Deamidation N	no	8.38	16.59	14.45
T̄TVSN̄VK	17377	5.53	1	828.3614	1		2	[1] Phosphoryl STY|[5] Deamidation N	no	9.21	179.44	147.58
TTVS̄NVKMDILSYM̄R	53753	5.29	1	1852.8481	2		1	[4] Phosphoryl STY|[14] Oxidation M	no	6.51	49.74	26.39
TTVSNVKMDILSYMRR	19420	5.74	3	1912.9597	2		0		no	6.79	670.76	791.02
VHDIRWNFEK	31730	5.71	1	1342.6691	3		0		no	2.98	204.94	29.29
VHDIRWNFEK	38174	5.71	1	1342.6695	2		3		no	4.78	249.36	137.81
WHHRTTVSN̄VK	15202	5.17	1	1364.7046	2		2	[9] Deamidation N	no	4.64	1379.34	1353.63
WHHRTTVSNVKM̄DILSYMR	16743	5.21	1	2389.1783	3		4	[12] Oxidation M	no	4.92	239.99	295.18
WHHRTTVSNVKM̄DILSYMR	92568	---	---	2389.1579	2		4	[12] Oxidation M	no	8.73	7.53	2.63
WNFEK	4427	6.32	7	722.3267	1		1		no	7.69	1957.51	2648.16
WNFEK	157	6.05	14	722.3417	1		1		no	7.69	2.85e+004	3.12e+004
WNFEK	124	6.55	4	722.3270	1		0		yes	7.62	4.52e+004	3.44e+004
WNFEK	52384	6.05	2	722.3246	1		0		no	3.53	37.71	21.70
WNFEK	61933	6.32	2	722.3479	2		0		no	2.56	38.23	34.52
WNFEK	73097	6.19	6	722.3254	1		0		no	3.60	28.00	4.33
WNFEK+	27233	0.00	1	704.3336	1		0	[C-term] neutral loss	yes	7.69	1939.47	1305.14
YTFLK	10505	0.00	1	670.3506	1		1		no	7.27	815.34	867.48
YVRPGGGFVPNFQLFEK	12354	5.56	1	1953.9956	2		2		no	7.06	795.95	28.95

Table 4.19. *A summary showing comparison of 8 ASL peptides absolute abundances between control and heart failure samples.*

Sequence	Peptide Ion	Score	Hits	Mass	Charge	Tags	Conflicts	Modifications	In quantitation	Drift time (ms)	Average Normalised Abundances	
											Control	HF
AELNFGAITLNSMDATSER	78769	5.71	1	2199.8766	3		0	[4] Deamidation N|[12] Phosphoryl STY|[16] Phosphoryl STY	yes	4.22	27.36	21.75
AVFMAETK	6144	5.07	1	895.4306	1		3		no	9.76	716.64	535.44
EFSFVQLSDAYSTGSSLMPQK	8234	5.60	1	2321.0732	3		0		yes	4.43	1766.22	4954.10
EFSFVQLSDAYSTGSSLMPQK	15868	5.60	1	2321.0401	2		1		no	8.38	861.67	2066.95
ELLR	5064	7.47	4	529.3249	1		53		no	5.89	782.95	853.34
ELLR	5499	7.25	2	529.3288	1		40		no	5.89	756.55	731.49
ELLR+	29178	0.00	4	511.3153	1		56	[C-term] neutral loss	no	5.61	97.47	104.21
FVGAVDPIMEK	25364	5.29	1	1204.6387	2		6		no	4.22	800.57	646.27
LLR	7497	0.00	1	400.2810	1		42		no	4.57	259.13	244.67
LLR	8098	0.00	2	400.2810	1		43		no	3.60	432.56	513.43
NDQVVTDLR	7166	4.76	1	1059.5332	2		2	[1] Deamidation N	no	3.67	7105.65	7433.31

4.4 Discussion

4.5 Hallmarks of heart failure

Neurohormones (SNS, RAAS and arginine vasopressin (AVP) are defence mechanisms triggered when the heart is struggling to pump blood to the rest of the body. They are normally stimulated when there is a decrease in cardiac output and arterial pressure. The activation of neurohormones has been summarised in section 1.2.2. In this research, Rab3 GTPase-activating protein non-catalytic subunit (RABGAP3) was higher in HF groups compared to the control group with prevalence in SHF. This protein is a key regulator of calcium mediated hormone and neurotransmitter exocytosis (Aligianis *et al.,* 2006) which suggests that its upregulation in the HF group is due to the abnormal pumping of the heart.

Reactive oxygen/Nitrogen species (ROS/RNS) (free radicals) are chemical reactive species containing oxygen or nitrogen for example hydroxyl radical, peroxides, superoxide and singlet oxygen. In the heart, these free radicals ($O_2^{\bullet -}$, $NO^{\bullet}$, $OONO^{-}$) are formed when eNOS generates NO for vascular toning but are scavenged by GPX3 which catalyses the reduction of H_2O_2 formed when $O_2^{\bullet}$ and $OONO^{-}$ react with ROOH thereby protecting cells against oxidative damage (Shiomi *et al.,* 2004). However, in heart failure group, ASL which catalyses the production of NO through L-arginine was upregulated in both HF groups compared to control group. This suggest that there was elevated NO levels in the epithelial cell. This could be due to;

- Since the activity of eNOS is inhibited by ROCK (upregulated in HF group), as a counter regulatory mechanism, more ASL was generated to catalyse the production of L-arginine from L-citrulline thus increasing the production of NO.
- Cardiac hypertrophy is associated with polyamine synthesis from ornithine. On the other hand, urea is a by-product of the production of ornithine from arginine produced from citrulline in the presence of ASL. Therefore, NO production might be a counter regulatory mechanism and the increase in ASL maybe because of polyamine synthesis.

Inflammation is very common in heart failure patients. It is often associated with response to infection but chemical and physical injury can trigger inflammation. Its role in HF has not been proven and remains a topic of ongoing research. However, Dick *et al.,* 2016 reported an

association of circulatory inflammatory cytokines with chronic HF but insists the causative role that inflammation plays in HF progression is not well defined. The secretion of pro inflammatory cytokines has been associated with myocardium with some evidence suggesting that catecholamine enhance this myocardial cytokine production due to myocardial injury (Anker *et al.,* 2004). The complement system plays a major role in regulation of inflammation to protect host against microbes and repair the tissue (Yang *et al.,* 2010). However, uncontrolled activation could increase inflammation, which could initiate several pathologies (Maciej et al., 2007), including the development of HF. In this research, C4b-binding protein alpha chain (C4BPA) was upregulated in both HF groups as compared to the control group suggesting that this complement protein could (Figure 4.19) have triggered the activation of the complement system leading to inflammation (Markiewski *et al.*, 2007).

Ischemia results from accumulation of fatty deposits in the walls of the arteries that supply blood to the heart. These fatty deposits develop into a plaque, which narrow the arteries and eventually block the flow of blood, thus ischemia. Reduced blood flow deprives oxygen and other nutrients to the heart muscle resulting to muscle injury. This muscle injury triggers a cascade of events, which modulate inflammatory response, which enhance recruitment of white blood cells (WBC) into the blood vessels. These WBC cells form abnormal foam cells, which initiate the formation of atherosclerosis lesions (Pfutzner *et al.,* 2010). Acyl-coenzyme A synthetase ACSM5, mitochondrial (ACSM5) has been associated with multiple risk factors of cardiovascular diseases including pathogenesis of hypertriglyceridemia, visceral obesity and hypertension (Iwai *et al.,* 2003). This protein was downregulated in the HF groups as compared to control group (Figure 4.19).

Apoptosis is programmed cell death that play role in activation of various pathological conditions including myocardial infarction (Kim *et al*., 2010), reperfusion injury and heart failure (Nam-Ho *et al*., 2010). Apoptosis is activated through caspases (a group of cysteinyl-aspartate-directed proteases) which are activated by proteolytic cleavage. Caspase substrates in the heart include α-actin, α/β-myosin heavy chain, myosin light chain I and II, tropomyosin, and cardiac troponins (Communal *et al.,* 2002). Cardiomyocyte apoptosis has been implicated with ischemia (van Empel *et al.,* 2005). According to Araujo *et al.,* 2014, QSOX1 protein may be involved in extra cellular matrix (ECM) maturation, redox reaction, increase cellular oxidative stress and the induction of apoptosis. ECM generation is an anti-

apoptotic event that acts as a protective mechanism. In this research, QSOX1 was upregulated in the HF groups compared to the control group. Apoptosis also plays a role in activation of the complement system, which might explain the upregulation of C4BPA in HF group (Figure 4.19).

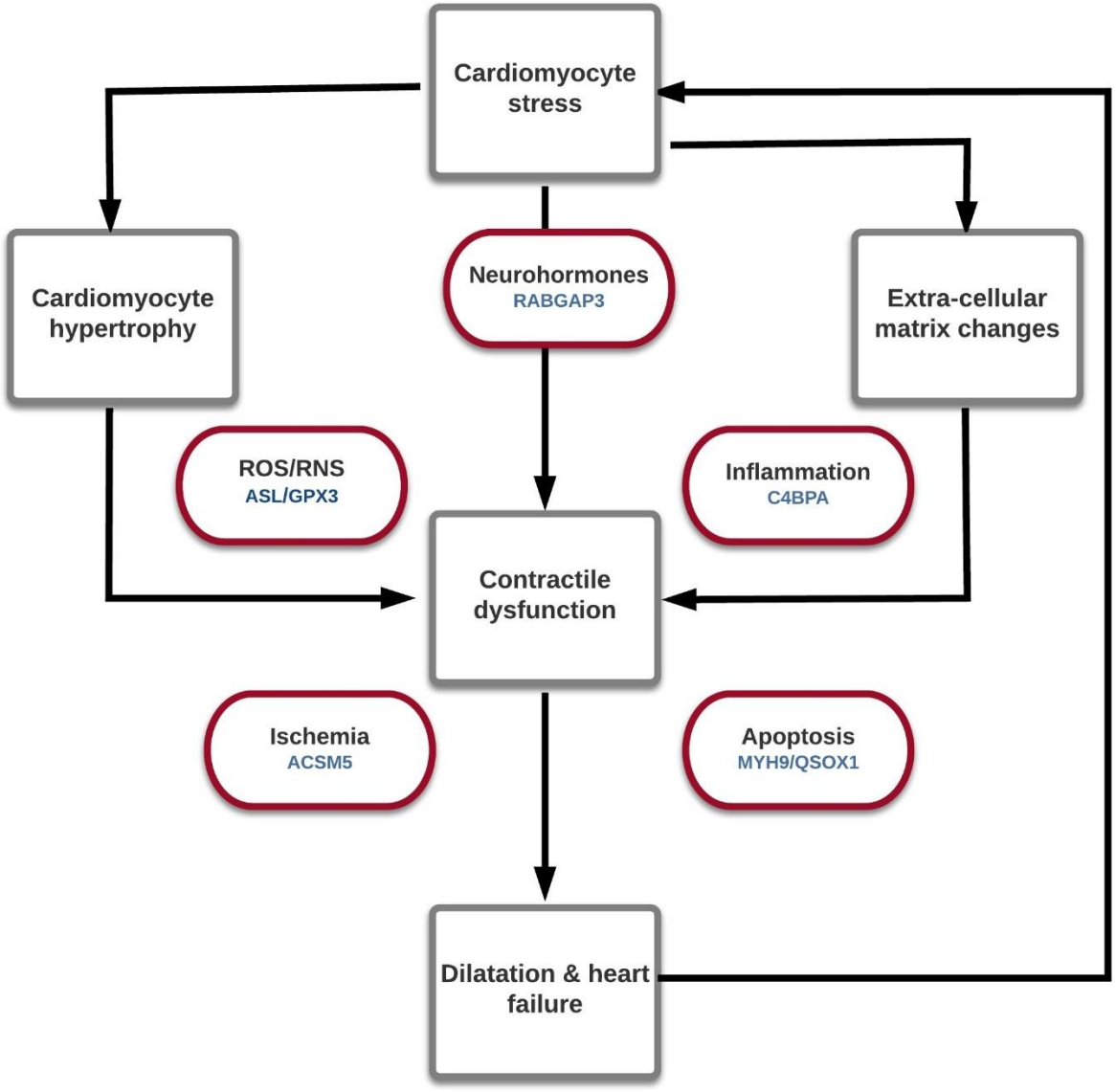


Figure 4.23. *Hallmarks of heart failure showing proteins associated with their progression. Rab3 GTPase-activating protein non-catalytic subunit (RAB3GAP3), C4b-binding protein alpha chain (C4BPA), Isoform 2 of Argininosuccinate lyase (ASL), Glutathione peroxidase 3 (GPX3), Acyl-coenzyme A synthetase ACSM5, mitochondrial (ACSM5), Myosin (MYH9) and Sulfhydryl oxidase 1 (QSOX).*

Several proteins that have relevance to cardiovascular diseases were identified. Some of these proteins have been reported to initiate or promote the hallmarks (Figure 4.23) of disease that eventually lead to heart failure. A potential mechanistic pathway is described that could form the basis of an exciting new area for target screening. It is feasible that when proteins that influence the contraction and relaxation of the heart have been altered then the heart will start failing. When the heart start failing, these proteins are leaked into the blood and can be identified as biomarkers.

The rationale in this is since ASL and GPX3 are associated with activities that causes the relaxation of the heart and have both been altered in heart failure group as compared to control group, they could be potential biomarkers of diastolic heart failure. ARHGAP29 and GEF are activators of the downstream effectors that regulate the activities of the proteins involved in either diastole or systole. Thus, ARHGAP29 and GEF cannot be specific for a particular heart failure group. Therefore, measuring the blood levels of ASL and GPX3 could be an indicator of disease progression. Thus, this technique could be used for early diagnosis of HFPEF.

In this research, the discovery studies required a small number of samples (90 samples) for HFPEF biomarkers discovery. However, translating to large-scale studies not only require a larger cohort but also highly accurate quantitation. Targeted LC-MS based assays are increasingly applied in the post-discovery proteomics area with emphasis on validation. For a long time, enzyme linked immunosorbent assays (ELISA) were used for validation. However, this technique could only be used where only a few biomarkers needed to be validated on a larger cohort. The challenge with ELISA is that each targeted protein or peptide requires an antibody and high quality ELISA assays are rare for these proteins (Haab *et al.,* 2006). ELISAs are also costly and time consuming which may involve development time of up to 2 years (Wang *et al.,* 2009).

Therefore, a multiple reaction monitoring method (MRM) (chapter 5) was developed in parallel to this clinical study to provide a basis for the validation of the findings of this research.

Chapter 5

QUANTITATION AND TRANSLATIONAL ANALYSIS

5 Quantitation and translational analysis of cardiovascular related proteins.

5.1 Introduction

In the past bioanalytical proteomics workflows consisted of the quantitative profiling of two or more samples using gel-based separation techniques (SDS-PAGE) with some type of optical read-out (Gorg *et al.,* 2004, 2009). This relied on detecting small changes in protein and peptide abundance because of their altered state. However, this was very challenging especially when detecting trace amounts of proteins. As a result, the benefits of mass spectrometry were readily recognised. These involved the development of in-gel digestion and peptide extraction techniques in combination with peptide mass fingerprint MALDI time-of-flight (Henzel *et al.,* 1993) and nano-electrospray tandem mass spectrometry (Wilm *et al.,* 1996). This entailed the use of one and two-dimensional nano scale LC chromatography which improved the identification of proteins. Meanwhile, continued high-resolution mass spectrometer developments, such as modes of acquisition, resolution, analyser types and combinations, speed and sensitivity, lead to the identification of more proteins and putative disease markers at increasingly faster identification rates. One of the current advances is the incorporation of orthogonal separation techniques within instruments designs, as for example, ionic gas phase based ion-mobility (Valentine *et al.,* 2006, Pringle *et al.,* 2007).

Quantitation in discovery LC-MS based proteomics experiments have been conducted in two ways: labelled or label free methods (section 1.3.5.3 and section 1.3.5.4 respectively). Between the two methods, Label-free have gained great popularity over the last decade with ion abundance based methods believed to provide the best accuracy and precision (Wang *et al.,* 2003, Silva *et al.,* 2005, Radulovic *et al.,* 2004). The increase in dynamic range and the freedom in experimental design with label free methods are remarkable compared to labelled LC-MS methods. In addition, the amounts of proteins identified can be estimated without the use of labelled isotope standards. Despite this remarkable improvement, we still face experimental variation in our analysis that affects the total observed error.

Therefore, the development of LC-MS acquisition methods that could provide both qualitative and quantitative information in a single experiment significantly improved protein detection and quantitation. Multiple/selected reaction monitoring (MRM/SRM) based assays have the potential to afford protein quantitation with the reproducibility and throughput

required in order to improve biomarker acceptance. MRM is a tandem MS/MS technique which is typically performed on triple quadruple MS (Anderson *et al.*, 2006, Keshishian *et al.*, 2007, Lange *et al*., 2008, Perterson *et al.*, 2012) and involves selection of a precursor ion which acts as an alternative to the protein of interest (Parker *et al.*, 2014). Recently, high resolution SRM have been arried out on Q-Exactive (PRM – Domon *et al*., 2015) and ToF (Heaney *et al.*, 2015). These types of experiments necessitate high-throughput, high sensitivity, high resolution, large dynamic range and excellent selectivity in a single assay.

Hypothesis is that we are able to analyse multiple cardiovascular disease related markers in neat human plasma using a combination of state of the art chromatographic technology coupled to highest specification mass spectrometry platforms.

Aims and objectives

- Compare tandem quadrupole and high-resolution mass spectrometers to show and distinguish the performance of these analyser types in targeted protein quantitation experiments.
- Use a small cohort of heart failure samples to test the feasibility and potential use of LC-MRM-MS for translational analysis.
- Examine thirteen proteins that are involved in biological pathways in heart failure.
- Use this targeted LC-MS based MRM assay to validate the findings of chapter 4.

5.2 Method

5.2.1 Sample Preparation

Human blood samples (20 mL) were obtained in EDTA blood tubes collected from a cohort of twenty healthy donors, twenty heart failure with preserved ejection fraction (HFPEF) patients, and twenty heart failure with reduced ejection fraction (HFREF) patients, following informed consent. The plasma was separated from blood via centrifugation at 15,000 g at 4°C for 30 mins using a refrigerated centrifuge. The plasma layer was separated from the buffy layer and red blood cells, and stored at -80 °C.

5.2.2 Patient's demographics

Both male and female subjects had a mean age of about 72 with over 25% being above the age of 67 (Table 5.1, Figure 5.1 and Figure 5.2). All HFPEF patients had an ejection fraction of ≥50% while the HFREF patients had an ejection fraction ≤40%.

Table 5.1. *A summary of the twenty healthy donors, twenty heart failure with preserved ejection fraction (HFPEF) patients, and twenty heart failure with reduced ejection fraction (HFREF) patients based on their age and sex.*

Sex (Female) Age (years)	**Healthy group**	**HFPEF**	**HFREF**
Mean	72.08	73.13	72.60
Median	72.00	75.50	67.00
Standard Deviation	4.80	7.40	7.00
Minimum	66.00	62.00	67.00
Maximum	80.00	82	83.00
Sex (Male) Age (years)			
Mean	76.86	70.83	71.93
Median	79.00	69.00	75.00
Standard Deviation	5.70	10.01	10.20
Minimum	67.00	48.00	42.00
Maximum	84.00	87.00	87.00

5.2.3 Protein digestion

An optimised protocol developed by Waters Corporation was used. For the comparative configuration part of the study, twenty microliters of undepleted human EDTA plasma (Sigma-Aldrich, St. Louis, MO) was diluted with 80 µL of 50 mM ammonium bicarbonate

solution and denatured in the presence of 10 µL of 1% RapiGest (Waters Corporation, Milford, MA) detergent solution at 80 °C for 45 mins. The plasma proteins were reduced with 100 mM dithiothreitol made up in 50 mM ABC and was used at a final concentration of 15 mM. These samples were vortexed and incubated at 60 °C for 30 mins. The samples were then alkylated in the dark with 200 mM iodoacetamide made up in 50 mM ABC to at a final concentration of 20 mM and incubated at ambient temperature for 30 mins. Proteolytic digestion was initiated by adding 40 µL of 1 µg/µL sequencing grade, modified trypsin (Promega, Madison, MI) and incubated overnight at 37 °C. Breakdown of the acid-labile detergent was achieved in the presence of 1% TFA at 37 °C for 45 mins. The peptide solutions were centrifuged at 13,000 rpm for 10 min, and the supernatants collected.

For the main part of the study, patient and donor plasma samples were digested as previously described with minor modifications (Daly *et al.,* 2014) RapiGest solution was then added to the samples to give a 0.1% final concentration and incubated at 80 °C for 45 mins. The samples were then reduced with 100 mM aqueous DTT solution added to give a final concentration of 5 mM prior to incubation at 60 °C for 30 mins. A 200 mM IAA solution was added to the samples to give a final concentration of 10 mM before incubation in the dark at room temperature for 30 mins. A trypsin solution of 1 µg/µL was added to the sample in a 1:50 w/w ratio and incubated at 37 °C overnight. Digestion was concluded, and RapiGest cleaved, with the addition of neat formic acid to the sample to give a final concentration of 0.5%. The plasma samples were centrifuged at 13,000 rpm for 10 mins to remove insoluble material, and the supernatants collected.

5.2.4 LC configurations

One-dimensional nanoscale LC separation of tryptic peptides was performed with an Acquity M-class system (Waters Corporation), equipped with a Symmetry C18 5 µm, 2 cm x 180 µm precolumn and an HSS T3 C18 1.8 µm, 25 cm x 75 µm analytical RP column (Waters Corporation). The samples were transferred with aqueous 0.1% (v/v) formic acid to the precolumn at a flow rate of 5 µl/min for 3 mins. Mobile phase A was water containing 0.1% (v/v) formic acid, whilst mobile phase B was acetonitrile containing 0.1% (v/v) formic acid. After desalting and preconcentration, the peptides were eluted from the precolumn to the analytical column and separated with a gradient of 3-40% mobile phase B over 90 mins at a flow rate of 300 nL/min, followed by a 2 mins column rinse with 85% of mobile phase B.

The columns were re-equilibrated at initial conditions for 20 mins. The analytical column temperature was maintained at 35 ºC.

Additional, higher throughput experiments were performed with 150 µm × 100 mm ionKey/MS micro-fluidics packed with BEH C18 1.7 µm (Chambers *et al.,* 2015). Gradient conditions were from 3-40% B gradient over 45 min at a flow rate of 1 µL/min, followed by a 6 min column wash with 85% of mobile phase B. The columns were re-equilibrated at initial conditions for 9 mins. The analytical column temperature was maintained at ambient temperature. Samples were injected/loaded directly on-column or using a precolumn configuration. In this instance, the precolumn was 5 cm x 300 µm id, packed with 5 µm Symmetry C18 and samples loaded with a flow rate of 15 µL/min for 1 min. Faster reversed phase gradient separation, both nanoscale LC and micro-fluidics based, were explored but not considered for detailed analysis as (isobaric) interferences or detection issues were readily observed, as exemplified in Supplementary Figure D-1, hampering comparative configuration analysis.

5.2.5 MS configurations

Multiple Reaction Monitoring (MRM) analysis was performed using two tandem quadrupole mass spectrometers, Xevo TQ-S and Xevo-TQ-S micro, and two hybrid quadrupole orthogonal acceleration time-of-flight (Q-Tof) platforms, Xevo G2-XS Q-ToF and Synapt G2-S*i* (Waters Corporation, Wilmslow, United Kingdom). The Synapt G2-S*i* instrument is equipped with a travelling wave ion, tri-wave ion guide, which comprises two stacked ring collision induced dissociation (CID) regions, separated by a travelling wave-guide that can be used for ion mobility separation. This configuration is described in detail elsewhere (Giles *et al.,* 2007; 2011). All experiments were performed in positive mode ESI. The ion source block temperatures and capillary voltages were kept constant for all instruments and set to 70 ºC and 3.2 kV. The N_2 cone gas flow and nanoflow gas pressure were 35 l/h and 0.2 bar, respectively, whereas the Ar collision gas flow equalled 2 ml/min. The quadrupole and time-of-flight analysers were externally calibrated with NaCsI mixtures from *m/z* 50 to 1990.

A number of instrument and analyser specific parameters, as well as the acquisition types, used in this study are presented in Table 5.2. Endogenous and SIL peptides were targeted by at least three transitions with a minimum of 10 data points over a chromatographic peak. Tandem quadrupole dwell and interscan delay times were automatically calculated by the

operating software based on a minimum number of data points specified at half height across a chromatographic peak. Collision energies were set at fixed values for the tandem quadrupole instruments and ramped for the time-of-flight instrument. In addition, for the time-of-flight based MRM acquisitions, integration and interscan delay times were manually set. Collision energies were ramped and initially calculated using the following regression equation: 0.034 times *m/z* + 3.314 eV, and further optimised by CID fragmentation evaluation obtained by repeat injections of SIL peptides in the absence of matrix. The MRM transitions for both instrument types are listed in Supplementary Table D-1.

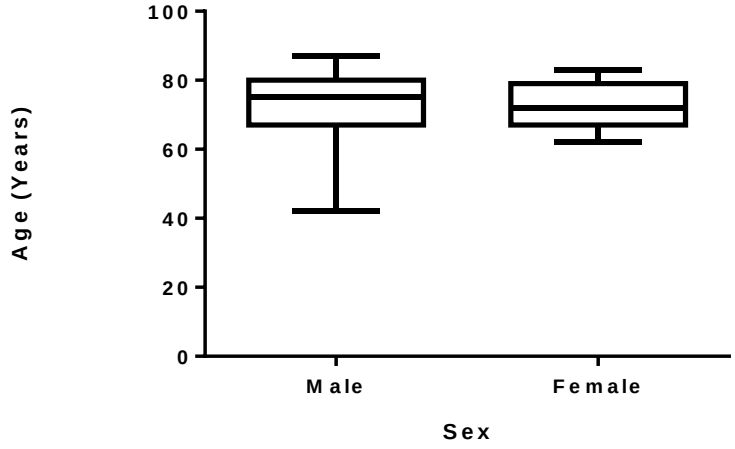


Figure 5.1. *Box plot analysis comparing the age of male and female subjects used in this study. There was no significance between the sex groups in the three cohort.*

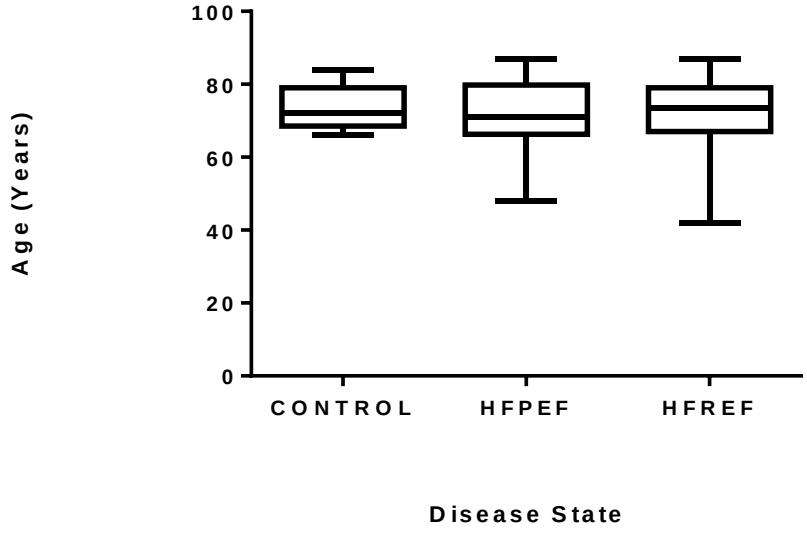


Figure 5.2. *Box plots showing differences in the age group versus disease state of all the 60 subjects used. There was no significance between the age groups in the three cohorts.*

Table 5.2*. Configuration and MRM method overview.*

configuration *instrument*	acquisition mode	Q2 (Da)/ ToF (FWHM) resolution	selectivity parameters
Tandem quadrupole			
Xevo TQ-S	MRM	0.7	t_r, transition
Xevo TQ-S micro	MRM	0.7	t_r, transition
Quadrupole time-of flight			
Xevo G2-XS QTof	ToF	25,000	t_r, transition, *m/z*
Synapt G2-Si	MRM/EDC	25,000	t_r, transition, *m/z*
	ToF	25,000	t_r, t_d, transition, *m/z*
	MRM/EDC	25,000	t_r, transition, *m/z*
	IM_p-MRM/EDC		
	MRM-IM_f		

Q1 resolution	0.7 Da for all instruments/geometries
ToF MRM	high-resolution separation and accurate mass detection fragment ions (Q1 isolation, IMS, followed by 'transfer' CID and EDC)
EDC	enhanced duty cycle (synchronization of pusher in QTof geometries of the time-of-flight analyzer for selected target *m/z* values/ranges as ions are released from the collision cell/region) [48]
IM_p-MRM	ion mobility separation peptide precursor ions pre-CID (Q1 isolation, IMS, followed by transfer CID and EDC)
MRM-IM_f	ion mobility separation fragment ions post-CID (Q1 isolation, trap CID, followed by IMS; affords synchronization of pusher in QTof geometries of the time-of-flight analyzer over the complete *m/z* range, providing near 100% duty cycle for all product ions) [51]
t_r	retention time (s)
t_d	drift time (s)

5.2.6 Experimental design

Fifteen stable isotope labeled (SIL) peptides (PepScan, Lelystad, Netherlands), representing putative blood-based cardiovascular disease protein biomarkers spanning over five orders of dynamic concentration range (Domanski *et al.*, 2012), were initially spiked as a dilution series from 6.25 amol to 12.5 fmol on-column, in tryptic digested EDTA human plasma. The SIL peptides and associated proteins of interest are shown in Table 5.3, including normal plasma protein and molar peptide concentration values (Anderson, 2005). In total, considering all LC-MS configurations and column/interface formats, 108 LC-MS

experiments were conducted. Extended dynamic range experiments were conducted with one of the SIL peptides at spike levels ranging from 7.5 amol to 1.5 pmol, in tryptic digested human plasma as well.

The healthy donor, HFPEF and HFREF plasma samples were spiked post-digestion at four individual different levels of 0.25, 0.5, 2 and 10 fmol each with the same fifteen SIL peptides, analysed separately (Figure 5.3), providing multiple, user selectable quantitation levels and surrogate technical replicates, mounting to a total of 240 LC-MS experiments (~ 10 days of measurement time). This was to ensure the concentrations were not over or under diluted and the right peptide signal from the samples was achieved (Figure 5.3).

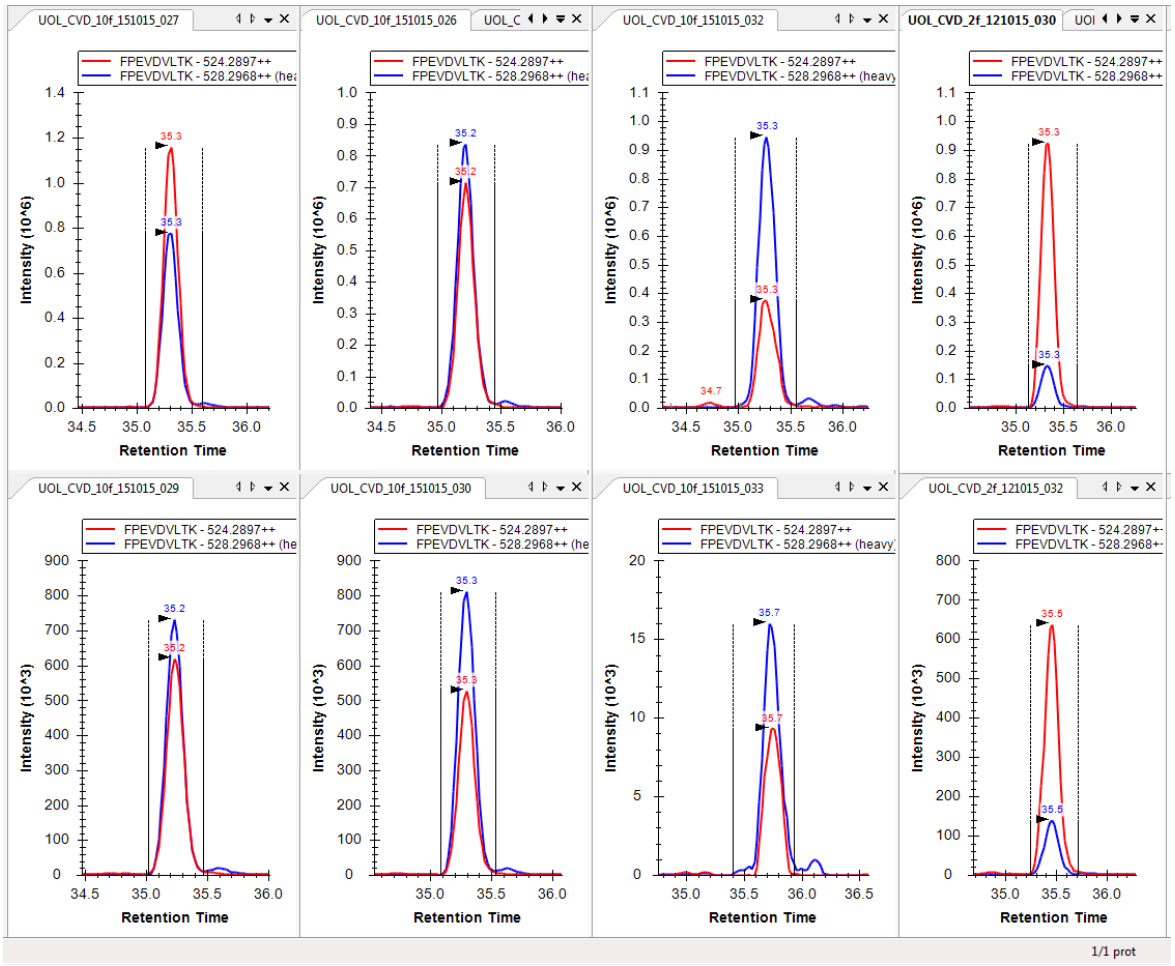


***Figure 5.3**. Plasma samples were spiked post-digestion at four individual different levels of 0.25, 0.5, 2 and 10 fmol each with the same fifteen SIL peptides. The above peaks show the integration of APOB_HUMAN (Red) spiked with 10fmols of the internal standard (Blue) using different biological samples.*

Table 5.3*. Target peptide sequences and proteins, including normal amount and molar concentration plasma levels.*

peptide sequence	entry name	normal plasma concentration (ng/ml)	peptide amount (nmol)/ ml plasma
LVNEVTEFA[K]	ALBU_HUMAN	41,000,000	620
ATEHLSTLSE[K]	APOA1_HUMAN	1,400,000	50
TGLQEVEV[K]	CO3_HUMAN	1,300,000	7.0
FPEVDVLT[K]	APOB_HUMAN	730,000	1.4
FQPTLLTLP[R]	IC1_HUMAN	137,000	2.6
TAAQNLYE[K]	APOC2_HUMAN	33,000	3.7
LGPLVEQG[R]	APOE_HUMAN	40,000	1.2
AAAATGTIFTF[R]	IPSP_HUMAN	4,400	0.10
EANYIGSD[K]	SAA1_HUMAN	4,000	0.34
ESDTSYVSL[K] GYSIFSYAT[K]	CRP_HUMAN	2,300	0.10
GFYFNKPTGYGSSS[R]	IGF1_HUMAN	144	0.019
LVNVVLGAHNV[R]	PRTN3_HUMAN	23	0.000949
ITLYG[R] SYPGLTSYLV[R]	NGAL_HUMAN	87	0.00423

[K] = $^{13}C_6{}^{15}N_2$ labelled; [R] = $^{13}C_6{}^{15}N_4$ labelled

ALBU_HUMAN = serum albumin; APOA1_HUMAN = Apolipoprotein A-1; APOB_HUMAN = Apolipoprotein B-100; APOC2_HUMAN = Apolipoprotein C-II; APOE_HUMAN = Apolipoprotein E; CO3_HUMAN = Complement C3; CRP_HUMAN = C-reactive protein; IGF1_HUMAN= Insulin-like growth factor I; PRTN3_HUMAN = Myeloblastin; NGAL_HUMAN = Neutrophil gelatinase-associated; IC1_HUMAN = Plasma protease C1 inhibitor; IPSP_HUMAN = Plasma serine protease inhibitor; SAA1_HUMAN = Serum amyloid A-1 protein

5.2.7 Informatics

Tandem quadrupole and high resolution Q-ToF LC-MS peptide MRM data were quantified with either TargetLynx (Waters Corporation) or Skyline (Maclean *et al.*, 2010), and analysed and visualized with Spotfire v6.0.0 (TIBCO software, Boston, MA). All statistical analyses were conducted with SIMCA (v14, MKS Umetrics AB, Umeå, Sweden) or IBM SPSS statistics v22 (IBM Corporation, Armonk, NY).

Orthogonal partial least squares-discriminant analyses (OPLS-DA) were performed on the candidate peptides as a ratio to their labelled standard and all data Pareto scaled. Sample runs that deviated significantly from the Hotelling's T^2 95% confidence interval were excluded and the model refitted. Peptides considered as contributing to the supervised separation of groupings were identified by consultation of the accompanying S-plot. OPLS-DA models were produced for control subjects *vs.* those with heart failure, and for heart failure with reduced ejection fracture *vs.* preserved ejection fraction.

Pair-wise comparisons for identified peptides of interest were performed by the Mann Whitney *U* test for independent samples. Logistic regressions were performed to calculate the probabilities of heart failure prediction for each individual peptide of interest, and as a combination of all these peptides. Receiver operator characteristic curves (ROCs) were produced using these probabilities and the areas under the curve (AUC) were calculated. All tests with a two-tailed p value of <0.05 were deemed as statistically significant.

5.3 Results

Although troponins and CRP have been discovered as good biomarkers for heart failure, it remains one of the biggest causes of mortality. This is because heart failure heart failure represents a heterogeneous clinical population for which biomarkers are sparse.

The causes of the disease also suggest that the types of changes observed are likely to be dependent on a number of factors. Thus, twenty healthy controls, twenty patients with heart failure with preserved ejection fraction (HFPEF) and twenty heart failure patients with reduced ejection fraction (HFREF) sample were used as proof of principal and test of overall sensitivity to classify this heterogeneous disease.

For this particular application, the microfluidics interface, equipped with a pre-column, was used in combination with one of the tandem quadrupole mass spectrometers. This was because it provided the best combination of throughput, loadability, sensitivity, precision and linearity. To increase chromatographic resolution, the micro-fluidics gradient conditions were set as previously described, from 2 to 30% B in 45 min at 35 °C.

Orthogonal partial least discriminant analysis (OPLS-DA) showed that disease and control samples could be classified using the data and results related to one of the SIL spike levels, as illustrated by the scores distribution/summary in Figure 5.4.

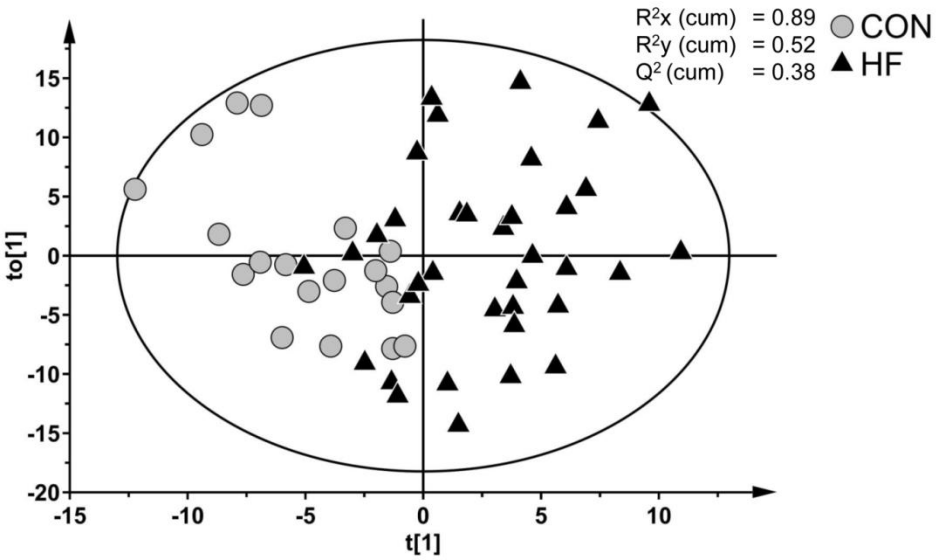


Figure 5.4*. Multivariate OPLS-DA analysis showing the separation and classification of patient and control samples. Circle = normal healthy patients; triangles = heart failure patients (HFPEF or HFREF). The model indicates that disease can be classified using a multiplexed tandem quadrupole LC-MRM-MS based assay.*

Partial separation of healthy controls and HF (combined HFPEF and HFREF) can be observed. Fit (R2) and prediction (Q2) values of 0.52 and 0.38 were obtained, respectively, for the developed model. ApoA1, CRP and plasma protease C1 inhibitor we found to be the main discriminators in the loadings. Further analysis (univariate analysis) of these three proteins showed significant changes in levels between the groups, as summarized in Figure 5.5A-C. In addition, a good discriminant power was obtained by combining these protein surrogate peptides, with an area under the receiver operating characteristic curve of 0.937 obtained as illustrated in Figure 5.5D. Previous work demonstrated that ApoA1 is potentially protective in HF, so the lower observed levels in the HF group would be consistent (Bhalla *et al.*, 2012).

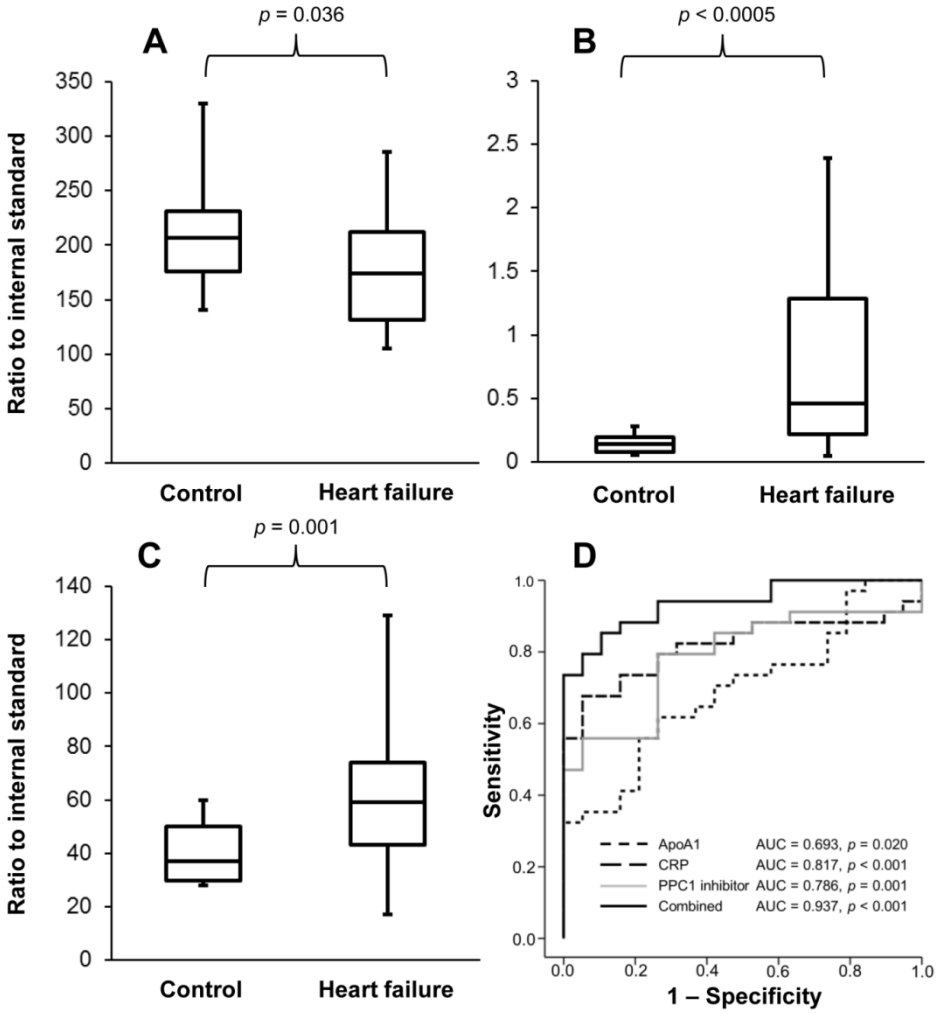


Figure 5.5*. Univariate analysis of ApoA1 (A), CRP (B) and Plasma Protease C1 Inhibitor (C) in HFPEF and HFREF and receiver operating curve performance analysis of peptide surrogates for Apo1, CRP and plasma protease C1 (D).*

A calibration line was constructed for each protein by altering the concentrations of a standard peptide and running with a fixed amount of labelled standard. Thus, a ratio of the analyte to internal standard was established for each concentration. For every sample, a fixed amount of SIL was added and ratio read to ascertain the concentration of the analyte (Figure 5.5). The absolute abundances of APOA1, CRP and Plasma Protease C1 inhibitor have been shown on Figure 5.6-Figure 5.8.

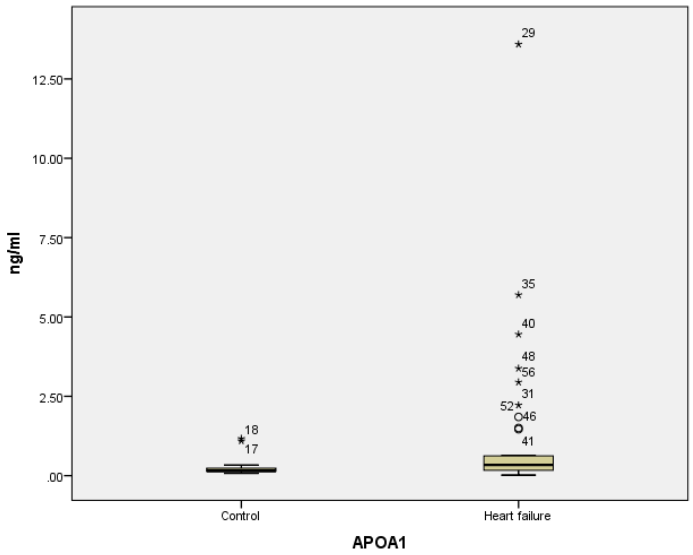


Figure 5.6. *Difference of the absolute abundance of protein APOA1 between the control (n=20) and heart failure (n=40) samples. These markers were established initially through multivariate analysis and tested in a univariate model. APOA1 had a significant change between control and heart failure cohort with a p value of 0.020. The numbers on the box blot represent outliers that were significant for the univariate analysis.*

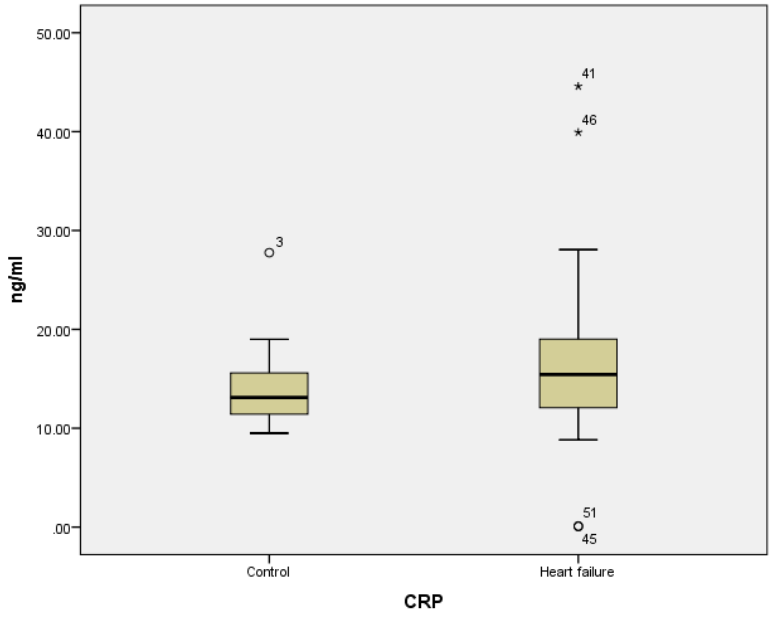


Figure 5.7*. Difference of the absolute abundance of protein CRP between the control (n=20) and heart failure (n=40) samples. These markers were established initially through multivariate analysis and tested in a univariate model. CRP had a significant change between control and heart failure cohort with a p value of <0.001. The numbers on the box blot represent outliers that were significant for the univariate analysis.*

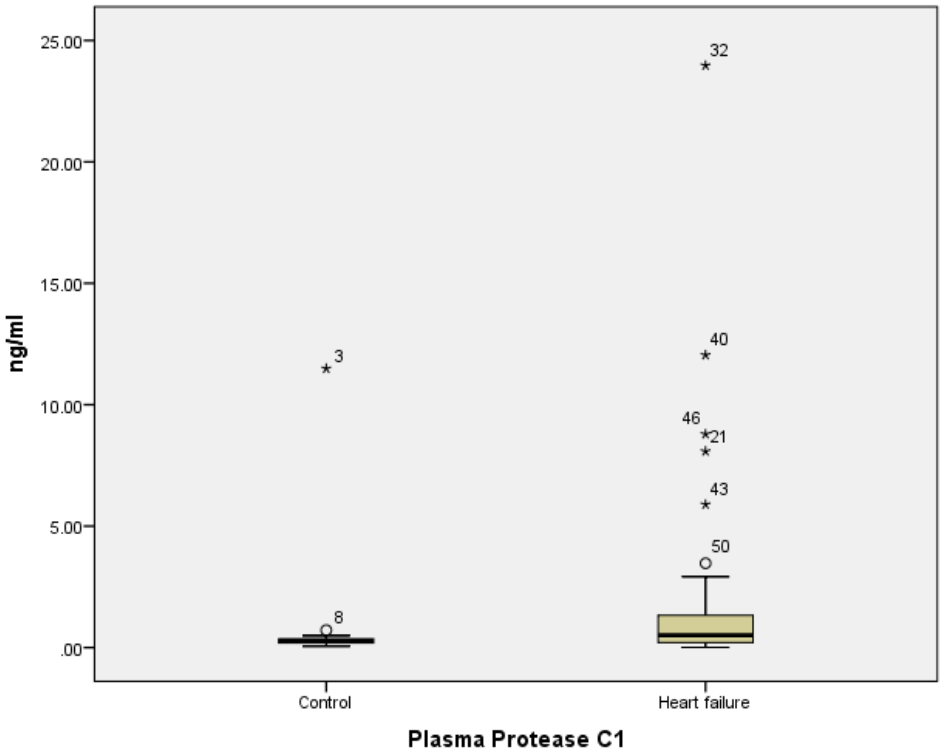


Figure 5.8. *Difference of the absolute abundance of protein plasma protease C1 between the control and heart failure samples. These markers were established initially through multivariate analysis and tested in a univariate model. Plasma protease C1had a significant change between control and heart failure cohort with a p value of <0.001. The numbers on the box blot represent outliers which were significant for the univariate analysis.*

Protease C1 has been associated with remodelling while CRP has been reported to be increased in heart failure (Anand *et al.*, 2005, Pfutzner *et al.*, 2010). Distinct separation between all three groups, namely control, HFPEF and HFREF, was not obtained with the selected subset of peptides, however, a partial separation model could be developed for HFREF and HFPEF, see Figure 5.9.

This project shows that when microfluidics interface, equipped with a pre-column, is used in combination with one of the tandem quadrupole mass spectrometers a classification model can be achieved using multiple analyte.

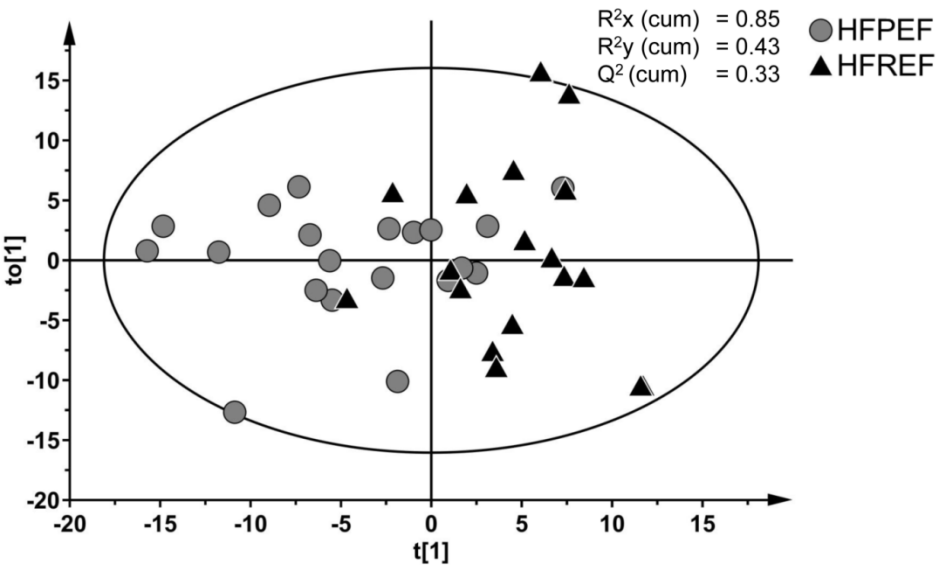


Figure 5.9. *Multivariate OPLS-DA analysis showing the separation and classification of HFPEF and HFREF samples. Circles = HFPEF patient samples; triangles = HFREF patient samples. The model indicates that disease can be partially classified using a multiplexed tandem quadrupole LC-MRM-MS based assay.*

Improvements to the model would necessitate a larger cohort of samples or the use of other peptides/proteins, which was outside the scope of this study. However, the strength of the model and the biological processes identified as contributing to the model indicate that this approach may yield successful opportunities for biomarker research.

5.3.1 Mobility enabled MRM methods

IM enabled oa-ToF MRM acquisition modes can provide either increased selectivity or sensitivity. These modes could be useful for lower level peptides that challenge assay specificity and are briefly described with example results presented. With the mobility enabled oa-ToF platform, the collision cell/mobility separation region comprises three stacked ring ion guide devices placed in series (tri-wave). The first region can be used to either trap ions or conduct CID fragmentation. Within the second region, the ions, either precursor or product ions, dependent on the use of the trap region, undergo mobility separation. In the third region, either the ions are transferred to the oa-ToF analyzer or CID fragmented. The device has been described in detail elsewhere (Chambers *et al.*, 2015, Giles *et al.*, 2004, Weaver *et al.*, 2007), as well as alternative uses of stacked ring ion guides, including electron transfer dissociation (Lermyte *et al.*, 2014) and top-down type studies

(Lermyte *et al.*, 2015). The two cases described here are graphically summarized in (Figure 5.10). The top (A) pane illustrates a case where ion mobility separation is achieved in the second region and CID conducted in the third region, aiming at achieving additional assay selectivity. The bottom (B) pane demonstrates ion mobility separation at the product ion level that is aimed at increasing sensitivity across the complete m/z fragment ion range by optimizing the duty cycle of the instrument. In short, product ions are trapped within the first region of the tri-wave device and gated into the high-pressure ion mobility region where they are separated according to their gas phase mobility, which is predominantly determined by mass, charge, size, and shape. As a result, fragment ions of the same mobility exit the cell as a series of compact packets. Hence, by synchronizing the pusher pulse that accelerates the fragment ions into the oa-ToF mass analyzer with the arrival of product ions into the pusher region, fragment ions are sequentially injected into the ToF analyzer with greatly enhanced duty cycle (85%) across the mass scale (Helm *et al.*, 2014).

The benefit of the latter case is illustrated in (Figure 5.11), showing typical sub 100 amol on-column results obtained with oa-ToF MRM data collected in normal MRM mode using EDC in the top (A) pane and with product ion IM separation in the bottom (B) pane. Data were normalized to the endogenous level to account for any experimental LC-MS variation and collected in so-called isotope stripping mode, whereby only isotopic information is stored for a predefined set of product ions, reducing file size significantly without losing quantitative information. As can be seen, the mobility enabled method provided better sensitivity. The effect of EDC can be noted as illustrated by the relative high intensity of the high mass versus the low mass product ions, which was aimed at sacrificing the low m/z product ion region nearly completely by promoting the oa-ToF duty cycle for the more specific high m/z product ions.

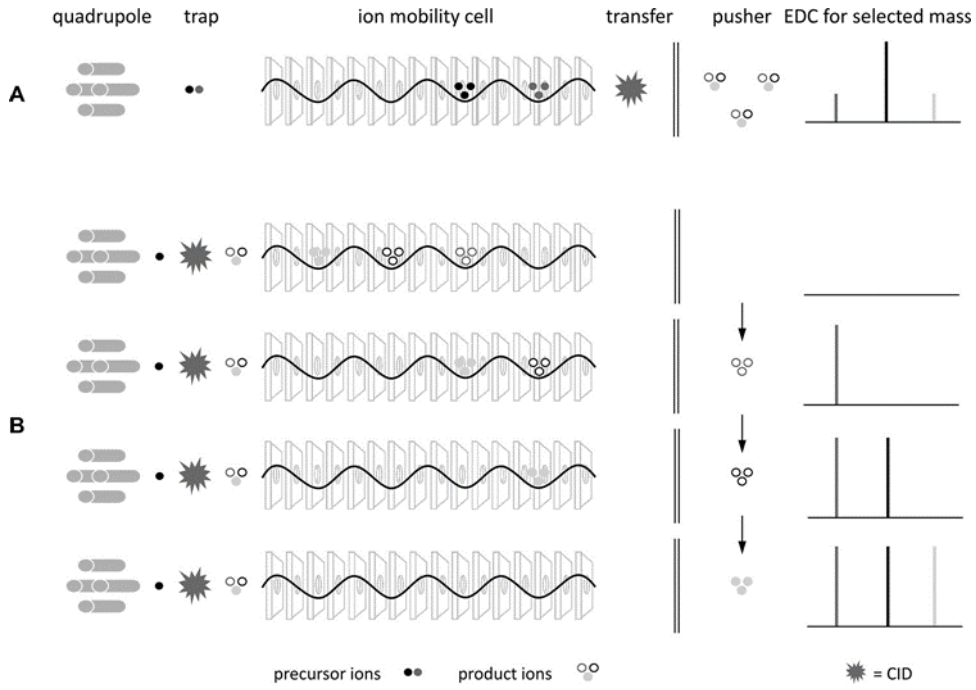


Figure 5.10 *Mobility-enabled oaToF MRM methods. (A) Precursor ion mobility separation followed by transfer CID including EDC for selected target masses and (B) trap CID followed by product ion mobility separation.*

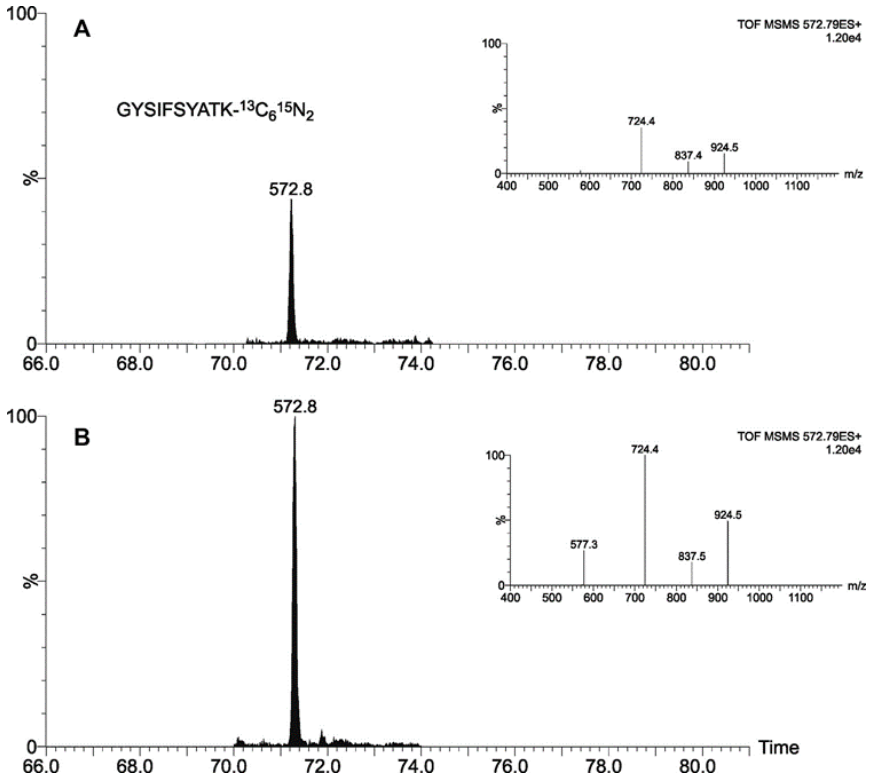


Figure 5.11. *Standard oa-ToF MRM with EDC (A) and trap CID followed by product ion mobility separation (B) for 62.5 amol of SIL-labeled GYSIFSYATK13C615N2 injected on-column, monitoring fragments y5 (m/z 577.3), y6 (m/z 724.4), y7 (m/z 837.5), and y8 (m/z (EDC) 924.*

5.4 Conclusions

LC-MRM-MS has been adopted and widely applied in translational biomarker studies due to the inherent speed of developing a multiplexed assay that can be deployed with relatively high sensitivity and throughput. In this study, building upon previous studies, the effect of both the LC scale and the choice of MS platform has been investigated. In biomarker discovery, many performance studies have been conducted using extensively high performance LC-MS systems. However, the technical and performance challenges in biomarker translation studies are different and require a robust and reliable platform, with good quantitative precision that can analyse much larger cohorts of samples. The results presented show that with regard to the LC component, despite nanoscale LC being the predominant separation technology in quantitative protein biomarker studies, we find that a larger 150 µm inner diameter scale microfluidic based system has the required sensitivity and quantitative performance, combined with throughput characteristics required for quantitative translational studies.

In the case of the MS platform, the use of higher resolution systems can have significant benefits in the targeted analysis of biomarkers in biological matrices. The elevated resolution can often be translated in increased specificity and therefore a more reliable measure of a peptide and hence proteins abundance. In this study, looking at a range of peptides in undepleted human plasma, the sensitivity was found to be very comparable between the modern tandem quadrupole and the quadrupole time-of-flight systems, but the combination of a high sensitivity tandem quadrupole with a microfluidic inlet provided the best coefficient of variation, throughput and sensitivity/signal-to-noise. It may be surprising, in some regards, that the elevated resolution of the second mass analyser of the time-of-flight systems does not translate to into improved quantitative performance, but in this study the peptides, even at levels close to their lower limit of quantitation provided a linear response with all MS platforms.

The preferred instrument configuration was used for the analysis of the selected peptides, and hence proteins, originating from heart failure disease patients, and compared to matched control samples. This preliminary study was designed to test the overall sensitivity and to classify this heterogeneous disease, resulting in a clear separation of the two disease groups from the control group and partial separation of the two disease phenotypes.

LC-MS technology continues to mature and the current technology has the required analytical performance to make large-scale biomarker translation studies a reality. This combined with the ability to build highly flexible, multiplexed assays will ensure that the technology will have an impact in the detection and monitoring of complex heterogeneous diseases.

Chapter 6

CONCLUSIONS AND FUTURE WORKS

6 General discussion

Research into HFPEF continues to grow as heart failure cases continue to increase (Ahmed *et al.,* 2014). HFPEF has proved to be as prevalent as HFREF in recent years and new guidelines have been put in place to define and diagnose HFPEF (Bhuiyan *et al.,* 2011). HFPEF is mostly seen in older population with women more affected than men especially those with a previous history of hypertension. As the population ages, HFPEF will inevitably continue to increase. The work described in this thesis aimed to develop a method for plasma proteomics to characterise (identify and quantify) the proteins observed in patient blood plasma that would enable us to investigate differences between controls, HFREF and HFPEF samples. These differences would be pushed forward as potential candidate biomarkers in HFPEF.

Heart failure is a complex clinical syndrome that involves multiple physiological processes as highlighted on Figure 4.23. Due to these physiological conditions, it has made it almost impossible to identify only one biomarker for heart failure. This is because each condition (e.g. inflammation) is enhanced by a different group of proteins that all together result in heart failure. As the population of patients with heart failure is increasing, their clinical management has also become very complex. These difficulties have been attributed to the changing profile of heart failure patients who in most cases are the older population, taking more medication and have multiple comorbidities (Tschope *et al.,* 2012, Dickson *et al.,* 2013). Lack of prediction of disease severity and mismatch between risk stratification and intensity of therapy has also been reported as a major concern in treating heart failure (Tariq *et al.,* 2014). Thus, identification of biomarkers that can allow measurements of the disease on a molecular level will greatly improve diagnosis, prognosis and selection of appropriate therapies for heart failure. In addition, with the new heart failure guidelines (Jessup *et al.,* 2009) and improved technology, patients will be able to be diagnosed effectively. Thus, this research explored the new ways of identifying new biomarkers for heart failure.

Many strategies have been used in heart failure research including gene therapy (Pleger *et al.,* 2013) that replaces the genetic material linked with heart failure with a normal gene (Sven *et al.,* 2013). Urinary proteomics has also been investigated for biomarkers discovery and diagnosis of heart failure (Rossing *et al.,* 2016). This research sought to incorporate novel strategies of plasma proteomics with the state of the art mass spectrometry to investigate the

proteins that have been altered in heart failure, mainly focusing on heart failure with preserved ejection fraction. The use of proteomics for plasma protein profiling using mass spectrometry has developed considerably over the years. In addition, the mass spectrometry platform performance enhancement such as ion mobility has enabled more accurate and reproducible quantitative data to be generated (Giles *et al.,* 2011). In the last decade, research involving biomarkers of heart failure has significantly increased. This is probably due to the high-throughput molecular biology techniques used that has enabled availability of rapid improvement of biomarker testing which has cut down the cost of analysis as Ahmad *et al.,* 2014 reported. Over 6500 publications on biomarkers in heart failure are now available on PubMed search (van Kimmenade *et al.,* 2012).

Despite these exponential gains in biomarkers of heart failure research in the last decade, heart failure cases are still on the rise. This could be due to lack of new biomarkers to diagnose and monitor the disease progression. The failure to find new biomarkers could be attributed to several issues including inconsistencies in research methodologies, insufficient study size and lack of clinical correlations. These issues have made it difficult to translate research into routine clinical practice that has resulted in delayed adoption of newly established biomarkers. Considering all the biomarkers of heart failure that have been discussed in literature, only B-type natriuretic peptide (BNP) and N-terminal proBNP (NT-proBNP) have been cleared for diagnosis of heart failure by food and drug development administration (FDA). On the other hand, beyond the established natriuretic peptides, highly sensitive troponins (Gaggin *et al.,* 2013), galectin-3 and serum ST2 have been cleared to be used for prognosis in heart failure (Ahmad *et al.,* 2014). However, their role in clinical use is poorly understood and more studies about them is needed.

Gerszten *et al.,* 2008 reiterates the need for novel biomarkers and claims that most of the plasma protein biomarkers for cardiovascular disease in clinical use today were developed as an extension of targeted physiological studies of pathways that had been identified previously. This include inflammation, coagulation and endothelial injury. They report that despite the success in using a single biomarker in disease diagnosis, combining individual biomarker improves the prediction of risk of an individual patient. In another study, 10 contemporary biomarkers of cardiovascular disease were monitored in 3,000 participants for almost 10 years. Most of the biomarkers measured including B-type natriuretic peptide, urinary albumin excretion, C-reactive protein, renin, homocysteine were significant

predictors of cardiovascular events (Wang *et al.,* 2006). When the biomarkers were combined, the risk prediction was improved slightly (Pepe *et al.,* 1997). In this research, we can also see a slight improvement in disease prediction when we combine the potential biomarkers. The downside of this strategy is that the existing candidate biomarkers are derived from pathways (e.g. inflammation) that have already been established. As a result, these biomarkers give predictive information that correlates with currently measured characteristics, thus, restricting their incremental predictive value (Ahmad *et al.,* 2014). This could be improved by identifying uncorrelated biomarkers along novel pathway. This thesis has proposed some novel pathways that could be associated with heart failure based on the proteins identified and their interactions in a molecular level.

6.1 Method development

Biomarker discovery with plasma proteomics requires a method that is reproducible, sensitive, has a high throughput, covers a wide range of the plasma proteome and can provide quantitative precision (Carr *et al.,* 2015). However, plasma samples have several challenges, the main one being the dynamic range of proteins dominated by HAPs (Tu *et al.,* 2010). These HAPs mask the identification of LAPs that could be potential biomarkers. Therefore, it is essential to eliminate these HAPs for analysis of the LAPs whilst maintaining minimal sample loss.

In this thesis, a method described in chapter 2 and chapter 3 was developed for protein fractionation, depletion and enrichment. Firstly, RP-RP fractionation with HPLC was used for sample fractionation prior to analysis with label free LC-MS. Sample fractionation was important because it enabled a greater peak capacity so that ions of low abundance could be resolved and thus observed. In addition, it reduced the dominance of high abundance peptide ions within the peak capacity space. With higher peak capacity, we could expand the dynamic range. In 2012, Zhu *et al* developed a strong anion exchange/RP method to deplete high abundance proteins in human plasma. In their research, they only identified 83 proteins. In the study described in chapter 3, only 91 proteins were identified. This poor return was perhaps due to poor separation of peptides and peak resolution thus defeating the aims of the experiment. This prompted the introduction of the second method with LRA affinity binding matrix. This matrix has a high affinity for lipoproteins and was used to pull out a number of

high abundance proteins including apolipoproteins and other lipoproteins. This LRA matrix also did not yield enough proteins as discussed in chapter 3 that prompted the development of M3 which was used for LAPs enrichment. Meanwhile, the LRA method was being optimised in parallel to the development of the M3 method.

In chapter 3 the enrichment of LAPs with M3 has been described which was later used for a pilot study as shown in chapter 4. This M3 method was very successful since it improved the protein identification considerably. Despite this success, this method had some setbacks. Firstly, it did not address the issue of throughput when large samples analysis were involved. This is because each sample required 5 fractions to be analysed in triplicate, thus totalling to fifteen 75 mins MS runs per sample. In a clinical study setting, this method would have been laborious. Secondly, despite the many proteins identified with M3, it was hard to identify fractions that had only HAPs or proteins of interest since they smeared across all fractions. This introduced reproducibility issues. Finally, due to the number of MS runs (15 per sample) and data sets generated it was time consuming, laborious and required complex informatic strategies. However, while the M3 method was falling out of favour, the LRA protocol was optimised.

Chapter 4 highlights the new development with LRA that was used to carry out clinical study with HF samples. The LRA method reduced the variations that were involved with M3 method significantly. The advantage of LRA is that there is no fractionation as such involved and the sample results in a single shot analysis. Thus, instead of analysing 15 fractions, each sample is analysed in triplicate on MS. In addition, immunodepletion that could have caused non-specific binding of LAPs and potential sample loss was not used in this protocol. In this protocol ammonium salt of deoxycholic acid (ADC) was introduced which enabled proteins to be eluted from the LRA matrix. This optimised LRA method not only identified many proteins but also LAPs at pg/mL level (cytokines e.g RhoA). Moreover, the proteins identified were associated with heart failure as shown in chapter 4, which could be potential biomarkers. Thus, this LRA method was reproducible, sensitive, had a high throughput, covered a wide range of plasma proteome and displayed quantitative precision. The LRA matrix is particularly inexpensive, especially when compared to mRP-C18 columns and equalizer beads.

Perhaps the most challenging bit in this thesis was data storage, processing and analysis. Firstly, considering the number of raw data files (15 per sample) generated with the M3

method which were approximately 5-8 GB each meant we had to have a large storage capacity for these files. This also meant that the computer that stored the data generated from mass spectrometer was cleared every time to avoid backlog that would otherwise stop the MS runs. Secondly, data processing with PLGS and Progenesis for the M3 raw data files was impossible due to their big size and volume. Therefore, data had to be processed in batches that could have resulted in some inconsistencies especially when 2 different computers were required due to storage space. Thirdly, data analysis required the use of several data mining software to attain confident results. This was particularly challenging because each software produced different results as highlighted on chapter 4. As a result, the most significant proteins from each software were chosen for the final analysis, with the most confident proteins the ones identified in all the software used (supplementary Table C-1).

6.2 Identification of putative prognostic biomarkers

Chapter 4 highlights 30 proteins that were differentially expressed between the control and HF groups. Their expression levels were shown including their molecular functions. Most of these proteins were involved in protein binding, catalytic activity and metal binding and were all thought to potentially relate to disease pathology. This section highlights the mechanistic plausibility of these proteins as potential biomarkers of heart failure.

6.2.1 Acyl-coenzyme A synthetase ACSM5, mitochondrial (ACSM5) ↓HF

Acyl-coenzyme A synthetase ACSM5, mitochondrial is an ATP-binding protein with medium-chain fatty acid (CoA ligase) located in the mitochondrion matrix. It is expressed in the kidney and liver. It has been associated with multiple risk factors of cardiovascular diseases including pathogenesis of hypertriglyceridemia, visceral obesity and hypertension (Iwai *et al.,* 2003). These risk factors are enhanced by SAH, an acyl-CoA synthetase influencing fatty acids metabolism. Fatty acid metabolism is a cornerstone of many diseases including obesity and cardiovascular diseases. Investigation by Ellis *et al.,* 2015 in animal studies showed that cardiac-specific overexpression of ACSL1 (Acyl-CoA synthase long

chain family member 1) resulted in cell death, physiological dysfunction and reduced lifespan. In this study, ACSM5 was downregulated in both HF groups.

6.2.2 Isoform 5 of Androgen-induced gene 1 protein (AIG 1) ↓HF

Isoform 5 of Androgen-induced gene 1 protein is located in the membrane of the cell and highly expressed in the heart, liver, kidney, ovary and testis. It is responsible for androgen-regulated growth of hair follicles. Its expression is higher in hair follicles in men than women (Seo *et al.,* 2001). Epidemiology studies suggest that males in their reproductive years are at higher risk of cardiovascular diseases as compared to age matched female (Liu *et al.,* 2003). This has been associated with many factors including estrogen protective effects in females and androgen impairment in males. According to Cai *et al.,* 2016 one of the possible mechanisms by which androgens acts on the cardiovascular system is through their role in modulation of endothelial functions (Jing-Jang *et al.,* 2016). Endothelial dysfunction has been reported to be the initial step in the development of atherosclerosis and cardiovascular diseases. In this research, the protein levels of AIG1 were lower in both HF groups and evidently higher in males than females that support Seo *et al.,* 2001 findings.

6.2.3 Rho GTPase-activating protein 29 (ARHGAP29) ↓HF

Rho GTPase-activating protein 29 is a metal-binding, zinc protein located in the cytosol. It is widely expressed in skeletal muscle and the heart but also liver and pancreas. It is responsible for GTPase activation by converting them to an inactive GDP-bound state. Rho GTPases are a subfamily of Ras superfamily proteins responsible for regulation of cell shape change, cell adhesion and migration and cytokinesis (Bustelo *et al.,* 2007). The outcome of activation of a single Rho GTPase is dependent on their downstream effectors. The downstream targets dictates whether Rho GTPase plays a role in cell morphology, vesicular trafficking or cell cycle control. As our results suggest, ARHGAP29 was downregulated in both HF groups implying it was involved in disease progression.

6.2.4 Isoform 2 of Argininosuccinate lyase (ASL) ↑HF

Isoform 2 of Argininosuccinate lyase is responsible for the synthesis of L-arginine and fumarate from argininosuccinate. L-arginine generates nitrogen oxide (NO) and L-citrulline through nitrogen oxide synthases (NOS). Citrulline can be recycled back to arginine by argininosuccinate synthatase (ASS) and ASL thus forming the citrulline-NO cycle (Figure 6.1) (Ayelet *et al.,* 2011). This sub pathway is part of the pathway urea cycle, which is itself part of Nitrogen metabolism. NO production activates intracellular soluble guanylate cyclase (sGC) which aids the production of 3'5'-cyclic guanosine monophosphate (cGMP) responsible for most of the physiological and pathological effects of NO (Dweik *et al.,* 2001, 2002). The effects are mediated through cGMP-gated channels, cGMP-regulated phosphodiesterases or cGMP-dependent protein kinases (Figure 6.1).

Nitric oxide (NO) is a colourless and odourless gas made up of nitrogen and oxygen. In the human body, it is an autocrine and paracrine signalling molecule responsible for smooth muscle relaxation, thus decreases blood pressure. In addition, by inhibiting smooth muscle cell proliferation and platelet aggregation (Garg *et al.,* 1989), it acts as a vasoprotector. It is also responsible for tumor cell lysis, bacterial killing and stimulation of hormonal release (neurotransmitter) (Simon *et al.,* 2013). NO has been reported to play major role in pulmonary hypertension (PAH) which is a condition that causes the failure of the right ventricle of the heart resulting to death (Dweik *et al.,* 2001). Reduced levels of exhaled NO has also been associated with increased pulmonary artery pressures (Machado *et al.,* 2004) suggesting that exhaled NO could be an important maker of disease. NO has also been reported to play a major part in lung biology and has been associated with Asthma, Cystic fibrosis, Bronchopulmonary dysplasia in addition to pulmonary hypertension (Gaston *et al.,* 2001, Barnes, 2010, Dweik *et al.,* 2002).

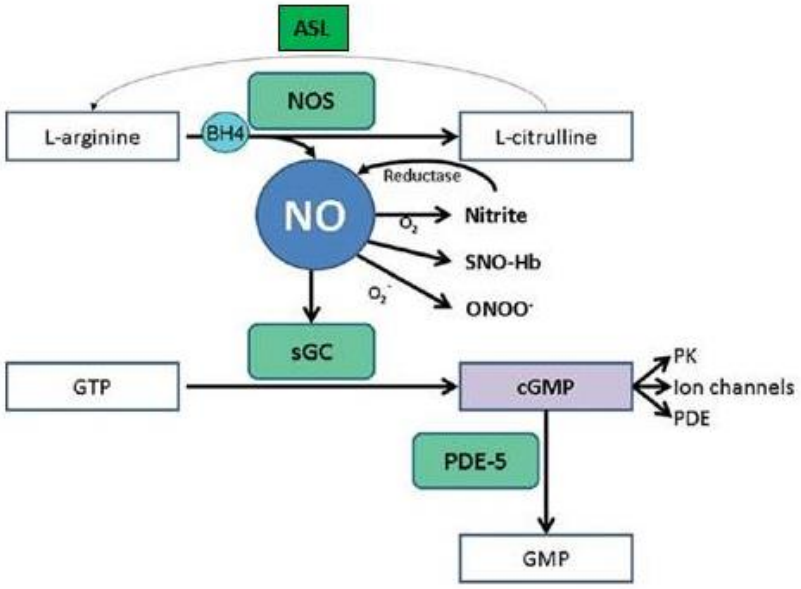


***Figure 6.1**. Nitrogen oxide (NO) synthesis and signalling pathways. NOS (nitrogen oxide synthases), (sGC) soluble guanylate cyclase, cGMP (cyclic guanosine monophosphate), PK (protein kinases), GTP (Guanosine triphosphate) and ONOO (peroxynitrite). (Adapted from Tonelli et al., 2013).*

6.2.5 C4b-binding protein alpha chain (C4BPA) ↑HF

C4b-binding protein alpha chain is a large (500 kDa) protein plasma glycoprotein produced in the liver (Martin *et al.*, 2009). It is responsible for controlling the classical pathway of complement activation and interacts with protein S and serum amyloid P component (SAP). It is expressed in the chylomicrons (Samstad *et al.*, 2014) in the plasma. Chylomicrons (ultra-low density lipoproteins) one of five types of lipoproteins responsible for lipid transportation from the intestines to other parts of the body. Research by Martin *et al.*, 2009 emphasises on the clinical relevance of complement activation and increased plasma levels of C4b-binding proteins as seen in Critical limb ischemia patients (CLI). They reported that CLI is a disease peripheral arterial disease caused by reduced blood flow to the lower limbs because of atherosclerosis. Atherosclerosis is caused by accumulation of lipids in the arteries that form plaque which obstruct the blood flow (Blasko *et al.*, 2008). C4b-binding protein was increased in both HF groups in this study. Previous unpublished work with the van Geest (VG) laboratory (University of Leicester) also showed a significant change of SAP in HF.

6.2.6 Complement factor D (CFD) ↑HF

Complement factor D is a signalling protein secreted by adipocytes and functions as an adipokine. It is responsible for the regulation of insulin secretion in mice. Mutation in this gene is associated with recurrent bacterial meningitis infection in homosapiens. Its activation has been reported to be an emerging player in the pathogenesis of cardiovascular disease (Carter *et al.,* 2012). CFD is also known as adipsin, C3 convertase activator or properdin factor D. It cleaves factor B when coupled with factor C3b to form C3bbb complex which is becomes the C3 convertase of the alternative pathway (Speidl *et al.,* 2011).This complement (C3) activation is responsible for endothelial cell activation, stimulation of cytokine release (Barratt-Due *et al*., 2012) from vascular smooth muscle cells and also promotion of plaque rapture which all leads to atherosclerosis (Carter *et al.,* 2012). Endothelial cell activation represents the beginning of the development of atherosclerosis. In this research, CFD levels were upregulated in both heart failure groups and down regulated in the control group that confirms the findings of Speidl *et al*., 2011 and Carter *et al.,* 2012.

6.2.7 Fatty acid-binding protein, heart (H-FABP) ↓HF

Heart fatty acid-binding protein, is a lipid binding protein located in the heart. It is responsible for the intracellular transport of long chain fatty acids and their acyl-CoA esters. This protein was recently reported as an ischemic biomarker for acute heart failure (Hoffmann *et al.,* 2015). According to Basak *et al.,* 2011, H-FABP is a small protein of about 15 kDa that helps in determining the kinetic release of cardiac biomarkers (Karbek *et al*., 2011). He reports that H-FABP appears earlier than the larger molecules (cardiac troponins) thus making it an important marker for acute coronary syndrome. In this study, H-FABP levels were down regulated in both HF groups (lowest in DHF) and upregulated in the control group.

6.2.8 Golgi-specific brefeldin A-resistance guanine nucleotide exchange factor 1 (GBF1) ↑HF

Golgi-specific brefeldin A-resistance guanine nucleotide exchange factor 1 is a lipid binding protein located in the cytoplasm, golgi apparatus, lipid droplet (LD) and the membrane. This

protein is expressed everywhere in the human body. It is involved in ARF guanyl-nucleotide exchange factor (GEF) activity. Metabolism of LDs has been reported to play a role in metabolic disease such as lipodystrophies and obesity (Walther *et al.*, 2012, Zechner *et al.*, 2012). Obesity in particular is a serious health concern due to its predisposition to serious diseases such as heart disease, diabetes, liver disease and cancer (Bozza *et al.,* 2010). The mechanisms linking LDs to secretory pathways are poorly understood. However, the role of Arf1 small protein in recruitment of LD-associated protein to LD surface has been reported (Guo *et al.,* 2008). In this research, this protein (GBF1) was significantly increased in the HF groups as compared to the control group.

6.2.9 Glutathione peroxidase 3 (GPX3) ↓HF

Glutathione peroxidase 3 is a protein secreted in plasma. It is responsible for protecting cells and enzymes from oxidative damage. It does this by acting as a catalyst in the reduction of hydrogen peroxide, lipid peroxides and organic hydro peroxide by glutathione. A research carried out by Forgione *et al.,* 2002 showed that reduced levels of GPX activity results in abnormal vascular and cardiac structure and function. In addition, Lapenna *et al.,* 1998 reported a correlation between reduced GPX activity and vascular injury when atherosclerotic plaques were excised from patients with carotid artery disease. When 50% of GPX was inhibited by drugs there was an increased influx of reactive oxygen species and low-density lipoprotein oxidation (Rosenblant *et al.,* 1998) that have previously been associated with endothelial cell activation.

6.2.10 Luc7-like protein 3 (LUC7L3) ↑HF

Luc7-like protein 3 is a DNA binding protein located in the nucleus speckle. It binds the cAMP regulatory element DNA sequence but has been reported to play a role in RNA splicing (Shipman *et al*., 2006). mRNA splicing of type five voltage gated cardiac sodium (Na^+) channel alpha subunit (SCN5A) is increased by heart failure. This Na^+ channel is located in cardiac muscle and is responsible for the initial upstroke of the action potential in an electrocardiogram. SHF has been reported to be associated with reduced cardiac voltage gated sodium (Na^+) channel current which have been implicated in the increased risk of

sudden death (Gao *et al*., 2013). When mRNA is spliced with SCN5A, HF is increased due to angiotensin II and hypoxia (GAO *et al*., 2011).

6.2.11 Neuropeptide FF receptor 2 (NPFFR2) ↓HF

NPFFR2 is a G-protein coupled receptor protein localised in the cell membrane. It is mainly expressed in the thymus, testis and small intestine but also in heart, lung, kidney and pancreas (Cikos *et al.,* 1999). The binding of synthetic FMRF-amide like ligands (human RFamide-related peptide-1) that diminishes cardiovascular performance can activate this receptor. Studies by Nichols *et al.,* 2012 report that human RF-amide-related peptide-1 (hRFRP-1) and its signalling pathway may provide targets to address cardiac dysfunction. Other studies have reported involvement of neuropeptide FF (NPFF) in generation of cardiovascular responses (Jhamandas *et al.,* 2013). Intrathecal and ICV administration of NPFF showed dose-dependent elevations in arterial blood pressure and heart rate (Fang *et al.,* 2010)

6.2.12 Sulfhydryl oxidase 1 (QSOX1) ↑HF

Sulfhydryl oxidase 1 is a FAD-linked quiescin/sulfhydryl oxidase that catalyses the oxidation of sulfhydryl groups in peptide and protein thiols to disulphides while reducing oxygen to hydrogen peroxide (Chakravarthi *et al.,* 2007). It is expressed in the heart, lung, liver, placenta, skeletal muscle and kidney. QSOX has been implicated in many activities including production of extracellular matrix, angiogenesis, protection from apoptosis, cell differentiation and progression of atherosclerosis (de Andrade *et al.,* 2011). QSOX expression explains its anti-oxidative role rather than pro-apoptotic that may explain the potential link between this protein and myocardial injury (Doehner, 2012). In addition, QSOX expression in animal model and human studies corresponded to pressure overload, hypertrophy and development of acute decompensated heart failure (Mebazaa *et al*., 2012). In my research, QSOX1 was overregulated in both heart failure groups and compared to the control group.

6.2.13 Rab3 GTPase-activating protein non-catalytic subunit (RAB3GAP2) ↑HF

Rab3 GTPase-activating protein non-catalytic subunit is a GTPase activating enzyme located in the cytoplasm. It is involved in regulating exocytosis of neurotransmitters and hormones. Rab3 GTPase activating complex specifically converts active Rab3-GTP to inactive form Rab3GDP (Nagano *et al.,* 1998). Involvement of this RAB3GAP3 protein with cardiovascular disease is poorly understood. However, its mutation has been reported to cause Martsolf syndrome. In this thesis, RAB3GAP3 was upregulated in HF groups and downregulated in control group.

6.2.14 Zinc finger protein 701 (ZNF701) ↑ HF

Zinc finger protein 701 is a DNA and metal binding protein found in the nucleus that may be involved in transcriptional regulation. The zinc finger proteins have been reported to highly abundant in human plasma yet they remain poorly characterised. These proteins interact with both DNA and RNA and have been described as multifunctional proteins that link transcription with post transcriptional processes which have been linked with numerous human diseases including cancer (Ladomery *et al*., 2002). The ZNF701 functions in particular has not been reported. However, its involvement in HF samples could have an implication in cardiovascular diseases. In this study, ZNF701 levels were upregulated in both HF group with higher levels observed in SHF.

6.2.15 Isoform 3 of Collagen alpha-1 (XVIII) chain (Col18a1)

Isoform 3 of Collagen alpha-1 (XVIII) chain is a metal binding, zinc protein involved in cell adhesion. It is an endostatin that potentially inhibits endothelial cell proliferation and angiogenesis. The angiogenesis inhibition could be through Col18a1 binding to the heparin sulfate proteoglycans which are involved in growth factor signalling. It is expressed in the liver, kidney, lung, skeletal muscle and testis. Mutation in Col18a1 gene has be associated with Knobloch syndrome 1 (KNO1) which is a disorder characterised by abnormalities of the eye (Sertie *et al.,* 1996). Its involvement in endothelial cell morphogenesis and appearance in HF samples strongly suggests its involvement in cardiovascular diseases. One of the roles of

endostatin is modulation of embryonic vascular development through increasing proliferation, migration and apoptosis. Col18a1 in particular has been shown to be distributed in epithelial and endothelial membranes (Lorenza *et al.,* 2005). One of the genes encoding it on chromosome 21 is associated with congenital heart disease phenotype observed in Down's syndrome. Investigation by Lorenza *et al.,* 2005 confirmed that Col18a1 is highly expressed in connective tissue core of endocardial cushions and forming atrio-ventricular valve leaflets (Carvalhaes *et al*., 2006). Col18a1 had very low protein amounts thus, could not be analysed on box plots.

6.3 Pathway analysis of some of the 30 proteins selected as potential biomarkers

Rho (member of the Ras family) has over 60 known downstream effectors that are responsible for regulation of GTPase proteins. Approximately 10 different mammalian Rho GTPases including Rho (A, B, C isoforms), Rac (1, 2, 3 isoforms) and Cdc42 have been identified (Bishop *et al.,* 2000). In addition, about 20 GTPase activating proteins (GAPs) responsible for increasing the intrinsic rate of GTP hydrolysis of GTPases have also been identified (Sagnier *et al.,* 1994). The action of GTPase is dependent on its downstream targets. RhoA is associated with the dynamic reorganisation actin cytoskeleton. On the other hand, actin together with myosin form the contractile filaments of muscle and are responsible for mobility and contraction of cells. Besides reorganisation of the cytoskeleton, RhoA is also responsible for activation of Rho kinases (ROCK). Upregulation of ROCK has been associated with many disease pathways. This mainly affects physiological processes associated with (myosin light chain phosphatase (MLCP) and epithelial nitric oxide synthase (eNOS).

Inhibition of eNOS (Takemoto *et al.,* 2002) impairs the relaxation of the muscles and causes vasoconstriction thus increasing vascular resistance and hypertension. When eNOS is partially uncoupled it results in the production of peroxynitrite ($ONOO^-$) through the reaction of superoxide anion (O_2^-) and NO. In addition, O_2^- in the presence of superoxide dismutase (SOD) forms hydrogen peroxide (H_2O_2). The H_2O_2 formed can be broken down to form water and alcohol (redox reaction) in the presence of glutathione peroxidase 3 (GPX3) enzyme or undergo Fenton reaction (Fe^{2+} + H_2O_2) to form free radicals ($OH^{\cdot}$) resulting to

apoptosis/necrosis. This process is a crucial step in the development of many cardiovascular diseases.

On the other hand, inhibition of myosin light chain (MLC) phosphatase activity in vascular smooth muscles (Nunes *et al.,* 2010) affects the activity of the myosin causing impaired contraction. Since the main substrate of ROCK is MLCP, this incriminates ROCK in mediating Ca^{2+} activity in the cell resulting in contraction by inhibition of MLCP. In addition, intracellular Ca^{2+}-calmodulin (CaM)-activates myosin light chain kinase (MLCK) which leads to activation of MLC, thus resulting in contraction (Nunes *et al.*, 2010). MLC can be phosphorylated directly from ROCK.

In this research, Rho GTPase-activating protein 29 (ARHGAP29) (Figure 4.19) was upregulated in the control group and downregulated in both heart failure groups suggesting that there exists adequate amounts of NO in the epithelium of the HF patients to compensate its reduced levels as a result of ROCK activation. It is also evident that MYH9 (Figure 4.19) responsible for muscle contraction was upregulated in control group and downregulated in the HF groups. In addition, ras interacting protein 1 (RASIP1) has been reported to recruit ARHGAP29 and supress the RhoA signalling and diminish ROCK (Myagmar *et al.,* 2005). The fact that MYH9 protein was downregulated in both HF groups suggests that its activities were suppressed in the HF group.

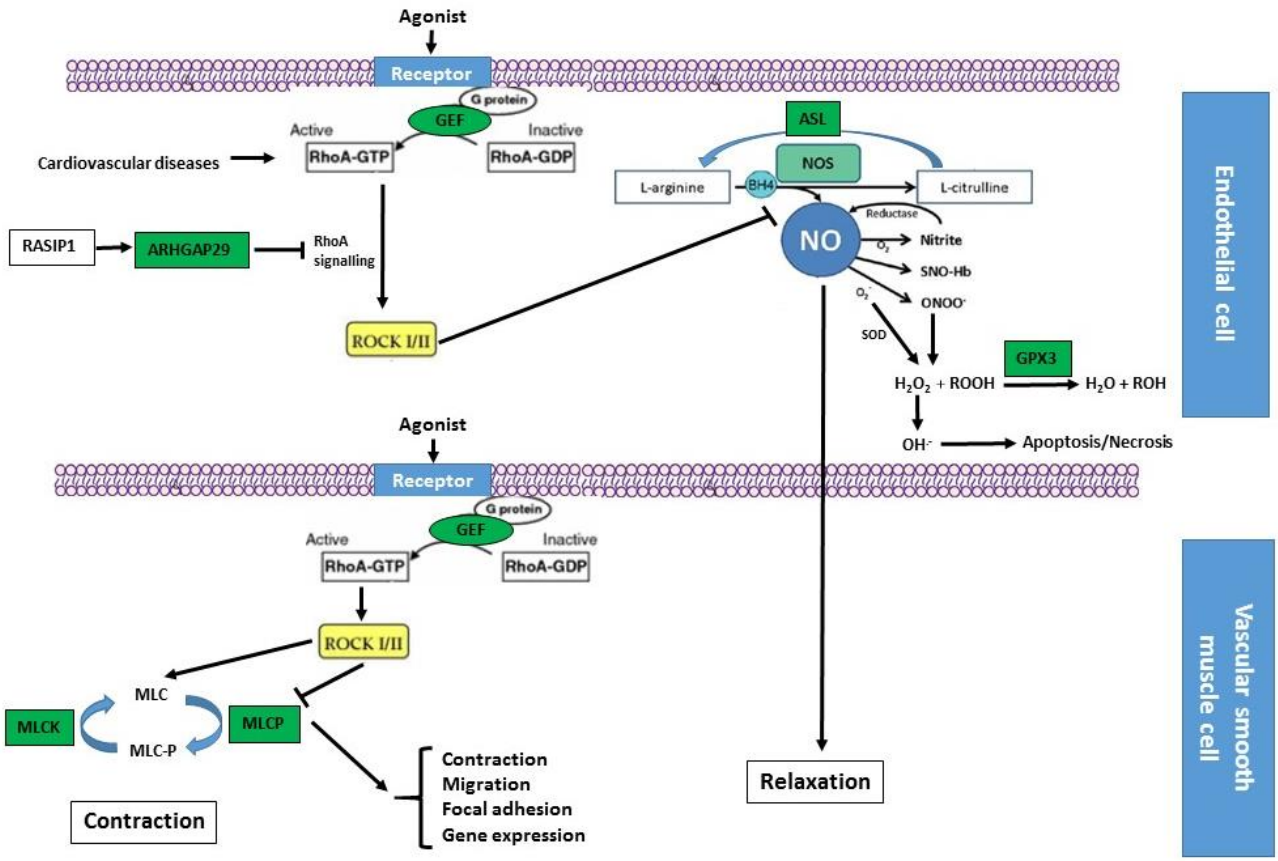


Figure 6.2. *Schematic illustration of Rho GTPase signalling pathway leading to actin reorganisation. Kinase phosphorylates and activate proteins while phosphatase dephosphorylates and inactivates proteins. Green boxes indicate proteins identified in this research. RhoA kinase (ROCK), myosin light chain kinase (MLCK), myosin light chain phosphatase (MLCP), argininosuccinate lyase (ASL), Rho GTPase-activating protein 29 (ARHGAP29), nitric oxide synthase (NOS), nitric oxide (NO), glutathione peroxidase 3 (GPX3) and guanine nucleotide exchange (GEF).*

The means of the protein amounts for the 2 proposed biomarker candidates (ASL and GPX3) have been highlighted on Table 4.15. The raw data shows that ASL was upregulated in heart failure cohort compared to control group while GPX3 was downregulated in heart failure cohort as compared to control group. The results from data analysis have also highlighted the same trend in protein expression levels of the 2 proteins in the 3 cohorts (Figure 4.19-Figure 4.21).

6.4 Summary

Three novel methods for biomarker discovery were developed. The M3 method provided a deeper coverage of the plasma proteome with over 1400 unique proteins (1%FDR) identified from 30 samples. Depending with the application, the M3 method could be used to improve throughput. The LRA method reduced the number of variations involved in sample preparation thus improved throughput considerably. Over 800 unique proteins (1%FDR) were identified from 90 samples each with a single fraction. This LRA method was robust, reproducible and improved throughput significantly. Moreover, potential biomarkers that could possibly lead to heart failure were identified with the LRA method.

Overall, 2 (ASL and GPX) out of 30 candidate biomarkers were selected for HFPEF due to their roles in heart muscle relaxation. The role of ASL in NO production is particularly very important in maintaining the homeostasis of the epithelium. Equally, GPX3 is also very important since it prevents premature cell death because of circulating free radicals produced in the blood. Therefore, it would be vital to measure the expression levels of these proteins in a larger cohort between HFPEF and HFREF. These could then be validated using the targeted LC-MS based assays developed in chapter 5 which could detect potential biomarkers in neat plasma. Should these proteins test positive, then a further clinical and multicentre validation will be carried out to boost its chances of translating into the clinical phase.

6.5 Future direction

Potential biomarkers of heart failure with preserved ejection fraction were identified in this research. However, due to the time constraints and funding limitations, validation of these proteins could not be carried out. In the past, ELISA has been used to validate potential biomarkers in samples. However, it comes with several limitations that have been discussed in chapter 4. Western blots have also been used to detect the proteins under investigation. The advantage of western blot is that the sensitivity of analysis can be increased substantially depending on the quality of the antibodies used. However, western blots can be costly and time consuming. Currently, targeted methods are becoming common as a validation technique due to their accuracy and precision. In chapter 5, a targeted method with multiple

reaction monitoring (MRM) was developed which we hope to utilise on a larger cohort of samples to validate our findings.

In this thesis, thirty significant proteins were identified as potential biomarkers due to their association with cardiovascular diseases. Some of these potential biomarkers played major roles in heart failure progression as shown on the pathway analysis on Figure 6.2. These proteins will be tested on a larger cohort. Firstly, a clinical design for controls, HFPEF and HFREF samples will be established. Stable isotope label (SIL) standards for the proteins mentioned on Figure 6.2 will be purchased and using the MRM method developed in chapter 5 test these proteins on a larger cohort. These plasma samples will be spiked with the SIL standards before running them on MS. Their retention times and intensities will be analysed in each sample. If these proteins test positive, a mechanistic study will be incorporated using *in vivo* model. In this *in vivo* study, different drugs will be tested on pathways of wild type and knockout mouse models and their cardiovascular activities assed. Some of the things that will be assessed will be the pumping function of the heart, hyper/hypotension and plaque development in the arteries.

Post-translational modification (PTM) like phosphorylation and glycosylation could also be investigated. In this thesis, the activities of myosin light chain kinase (MLCK) and myosin light chain phosphatase (MLCP) have been mentioned on the pathway analysis. PTMs transform the structure of the proteins thereby influencing and regulating their functions. This transformation may change the structure of the protein, direct cellular trafficking and influence proteolytic breakdown that are common mechanisms of disease formation. Therefore, investigating proteins that have PTMs could be useful in finding new biomarkers of a disease.

Chapter Seven

Appendix

7 Appendix

7.1 Appendix A

List of the proteins identified with M3 and LRA.

Table A- 1. *M3- 141 unique proteins (Depleted sample)*

14-3-3 protein zeta/delta	Insulin-like growth factor-binding protein 2	ProSAAS
Acyl-coenzyme A synthetase ACSM6, mitochondrial	Insulin-like growth factor-binding protein 4	Protein fem-1 homolog C
Adiponectin	Insulin-like growth factor-binding protein 6	Protein HEG homolog 1
ADP-ribosyl cyclase/cyclic ADP-ribose hydrolase 2	Integrator complex subunit 2	Protein phosphatase 1 regulatory subunit 3A
Adseverin	Intercellular adhesion molecule 1	Protein RER1
Alpha-actinin-3	Interleukin-1 receptor accessory protein	Protein spire homolog 2
Alpha-tectorin	Iron/zinc purple acid phosphatase-like protein	Protein-lysine 6-oxidase
Aminopeptidase N	Keratin, type I cytoskeletal 17	Protocadherin alpha-3
Angiotensin-converting enzyme	Keratin, type I cytoskeletal 20	Protocadherin alpha-7
Ashwin	Keratin-like protein KRT222	Putative IQ motif and ankyrin repeat domain-containing protein LOC642574
ATP-binding cassette sub-family A member 2	Leucine-rich repeat and coiled-coil domain-containing protein 1	Putative tenascin-XA
Beta-2-microglobulin	Leucine-rich repeat-containing protein 28	Putative zinc-alpha-2-glycoprotein-like 1
BUD13 homolog	Leucine-rich repeat-containing protein 75A	RAC-alpha serine/threonine-protein kinase
Cadherin-1	LIM domain kinase 1	Receptor-type tyrosine-protein phosphatase eta
Cadherin-13	Lipopolysaccharide-binding protein	Reelin
CAP-Gly domain-containing linker protein 4	L-lactate dehydrogenase A chain	Retinoic acid receptor responder protein 2
Carbohydrate sulfotransferase 10	L-selectin	SAFB-like transcription modulator
Cartilage acidic protein 1	Lysine-specific demethylase 5C	Secreted Ly-6/uPAR-related protein 1
Cartilage oligomeric matrix protein	Mannan-binding lectin serine protease 1	Serine/threonine-protein kinase D1
Cell surface glycoprotein MUC18	Mannose-binding protein C	Serine/threonine-protein kinase RIO3
Centromere protein R	Mannosyl-oligosaccharide 1,2-alpha-mannosidase IA	Serine/threonine-protein phosphatase 6 regulatory subunit 3

Coagulation factor VII	Monocyte differentiation antigen CD14	Solute carrier family 25 member 45
Coagulation factor XII	MORF4 family-associated protein 1-like 1	Striatin-interacting protein 2
Coiled-coil domain-containing protein 126	Multiple epidermal growth factor-like domains protein 8	SWI/SNF-related matrix-associated actin-dependent regulator of chromatin subfamily A member 5
Collagen alpha-1(XVIII) chain	Multiple inositol polyphosphate phosphatase 1	Target of Nesh-SH3
Complement C1q subcomponent subunit A	N-acetyllactosaminide beta-1,3-N-acetylglucosaminyltransferase	Tenascin-X
CREB-regulated transcription coactivator 2	NACHT, LRR and PYD domains-containing protein 5	Testis-expressed sequence 10 protein
Cytosolic carboxypeptidase 1	Nebulin	Thrombospondin-4
Desmoglein-2	Neural cell adhesion molecule L1-like protein	Tigger transposable element-derived protein 7
Dipeptidase 2	Neuronal pentraxin receptor	Tissue alpha-L-fucosidase
DNA dC->dU-editing enzyme APOBEC-3D	Neuropeptide S receptor	Titin
Dopamine beta-hydroxylase	Nuclear pore complex protein Nup107	Transcriptional regulator ERG
E3 ubiquitin-protein ligase RNF123	Olfactory receptor 8D4	Transforming growth factor-beta-induced protein ig-h3
ELMO domain-containing protein 2	Olfactory receptor 8J1	Transmembrane 9 superfamily member 1
Exportin-7	Oncoprotein-induced transcript 3 protein	Tryptophan 5-hydroxylase 2
Flotillin-1	Peptidase inhibitor 16	Uncharacterized protein KIAA1841
Formin-2	Peripherin-2	UPF0586 protein C9orf41
Fukutin-related protein	Plastin-2	Vascular cell adhesion protein 1
Gamma-glutamyl hydrolase	Platelet glycoprotein Ib alpha chain	Voltage-dependent calcium channel subunit alpha-2/delta-1
Glutathione S-transferase omega-1	Platelet glycoprotein V	Voltage-dependent P/Q-type calcium channel subunit alpha-1A
Golgin subfamily A member 3	Plexin domain-containing protein 2	Volume-regulated anion channel subunit LRRC8A
Homeobox protein cut-like 1	Plexin-A2	WASH complex subunit 7
Hsc70-interacting protein	Polyprenol reductase	Xaa-Pro aminopeptidase 2
ICOS ligand	Pre-mRNA-processing factor 6	Xaa-Pro dipeptidase
Ig lambda-7 chain C region	Probable RNA-binding protein 23	Zinc finger protein 132
IgGFc-binding protein	Procollagen C-endopeptidase enhancer 1	Zinc finger protein 155
Inositol 1,4,5-trisphosphate receptor type 3	Proprotein convertase subtilisin/kexin type 9	Zinc finger protein 333

Table A- 2*. M3- 112 unique proteins (Crude plasma-undepleted sample)*

Abelson tyrosine-protein kinase 2	Ig heavy chain V-III region JON	Probable E3 ubiquitin-protein ligase HERC1
Acetyl-CoA carboxylase 1	Ig heavy chain V-III region LAY	Probable guanine nucleotide exchange factor MCF2L2
Adenomatous polyposis coli protein	Ig kappa chain V-I region CAR	Profilin-1
ADP/ATP translocase 2	Ig kappa chain V-I region DEE	Protein FAM83B
Alpha-fetoprotein	Ig kappa chain V-I region EU	Protein NipSnap homolog 2
Alpha-N-acetylneuraminide alpha-2,8-sialyltransferase	Ig kappa chain V-I region Ka	Protein phosphatase 1 regulatory inhibitor subunit 16B
Ankyrin repeat and BTB/POZ domain-containing protein 2	Ig kappa chain V-I region WEA	Protein transport protein Sec23B
Antileukoproteinase	Ig kappa chain V-III region CLL	Putative methyltransferase-like protein 21E pseudogene
Apolipoprotein L2	Ig kappa chain V-III region IARC/BL41	Putative uncharacterized protein C10orf113
Arf-GAP with Rho-GAP domain, ANK repeat and PH domain-containing protein 2	Ig kappa chain V-III region POM	RCC1 and BTB domain-containing protein 2
Cardiomyopathy-associated protein 5	Ig kappa chain V-IV region Len	Retinal-specific ATP-binding cassette transporter
Centromere protein N	Ig kappa chain V-IV region STH	Roundabout homolog 4
Centrosomal protein of 83 kDa	Ig lambda chain V region 4A	Serine/threonine-protein phosphatase 4 regulatory subunit 3B
Coatomer subunit alpha	Ig lambda chain V-I region EPS	Serum amyloid A-4 protein
Coiled-coil domain-containing protein 144B	Ig lambda chain V-I region NEW	Slit homolog 2 protein
Coiled-coil domain-containing protein 158	Ig lambda chain V-IV region Bau	Sodium channel and clathrin linker 1
Coiled-coil domain-containing protein 50	Immunoglobulin lambda-like polypeptide 1	Solute carrier family 12 member 1
Collagen alpha-1(XI) chain	Immunoglobulin lambda-like polypeptide 5	Sorting nexin-1
Cubilin	Integrin beta-2	Sorting nexin-30
Cystatin-C	Islet cell autoantigen 1	Sperm-specific antigen 2
Cytochrome P450 4F12	Lck-interacting transmembrane adapter 1	Structural maintenance of chromosomes protein 1A
DNA polymerase kappa	Leucine-rich repeat-containing protein 23	Sucrase-isomaltase, intestinal
Double homeobox protein A	Lipid phosphate phosphatase-related protein type 3	Synaptotagmin-like protein 2
Dynein heavy chain domain-containing protein 1	Macoilin	T-cell activation Rho GTPase-activating protein
E3 ISG15--protein ligase HERC5	Metalloproteinase inhibitor 1	Testis-specific gene 10 protein
E3 ubiquitin-protein ligase CBL-B	Myoneurin	Thyroid hormone receptor alpha
E3 ubiquitin-protein ligase RAD18	NAD-dependent protein deacetylase sirtuin-7	Transformation/transcription domain-associated protein
Extracellular matrix protein FRAS1	Nesprin-1	Translation initiation factor IF-3, mitochondrial
Ferric-chelate reductase 1	NKG2-C type II integral membrane protein	Tripartite motif-containing protein 14
Fibrocystin	Non-secretory ribonuclease	tRNA pseudouridine synthase

		A, mitochondrial
Glutamyl aminopeptidase	Nucleolar pre-ribosomal-associated protein 1	Tubulin-specific chaperone C
Growth factor receptor-bound protein 10	Olfactory receptor 1F1	Tyrosine-protein phosphatase non-receptor type 4
Guanine nucleotide-binding protein G(t) subunit alpha-2	Olfactory receptor 5D14	Ubiquitin-protein ligase E3C
Ig heavy chain V-I region HG3	Olfactory receptor 5K4	Zinc finger and BTB domain-containing protein 41
Ig heavy chain V-I region V35	Phosphoinositide 3-kinase regulatory subunit 6	Zinc finger MYM-type protein 6
Ig heavy chain V-II region COR	Phospholipid transfer protein	Zinc finger protein 778
Ig heavy chain V-II region OU	Platelet endothelial aggregation receptor 1	
Ig heavy chain V-III region GAR	Pleckstrin homology domain-containing family G member 2	

Table A- 3*. M3- 77 unique proteins (20% AS precipitate sample)*

39S ribosomal protein L10, mitochondrial	Host cell factor 2	Protein phosphatase 1 regulatory subunit 3F
40S ribosomal protein S18	Ig heavy chain V-II region ARH-77	Protein TMEPAI
Alanyl-tRNA editing protein Aarsd1	Ig heavy chain V-III region HIL	Protein-S-isoprenylcysteine O-methyltransferase
Anthrax toxin receptor-like	Ig kappa chain V-I region Roy	Protocadherin beta-15
Apolipoprotein(a)	Ig kappa chain V-III region NG9	Putative uncharacterized protein encoded by LINC00114
Autophagy-related protein 2 homolog B	Inositol hexakisphosphate kinase 3	Ras and Rab interactor 3
Beta-glucuronidase	Interleukin-22	Receptor-type tyrosine-protein phosphatase N2
Bromodomain adjacent to zinc finger domain protein 1A	Keratin, type I cytoskeletal 16	Regulator of G-protein signaling 8
BTB/POZ domain-containing protein KCTD4	LIM/homeobox protein Lhx8	Rho guanine nucleotide exchange factor 18
Catalase	MAX gene-associated protein	RNA exonuclease 1 homolog
Collectin-12	MBT domain-containing protein 1	Segment polarity protein dishevelled homolog DVL-2
Cyclin-dependent kinase 12	Menin	Serine/threonine-protein kinase MRCK alpha
Cytochrome P450 1B1	Msx2-interacting protein	Serine/threonine-protein kinase mTOR
Deoxycytidine kinase	Nesprin-2	Sialoadhesin
Dynein heavy chain 1, axonemal	Ninein-like protein	Signal transducer and activator of transcription 5B
E3 ubiquitin-protein ligase Itchy homolog	O-acetyl-ADP-ribose deacetylase MACROD2	Sorting nexin-2
Endoplasmic reticulum aminopeptidase 1	Olfactory receptor 5AU1	Spermatogenesis-associated protein 31A4
Ephrin type-B receptor 3	Olfactory receptor 5D13	Spermidine/spermine N(1)-acetyltransferase-like protein 1
Epiphycan	Pappalysin-2	Supervillin
Fas apoptotic inhibitory molecule 3	Phosducin-like protein 2	Testis-expressed sequence 15 protein
F-box/WD repeat-containing protein 1A	Protein DBF4 homolog A	Transcription factor MafA
Ficolin-2	Protein FAM110D	Transmembrane protein 198
Gap junction alpha-10 protein	Protein FAM208A	Uncharacterized protein C16orf71
Glutamate receptor-interacting protein 1	Protein FAM214A	Zinc finger protein 62 homolog
HLA class II histocompatibility antigen, DRB1-15 beta chain	Protein kinase C gamma type	
HLA class II histocompatibility antigen, DRB1-4 beta chain	Protein MAATS1	

Table A- 4*. M3- 128 common proteins between the three methods*

Afamin	Complement component C8 beta chain	Ig mu chain C region
Alcohol dehydrogenase 1	Complement component C8 gamma chain	Ig mu heavy chain disease protein
Alpha-1-acid glycoprotein 1	Complement component C9	Immunoglobulin J chain
Alpha-1-acid glycoprotein 2	Complement factor B	Insulin-like growth factor-binding protein complex acid labile subunit
Alpha-1-antichymotrypsin	Complement factor H	Inter-alpha-trypsin inhibitor heavy chain H1
Alpha-1-antitrypsin	Complement factor H-related protein 1	Inter-alpha-trypsin inhibitor heavy chain H2
Alpha-1B-glycoprotein	Complement factor I	Inter-alpha-trypsin inhibitor heavy chain H3
Alpha-2-antiplasmin	Corticosteroid-binding globulin	Inter-alpha-trypsin inhibitor heavy chain H4
Alpha-2-HS-glycoprotein	EGF-containing fibulin-like extracellular matrix protein 1	Keratin, type I cytoskeletal 10
Alpha-2-macroglobulin	Extracellular matrix protein 1	Keratin, type I cytoskeletal 9
Angiotensinogen	Fibrinogen alpha chain	Keratin, type II cytoskeletal 1
Antithrombin-III	Fibrinogen beta chain	Keratin, type II cytoskeletal 2 epidermal
Apolipoprotein A-I	Fibrinogen gamma chain	Kininogen-1
Apolipoprotein A-II	Fibronectin	Leucine-rich alpha-2-glycoprotein
Apolipoprotein A-IV	Fibulin-1	Lumican
Apolipoprotein C-II	Ficolin-3	N-acetylmuramoyl-L-alanine amidase
Apolipoprotein C-III	Galectin-3-binding protein	Phosphatidylcholine-sterol acyltransferase
Apolipoprotein E	Gelsolin	Phosphatidylinositol-glycan-specific phospholipase D
Apolipoprotein L1	Glutathione peroxidase 3	Pigment epithelium-derived factor
Apolipoprotein M	Haptoglobin	Plasma kallikrein
Attractin	Haptoglobin-related protein	Plasma protease C1 inhibitor
Beta-2-glycoprotein 1	Hemoglobin subunit alpha	Plasminogen
Biotinidase	Hemoglobin subunit beta	Protein AMBP
C4b-binding protein alpha chain	Hemoglobin subunit delta	Prothrombin
C4b-binding protein beta chain	Hemopexin	Retinol-binding protein 4
Carboxypeptidase B2	Heparin cofactor 2	Selenoprotein P
Carboxypeptidase N catalytic chain	Hepatocyte growth factor activator	Serotransferrin
Carboxypeptidase N subunit 2	Hepatocyte growth factor-like protein	Serum albumin
Ceruloplasmin	Histidine-rich glycoprotein	Serum amyloid P-component
Clusterin	Ig alpha-1 chain C region	Serum paraoxonase/arylesterase 1
Coagulation factor V	Ig alpha-2 chain C region	Sialic acid-binding Ig-like lectin 16
Coagulation factor X	Ig delta chain C region	Tetranectin
Complement C1q subcomponent subunit B	Ig gamma-1 chain C region	Tetratricopeptide repeat protein 37
Complement C1q subcomponent subunit C	Ig gamma-2 chain C region	Thyroxine-binding globulin
Complement C1r	Ig gamma-3 chain C region	Transthyretin

subcomponent		
Complement C1r subcomponent-like protein	Ig gamma-4 chain C region	Vitamin D-binding protein
Complement C1s subcomponent	Ig heavy chain V-III region TEI	Vitamin K-dependent protein S
Complement C2	Ig heavy chain V-III region TIL	Vitronectin
Complement C3	Ig kappa chain C region	von Willebrand factor
Complement C4-A	Ig kappa chain V-III region SIE	Zinc-alpha-2-glycoprotein
Complement C5	Ig lambda chain V-I region WAH	
Complement component C6	Ig lambda chain V-III region LOI	
Complement component C7	Ig lambda chain V-IV region Hil	
Complement component C8 alpha chain	Ig lambda-1 chain C regions	

Table A- 5. *M3- 46 common proteins between depleted and crude plasma*

Platelet basic protein	Vitamin K-dependent protein Z	Carbonic anhydrase 1
Coagulation factor IX	Basement membrane-specific heparan sulfate proteoglycan core protein	Lysozyme C
Ribonuclease 4	Insulin-like growth factor-binding protein 3	Vitamin K-dependent protein C
Dedicator of cytokinesis protein 11	Pantetheinase	Ig kappa chain V-I region AG
Endothelial protein C receptor	Properdin	Beta-1,4-galactosyltransferase 5
Poliovirus receptor	Hyaluronan-binding protein 2	Kallistatin
Lysosome-associated membrane glycoprotein 1	Coagulation factor XI	Intercellular adhesion molecule 2
Membrane primary amine oxidase	Sex hormone-binding globulin	Sulfhydryl oxidase 1
CD44 antigen	Protein Z-dependent protease inhibitor	Low affinity immunoglobulin gamma Fc region receptor III-A
Fetuin-B	Prolow-density lipoprotein receptor-related protein 1	Insulin-like growth factor I
Plasma serine protease inhibitor	Macrophage colony-stimulating factor 1 receptor	Actin, cytoplasmic 1
Vasorin	A disintegrin and metalloproteinase with thrombospondin motifs 13	Alpha-actinin-1
Insulin-like growth factor II	Alpha-N-acetylglucosaminidase	Lysosome-associated membrane glycoprotein 2
Apolipoprotein B-100	Complement factor D	Prostaglandin-H2 D-isomerase
ADAMTS-like protein 4	Mannan-binding lectin serine protease 2	
Proteoglycan 4	Neutrophil defensin 1	

Table A- 6. *LRA Matrix proteins (224)*

abhydrolase domain-containing proten 12B	Extracellular matrix protein 1	Inter-alpha-trypsin inhibitor heavy chain H1
Acetylcholine receptor subunit alpha	FAD-dependent oxidoreductase domain-containing protein 1	Inter-alpha-trypsin inhibitor heavy chain H2
Actin, cytoplasmic 1	Fetuin-B	Inter-alpha-trypsin inhibitor heavy chain H3
Adenylate kinase 9	Fibrinogen alpha chain	Inter-alpha-trypsin inhibitor heavy chain H4
Afamin	Fibrinogen beta chain	Kallistatin
Alcohol dehydrogenase 1	Fibrinogen gamma chain	Keratin, type II cytoskeletal 1
Alpha-1-acid glycoprotein 1	fibronectin	Keratin, type II cytoskeletal 6A
Alpha-1-acid glycoprotein 2	Fibulin-1	kininogen-1
Alpha-1-antichymotrypsin	Ficolin-3	Leucine-rich alpha-2-glycoprotein
alpha-1-antitrypsin	gamma-tubulin complex component 4	Leucine-rich repeat and calponin homology domain-containing protein 2
Alpha-1B-glycoprotein	Gelsolin	lipopolysaccharide-binding protein
Alpha-2-antiplasmin	Glutathione peroxidase 3	Lumican
Alpha-2-HS-glycoprotein	Haptoglobin	lysozyme c
alpha-2-macroglobulin	Haptoglobin-related protein	Malate dehydrogenase, cytoplasmic
Alpha-actinin-2	Hemoglobin subunit alpha	Mannan-binding lectin serine protease 1
Angiotensinogen	Hemoglobin subunit beta	Mannose-binding protein C
Antithrombin-III	Hemopexin	MBT domain-containing protein 1
Apolipoprotein A-I	Heparin cofactor 2	metabotropic glutamate receptor 6
Apolipoprotein A-II	Hepatocyte growtH factor activator	Monocyte differentiation antigen CD14
Apolipoprotein A-IV	Hepatocyte growth factor-like protein	Myosin-9
apolipoprotein B-100	Hephaestin	N-acetylmuramoyl-L-alanine amidase
Apolipoprotein C-I	Histidine-rich glycoprotein	Nebulin
Apolipoprotein C-II	Hyaluronan-binding protein 2	Neural cell adhesion molecule L1-like protein
Apolipoprotein C-III	Ig alpha-1 chain C region	neuronal pentraxin receptor
Apolipoprotein C-IV	Ig alpha-2 chain C region	phosphatidylcholine-sterol acyltransferase
apolipoprotein D	Ig delta chain C region	Phosphatidylinositol-glycan-specific phospholipase D

Apolipoprotein E	Ig gamma-1 chain C region	phospholipid transfer protein
Apolipoprotein L1	Ig gamma-2 chain C region	Pigment epithelium-derived factor
apolipoprotein M	Ig gamma-3 chain C region	Plasma kallikrein
apolipoprotein(a)	Ig gamma-4 chain C region	Plasma protease C1 inhibitor
ATP-binding cassette sub-family B member 9	Ig heavy chain V-I region HG3	Plasma serine protease inhibitor
Attractin	Ig heavy chain V-I region V35	Plasminogen
Beta-2-glycoprotein 1	Ig heavy chain V-II region ARH-77	plectin
Beta-2-microglobulin	Ig heavy chain V-II region OU	profilin-1
Biotinidase	Ig heavy chain V-II region WAH	prolyl 4-hydroxylase subunit alpha-1
C4b-binding protein alpha chain	Ig heavy chain V-III region BUT	prostaglandin F2 receptor negative regulator
C4b-binding protein beta chain	Ig heavy chain V-III region GAL	Protein AMBP
Carboxypeptidase B2	Ig heavy chain V-III region GAR	Protein bicaudal D homolog 2
Carboxypeptidase N catalytic chain	Ig heavy chain V-III region JON	Protein diaphanous homolog 2
Carboxypeptidase n subunit 2	Ig heavy chain V-III region TEI	Protein Z-dependent protease inhibitor
Cartilage oligomeric matrix protein	Ig heavy chain V-III region TIL	Prothrombin
casein kinase II subunit alpha'	Ig heavy chain V-III region TUR	Proto-oncogene tyrosine-protein kinase receptor Ret
CD5 antigen-like	Ig heavy chain V-III region VH26	Putative glycerol kinase 3
Ceruloplasmin	Ig heavy chain V-III region WEA	Putative tenascin-XA
Charged multivesicular body protein 7	Ig kappa chain C region	Ras and Rab interactor 3
cholinesterase	Ig kappa chain V-I region AG	Retinol-binding protein 4
cilia- and flagella-associated protein 221	Ig kappa chain V-I region BAN	RUN domain-containing protein 3A
citrate synthase, mitochondrial	Ig kappa chain V-I region CAR	Selenoprotein P
Clusterin	Ig kappa chain V-I region EU	Serotransferrin
Coagulation factor IX	Ig kappa chain V-I region Ka	Serpin I2
Coagulation factor V	Ig kappa chain V-I region Mev	Serum albumin
Coagulation factor X	Ig kappa chain V-I region Ni	Serum amyloid A-4 protein
Coagulation factor XI	Ig kappa chain V-I region Roy	Serum paraoxonase/arylesterase 1
Coagulation factor XII	Ig kappa chain V-I region WEA	Sialic acid-binding Ig-like lectin 16

Coagulation factor XIII A chain	Ig kappa chain V-I region Wes	Insulin-like growth factor-binding protein 3
Coagulation factor XIII B chain	Ig kappa chain V-II region MIL	Insulin-like growth factor-binding protein complex acid labile subunit
Collagen alpha-2(IX) chain	Ig kappa chain V-II region RPMI 6410	Smith-Magenis syndrome chromosomal region candidate gene 8 protein
Complement C1q subcomponent subunit A	Ig kappa chain V-II region TEW	Solute carrier family 25 member 40
Complement C1q subcomponent subunit B	Ig kappa chain V-III region CLL	Spermatid-specific manchette-related protein 1
Complement C1q subcomponent subunit C	Ig kappa chain V-III region NG9	Talin-1
Complement C1r subcomponent	Ig kappa chain V-III region SIE	TBC1 domain family member 10A
Complement C1s subcomponent	Ig kappa chain V-III region VG	telomere length regulation protein TEL2 homolog
complement C2	Ig kappa chain V-III region VH	Telomere-associated protein RIF1
Complement C3	Ig kappa chain V-III region WOL	tenascin-X
Complement C4-A	Ig kappa chain V-IV region Len	Tetranectin
Complement C4-B	Ig lambda chain V-I region NEW	thyroxine-binding globulin
Complement C5	Ig lambda chain V-I region WAH	Transforming growth factor-beta-induced protein ig-h3
Complement component c6	Ig lambda chain V-III region LOI	Transthyretin
Complement component C7	Ig lambda chain V-III region SH	Tudor domain-containing protein 15
Complement component C8 alpha chain	Ig lambda chain V-IV region Hil	tuftelin
Complement component C8 beta chain	Ig lambda chain V-VI region SUT	Tumor necrosis factor alpha-induced protein 8-like protein 3
Complement component C8 gamma chain	Ig lambda-1 chain C regions	vasorin
complement component C9	Ig lambda-2 chain C regions	vitamin D-binding protein
Complement factor B	Ig mu chain C region	Vitamin K-dependent protein C
complement factor H	Ig MU heavy chain disease protein	Vitamin K-dependent protein S
Complement factor H-related protein 1	Immunoglobulin J chain	Vitronectin
Complement factor H-related protein 2	Immunoglobulin lambda-like polypeptide 1	WD repeat-containing protein WRAP73
Complement factor I	Inactive tyrosine-protein kinase transmembrane receptor ROR1	Zinc-alpha-2-glycoprotein
corticosteroid-binding globulin	Insulin-like growth factor II	

Table A- 7. *An illustration of proteins eluting from a single fraction.*

	Fraction 1			Fraction 2			Fraction 3			Fraction 4			Fraction 5		
Protein Names	**Rep. 1**	**Rep. 2**	**Rep. 3**	**Rep. 1**	**Rep. 2**	**Rep. 3**	**Rep. 1**	**Rep. 2**	**Rep. 3**	**Rep. 1**	**Rep. 2**	**Rep. 3**	**Rep. 1**	**Rep. 2**	**Rep. 3**
Alpha-2-HS-glycoprotein	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
Inter-alpha-trypsin inhibitor heavy chain H1	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
Retinol-binding protein 4	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
Alpha-1B-glycoprotein	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
Plasminogen	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
Hemopexin	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
Alpha-1-acid glycoprotein 2	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
Serum albumin	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
Fibrinogen beta chain	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
Apolipoprotein A-I	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
Gelsolin	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
Ig kappa chain C region	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
Vasorin							X	X	X						
Cystatin-C							X	X	X						
Plasma serine protease inhibitor							X	X	X						
Ig kappa chain V-I region WEA							X	X	X						
Membrane primary amine oxidase							X	X	X						
Ribonuclease 4							X	X	X						
Ig kappa chain V-III region IARC/BL41							X	X	X						
Sex hormone-binding globulin							X	X	X						
Kallistatin							X	X	X						
Pantetheinase							X	X	X						
Macrophage colony-stimulating factor 1 receptor							X	X	X						

7.2 Appendix B

List of proteins identified in the pilot study with M3

Table B-1. *Total list of proteins (820) obtained with M3 when the proteins not quantified and top 20 HAPs were filtered. A list of 173 most differentiate proteins between controls, HFPEF and HFREF ($p < 0.005$) are highlighted in green.*

Protein
Retinol.binding.protein.4
Complement.component.C7
Hemopexin
Complement.C1s.subcomponent
Alpha.1B.glycoprotein
Alpha.1.antichymotrypsin
Complement.C4.B
Ig.gamma.3.chain.C.region
Plasma.protease.C1.inhibitor
complement.C2
Protein.AMBP
Complement.C1r.subcomponent
Zinc.alpha.2.glycoprotein
Ceruloplasmin
fibronectin
Inter.alpha.trypsin.inhibitor.heavy.chain.H4
Ankyrin.repeat.domain.containing.protein.36A
Hepatocyte.growtH.factor.activator
Alpha.2.HS.glycoprotein
Extracellular.matrix.protein.1
complement.factor.H
Antithrombin.III
N.acetylmuramoyl.L.alanine.amidase
Vitronectin
Inter.alpha.trypsin.inhibitor.heavy.chain.H2
POTE.ankyrin.domain.family.member.E
corticosteroid.binding.globulin
Kallistatin
Pigment.epithelium.derived.factor
Isoform.C.of.Fibulin.1
Afamin
Carboxypeptidase.n.subunit.2
Leucine.rich.alpha.2.glycoprotein
Complement.factor.H.related.protein.5
Inter.alpha.trypsin.inhibitor.heavy.chain.H1
Complement.factor.H.related.protein.1
L.selectin
Alpha.2.antiplasmin
Coagulation.factor.V
C4b.binding.protein.alpha.chain
Coagulation.factor.XIII.B.chain
Complement.C1r.subcomponent.like.protein
Glutathione.peroxidase.3
Carboxypeptidase.N.catalytic.chain
Beta.2.microglobulin
Hemoglobin.subunit.delta
Biotinidase
cholinesterase
Complement.factor.B
Insulin.like.growth.factor.II
Heparin.cofactor.2
Hemoglobin.subunit.gamma.1
Beta.2.glycoprotein.1
complement.component.C9
Complement.component.C8.gamma.chain
Complement.C5
Complement.factor.I
Isoform.Short.of.Complement.factor.H.related.protein.2
Hemoglobin.subunit.alpha
Lysosome.associated.membrane.glycoprotein.1
Phosphatidylinositol.glycan.specific.phospholipase.D
Hemoglobin.subunit.beta
Hepatocyte.growth.factor.like.protein
Keratin..type.II.cytoskeletal.1
Prostaglandin.H2.D.isomerase
vitamin.D.binding.protein
Attractin
Coagulation.factor.XIII.A.chain
Complement.C1q.subcomponent.subunit.C
Plasminogen
Monocyte.differentiation.antigen.CD14
Haptoglobin.related.protein
Tetranectin
vasorin
Lymphatic.vessel.endothelial.hyaluronic.acid.receptor.1
Actin..alpha.skeletal.muscle
Beta.Ala.His.dipeptidase
Coagulation.factor.X
Angiotensinogen
Apolipoprotein.A.IV
Isoform.2.of.Inter.alpha.trypsin.inhibitor.heavy.chain.H4
Ficolin.3
Hemoglobin.subunit.epsilon
Isoform.4.of.Dynein.heavy.chain.14..axonemal
Isoform.2.of.Alpha.1B.glycoprotein
Transgelin.2
Complement.factor.D
Platelet.glycoprotein.Ib.alpha.chain
Histidine.rich.glycoprotein
Carboxypeptidase.B2
Complement.factor.H.related.protein.2
Complement.component.C8.beta.chain
Ig.alpha.1.chain.C.region
Metalloproteinase.inhibitor.1
Ig.delta.chain.C.region
Fetuin.B

POTE.ankyrin.domain.family.member.F
Ig.MU.heavy.chain.disease.protein
BET1.like.protein
Fibulin.1
Selenoprotein.P
Plastin.2
Protein.Z.dependent.protease.inhibitor
Complement.C4.A
Peptidase.inhibitor.16
Dermcidin
Actin..cytoplasmic.1
Prothrombin
Plasma.kallikrein
Complement.component.C8.alpha.chain
Thrombospondin.4
Mannose.binding.protein.C
Gelsolin
Trypsin.1
Keratin..type.I.cytoskeletal.10
Insulin.like.growth.factor.binding.protein.complex.acid.labile.subunit
Vitamin.K.dependent.protein.S
CD44.antigen
Endothelial.protein.C.receptor
lipopolysaccharide.binding.protein
Mannan.binding.lectin.serine.protease.1
lysozyme.c
Dedicator.of.cytokinesis.protein.10
Complement.C1q.subcomponent.subunit.A
Fibrinogen.gamma.chain
Immunoglobulin.lambda.like.polypeptide.5
Complement.component.c6
apolipoprotein.M
Coagulation.factor.XII
Microtubule.associated.tumor.suppressor.candidate.2
Lumican
Ig.alpha.2.chain.C.region
Ig.kappa.chain.V.III.region.SIE
tenascin.X.1
kininogen.1
Coagulation.factor.IX

keratin..type.II.cytoskeletal.6B
Complement.C1q.subcomponent.subunit.B
alpha.enolase
Vitamin.K.dependent.protein.C
Vitamin.K.dependent.protein.Z
Isoform.4.of.CDK5.regulatory.subunit.associated.protein.2
Serum.paraoxonase.arylesterase.1
Clusterin
Ig.heavy.chain.V.III.region.BRO
Cystatin.C
CDK5.regulatory.subunit.associated.protein.2
Isoform.2.of.Keratin..type.I.cuticular.Ha6
Insulin.like.growth.factor.binding.protein.3
Complement.factor.H.related.protein.3
keratin..type.II.cytoskeletal.5
Macrophage.colony.stimulating.factor.1.receptor
thyroxine.binding.globulin
Isoform.2.of.Complement.C2
Beta.actin.like.protein.2
POTE.ankyrin.domain.family.member.J
Isoform.4.of.Inter.alpha.trypsin.inhibitor.heavy.chain.H4
apolipoprotein.B.100
Keratin..type.I.cytoskeletal.9
Apolipoprotein.F
forkhead.box.protein.P2
Ig.lambda.chain.V.IV.region.Hil
Cofilin.1
Isoform.2.of.14.3.3.protein.sigma
Isoform.SV.of.14.3.3.protein.epsilon
Isoform.2.of.Haptoglobin.related.protein
Keratin..type.II.cytoskeletal.2.epidermal
L.lactate.dehydrogenase.A.chain
Ig.lambda.chain.V.I.region.HA
POTE.ankyrin.domain.family.member.I
Apolipoprotein.C.II
Cartilage.oligomeric.matrix.protein
Isoform.2.of.Pregnancy.zone.protein
Ig.kappa.chain.C.region
Ig.mu.chain.C.region

Plasminogen.like.protein.B
CD5.antigen.like
Isoform.2.of.Complement.factor.H.related.protein.3
Platelet.basic.protein
Keratin..type.II.cytoskeletal.1b
Apolipoprotein.C.I
SH3.domain.binding.protein.2
X14.3.3.protein.theta
Dopamine.beta.hydroxylase
Ig.gamma.4.chain.C.region
Serum.amyloid.P.component
Mannan.binding.lectin.serine.protease.2
insulin.like.growth.factor.binding.protein.4
Galectin.3.binding.protein
Isoform.2.of.Vitamin.D.binding.protein
Putative.beta.actin.like.protein.3
Lysosome.associated.membrane.glycoprotein.2
Actin..alpha.cardiac.muscle.1
tenascin.X
Isoform.7.of.Nesprin.1
Serotransferrin
Hemoglobin.subunit.zeta
profilin.1
C4b.binding.protein.beta.chain
Intercellular.adhesion.molecule.1
Multiple.inositol.polyphosphate.phosphatase.1
Intercellular.adhesion.molecule.2
xaa.Pro.dipeptidase
Apolipoprotein.E
Putative.macrophage.stimulating.1.like.protein
Putative.zinc.alpha.2.glycoprotein.like.1
Receptor.type.tyrosine.protein.phosphatase.eta
phosphatidylinositol.5.phosphate.4.kinase.type.2.alpha
Ig.lambda.chain.V.III.region.LOI
Sex.hormone.binding.globulin
Tyrosine.protein.kinase.transmembrane.receptor.ROR1
Interleukin.1.receptor.accessory.protein
Von.Willebrand.factor
Immunoglobulin.J.chain
CD109.antigen

Properdin
Pregnancy.zone.protein
Membrane.primary.amine.oxidase
Ig.gamma.2.chain.C.region
apolipoprotein.a.
BTB.POZ.domain.containing.protein.KCTD9
angiogenin
X14.3.3.protein.eta
Putative.GED.domain.containing.protein.DNM1P34
Actin..cytoplasmic.2
Isoform.2.of.Insulin.like.growth.factor.binding.protein.complex.acid.labile.subunit
Alpha.actinin.4
Isoform.2.of.Gelsolin
Talin.1
Alpha.actinin.1
Ig.lambda.7.chain.C.region
Ig.lambda.6.chain.C.region
Transient.receptor.potential.cation.channel.subfamily.M.member.8
Vascular.non.inflammatory.molecule.2
CD180.antigen
Sperm.protein.associated.with.the.nucleus.on.the.X.chromosome.N3
Interleukin.6.receptor.subunit.beta
Ig.kappa.chain.V.III.region.TI
X14.3.3.protein.beta.alpha
Keratin.associated.protein.25.1
Casein.kinase.II.subunit.beta
Hyaluronidase.1
Calmodulin.like.protein.3
Ig.kappa.chain.V.III.region.GOL
Ig.kappa.chain.V.IV.region.Len
Ig.kappa.chain.V.I.region.Lay
Ig.heavy.chain.V.II.region.ARH.77
Ig.kappa.chain.V.III.region.HAH
Lithostathine.1.alpha
Isoform.2.of.Neurofilament.heavy.polypeptide
Ig.kappa.chain.V.I.region.Hau
GAS2.like.protein.2
Ig.kappa.chain.V.I.region.Ni
Non.secretory.ribonuclease
Keratin..type.I.cytoskeletal.26
Isoform.5.of.T.lymphocyte.activation.antigen.CD86
Ig.lambda.chain.V.region.4A
Vascular.non.inflammatory.molecule.3
alpha.N.acetylglucosaminidase
THO.complex.subunit.7.homolog
Isoform.LMW.of.Kininogen.1
Isoform.4.of.Tenascin.X
Isoform.4.of.Sex.hormone.binding.globulin
Calmodulin
Isoform.2.of.Matrix.metalloproteinase.19
Actin..aortic.smooth.muscle
Plasminogen.like.protein.A
Keratin..type.I.cytoskeletal.27
Keratin..type.I.cytoskeletal.20
Tropomyosin.alpha.4.chain
Isoform.3.of.Complement.C2
Isoform.2.of.Keratin.like.protein.KRT222
serine.threonine.protein.kinase.3
Isoform.2.of.Keratin..type.II.cytoskeletal.73
Keratin..type.II.cytoskeletal.75
phospholipid.transfer.protein
Protein.N.lysine.methyltransferase.METTL20
Isoform.3.of.Insulin.like.growth.factor.II
Coagulation.factor.XI
Serum.amyloid.A.1.protein
Isoform.2.of.Clusterin
Transferrin.receptor.protein.1
adiponectin
Keratin..type.II.cytoskeletal.3
IgGFc.binding.protein
Ig.lambda.chain.V.III.region.SH
cadherin.13
Pantetheinase
Procollagen.C.endopeptidase.enhancer.1
Sulfhydryl.oxidase.1
SH3.domain.binding.glutamic.acid.rich.like.protein.3
Ig.lambda.2.chain.C.regions
Serum.paraoxonase.lactonase.3
Serum.amyloid.A.4.protein
coagulation.factor.VII
Matrix.metalloproteinase.19
thrombospondin.1
C.reactive.protein
Isoform.2.of.Mannan.binding.lectin.serine.protease.1
phosphatidylcholine.sterol.acyltransferase
Bone.marrow.proteoglycan
Cathepsin.D
Neutrophil.defensin.1
Platelet.factor.4
Ig.kappa.chain.V.III.region.WOL
Neuropilin.1
Protein.S100.A8
Plexin.domain.containing.protein.2
Vascular.cell.adhesion.protein.1
Keratin..type.II.cytoskeletal.6A
Ig.gamma.1.chain.C.region
Keratin..type.II.cytoskeletal.74
X14.3.3.protein.gamma
Keratin..type.II.cytoskeletal.79
Keratin..type.II.cytoskeletal.72
Isoform.3.of.CDK5.regulatory.subunit.associated.protein.2
keratin..type.I.cytoskeletal.28
Multimerin.1
Isoform.Calbrain.of.Calcium.binding.protein.1
insulin.like.growth.factor.binding.protein.6
Isoform.2.of.Attractin
Isoform.4.of.Extracellular.matrix.protein.1
Mitochondrial.ribosome.associated.GTPase.1
Protein.S100.A9
Symplekin
Ig.heavy.chain.V.III.region.23
Charged.multivesicular.body.protein.4a
Uteroglobin
Ig.heavy.chain.V.III.region.WEA
Ig.kappa.chain.V.II.region.Cum
Probable.E3.ubiquitin.protein.ligase.HERC4
Kinesin.like.protein.KIF21B
Keratin..type.I.cytoskeletal.25
Protein.HEG.homolog.1
alpha.lactalbumin
Isoform.4.of.Mannan.binding.lectin.serine.protease.1

Keratin..type.I.cytoskeletal.16
Putative.tenascin.XA
Keratin..type.I.cytoskeletal.17
Apolipoprotein.L3
Ubiquitin.carboxyl.terminal.hydrolase.4
Ig.kappa.chain.V.II.region.RPMI.6410
coactosin.like.protein
Adenosine.deaminase.CECR1
Tropomyosin.beta.chain
Cysteine.rich.secretory.protein.3
Isoform.2.of.Carboxypeptidase.B2
Apolipoprotein.C.IV
Leucine.rich.repeat.and.IQ.domain.containing.protein.3
Isoform.2.of.Nuclear.mitotic.apparatus.protein.1
L.lactate.dehydrogenase.C.chain
Dual.oxidase.1
Ig.heavy.chain.V.III.region.TIL
Microtubule.associated.serine.threonine.protein.kinase.4
Centrosomal.protein.of.57.kDa
Cadherin.related.family.member.2
AMP.deaminase.3
Ig.kappa.chain.V.I.region.AG
Ig.kappa.chain.V.III.region.VG
Coiled.coil.domain.containing.protein.85B
CD59.glycoprotein
carbonic.anhydrase.1
Nesprin.2
Isoform.E.of.Proteoglycan.4
Rho.guanine.nucleotide.exchange.factor.5
P2X.purinoceptor.7
DNA.excision.repair.protein.ERCC.6
OTU.domain.containing.protein.5
Kielin.chordin.like.protein
Angiotensin.converting.enzyme
catalase
Mediator.of.RNA.polymerase.II.transcription.subunit.23
Protein.O.mannose.kinase
Isoform.6.of.Myomegalin
Transcription.factor.TFIIIB.component.B.homolog
ribonucleoside.diphosphate.reductase.subunit.M2

Roquin.1
U4.U6.U5.tri.snRNP.associated.protein.1
Olfactory.receptor.52B4
Leucine.rich.repeat.LGI.family.member.2
Coiled.coil.domain.containing.protein.121
Ras.related.protein.Rab.43
Isoform.8.of.Testican.3
Isoform.3.of.Myocardial.zonula.adherens.protein
Delta.3.5..Delta.2.4..dienoyl.CoA.isomerase..mitochondrial
ATP.binding.cassette.sub.family.A.member.13
Lymphoid.specific.helicase
Isoform.17.of.CD44.antigen
Calcium.binding.protein.39.like
Far.upstream.element.binding.protein.1
Isoform.4.of.Centrosomal.protein.of.57.kDa
activator.of.90.kDa.heat.shock.protein.ATPase.homolog.1
Isoform.3.of.Homocysteine.responsive.endoplasmic.reticulum.resident.ubiquitin.like.domain.member.1.protein
Zinc.finger.protein.469
Butyrophilin.subfamily.2.member.A1
Fanconi.anemia.group.I.protein
Isoform.7.of.Z.DNA.binding.protein.1
Isoform.3.of.Pre.mRNA.processing.factor.40.homolog.B
Histone.lysine.N.methyltransferase.SETMAR
Protein.O.mannosyl.transferase.1
X39S.ribosomal.protein.L41..mitochondrial
Multiple.epidermal.growth.factor.like.domains.protein.8
Isoform.Gamma.of.Glycogenin.2
WD40.repeat.containing.protein.SMU1
EF.hand.calcium.binding.domain.containing.protein.11
Isoform.2.of.Extracellular.matrix.protein.1
HIRA.interacting.protein.3
Isoform.2.of.Nipped.B.like.protein
DNA.binding.protein.SATB1
Thrombospondin.type.1.domain.containing.protein.1
CUB.and.sushi.domain.containing.protein.2

Ankyrin.and.armadillo.repeat.containing.protein
transmembrane.protein.151B
X1.phosphatidylinositol.4.5.bisphosphate.phosphodiesterase.eta.2
Uncharacterized.protein.C11orf65
Trinucleotide.repeat.containing.gene.6A.protein
Ig.lambda.chain.V.II.region.BUR
Ubiquitin.carboxyl.terminal.hydrolase.2
Gamma.glutamyltransferase.light.chain.2
Isoform.4.of.Acyl.CoA.dehydrogenase.family.member.10
Isoform.2.of.Sodium.driven.chloride.bicarbonate.exchanger
Collagen.alpha.1.VI..chain
Isoform.3.of.Pyruvate.dehydrogenase.E1.component.subunit.beta..mitochondrial
Isoform.2.of.Multimerin.1
Fc.receptor.like.protein.5
Vacuolar.protein.sorting.associated.protein.37B
Patatin.like.phospholipase.domain.containing.protein.4
IQ.domain.containing.protein.G
Granulocyte.macrophage.colony.stimulating.factor
Ig.heavy.chain.V.I.region.HG3
Mucin.15
Isoform.3.of.Adseverin
X55.kDa.erythrocyte.membrane.protein
Protein.Tob2
Isoform.3.of.Interleukin.6.receptor.subunit.beta
Isoform.4.of.Target.of.Nesh.SH3
Isoform.2.of.Magnesium.transporter.NIPA1
Mitogen.activated.protein.kinase.kinase.kinase.19
Manganese.dependent.ADP.ribose.CDP.alcohol.diphosphatase
Nucleoside.diphosphate.linked.moiety.X.motif.19..mitochondrial
Stanniocalcin.1
Protein.FAM3D
Beta.galactosidase.1.like.protein.3
Complement.component.C1q.receptor
GPN.loop.GTPase.2
LIM.homeobox.protein.Lhx4
Vesicle.transport.protein.SEC20
Hydrolethalus.syndrome.protein.1

Lysine.specific.demethylase.3B
RUN.domain.containing.protein.3A
Isoform.2.of.Rhomboid.related.protein.4
sestrin.3
Carbonyl.reductase..NADPH..1
Cadherin.2
Dual.specificity.mitogen.activated.protein.kinase.kinase.4
DC.STAMP.domain.containing.protein.1
Isoform.2.of.Tropomyosin.alpha.4.chain
Isoform.5.of.Synaptotagmin.like.protein.2
Methylcrotonoyl.CoA.carboxylase.beta.chain..mitochondrial
X39S.ribosomal.protein.L15..mitochondrial
Isoform.2.of.Trypsin.3
Type.II.inositol.3.4.bisphosphate.4.phosphatase
Isoform.2.of.Rab11.family.interacting.protein.3
Glycine.receptor.subunit.alpha.3
Isoform.2.of.C4b.binding.protein.beta.chain
Isoform.2.of.Alpha.1.3.mannosyl.glycoprotein.4.beta.N.acetylglucosaminyltransferase.A
vezatin
Anion.exchange.protein.3
Calcium.homeostasis.modulator.protein.2
Superoxide.dismutase..Cu.Zn.
Stimulated.by.retinoic.acid.gene.8.protein.homolog
CD166.antigen
Isoform.4.of.Cytoplasmic.dynein.2.light.intermediate.chain.1
La.related.protein.1
X3..2...5..bisphosphate.nucleotidase.1
ZBED6.C.terminal.like.protein
protein.SEC13.homolog
Transient.receptor.potential.cation.channel.subfamily.M.member.1
Zinc.finger.matrin.type.protein.2
Isoform.2.of.Tau.tubulin.kinase.1
Putative.inactive.phosphatidylinositol.4.kinase.alpha.like.protein.P1
Dual.serine.threonine.and.tyrosine.protein.kinase
RNA.binding.protein.4B
Isoform.2.of.Clavesin.1
testis.expressed.sequence.12.protein
Cysteine.rich.and.transmembrane.domain.containing.protein.1
Isoform.5.of.Hereditary.hemochromatosis.protein
Isoform.4.of.Probable.E3.ubiquitin.protein.ligase.HERC4
Glycosyl.phosphatidylinositol.anchored.molecule.like.protein
Isoform.4.of.Protein.PAT1.homolog.1
Uncharacterized.protein.ENSP00000372125
Polyadenylate.binding.protein.interacting.protein.2B
solute.carrier.family.22.member.2
Ig.kappa.chain.V.I.region.BAN
POTE.ankyrin.domain.family.member.D
RNA.polymerase.II.associated.factor.1.homolog
pre.mRNA.processing.factor.40.homolog.B
X4F2.cell.surface.antigen.heavy.chain
Serpin.A9
Poliovirus.receptor
X5.azacytidine.induced.protein.2
Putative.protein.ATP11AUN
Isoform.2.of.Protocadherin.gamma.B2
Isoform.3.of.Tropomyosin.alpha.3.chain
Ankyrin.repeat.and.SOCS.box.protein.14
Cleavage.stimulation.factor.subunit.3
HLA.class.I.histocompatibility.antigen..B.41.alpha.chain
Transmembrane.protein.88
Isoform.2.of.Polycystic.kidney.disease.protein.1.like.2
Latexin
Testis.specific.serine.threonine.protein.kinase.6
Adrenodoxin.like.protein..mitochondrial
Rho.related.GTP.binding.protein.RhoQ
Protocadherin.gamma.B1
Isoform.1.of.Fragile.X.mental.retardation.protein.1
phosphopantothenoylcysteine.decarboxylase
Isoform.S.CaBP2.of.Calcium.binding.protein.2
Isoform.4.of.C.C.motif.chemokine.4.like
N.acetylaspartylglutamate.synthase.A
Secreted.and.transmembrane.protein.1
Lysophosphatidylcholine.acyltransferase.2
Transmembrane.protein.70..mitochondrial
Ribosyldihydronicotinamide.dehydrogenase..quinone.
Isoform.3.of.DNA.directed.RNA.polymerase.III.subunit.RPC9
Ig.heavy.chain.V.III.region.TUR
Metallo.beta.lactamase.domain.containing.protein.1
ICOS.ligand
T.cell.leukemia.lymphoma.protein.1A
Rab11.family.interacting.protein.1
Isoform.3.of.Dual.serine.threonine.and.tyrosine.protein.kinase
Calumenin
Isoform.2.of.Alpha.actinin.1
Tyrosine.protein.phosphatase.non.receptor.type.3
Protein.Hikeshi
Leucine.rich.repeat.containing.protein.1
Probable.methyltransferase.like.protein.15
peptidoglycan.recognition.protein.1
N.G..N.G..dimethylarginine.dimethylaminohydrolase.2
Cas.scaffolding.protein.family.member.4
Ig.lambda.chain.V.V.region.DEL
Isoform.2.of.Podoplanin
flavin.reductase..NADPH.
Ig.heavy.chain.V.II.region.NEWM
WNT1.inducible.signaling.pathway.protein.1
Glutamate..cysteine.ligase.regulatory.subunit
Leucine.rich.repeat.containing.protein.16A
Phosducin.like.protein.2
N.terminal.kinase.like.protein
Olfactory.receptor.4C12
pikachurin
Inhibitor.of.growth.protein.3
Ig.kappa.chain.V.I.region.Roy
HLA.class.I.histocompatibility.antigen..B.7.alpha.chain
Protein.phosphatase.inhibitor.2
Isoform.3.of.CMRF35.like.molecule.1
Isoform.2.of.Pulmonary.surfactant.associated.protein.A1

Protein.Fam9b
Isoform.3.of.Mth938.domain.containing.protein
Putative.uncharacterized.protein.encoded.by.LINC00527
Synaptogyrin.1
Ecto.ADP.ribosyltransferase.3
Isoform.3.of.Target.of.Nesh.SH3
Isoform.CAE3.of.Anion.exchange.protein.3
Isoform.2.of.Cell.cycle.checkpoint.protein.RAD1
Isoform.7.of.Leucine.zipper.putative.tumor.suppressor.1
Protein.slowmo.homolog.1
Magnesium.transporter.NIPA4
argininosuccinate.lyase
tubulin.alpha.3C.D.chain
Neuromedin.U.receptor.2
Isoform.3.of.IQ.domain.containing.protein.K
Putative.chemokine.related.protein.B42
Retinoic.acid.early.transcript.1L.protein
transcription.initiation.factor.TFIID.subunit.1.like
Multidrug.and.toxin.extrusion.protein.1
Alanine.and.arginine.rich.domain.containing.protein
zinc.finger.protein.24
DPH3.homolog
NLR.family.CARD.domain.containing.protein.4
E3.ubiquitin.protein.ligase.RNF180
Isoform.3.of.Vacuolar.fusion.protein.MON1.homolog.A
Protein.tyrosine.phosphatase.domain.containing.protein.1
Isoform.2.of.Signal.regulatory.protein.beta.2
Interleukin.32
Antileukoproteinase
TP53.regulated.inhibitor.of.apoptosis.1
Protein.spinster.homolog.1
Mesoderm.induction.early.response.protein.3
Isoform.Mdm2.D.of.E3.ubiquitin.protein.ligase.Mdm2
Asialoglycoprotein.receptor.2
Protein.HIDE1
ribulose.phosphate.3.epimerase
rab.proteins.geranylgeranyltransferase.component.A.1

Isoform.2.of.Homeobox.protein.Hox.A10
Mitogen.activated.protein.kinase.kinase.kinase.12
Leukocyte.cell.derived.chemotaxin.2
MAP.kinase.activated.protein.kinase.3
Fibronectin.type.III.domain.containing.protein.8
Protein.FAM71E1
Granzyme.H
Leukocyte.immunoglobulin.like.receptor.subfamily.B.member.3
Neuroserpin
Isoform.2.of.Keratin..type.II.cytoskeletal.78
Isoform.2.of.Leucine.rich.repeat.and.death.domain.containing.protein.1
Cytochrome.c.oxidase.assembly.protein.COX15.homolog
delphilin
DNA.binding.protein.RFXANK
O.acetyl.ADP.ribose.deacetylase.MACROD2
Isoform.2.of.Cilia..and.flagella.associated.protein.52
Proepiregulin
Transmembrane.4.L6.family.member.4
Ig.heavy.chain.V.III.region.BUT
Protein.PAT1.homolog.1
Sorting.nexin.24
Isoform.2.of.Mucin.1
Cytochrome.b.reductase.1
Glycophorin.B
Ig.kappa.chain.V.I.region.Bi
Isoform.2.of.EP300.interacting.inhibitor.of.differentiation.1
Protein.FAM166B
Isoform.3.of.Uncharacterized.protein.C12orf40
Isoform.2.of.Transmembrane.protein.235
Isoform.2.of.Serine.hydrolase.like.protein.2
Secretoglobin.family.3A.member.1
Kelch.like.protein.42
Transmembrane.and.coiled.coil.domain.containing.protein.2
Nucleoporin.p54
Putative.uncharacterized.protein.PRO0461
Putative.protein.GATS
Isoform.5.of.Protein.Mdm4

Isoform.5.of.LEM.domain.containing.protein.1
B.cell.CLL.lymphoma.7.protein.family.member.B
Alpha.amylase.2B
Isoform.4.of.Mucosal.addressin.cell.adhesion.molecule.1
Carbohydrate.sulfotransferase.6
Isoform.2.of.Src.kinase.associated.phosphoprotein.1
uncharacterized.protein.MIR1.1HG
Tumor.necrosis.factor.ligand.superfamily.member.6
Isoform.6.of.CMRF35.like.molecule.1
Aquaporin.7
ubiquitin.conjugating.enzyme.E2.H
Isoform.2.of.Sprouty.related..EVH1.domain.containing.protein.3
Transmembrane.protein.179B
Isoform.11.of.Peroxisomal.N.1..acetyl.spermine.spermidine.oxidase
Eukaryotic.translation.initiation.factor.4E.type.3
Lithostathine.1.beta
Fibroblast.growth.factor.20
ADP.ribosylation.factor.like.protein.4A
Dysbindin.domain.containing.protein.1
Trefoil.factor.1
Isoform.2.of.Tumor.necrosis.factor.receptor.superfamily.member.1B
GSK3.beta.interaction.protein
Acetylserotonin.O.methyltransferase
Isoform.2.of.Tumor.necrosis.factor.ligand.superfamily.member.6
Protein.tyrosine.phosphatase.receptor.type.C.associated.protein
Isoform.2.of.Calcium.binding.protein.4
putative.DPH3.homolog.B
Alkaline.ceramidase.2
Isoform.2.of.Syntaxin.1B
Isoform.3.of.WAP.four.disulfide.core.domain.protein.2
peptide.YY
Isoform.2.of.La.related.protein.1B
lysosomal.amino.acid.transporter.1.homolog
cholecystokinin
Isoform.3.of.Motile.sperm.domain.containing.protein.1
Salivary.acidic.proline.rich.phosphoprotein.1.2

X14.3.3.protein.epsilon
Isoform.8.of.Myomegalin
Isoform.5.of.Double.stranded.RNA.specific.editase.1
armadillo.repeat.containing.protein.1
CDC42.small.effector.protein.2
Isoform.2.of.WAP.four.disulfide.core.domain.protein.2
Uroplakin.2
Lipoma.HMGIC.fusion.partner
Isoform.2.of.Tropomyosin.alpha.3.chain
EP300.interacting.inhibitor.of.differentiation.1
Kallikrein.7
myosin.regulatory.light.chain.2..ventricular.cardiac.muscle.isoform
HIG1.domain.family.member.1C
Putative.ubiquitin.conjugating.enzyme.E2.D2.like.protein
Isoform.4.of.Leucine.rich.repeat.containing.protein.16A
Keratin..type.I.cytoskeletal.18
Proteoglycan.4
actin..gamma.enteric.smooth.muscle
Keratin..type.I.cytoskeletal.15
Tripartite.motif.containing.protein.15
Keratin..type.I.cytoskeletal.19
Keratin..type.II.cytoskeletal.71
Neural.cell.adhesion.molecule.L1.like.protein
Isoform.2.of.NXPE.family.member.1
Isoform.2.of.Protein.SOGA1
Tripartite.motif.containing.protein.42
Trypsin.3
Plasma.serine.protease.inhibitor
Insulin.like.growth.factor.binding.protein.2
Myomegalin
Keratin..type.II.cytoskeletal.7
Isoform.3.of.Myomegalin
Apolipoprotein.C.III
Keratin..type.II.cytoskeletal.2.oral
Isoform.2.of.CDK5.regulatory.subunit.associated.protein.2
Carbonic.anhydrase.2
Keratin..type.II.cytoskeletal.8
aminopeptidase.N
X14.3.3.protein.zeta.delta

D.amino.acid.oxidase
apolipoprotein.D
Apolipoprotein.L1
L.lactate.dehydrogenase.B.chain
Transforming.growth.factor.beta.induced.protein.ig.h3
Hyaluronan.binding.protein.2
Keratin..type.II.cytoskeletal.73
Isoform.3.of.Keratin..type.I.cytoskeletal.13
NACHT.domain..and.WD.repeat.containing.protein.1
Nephrocystin.4
Probable.E3.ubiquitin.protein.ligase.HECTD4
Putative.ankyrin.repeat.domain.containing.protein.31
Periaxin
Uncharacterized.protein.C10orf131
X2.hydroxyacylsphingosine.1.beta.galactosyltransferase
Protein.zwilch.homolog
Isoform.2.of.Phosphatidylinositol.5.phosphate.4.kinase.type.2.gamma
Mixed.lineage.kinase.domain.like.protein
Uncharacterized.protein.KIAA0825
Isoform.2.of.Fibronectin
Isoform.2.of.Neuroblast.differentiation.associated.protein.AHNAK
E3.ubiquitin.protein.ligase.TRIM58
Homeobox.protein.Hox.C10
PHD.finger.protein.13
cingulin
Isoform.3.of.Xaa.Pro.dipeptidase
ADAMTS.like.protein.4
Isoform.2.of.Actin..gamma.enteric.smooth.muscle
heat.shock.factor.protein.4
Ribonuclease.4
Sprouty.related..EVH1.domain.containing.protein.3
Homeobox.protein.Hox.A10
Isoform.2.of.Collagen.alpha.1.XVIII..chain
Fat.storage.inducing.transmembrane.protein.1
Chitotriosidase.1
Ig.heavy.chain.V.III.region.POM
X78.kDa.glucose.regulated.protein
Pancreatic.alpha.amylase

Ig.lambda.chain.V.I.region.WAH
Noelin
Kelch.like.protein.20
Serum.amyloid.A.2.protein
Zinc.finger.protein.862
Interleukin.7
epidermal.growth.factor.receptor
Isoform.4.of.L.lactate.dehydrogenase.A.chain
X15.hydroxyprostaglandin.dehydrogenase..NAD....
Keratin..type.II.cytoskeletal.78
Ig.kappa.chain.V.IV.region
Isoform.2.of.Peptidase.inhibitor.16
ADSEVERIN
proteoglycan.3
DBIRD.complex.subunit.ZNF326
Zinc.transporter.10
hyaluronidase.3
Ig.kappa.chain.V.III.region.CLL
Keratin..type.II.cytoskeletal.6C
Hemogen
Tropomyosin.alpha.1.chain
Integrator.complex.subunit.5
X14.3.3.protein.sigma
Protein.FAM9C
ER.membrane.protein.complex.subunit.7
Isoform.5.of.Protein.FRA10AC1
X2.5.oligoadenylate.synthase.2
PHD.finger.protein.6
Testis.expressed.sequence.2.protein
Synaptojanin.2.binding.protein
Cytochrome.c.oxidase.assembly.factor.6.homolog
Reticulocalbin.1
Cell.surface.glycoprotein.MUC18
Cadherin.5
Inter.alpha.trypsin.inhibitor.heavy.chain.H3
Coiled.coil.domain.containing.protein.126
Ectonucleotide.pyrophosphatase.phosphodiesterase.family.member.2
X1.phosphatidylinositol.3.phosphate.5.kinase
Ig.heavy.chain.V.III.region.GAL
Di.N.acetylchitobiase
Retinoic.acid.receptor.responder.protein.2

Isoform.3.of.Keratin..type.II.cytoskeletal.72
Isoform.2.of.Sex.hormone.binding.globulin
Keratin..type.I.cytoskeletal.13
Isoform.3.of.Charged.multivesicular.body.protein.3
fructose.bisphosphate.aldolase.A
Ribonuclease.pancreatic
dipeptidase.2
EGF.containing.fibulin.like.extracellular.matrix.protein.1
E3.ubiquitin.protein.ligase.RNF149
Zinc.finger.protein.basonuclin.2
Extracellular.superoxide.dismutase..Cu.Zn.
Doublesex..and.mab.3.related.transcription.factor.2

Isoform.3.of.Dedicator.of.cytokinesis.protein.10
Ig.kappa.chain.V.III.region.NG9
Reticulon.4.receptor.like.2
Calcium.binding.protein.1
Ig.heavy.chain.V.III.region.WAS
Keratin..type.I.cytoskeletal.12
Roundabout.homolog.4
glyceraldehyde.3.phosphate.dehydrogenase
Dynein.heavy.chain.domain.containing.protein.1
Ig.heavy.chain.V.III.region.TEI
Isoform.2.of.Complement.C4.A
Ig.kappa.chain.V.II.region.TEW
keratin..type.I.cytoskeletal.24

TBC1.domain.family.member.2A
Keratin..type.II.cytoskeletal.4
Isoform.2.of.Putative.macrophage.stimulating.1.like.protein
Peroxiredoxin.2
Target.of.Nesh.SH3
Ras.related.GTP.binding.protein.A
Isoform.Short.of.14.3.3.protein.beta.alpha
Keratin..type.I.cytoskeletal.14
Dynein.heavy.chain.5..axonemal
Ig.lambda.1.chain.C.regions
GRIP.and.coiled.coil.domain.containing.protein.2

Table B-2. *Total list of proteins (664) obtained with LRA when the proteins not quantified and top 20 HAPs were filtered.*

14-3-3 protein theta YWHAQ
14-3-3 protein zeta/delta YWHAZ
1-phosphatidylinositol 4,5-bisphosphate phosphodiesterase delta-4 PLCD4
39S ribosomal protein L1, mitochondrial MRPL1
3-oxo-5-beta-steroid 4-dehydrogenase AKR1D1
40S ribosomal protein S14 RPS14
40S ribosomal protein S27-like RPS27L
40S ribosomal protein S3 RPS3
4-trimethylaminobutyraldehyde dehydrogenase ALDH9A1
60S ribosomal protein L21 RPL21
72 kDa type IV collagenase MMP2
Acetoacetyl-CoA synthetase AACS
Acid-sensing ion channel 5 ASIC5
Actin, aortic smooth muscle ACTA2
Actin, cytoplasmic 1 ACTB
Acyl-coenzyme A synthetase ACSM1, mitochondrial ACSM1
Acyl-coenzyme A synthetase ACSM5, mitochondrial ACSM5
Acylpyruvase FAHD1, mitochondrial FAHD1
Adenosine deaminase CECR1 CECR1
Adipocyte plasma membrane-associated protein APMAP
Adiponectin ADIPOQ
ADP-ribosylation factor-like protein 8A ARL8A
Afamin AFM
Alcohol dehydrogenase 4 ADH4
Alpha-1-acid glycoprotein 1 ORM1
Alpha-1-acid glycoprotein 2 ORM2
Alpha-1-antichymotrypsin SERPINA3
Alpha-1B-glycoprotein A1BG
Alpha-2-antiplasmin SERPINF2
Alpha-2-HS-glycoprotein AHSG
Alpha-actinin-1 ACTN1
Alpha-actinin-4 ACTN4
Alpha-enolase ENO1
Alpha-ketoglutarate-dependent dioxygenase alkB homolog 2 ALKBH2

Anaphase-promoting complex subunit 16 ANAPC16
Anaphase-promoting complex subunit CDC26 CDC26
Angiogenin ANG
Angiotensinogen AGT
Ankyrin repeat domain-containing protein 11 ANKRD11
Ankyrin repeat domain-containing protein 12 ANKRD12
Ankyrin repeat domain-containing protein 36A ANKRD36
Ankyrin repeat domain-containing protein 36B ANKRD36B
Ankyrin repeat domain-containing protein 36C ANKRD36C
Annexin A1 ANXA1
Annexin A4 ANXA4
Antithrombin-III SERPINC1
AP-1 complex subunit gamma-1 AP1G1
Apoptotic chromatin condensation inducer in the nucleus ACIN1
Armadillo repeat-containing X-linked protein 3 ARMCX3
ATPase family AAA domain-containing protein 5 ATAD5
ATP-dependent RNA helicase DDX55 DDX55
Axin interactor, dorsalization-associated protein AIDA
Band 3 anion transport protein SLC4A1
Bardet-Biedl syndrome 7 protein BBS7
Beta-2-glycoprotein 1 APOH
Beta-2-microglobulin B2M
Beta-actin-like protein 2 ACTBL2
Beta-Ala-His dipeptidase CNDP1
Bicaudal D-related protein 2 CCDC64B
Biliverdin reductase A BLVRA
Blood vessel epicardial substance BVES
BTB/POZ domain-containing adapter for CUL3-mediated RhoA degradation protein 3 KCTD10
C4b-binding protein alpha chain C4BPA
C4b-binding protein beta chain C4BPB
Cadherin-5 CDH5
Calcium-binding protein 1 CABP1
Calcium-binding protein 5 CABP5

Calcium-dependent secretion activator 1 CADPS
Calcium-dependent secretion activator 2 CADPS2
Calcyphosin-2 CAPS2
Calmodulin-like protein 3 CALML3
Calmodulin-like protein 4 CALML4
Calmodulin-like protein 6 CALML6
Calpain-1 catalytic subunit CAPN1
cAMP-dependent protein kinase inhibitor alpha PKIA
Carbonic anhydrase 2 CA2
Carboxypeptidase B2 CPB2
Carboxypeptidase N catalytic chain CPN1
Carboxypeptidase N subunit 2 CPN2
Cartilage acidic protein 1 CRTAC1
Cartilage oligomeric matrix protein COMP
Casein kinase II subunit alpha' CSNK2A2
Caspase recruitment domain-containing protein 9 CARD9
Caspase-5 CASP5
Caspase-6 CASP6
Catenin beta-1 CTNNB1
CCA tRNA nucleotidyltransferase 1, mitochondrial TRNT1
CD5 antigen-like CD5L
Centromere protein S APITD1
Centrosomal protein of 41 kDa CEP41
Charged multivesicular body protein 3 CHMP3
Chitinase-3-like protein 2 CHI3L2
Cholesteryl ester transfer protein CETP
Chromobox protein homolog 3 CBX3
Cingulin CGN
Cofilin-1 CFL1
Coiled-coil domain-containing protein 126 CCDC126
Collagen alpha-3(VI) chain COL6A3
Collectin-11 COLEC11
Complement factor B CFB
Complement factor D CFD

Complement factor H CFH
Complement factor H-related protein 1 CFHR1
Complement factor H-related protein 2 CFHR2
Complement factor H-related protein 3 CFHR3
Complement factor H-related protein 5 CFHR5
Complement factor I CFI
Connector enhancer of kinase suppressor of ras 1 CNKSR1
Contactin-associated protein-like 2 CNTNAP2
Contactin-associated protein-like 3B CNTNAP3B
Coordinator of PRMT5 and differentiation stimulator COPRS
COX assembly mitochondrial protein 2 homolog CMC2
COX assembly mitochondrial protein homolog CMC1
C-type lectin domain family 4 member D CLEC4D
Cullin-4A CUL4A
Cullin-4B CUL4B
Cullin-associated NEDD8-dissociated protein 1 CAND1
Cyclic AMP-responsive element-binding protein 3 CREB3
Cystatin-C CST3
Cytochrome b-c1 complex subunit 7 UQCRB
Cytochrome P450 4Z1 CYP4Z1
Cytosolic carboxypeptidase 2 AGBL2
DC-STAMP domain-containing protein 1 DCST1
Dedicator of cytokinesis protein 10 DOCK10
Dedicator of cytokinesis protein 3 DOCK3
DNA excision repair protein ERCC-6 ERCC6
DNA polymerase alpha catalytic subunit POLA1
DNA polymerase kappa POLK
DNA-directed RNA polymerase I subunit RPA43 TWISTNB
DnaJ homolog subfamily B member 3 DNAJB3
Doublesex- and mab-3-related transcription factor 2 DMRT2
dTDP-D-glucose 4,6-dehydratase TGDS
Dual oxidase 2 DUOX2
Dynamin-binding protein DNMBP
Dynein heavy chain 12, axonemal DNAH12
Dynein heavy chain 17, axonemal DNAH17
Dynein heavy chain domain-containing protein 1 DNHD1
Dystonin DST
Dystrophin DMD
E3 ubiquitin-protein ligase BRE1B RNF40
E3 ubiquitin-protein ligase TRIM13 TRIM13
E3 UFM1-protein ligase 1 UFL1
Ectonucleotide pyrophosphatase/phosphodiesterase family member 2 ENPP2
EGF-containing fibulin-like extracellular matrix protein 1 EFEMP1
Elongation factor Ts, mitochondrial TSFM
Ephrin type-A receptor 7 EPHA7
Epididymal secretory glutathione peroxidase GPX5
Epididymal secretory protein E3-alpha EDDM3A
Epoxide hydrolase 1 EPHX1
Erythrocyte band 7 integral membrane protein STOM
Extracellular matrix protein 1 ECM1
Factor in the germline alpha FIGLA
Fatty acid-binding protein, heart FABP3
F-box/WD repeat-containing protein 12 FBXW12
Ferrochelatase, mitochondrial FECH
Fetuin-B FETUB
Fibrinogen-like protein 1 FGL1
Fibroblast growth factor 6 FGF6
Fibroleukin FGL2
Fibronectin FN1
Fibulin-1 FBLN1
Ficolin-2 FCN2
Ficolin-3 FCN3
Filamin-A FLNA
Formimidoyltransferase-cyclodeaminase FTCD
Frataxin, mitochondrial FXN
Frizzled-6 FZD6
Fructose-bisphosphate aldolase B ALDOB
FYVE, RhoGEF and PH domain-containing protein 2 FGD2
Galactose-1-phosphate uridylyltransferase GALT
Galectin-3-binding protein LGALS3BP
Gamma-tubulin complex component 3 TUBGCP3
Gelsolin GSN
Glutathione peroxidase 3 GPX3
Glutathione S-transferase theta-1 GSTT1
Glyceraldehyde-3-phosphate dehydrogenase GAPDH
Glycine N-acyltransferase-like protein 1 GLYATL1
Glycine N-acyltransferase-like protein 2 GLYATL2
Glycoprotein hormone beta-5 GPHB5
Glyoxalase domain-containing protein 5 GLOD5
Golgin subfamily A member 6-like protein 9 GOLGA6L9
Golgi-specific brefeldin A-resistance guanine nucleotide exchange factor 1 GBF1
Graves disease carrier protein SLC25A16
GRIP and coiled-coil domain-containing protein 2 GCC2
Guanine nucleotide-binding protein G(T) subunit gamma-T1 GNGT1
H/ACA ribonucleoprotein complex subunit 3 NOP10
Haptoglobin-related protein HPR
HCLS1-associated protein X-1 HAX1
Hemogen HEMGN
Hemopexin HPX
Heparin cofactor 2 SERPIND1
Hepatocyte growth factor activator HGFAC
Hepatocyte growth factor-like protein MST1
Heterogeneous nuclear ribonucleoprotein H3 HNRNPH3
Hippocalcin-like protein 4 HPCAL4
HIRA-interacting protein 3 HIRIP3
Histidine-rich glycoprotein HRG
Histidine--tRNA ligase, cytoplasmic HARS
Histone H2A type 1-H HIST1H2AH
Histone H3.1t HIST3H3
Histone H4-like protein type G HIST1H4G
Histone-lysine N-methyltransferase SUV39H1 SUV39H1
Hyaluronan-binding protein 2 HABP2
Hydrocephalus-inducing protein homolog HYDIN

Hydrolethalus syndrome protein 1 HYLS1
Importin-4 IPO4
Inhibin beta C chain INHBC
Inositol 1,4,5-trisphosphate receptor type 2 ITPR2
Inositol polyphosphate multikinase IPMK
Insulin-like growth factor II IGF2
Insulin-like growth factor-binding protein 3 IGFBP3
Insulin-like growth factor-binding protein complex acid labile subunit IGFALS
Integrin alpha-IIb ITGA2B
Integrin beta-3 ITGB3
Integrin-linked protein kinase ILK
Inter-alpha-trypsin inhibitor heavy chain H1 ITIH1
Inter-alpha-trypsin inhibitor heavy chain H2 ITIH2
Inter-alpha-trypsin inhibitor heavy chain H3 ITIH3
Inter-alpha-trypsin inhibitor heavy chain H4 ITIH4
Interferon epsilon IFNE
Inversin INVS
Iporin RUSC2
IQ domain-containing protein F1 IQCF1
Iron-sulfur cluster co-chaperone protein HscB, mitochondrial HSCB
Isoform 2 of 39S ribosomal protein L35, mitochondrial MRPL35
Isoform 2 of Alpha-1B-glycoprotein A1BG
Isoform 2 of Ankyrin repeat domain-containing protein 24 ANKRD24
Isoform 2 of Argininosuccinate lyase ASL
Isoform 2 of Calcyphosin-2 CAPS2
Isoform 2 of Clusterin CLU
Isoform 2 of Coiled-coil domain-containing protein 17 CCDC17
Isoform 2 of Coiled-coil domain-containing protein 28B CCDC28B
Isoform 2 of C-reactive protein CRP
Isoform 2 of C-type lectin domain family 2 member A CLEC2A
Isoform 2 of Dedicator of cytokinesis protein 7 DOCK7
Isoform 2 of Derlin-1 DERL1
Isoform 2 of DnaJ homolog subfamily A member 4 DNAJA4
Isoform 2 of Fermitin family homolog 3 FERMT3
Isoform 2 of Fibronectin FN1
Isoform 2 of Filamin A-interacting protein 1-like FILIP1L
Isoform 2 of Gelsolin GSN
Isoform 2 of Gem-associated protein 2 GEMIN2
Isoform 2 of GMP reductase 2 GMPR2
Isoform 2 of G-protein coupled receptor 161 GPR161
Isoform 2 of HLA class I histocompatibility antigen, A-11 alpha chain HLA-A
Isoform 2 of Insulin-like growth factor II IGF2
Isoform 2 of La-related protein 1B LARP1B
Isoform 2 of Liprin-beta-1 PPFIBP1
Isoform 2 of Mannan-binding lectin serine protease 1 MASP1
Isoform 2 of MDS1 and EVI1 complex locus protein EVI1 MECOM
Isoform 2 of Meiotic recombination protein REC8 homolog REC8
Isoform 2 of MIF4G domain-containing protein MIF4GD
Isoform 2 of MORN repeat-containing protein 3 MORN3
Isoform 2 of Mucolipin-2 MCOLN2
Isoform 2 of Myoneurin MYNN
Isoform 2 of Neuroblast differentiation-associated protein AHNAK AHNAK
Isoform 2 of Neuronal acetylcholine receptor subunit beta-4 CHRNB4
Isoform 2 of Neuropilin and tolloid-like protein 2 NETO2
Isoform 2 of NF-kappa-B essential modulator IKBKG
Isoform 2 of Nuclear receptor coactivator 7 NCOA7
Isoform 2 of Nuclear-interacting partner of ALK ZC3HC1
Isoform 2 of Nucleolar protein 16 NOP16
Isoform 2 of Palmitoyltransferase ZDHHC15 ZDHHC15
Isoform 2 of Pantothenate kinase 1 PANK1
Isoform 2 of PCNA-interacting partner PARPBP
Isoform 2 of PDZ and LIM domain protein 4 PDLIM4
Isoform 2 of PRELI domain-containing protein 2 PRELID2
Isoform 2 of Pre-mRNA-processing factor 40 homolog B PRPF40B
Isoform 2 of Probable histone-lysine N-methyltransferase PRDM7 PRDM7
Isoform 2 of Programmed cell death protein 6 PDCD6
Isoform 2 of Proline-serine-threonine phosphatase-interacting protein 1 PSTPIP1
Isoform 2 of Quinone oxidoreductase CRYZ
Isoform 2 of Regulator of G-protein signaling 3 RGS3
Isoform 2 of Regulator of G-protein signaling 4 RGS4
Isoform 2 of Rhombotin-2 LMO2
Isoform 2 of RNA-binding protein 26 RBM26
Isoform 2 of Shugoshin-like 2 SGOL2
Isoform 2 of Sodium channel and clathrin linker 1 SCLT1
Isoform 2 of SOSS complex subunit B2 NABP1
Isoform 2 of Splicing regulatory glutamine/lysine-rich protein 1 SREK1
Isoform 2 of Tetraspanin-14 TSPAN14
Isoform 2 of Transcription factor IIIB 90 kDa subunit BRF1
Isoform 2 of Tripartite motif-containing protein 72 TRIM72
Isoform 2 of Tropomyosin alpha-1 chain TPM1
Isoform 2 of Tropomyosin beta chain TPM2
Isoform 2 of Uncharacterized protein C14orf80 C14orf80
Isoform 2 of Uncharacterized protein C19orf60 C19orf60
Isoform 2 of Uncharacterized protein C9orf153 C9orf153
Isoform 2 of Vacuolar protein sorting-associated protein 13B VPS13B
Isoform 2 of WD repeat-containing protein 62 WDR62
Isoform 2 of Williams-Beuren syndrome chromosomal region 28 protein WBSCR28
Isoform 2 of Zinc finger CCCH-type antiviral protein 1-like ZC3HAV1L
Isoform 2 of Zinc finger MYM-type protein 3 ZMYM3
Isoform 2 of Zinc finger protein 177 ZNF177
Isoform 2 of Zinc finger protein 396 ZNF396
Isoform 2 of Zinc finger protein 91 ZNF91
Isoform 3 of Adenylate kinase 2, mitochondrial AK2

Isoform 3 of Adenylate kinase isoenzyme 5 AK5
Isoform 3 of Anion exchange protein 3 SLC4A3
Isoform 3 of Cadherin-13 CDH13
Isoform 3 of Cadherin-23 CDH23
Isoform 3 of Clusterin CLU
Isoform 3 of CMRF35-like molecule 8 CD300A
Isoform 3 of Coatomer subunit epsilon COPE
Isoform 3 of Collagen alpha-1(XVIII) chain COL18A1
Isoform 3 of Complement C1q tumor necrosis factor-related protein 6 C1QTNF6
Isoform 3 of Dedicator of cytokinesis protein 4 DOCK4
Isoform 3 of Egl nine homolog 1 EGLN1
Isoform 3 of Germinal center-associated signaling and motility protein GCSAM
Isoform 3 of Interleukin-12 receptor subunit beta-2 IL12RB2
Isoform 3 of Kv channel-interacting protein 1 KCNIP1
Isoform 3 of Membrane-spanning 4-domains subfamily A member 3 MS4A3
Isoform 3 of p53-regulated apoptosis-inducing protein 1 TP53AIP1
Isoform 3 of Peptidyl-prolyl cis-trans isomerase NIMA-interacting 4 PIN4
Isoform 3 of Protein pitchfork PIFO
Isoform 3 of Protrudin ZFYVE27
Isoform 3 of Putative postmeiotic segregation increased 2-like protein 3 PMS2P3
Isoform 3 of Ras-related protein Rap-1b RAP1B
Isoform 3 of SET domain-containing protein 4 SETD4
Isoform 3 of Vascular cell adhesion protein 1 VCAM1
Isoform 3 of V-type proton ATPase subunit E 1 ATP6V1E1
Isoform 4 of Abhydrolase domain-containing protein 12B ABHD12B
Isoform 4 of CD99 antigen-like protein 2 CD99L2
Isoform 4 of Dynein heavy chain 14, axonemal DNAH14
Isoform 4 of Inter-alpha-trypsin inhibitor heavy chain H4 ITIH4
Isoform 4 of Interleukin-1 receptor antagonist protein IL1RN
Isoform 4 of Ly6/PLAUR domain-containing protein 1 LYPD1

Isoform 4 of Major histocompatibility complex class I-related gene protein MR1
Isoform 4 of Myosin-11 MYH11
Isoform 4 of NF-kappa-B inhibitor-interacting Ras-like protein 2 NKIRAS2
Isoform 4 of Nuclear factor of activated T-cells, cytoplasmic 4 NFATC4
Isoform 4 of Protein PAT1 homolog 1 PATL1
Isoform 4 of RUN domain-containing protein 3A RUNDC3A
Isoform 4 of TBC1 domain family member 7 TBC1D7
Isoform 4 of Unconventional myosin-XIX MYO19
Isoform 4 of Zinc finger protein 195 ZNF195
Isoform 5 of Androgen-induced gene 1 protein AIG1
Isoform 5 of C-type lectin domain family 12 member A CLEC12A
Isoform 5 of Fibronectin FN1
Isoform 5 of LIM and senescent cell antigen-like-containing domain protein 1 LIMS1
Isoform 5 of Liprin-beta-1 PPFIBP1
Isoform 5 of Obscurin OBSCN
Isoform 5 of Synaptotagmin-like protein 2 SYTL2
Isoform 6 of Kv channel-interacting protein 2 KCNIP2
Isoform 6 of SWI/SNF-related matrix-associated actin-dependent regulator of chromatin subfamily E member 1 SMARCE1
Isoform 7 of Dual specificity protein phosphatase 13 isoform A DUSP13
Isoform 7 of Nesprin-1 SYNE1
Isoform 7 of Prolactin receptor PRLR
Isoform 8 of Dual specificity protein phosphatase 13 isoform A DUSP13
Isoform 8 of Neurofascin NFASC
Isoform A of Proline/serine-rich coiled-coil protein 1 PSRC1
Isoform B of Coagulation factor VII F7
Isoform C of Calpain-10 CAPN10
Isoform C of Caveolin-2 CAV2
Isoform C of Fibulin-1 FBLN1
Isoform C of Proteoglycan 4 PRG4
Isoform C of Trypsin-3 PRSS3
Isoform HNF4-Alpha-8 of Hepatocyte nuclear factor 4-alpha

HNF4A
Isoform LMW of Kininogen-1 KNG1
Isoform PDE1B2 of Calcium/calmodulin-dependent 3',5'-cyclic nucleotide phosphodiesterase 1B PDE1B
Isoform SV of 14-3-3 protein epsilon YWHAE
Isopentenyl-diphosphate delta-isomerase 2 IDI2
Kallistatin SERPINA4
Katanin p60 ATPase-containing subunit A1 KATNA1
KDEL motif-containing protein 1 KDELC1
Kelch-like protein 20 KLHL20
Killer cell lectin-like receptor subfamily B member 1 KLRB1
Kinesin-like protein KIF27 KIF27
Kininogen-1 KNG1
L-2-hydroxyglutarate dehydrogenase, mitochondrial L2HGDH
Lactosylceramide 4-alpha-galactosyltransferase A4GALT
Leucine-rich repeat-containing protein 14 LRRC14
Leucine-rich repeat-containing protein 23 LRRC23
Leucine-rich repeat-containing protein 59 LRRC59
Leukotriene B4 receptor 1 LTB4R
Lipopolysaccharide-binding protein LBP
LisH domain-containing protein FOPNL FOPNL
L-selectin SELL
LSM domain-containing protein 1 LSMD1
Luc7-like protein 3 LUC7L3
Lumican LUM
Ly-6/neurotoxin-like protein 1 LYNX1
Lysine-specific demethylase 3B KDM3B
Lysophosphatidic acid receptor 1 LPAR1
Lysozyme C LYZ
Macrophage colony-stimulating factor 1 receptor CSF1R
Mannan-binding lectin serine protease 1 MASP1
Mannan-binding lectin serine protease 2 MASP2
Mannose-binding protein C MBL2
MARVEL domain-containing protein 3 MARVELD3
Mediator of RNA polymerase II transcription subunit 30 MED30

Membrane-associated progesterone receptor component 2 PGRMC2
Metallo-beta-lactamase domain-containing protein 2 MBLAC2
Methyltransferase-like protein 14 METTL14
Microcephalin MCPH1
Microtubule-associated protein 2 MAP2
Microtubule-associated serine/threonine-protein kinase 4 MAST4
Mimitin, mitochondrial NDUFAF2
Minor histocompatibility protein HA-1 HMHA1
Mitochondrial import inner membrane translocase subunit Tim9 TIMM9
Mitochondrial inner membrane protease ATP23 homolog XRCC6BP1
Mitochondrial thiamine pyrophosphate carrier SLC25A19
Mixed lineage kinase domain-like protein MLKL
MOB kinase activator 1A MOB1A
Monocyte differentiation antigen CD14 CD14
Mucin-like protein 1 MUCL1
Multimerin-1 MMRN1
Mycophenolic acid acyl-glucuronide esterase, mitochondrial ABHD10
Myelin expression factor 2 MYEF2
Myeloid differentiation primary response protein MyD88 MYD88
Myogenic factor 6 MYF6
Myomegalin PDE4DIP
Myosin light chain 6B MYL6B
Myosin light polypeptide 6 MYL6
Myosin regulatory light chain 12A MYL12A
Myosin-14 MYH14
Myosin-9 MYH9
Myotilin MYOT
N6-adenosine-methyltransferase 70 kDa subunit METTL3
N-acetylmuramoyl-L-alanine amidase PGLYRP2
NACHT and WD repeat domain-containing protein 1 NWD1
NACHT, LRR and PYD domains-containing protein 4 NLRP4
NADH dehydrogenase [ubiquinone] 1 beta subcomplex subunit 9 NDUFB9

Nck-associated protein 5 NCKAP5
Negative elongation factor E NELFE
Nesprin-2 SYNE2
Neuropeptide FF receptor 2 NPFFR2
Ninjurin-2 NINJ2
NLR family CARD domain-containing protein 4 NLRC4
Nuclear cap-binding protein subunit 2-like NCBP2L
Nucleoporin-like protein 2 NUPL2
Opalin OPALIN
Oral-facial-digital syndrome 1 protein OFD1
Origin recognition complex subunit 6 ORC6
Osteopontin SPP1
Out at first protein homolog OAF
Pachytene checkpoint protein 2 homolog TRIP13
Palladin PALLD
Peptide YY PYY
Peroxiredoxin-1 PRDX1
Peroxiredoxin-2 PRDX2
Peroxisomal biogenesis factor 19 PEX19
Phosducin-like protein 2 PDCL2
Phosphatidylcholine-sterol acyltransferase LCAT
Phosphatidylinositol 4-kinase alpha PI4KA
Phosphatidylinositol 5-phosphate 4-kinase type-2 alpha PIP4K2A
Phosphatidylinositol-glycan-specific phospholipase D GPLD1
Phospholipid transfer protein PLTP
Phosphomevalonate kinase PMVK
Pigment epithelium-derived factor SERPINF1
Plasma kallikrein KLKB1
Plasma protease C1 inhibitor SERPING1
Plasma serine protease inhibitor SERPINA5
Plasminogen PLG
Plasminogen-like protein A PLGLA
Plasminogen-like protein B PLGLB1
Plastin-2 LCP1
Platelet basic protein PPBP
Platelet factor 4 PF4
Platelet factor 4 variant PF4V1

Platelet glycoprotein Ib beta chain GP1BB
Pleckstrin homology domain-containing family J member 1 PLEKHJ1
Pleckstrin homology-like domain family B member 3 PHLDB3
Poly(A) polymerase alpha PAPOLA
Polymeric immunoglobulin receptor PIGR
Potassium voltage-gated channel subfamily A member 5 KCNA5
POTE ankyrin domain family member E POTEE
POTE ankyrin domain family member F POTEF
POTE ankyrin domain family member I POTEI
PRAME family member 14 PRAMEF14
Prefoldin subunit 1 PFDN1
Pregnancy zone protein PZP
Pre-mRNA-splicing factor 18 PRPF18
Prenylcysteine oxidase 1 PCYOX1
Probable E3 ubiquitin-protein ligase HECTD4 HECTD4
Probable E3 ubiquitin-protein ligase HERC4 HERC4
Probable E3 ubiquitin-protein ligase MID2 MID2
Probable N-acetyltransferase 8 NAT8
Probable proline--tRNA ligase, mitochondrial PARS2
Probable serine carboxypeptidase CPVL CPVL
Procollagen C-endopeptidase enhancer 1 PCOLCE
Profilin-1 PFN1
Prolactin-inducible protein PIP
Prolyl 4-hydroxylase subunit alpha-3 P4HA3
Properdin CFP
Prostaglandin E synthase 2 PTGES2
Prostaglandin-H2 D-isomerase PTGDS
Protein AMBP AMBP
Protein disulfide-isomerase-like protein of the testis PDILT
Protein FAM163A FAM163A
Protein FAM212A FAM212A
Protein FAM223A FAM223A
Protein FAM45A FAM45A
Protein FAM9A FAM9A
Protein lunapark LNP

Protein Mdm4 MDM4
Protein NYNRIN NYNRIN
Protein phosphatase 1 regulatory subunit 3D PPP1R3D
Protein phosphatase 1B PPM1B
Protein piccolo PCLO
Protein ripply2 RIPPLY2
Protein ripply3 RIPPLY3
Protein S100-A16 S100A16
Protein TBATA TBATA
Protein Z-dependent protease inhibitor SERPINA10
Proteoglycan 4 PRG4
Prothrombin F2
PTB domain-containing engulfment adapter protein 1 GULP1
Pulmonary surfactant-associated protein B SFTPB
Putative ADP-ribosylation factor-like protein 5C ARL5C
Putative ankyrin repeat domain-containing protein 31 ANKRD31 PE=5
Putative beta-actin-like protein 3 POTEKP PE=5
Putative ciliary rootlet coiled-coil protein-like 2 protein CROCCP3 PE=5
Putative deoxyribonuclease TATDN3 TATDN3
Putative golgin subfamily A member 6-like protein 4 GOLGA6L4 PE=5
Putative GRINL1B complex locus protein 2 GCOM2 PE=5
Putative macrophage stimulating 1-like protein MST1L
Putative protein arginine N-methyltransferase 10 PRMT10
Putative ubiquitin-conjugating enzyme E2 N-like UBE2NL
Putative uncharacterized protein encoded by LINC00324 LINC00324 PE=5
Putative uncharacterized protein FLJ25328
Putative uncharacterized protein KIRREL3-AS3 KIRREL3-AS3 PE=5
Putative V-set and immunoglobulin domain-containing-like protein IGHV4OR15-8 IGHV4OR15-8 PE=5
Putative zinc finger protein LOC730110 PE=5
Putative zinc-alpha-2-glycoprotein-like 2 PE=5
Pyruvate kinase PKM PKM

Rab11 family-interacting protein 2 RAB11FIP2
Rab3 GTPase-activating protein non-catalytic subunit RAB3GAP2
Rab-like protein 5 RABL5
Ras-related protein Rab-1A RAB1A
Ras-related protein Rab-33A RAB33A
Ras-related protein Rap-1A RAP1A
Ras-specific guanine nucleotide-releasing factor RalGPS2 RALGPS2
Regulating synaptic membrane exocytosis protein 1 RIMS1
Regulator of G-protein signaling 3 RGS3
Regulator of microtubule dynamics protein 3 RMDN3
Relaxin receptor 2 RXFP2
Retinitis pigmentosa 1-like 1 protein RP1L1
Retinol-binding protein 4 RBP4
Rho GTPase-activating protein 1 ARHGAP1
Rho GTPase-activating protein 29 ARHGAP29
Rho GTPase-activating protein 32 ARHGAP32
Rho-related GTP-binding protein RhoC RHOC
Ribosomal protein S6 kinase delta-1 RPS6KC1
RING-box protein 2 RNF7
RNA polymerase-associated protein CTR9 homolog CTR9
RNA-binding motif protein, X chromosome RBMX
Scavenger receptor cysteine-rich domain-containing group B protein SRCRB4D
Schlafen family member 5 SLFN5
Sciellin SCEL
SEC14-like protein 1 SEC14L1
Secretoglobin family 3A member 1 SCGB3A1
Secretoglobin family 3A member 2 SCGB3A2
Selenide, water dikinase 2 SEPHS2
Selenoprotein P SEPP1
Sepiapterin reductase SPR
Septin-6 SEPT6
Serine/arginine-rich splicing factor 2 SRSF2
Serine/threonine-protein kinase 35 STK35
Serine/threonine-protein kinase Nek9 NEK9

Serine/threonine-protein kinase PDIK1L PDIK1L
Serine/threonine-protein kinase PLK2 PLK2
Serum amyloid A-1 protein SAA1
Serum amyloid A-2 protein SAA2
Serum amyloid A-4 protein SAA4
Serum amyloid P-component APCS
Serum deprivation-response protein SDPR
Serum paraoxonase/arylesterase 1 PON1
Serum paraoxonase/lactonase 3 PON3
Sex hormone-binding globulin SHBG
Sialic acid-binding Ig-like lectin 16 SIGLEC16
Sideroflexin-1 SFXN1
Spermatid-associated protein SPERT
Spermatogenesis-associated protein 22 SPATA22
Spermatogenesis-associated protein 5-like protein 1 SPATA5L1
Spindle and kinetochore-associated protein 2 SKA2
Splicing factor 3B subunit 1 SF3B1
Squamous cell carcinoma antigen recognized by T-cells 3 SART3
StAR-related lipid transfer protein 9 STARD9
Stathmin-2 STMN2
Stimulated by retinoic acid gene 8 protein homolog STRA8
Structural maintenance of chromosomes protein 1B SMC1B
Sulfhydryl oxidase 1 QSOX1
Sulfotransferase 1C3 SULT1C3
Synaptojanin-1 SYNJ1
Tafazzin TAZ
Talin-1 TLN1
TBC1 domain family member 2A TBC1D2
T-complex protein 1 subunit gamma CCT3
Tenascin-X TNXB
Testis- and ovary-specific PAZ domain-containing protein 1 TOPAZ1
Testis-specific chromodomain protein Y 2 CDY2A
Tetranectin CLEC3B
Tetratricopeptide repeat protein 40 TTC40
THAP domain-containing protein 4 THAP4

Thrombospondin-1 THBS1
Tigger transposable element-derived protein 5 TIGD5
TNFAIP3-interacting protein 1 TNIP1
Torsin-3A TOR3A
Trafficking kinesin-binding protein 1 TRAK1
Transcription termination factor 1 TTF1
Transcriptional repressor protein YY1 YY1
Transforming growth factor-beta-induced protein ig-h3 TGFBI
Transgelin-2 TAGLN2
Transient receptor potential cation channel subfamily M member 8 TRPM8
Translation initiation factor eIF-2B subunit gamma EIF2B3
Translationally-controlled tumor protein TPT1
Transmembrane protein 242 TMEM242
Transthyretin TTR
Trichoplein keratin filament-binding protein TCHP
Tripartite motif-containing protein 15 TRIM15
Tripartite motif-containing protein 51 TRIM51
Triple functional domain protein TRIO
Tropomyosin alpha-3 chain TPM3
Tropomyosin alpha-4 chain TPM4
Tubulin alpha-1A chain TUBA1A
Tubulin alpha-1B chain TUBA1B
Tubulin alpha-3C/D chain TUBA3C
Tubulin alpha-4A chain TUBA4A
Tubulin alpha-8 chain TUBA8
Tubulin beta-1 chain TUBB1
Ubiquitin carboxyl-terminal hydrolase 12 USP12
Ubiquitin carboxyl-terminal hydrolase 49 USP49
Ubiquitin-like protein 3 UBL3
Uncharacterized aarF domain-containing protein kinase 4 ADCK4
Uncharacterized aarF domain-containing protein kinase 5 ADCK5
Uncharacterized protein C10orf131 C10orf131
Uncharacterized protein C11orf70 C11orf70
Uncharacterized protein C14orf93 C14orf93

Uncharacterized protein C15orf26 C15orf26
Uncharacterized protein C17orf66 C17orf66
Uncharacterized protein C18orf63 C18orf63
Uncharacterized protein C4orf21 C4orf21
Uncharacterized protein C9orf163 C9orf163
Uncharacterized protein KIAA0753 KIAA0753
UPF0574 protein C9orf169 C9orf169
Uridine-cytidine kinase 1 UCK1
Vasohibin-2 VASH2
Vasorin VASN
Very long-chain acyl-CoA synthetase SLC27A2
Vinculin VCL
Vitamin D-binding protein GC
Vitamin K-dependent protein C PROC
Vitamin K-dependent protein S PROS1
Vitamin K-dependent protein Z PROZ
Vitelline membrane outer layer protein 1 homolog VMO1
Vitronectin VTN
Voltage-dependent calcium channel gamma-2 subunit CACNG2
von Willebrand factor VWF
WD repeat-containing protein 55 WDR55
WD repeat-containing protein 96 WDR96
WW domain-binding protein 4 WBP4
WW domain-containing oxidoreductase WWOX
Zinc finger CCCH domain-containing protein 13 ZC3H13
Zinc finger protein 208 ZNF208
Zinc finger protein 30 ZNF30
Zinc finger protein 469 ZNF469
Zinc finger protein 681 ZNF681
Zinc finger protein 701 ZNF701
Zinc finger protein 789 ZNF789
Zinc finger protein 99 ZNF99
Zinc finger protein basonuclin-2 BNC2
Zinc-alpha-2-glycoprotein AZGP1

7.3 Appendix C

List of significant proteins obtained from RapidMiner, SPSS and SIMCA.

Table C-1. *Thirty significant proteins identified with RapidMiner, SPSS and SIMCA.*

Protein description	RapidMiner	SPSS	SIMCA
Isoform 2 of Argininosuccinate lyase	✓	✓	✓
Doublesex- and mab-3-related transcription factor 2	✓	✓	
Nuclear cap-binding protein subunit 2-like	✓	✓	
Heterogeneous nuclear ribonucleoprotein H3	✓		✓
3-oxo-5-beta-steroid 4-dehydrogenase	✓		
C4b-binding protein alpha chain	✓		
Complement factor D	✓		
Fatty acid-binding protein, heart	✓		
Glutathione peroxidase 3	✓		
Isoform 5 of Androgen-induced gene 1 protein	✓		
Isoform 6 of Kv channel-interacting protein 2	✓		
LSM domain-containing protein 1	✓		
Mucin-like protein 1	✓		
Myosin-9	✓		
Zinc finger protein 701	✓		
Histone H2A type 1-H		✓	
Chromobox protein homolog 3		✓	
Isoform 3 of Collagen alpha-1(XVIII) chain		✓	
Tubulin alpha-3C/D chain		✓	
Luc7-like protein 3		✓	
Acyl-coenzyme A synthetase ACSM5, mitochondrial		✓	
Isoform 2 of Programmed cell death protein 6		✓	
Rho GTPase-activating protein 29			✓
Zinc finger protein basonuclin-2			✓
Golgi-specific brefeldin A-resistance guanine nucleotide exchange factor 1			✓
Neuropeptide FF receptor 2			✓
Rab3 GTPase-activating protein non-catalytic subunit			✓
Putative zinc finger protein LOC730110			✓
GRIP and coiled-coil domain-containing protein 2			✓
Sulfhydryl oxidase 1			✓

Table C-2. *Total list of 80 significant proteins obtained with LRA after one way ANOVA.*

Isoform 2 of Argininosuccinate lyase ASL
Rho GTPase-activating protein 29 ARHGAP29 1 2
Tetratricopeptide repeat protein 40 TTC40 2 3
Chromobox protein homolog 3 CBX3 1 4
Golgi-specific brefeldin A-resistance guanine nucleotide exchange factor 1 GBF1 1 2
Isoform 3 of Germinal center-associated signaling and motility protein GCSAM
Luc7-like protein 3 LUC7L3 1 2
Tropomyosin alpha-3 chain TPM3 1 2
Isoform 4 of CD99 antigen-like protein 2 CD99L2
Extracellular matrix protein 1 ECM1 1 2
Isoform 2 of C-reactive protein CRP
Mucin-like protein 1 MUCL1 1 1
Zinc finger protein basonuclin-2 BNC2 1 1
Doublesex- and mab-3-related transcription factor 2 DMRT2 2 2
Zinc finger protein 701 ZNF701 2 2
Heterogeneous nuclear ribonucleoprotein H3 HNRNPH3 1 2
Isoform 3 of Coatomer subunit epsilon COPE
Pregnancy zone protein PZP 1 4
Nuclear cap-binding protein subunit 2-like NCBP2L 3 1
Isoform 2 of Programmed cell death protein 6

PDCD6
Histone H2A type 1-H HIST1H2AH 1 3
POTE ankyrin domain family member F POTEF 1 2
Phosphatidylinositol-glycan-specific phospholipase D GPLD1 1 3
POTE ankyrin domain family member I POTEI 3 1
Probable serine carboxypeptidase CPVL CPVL 1 2
Leucine-rich repeat-containing protein 23 LRRC23 2 2
Neuropeptide FF receptor 2 NPFFR2 1 2
Putative beta-actin-like protein 3 POTEKP 5 1
Isoform 2 of Regulator of G-protein signaling 3 RGS3
Hyaluronan-binding protein 2 HABP2 1 1
Carbonic anhydrase 2 CA2 1 2
Regulator of G-protein signaling 3 RGS3 1 2
3-oxo-5-beta-steroid 4-dehydrogenase AKR1D1 1 1
GRIP and coiled-coil domain-containing protein 2 GCC2 1 4
Isoform 4 of Myosin-11 MYH11
LSM domain-containing protein 1 LSMD1 1 1
Oral-facial-digital syndrome 1 protein OFD1 1 1
Isoform 4 of Major histocompatibility complex class I-related gene protein MR1
Peptide YY PYY 1 3

Rab3 GTPase-activating protein non-catalytic subunit RAB3GAP2 1 1
Alpha-1-antichymotrypsin SERPINA3 1 2
Zinc-alpha-2-glycoprotein AZGP1 1 2
Tubulin alpha-1B chain TUBA1B 1 1
Tubulin beta-1 chain TUBB1 1 1
Ninjurin-2 NINJ2 2 1
Hemopexin HPX 1 2
Nesprin-2 SYNE2 1 3
Putative macrophage stimulating 1-like protein MST1L 2 2
Isoform 2 of Coiled-coil domain-containing protein 28B CCDC28B
Isoform 3 of Protein pitchfork PIFO
Isoform 2 of GMP reductase 2 GMPR2
F-box/WD repeat-containing protein 12 FBXW12 2 2
Fatty acid-binding protein, heart FABP3 1 4
Out at first protein homolog OAF 2 1
Spermatogenesis-associated protein 5-like protein 1 SPATA5L1 1 2
Calcium-dependent secretion activator 2 CADPS2 1 2
Scavenger receptor cysteine-rich domain-containing group B protein SRCRB4D 2 1
Isoform 5 of Androgen-induced gene 1 protein AIG1
Protein ripply2 RIPPLY2 2 1
Contactin-associated protein-like 3B

CNTNAP3B 2 2	1 1	Alpha-2-antiplasmin SERPINF2 1 3
Histidine--tRNA ligase, cytoplasmic HARS 1 2	ATPase family AAA domain-containing protein 5 ATAD5 1 4	Isoform 2 of PDZ and LIM domain protein 4 PDLIM4
Isoform 3 of Collagen alpha-1(XVIII) chain COL18A1	Fructose-bisphosphate aldolase B ALDOB 1 2	Plasminogen-like protein B PLGLB1 1 1
Importin-4 IPO4 1 2	Interferon epsilon IFNE 2 1	Mitochondrial thiamine pyrophosphate carrier SLC25A19 1 1
Trafficking kinesin-binding protein 1 TRAK1 1 1	Hydrocephalus-inducing protein homolog HYDIN 1 3	60S ribosomal protein L21 RPL21 1 2
Isoform 7 of Prolactin receptor PRLR	Protein FAM163A FAM163A 2 1	
Putative zinc finger protein LOC730110 5 3	72 kDa type IV collagenase MMP2 1 2	
Isoform 2 of Tetraspanin-14 TSPAN14	Isoform 2 of WD repeat-containing protein 62 WDR62	
Complement factor H-related protein 2 CFHR2		

7.3.1 Significant proteins identified between control and heart failure groups.

GRIP and coiled-coil domain-containing protein 2 (GCC2) is an identical protein binding protein located in the cytoplasm and golgi apparatus. It expression is ubiquitous and plays a role in recycling of the mannose 6-phosphate receptor from the late endosomes to the TGN (Reddy *et al.*, 2006).

Heterogeneous nuclear ribonucleoprotein H3 (HNRNPH3) is located in the nucleus and organelle lumen of the cell. It is responsible for RNA binding, the splicing process and participates in early heat shock-induced splicing arrest. It is a protein-binding gene associated with lymphatic system cancer.

Tubulin alpha-3C/D chain (TUBA3C) is a GTP-binding and nucleotide binding protein located in the cytoplasm that is expressed in the testis. It is a major constituent of microtubules

3-oxo-5-beta-steroid 4-dehydrogenase (AKR1D1) is located in the cytoplasm. It is an oxidoreductase enzyme that catalyses the reduction of progesterone, androstenedione, 17-alpha-hydroxyprogesterone and testosterone to 5-beta-reduced metabolites. This protein is

highly expressed in the liver and mutation in this gene has been associated with congenital bile acid synthesis defect 2 which is a condition characterised by jaundice, intrahepatic cholestasis and hepatic failure. Relationship between liver failure and heart failure is rarely documented. However, Saner et al. 2009 report that there is increasing evidence that when there is reduced cardiac output due to heart failure it causes reduced hepatic blood flow leading to hepatic failure. In addition, pulmonary conditions such as sleep apnea and chronic obstructive pulmonary disease (COPD) are risk factors for hepatic failure due to passive congestion of the liver because of heart failure (Giallourakis 2013). This protein was overexpressed on SHF patients.

7.4 Appendix D

Targeted MRM overview of the transitions used.

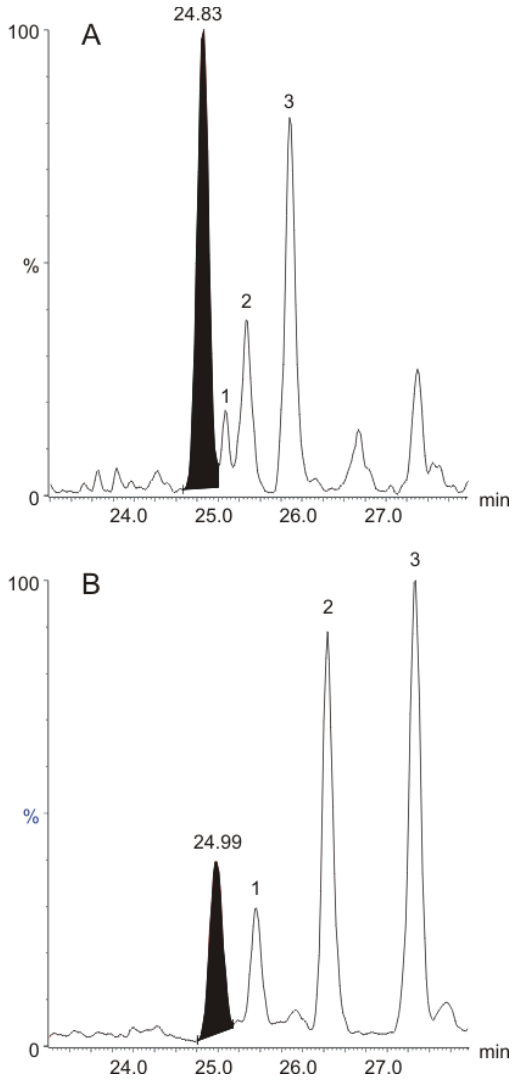


Figure D-1. *Development of chromatographic condition for LGPLVEQGR-$^{13}C_6{}^{15}N_4$ (summed MRM transitions 489.784 > 499.25 + 489.784 > 598.318 + 489.784 > 711.402 + 489.784 > 808.455) contrasting generic (A) and optimized (B) chromatographic conditions, temperature and gradient slope. The peak of interest is retention time (min) annotated. The '1', '2' and '3' components are isobaric species with one or more shared transitions. See section Experimental Conditions for final chromatographic conditions.*

Table D-1. *MRM transition overview tandem quadrupole and oa-ToF MRM experiments. [K] = $^{13}C_6{}^{15}N_4$ labeled; [R] = $^{13}C_6{}^{15}N_2$ labelled; *oa-ToF MRM based acquisitions only*

peptide sequence	precursor *m/z* [charge]	fragment *m/z* [fragment ion type]					EDC *m/z**
		1	2	3	4	5	
LVNEVTEFAK	575.3 [2+]	595.3 [y5]	694.4 [y6]	823.4 [y7]	937.5 [y8]		937.5
LVNEVTEFA[K]	579.3 [2+]	603.3 [y5]	702.4 [y6]	831.4 [y7]	945.5 [y8]		945.5
ATEHLSTLSEK	608.3 [2+]	664.4 [y6]	777.4 [y7]	914.5 [y8]			777.4
ATEHLSTLSE[K]	612.3 [2+]	672.4 [y6]	785.4 [y7]	922.5 [y8]			785.5
TGLQEVEVK	501.8 [2+]	603.3 [y5]	731.4 [y6]	844.5 [y7]	901.5 [y8]		731.4
TGLQEVEV[K]	505.8 [2+]	611.3 [y5]	739.4 [y6]	852.5 [y7]	909.5 [y8]		739.4
FPEVDVLTK	524.3 [2+]	575.3 [y5]	674.4 [y6]	803.5 [y7]	900.5 [y8]		674.4
FPEVDVLT[K]	528.3 [2+]	583.4 [y5]	682.4 [y6]	811.5 [y7]	908.5 [y8]		682.4
FQPTLLTLPR	593.4 [2+]	599.4 [y5]	712.5 [y6]	813.5 [y7]	910.6 [y8]		910.6
FQPTLLTLP[R]	598.4 [2+]	609.4 [y5]	722.5 [y6]	823.5 [y7]	920.6 [y8]		920.6
TAAQNLYEK	519.3 [2+]	552.3 [y4]	666.3 [y5]	794.4 [y6]	865.4 [y7]		865.4
TAAQNLYE[K]	523.3 [2+]	560.3 [y4]	674.4 [y5]	802.4 [y6]	873.5 [y7]		873.5
LGPLVEQGR	484.8 [2+]	489.2 [y4]	588.3 [y5]	701.4 [y6]	798.4 [y7]	855.5 [y8]	701.4
LGPLVEQG[R]	489.8 [2+]	499.2 [y4]	598.3 [y5]	711.4 [y6]	808.5 [y7]	865.5 [y8]	711.4
AAAATGTIFTFR	613.8 [2+]	683.4 [y5]	784.4 [y6]	841.5 [y7]	942.5 [y8]	1013.6 [y9]	942.5
AAAATGTIFTF[R]	618.8 [2+]	693.4 [y5]	794.4 [y6]	851.5 [y7]	952.5 [y8]	1023.6 [y9]	952.5
EANYIGSDK	498.7 [2+]	519.3 [y5]	682.3 [y6]	796.4 [y7]			682.3
EANYIGSD[K]	502.7 [2+]	527.3 [y5]	690.4 [y6]	804.4 [y7]			690.4
ESDTSYVSLK	564.8 [2+]	609.4 [y5]	696.4 [y6]	797.4 [y7]	912.5 [y8]		696.4
ESDTSYVSL[K]	568.8 [2+]	617.4 [y5]	704.4 [y6]	805.5 [y7]	920.5 [y8]		704.4
GYSIFSYATK	568.8 [2+]	569.3 [y5]	716.4 [y6]	829.4 [y7]	916.5 [y8]		916.5
GYSIFSYAT[K]	572.8 [2+]	577.3 [y5]	724.4 [y6]	837.5 [y7]	924.5 [y8]		924.5
GFYFNKPTGYGSSSR	834.4 [2+]	911.4 [y9]					911.4
GFYFNKPTGYGSSS[R]	839.4 [2+]	921.4 [y9]					921.4
LVNVVLGAHNVR	645.9 [2+]	653.3 [y6]	766.4 [y7]	865.5 [y8]	964.6 [y9]	1078.6 [y10]	766.4
LVNVVLGAHNV[R]	650.9 [2+]	663.4 [y6]	776.4 [y7]	875.5 [y8]	974.6 [y9]	1088.6 [y10]	776.4
ITLYGR	361.7 [2+]	395.2 [y3]	508.3 [y4]	609.3 [y5]			609.3
ITLYG[R]	366.7 [2+]	405.2 [y3]	518.3 [y4]	619.3 [y5]			619.3
SYPGLTSYLVR	628.3 [2+]	637.4 [y5]	738.4 [y6]	851.5 [y7]	908.5 [y8]	1005.6 [y9]	908.5
SYPGLTSYLV[R]	633.3 [2+]	647.4 [y5]	748.4 [y6]	861.5 [y7]	918.5 [y8]	1015.6 [y9]	918.5

7.5 Ethical approval

NHS
Health Research Authority

NRES Committee East Midlands - Nottingham 1
The Old Chapel
Royal Standard Place
Nottingham
NG1 6FS

Telephone: 0115 8839309
Facsimile: 0115 8839924

24 July 2012

Professor Leong L Ng
Professor of Medicine & Therapeutics
University of Leicester
Cardiovascular Sciences
Clinical Sciences Building
Leicester Royal Infirmary
LE2 7LX

Dear Professor Ng

Full title of study:	**Sample and data collection for diastolic heart failure study: Biomarkers for Diastolic Heart Failure**
REC reference number:	**12/EM/0222**

Thank you for your letter of 10th July 2012. I can confirm the REC has received the documents listed below as evidence of compliance with the approval conditions detailed in our letter dated 12 June 2012. Please note these documents are for information only and have not been reviewed by the committee.

Documents received

The documents received were as follows:

Document	*Version*	*Date*
Covering Letter		10 July 2012
Advertisement	2	10 July 2012
Participant Consent Form: Consent Form for Healthy Controls	1.0	10 July 2012
Participant Information Sheet: Healthy Volunteer Information Sheet	2.0	10 July 2012
Participant Information Sheet: Information Sheet for Patients with Weakened Heart Function	2.0	10 July 2012

You should ensure that the sponsor has a copy of the final documentation for the study. It is the sponsor's responsibility to ensure that the documentation is made available to R&D offices at all participating sites.

12/EM/0222	**Please quote this number on all correspondence**

Yours sincerely

W Rees

Wendy Rees
Assistant Committee Co-ordinator

NHS
Health Research Authority

NRES Committee East Midlands - Nottingham 1
The Old Chapel
Royal Standard Place
Nottingham
NG1 6FS

Telephone: 0115 8839436
Facsimile: 0115 8839924

18 June 2012

Professor Leong L Ng
Professor of Medicine & Therapeutics
University of Leicester
Cardiovascular Sciences
Clinical Sciences Building
Leicester Royal Infirmary
LE2 7LX

Dear Professor Ng

Study title:	**Sample and data collection for diastolic heart failure study: Biomarkers for Diastolic Heart Failure**
REC reference:	**12/EM/0222**

The Research Ethics Committee reviewed the above application at the meeting held on 12 June 2012. Thank you for sending Professor Ian Squire to discuss the application.

Ethical opinion

- The Committee advised it was an interesting study.
- The Committee asked the Researcher whether they would be informing GP's of the results of the ECHO and Spirometry. The Researcher confirmed they would be.
- The Committee explained they are concerned about the use of the term "would you like a heart MOT" in the poster as it could be coercive. The Researcher will remove it from the poster. The researcher explained he is always aware as to how he approaches healthy volunteers.
- The Researcher was advised the Spiometry test is not mentioned in the poster for healthy volunteers.
- The Committee advised the Researcher there are some spelling mistakes in the Participant Information Sheet.
- The Researcher was asked whether he would be looking into the medical records of healthy controls. The Researcher explained there would be no need to look at their records.
- The Researcher was asked which Dye would be used for both the MRI and ECHO. The Researcher advised he did not know that information.
- The Committee explained the risks of using Dye are not mentioned in the Patient Information Sheet. In addition the term harmless cannot be used.
- The Committee asked whether healthy controls would have their renal function checked as the Dye can damage the kidneys in people with previous kidney damage.
- The Committee explained the Participant Information Sheet should state there is a risk of discovering incidental findings.
- The Committee asked the Researcher whether the Participants GP would be informed of trivial findings. The Researcher explained they would respond to the

8 Reference

ABRAHAM, W.T. and KRUM, H., 2007. Heart failure: a practical approach to treatment. New York: McGraw-Hill.

AFSINOKTAY, A. and SHAH, S.J., 2013. Diagnosis and Management of Heart Failure with Preserved Ejection Fraction: 10 Key Lessons. *Current cardiology reviews*.

AHMAD, T., FIUZAT, M., PENCINA, M.J., GELLER, N.L., ZANNAD, F., CLELAND, J.G.F., SNIDER, J.V., BLANKENBERG, S., ADAMS, K.F., REDBERG, R.F., KIM, J.B., MASCETTE, A., MENTZ, R.J., O&AMP, APOS, CONNOR, C.M., FELKER, G.M. and JANUZZI, J.L., 2014. Charting a Roadmap for Heart Failure Biomarker Studies. *JACC: Heart Failure,* **2**(5), pp. 477-488.

AHMED, S.H., CLARK, L.L., PENNINGTON, W.R., WEBB, C.S., BONNEMA, D.D., LEONARDI, A.H., MCCLURE, C.D., SPINALE, F.G. and ZILE, M.R., 2006. Matrix metalloproteinases/tissue inhibitors of metalloproteinases: relationship between changes in proteolytic determinants of matrix composition and structural, functional, and clinical manifestations of hypertensive heart disease. *Circulation*, 113(17), pp. 2089-2096.

ALIGIANIS, I.A., MORGAN, N.V., MIONE, M., JOHNSON, C.A., ROSSER, E., HENNEKAM, R.C., ADAMS, G., TREMBATH, R.C., PILZ, D.T., STOODLEY, N., MOORE, A.T., WILSON, S. and MAHER, E.R., 2006. Mutation in Rab3 GTPase-Activating Protein (RAB3GAP) Non-catalytic Subunit in a Kindred with Martsolf Syndrome. *The American Journal of Human Genetics*, 78(4), pp. 702-707.

ASHCROFT, E. A., 2016. An Introduction to Mass Spectrometry. [ONLINE] Available at: http://www.astbury.leeds.ac.uk/facil/MStut/mstutorial.htm. [Accessed 30 August 2016].

ALTIERI, P., GHIGLIOTTI, G., BRUNELLI, C., GARIBALDI, S., NICOLINO, A., UBALDI, S., SPALLAROSSA, P., OLIVOTTI, L., ROSSETTIN, P. and BARSOTTI, A., 2003. Metalloproteinases 2 and 9 are increased in plasma of patients with heart failure. *European journal of clinical investigation*, 33(8), pp. 648-656.

ANAND, I.S., LATINI, R., FLOREA, V.G., KUSKOWSKI, M.A., RECTOR, T., MASSON, S., SIGNORINI, S., MOCARELLI, P., HESTER, A., GLAZER, R., COHN, J.N. and VAL-HEFT INVESTIGATORS, 2005. C-reactive protein in heart failure: prognostic value and the effect of valsartan. *Circulation*, 112(10), pp. 1428-1434.

ANDERSON, L. and HUNTER, C.L., 2006; 2005. Quantitative Mass Spectrometric Multiple Reaction Monitoring Assays for Major Plasma Proteins. *Molecular & Cellular Proteomics*, 5(4), pp. 573-588.

ANDERSON, N.L., 2010. The clinical plasma proteome: a survey of clinical assays for proteins in plasma and serum. *Clinical chemistry*, 56(2), pp. 177-185.

ANDERSON, N.L. and ANDERSON, N.G., 2002. The human plasma proteome: history, character, and diagnostic prospects. *Molecular & cellular proteomics*, 1(11), pp. 845-867.

ANDERSON, N.L., POLANSKI, M., PIEPER, R., GATLIN, T., TIRUMALAI, R.S., CONRADS, T.P., VEENSTRA, T.D., ADKINS, J.N., POUNDS, J.G., FAGAN, R. and LOBLEY, A., 2004. The human plasma proteome: a non-redundant list developed by combination of four separate sources. *Molecular & cellular proteomics*, 3(4), pp. 311-326.

ANKER, S.D. and VON HAEHLING, S., 2004. Inflammatory mediators in chronic heart failure: an overview. Heart (British Cardiac Society), 90(4), pp. 464-470.

ANNA BIERNACKA AND NIKOLAOS G FRANGOGIANNIS, 2011. Aging and Cardiac Fibrosis. 2(2): 158–173.

APOSTOLAKIS, S., LIP, G.Y. and SHANTSILA, E., 2010. Monocytes in heart failure: relationship to a deteriorating immune overreaction or a desperate attempt for tissue repair? *Cardiovascular research*, 85(4), pp. 649-660.

BARNES, P., 2010. Exhaled nitric oxide in pulmonary diseases: a comprehensive review. *Chest*, 138(3), pp. 682-692.

BARRATT-DUE, A., PISCHKE, S.E., BREKKE, O., THORGERSEN, E.B., NIELSEN, E.W., ESPEVIK, T., HUBER-LANG, M. and MOLLNES, T.E., 2012. Bride and groom in systemic inflammation--the bells ring for complement and Toll in cooperation. *Immunobiology*, 217(11), pp. 1047-1056.

BHALLA, V., GEORGIOPOULOU, V., GUPTA, D., MARTI, C.N., COLE, R.T., KALOGEROPOULOS, A., TANG, W.H.W. and BUTLER, J., 2012. Apolipoprotein A-1 Levels and Heart Failure Outcomes. *Journal of cardiac failure*, 18(8), pp. S71-S72.

BHUIYAN, T. and MAURER, M.S., 2011. Heart Failure with Preserved Ejection Fraction: Persistent Diagnosis, Therapeutic Enigma. Current Cardiovascular Risk Reports, 5(5), pp. 440-449.

BISHOP, A.L. and HALL, A., 2000. Rho GTPases and their effector proteins. *The Biochemical journal*, 348 Pt 2(2), pp. 241-255.

BLASKÓ, B., KOLKA, R., THORBJORNSDOTTIR, P., SIGURÐARSON, S.T., SIGURÐSSON, G., RÓNAI, Z., SASVÁRI-SZÉKELY, M., BÖÐVARSSON, S., THORGEIRSSON, G., PROHÁSZKA, Z., KOVÁCS, M., FÜST, G. and ARASON, G.J., 2008; 2007. Low complement C4B gene copy number predicts short-term mortality after acute myocardial infarction. *International immunology*, 20(1), pp. 31-37.

BLOOMER, L.D.S., BOWN, M.J. and TOMASZEWSKI, M., 2012. Sexual dimorphism of abdominal aortic aneurysms: A striking example of "male disadvantage" in cardiovascular disease. *Atherosclerosis*, 225(1), pp. 22-28.

BOICHENKO, A., GOVORUKHINA, N., VAN DER ZEE, A.G.J. and BISCHOFF, R., 2013. Simultaneous serum desalting and total protein determination by macroporous reversed-phase chromatography. *Analytical and Bioanalytical Chemistry*, 405(10), pp. 3195-3203.

BORBELY, A., VAN DER VELDEN, J., PAPP, Z., BRONZWAER, J.G., EDES, I., STIENEN, G.J. and PAULUS, W.J., 2005. Cardiomyocyte stiffness in diastolic heart failure. *Circulation*, 111(6), pp. 774-781.

BOSCHETTI, E., LOMAS, L., CITTERIO, A. and RIGHETTI, P.G., 2007. Romancing the "hidden proteome", Anno Domini two zero zero seven. *Journal of Chromatography A*, 1153(1), pp. 277-290.

BOWLING, C., FELLER, M.A., MUJIB, M., ZHANG, Y., EKUNDAYO, O., SANDERS, P., FONAROW, G., ALLMAN, R.M. and AHMED, A., 2011. Chronic Kidney Disease and Incident Heart Failure in Older Adults: Predictor but Not a Risk Factor? Journal of the American Geriatrics Society; J.Am.Geriatr.Soc. 59, pp. S107-S107.

BOZZA, P.T. and VIOLA, J.P.B., 2010. Lipid droplets in inflammation and cancer. *Prostaglandins, Leukotrienes and Essential Fatty Acids (PLEFA)*, 82(4), pp. 243-250.

BRAUNWALD, E., 2013. Research Advances in Heart Failure: A Compendium. *Circulation research*, 113(6), pp. 633-645.

BRAUNWALD, E., 2008. Biomarkers in heart failure. *The New England journal of medicine*, 358(20), pp. 2148-2159.

BRILLA, C.G., FUNCK, R.C. and RUPP, H., 2000. Lisinopril-mediated regression of myocardial fibrosis in patients with hypertensive heart disease. *Circulation*, 102(12), pp. 1388-1393.

BURGOYNE, J.R., MONGUE-DIN, H., EATON, P. and SHAH, A.M., 2012. Redox Signaling in Cardiac Physiology and Pathology. Circulation research, 111(8), pp. 1091-1106.

BUSTELO, X.R., SAUZEAU, V. and BERENJENO, I.M., 2007. GTP-binding proteins of the Rho/Rac family: regulation, effectors and functions in vivo. BioEssays: news and reviews in molecular, cellular and developmental biology, 29(4), pp. 356-370.

CARTER, A.M., 2012. Complement activation: an emerging player in the pathogenesis of cardiovascular disease. *Scientifica*, 2012, pp. 402783.

CARVALHAES, L.S., GERVÁSIO, O.L., GUATIMOSIM, C., HELJASVAARA, R., SORMUNEN, R., PIHLAJANIEMI, T. and KITTEN, G.T., 2006. Collagen XVIII/endostatin is associated with the epithelial-mesenchymal transformation in the atrioventricular valves during cardiac development. Developmental dynamics: an official publication of the American Association of Anatomists, 235(1), pp. 132-142.

CHAKRAVARTHI, S., JESSOP, C.E., WILLER, M., STIRLING, C.J. and BULLEID, N.J., 2007. Intracellular catalysis of disulfide bond formation by the human sulfhydryl oxidase, QSOX1. *The Biochemical journal*, 404(3), pp. 403-411.

CHAMBERS, E. E., LAME, M. E., RAINVILLE, P. D., MURPHY, J. 2015 Practical applications of integrated microfluidics for peptide quantification. *Bioanalysis, 7*, 857–867.

CHATTERJEE, K. and MASSIE, B., 2007. Systolic and diastolic heart failure: differences and similarities. *Journal of cardiac failure*, 13(7), pp. 569-576.

ČIKOŠ, Š., GREGOR, P. and KOPPEL, J., 1999. Sequence and Tissue Distribution of a Novel G-Protein-Coupled Receptor Expressed Prominently in Human Placenta. Biochemical and biophysical research communications, 256(2), pp. 352-356.

COMMUNAL, C., SUMANDEA, M., TOMBE, P.D., NARULA, J., SOLARO, R.J. and HAJJAR, R.J., 2002. Functional Consequences of Caspase Activation in Cardiac Myocytes. Proceedings of the National Academy of Sciences of the United States of America, 99(9), pp. 6252-6256.

COOK, C., COLE, G., ASARIA, P., JABBOUR, R. and FRANCIS, D.P., 2014. The annual global economic burden of heart failure. *International journal of cardiology*, 171(3), pp. 368-376.

COTTER, R.J., 1999. The new time-of- flight mass spectrometry. Analytical Chemistry, 71(13), pp. 445A.

CRAIG, R., CORTENS, J.C., FENYO, D. and BEAVIS, R.C., 2006. Using annotated peptide mass spectrum libraries for protein identification. *Journal of proteome research*, 5(8), pp. 1843-1849.

CUMMINGS, L.J., SNYDER, M.A. and BRISACK, K., 2009. Protein chromatography on hydroxyapatite columns. *Methods in enzymology*, 463, pp. 387-404.

CUSPIDI, C., CIULLA, M. and ZANCHETTI, A., 2006. Hypertensive myocardial fibrosis. Nephrology, dialysis, transplantation: official publication of the European Dialysis and Transplant Association - *European Renal Association*, 21(1), pp. 20-23.

DALY, C.E., NG, L.L., HAKIMI, A., WILLINGALE, R. and JONES, D.J.L., 2014. Qualitative and quantitative characterization of plasma proteins when incorporating traveling wave ion mobility into a liquid chromatography- mass spectrometry workflow for biomarker discovery: use of product ion quantitation as an alternative data analysis tool for label free quantitation. *Analytical Chemistry*, 86(4), pp. 1972.

DAYON, L. and KUSSMANN, M., 2013. Proteomics of human plasma: A critical comparison of analytical workflows in terms of effort, throughput and outcome. *EuPA Open Proteomics*, 1(0), pp. 8-16.

DE ANDRADE, C.R., STOLF, B.S., DEBBAS, V., ROSA, D.S., KALIL, J., COELHO, V. and LAURINDO, F.R.M., 2011. Quiescin sulfhydryl oxidase (QSOX) is expressed in the human atheroma core: possible role in apoptosis. *In Vitro Cellular & Developmental Biology - Animal*, 47(10), pp. 716-727.

DE HOFFFMAN, E. AND STROOBANT V. 2007. Mass Spectrometry: Principles and Applications, 3rd Edition, Wiley.

DEMELLO, W.C. and FROHLICH, E.D., 2009. Renin angiotensin system and cardiovascular disease. New York: Humana Press.

DICK, S.A. and EPELMAN, S., 2016. Chronic Heart Failure and Inflammation: What Do We Really Know? *Circulation research*, 119(1), pp. 159-176.

DICKSON, V., VICTORIA, BUCK, V., HARLEAH and RIEGEL, V., BARBARA, 2013. Multiple Comorbid Conditions Challenge Heart Failure Self- Care by Decreasing Self-Efficacy. *Nursing research*, **62**(1), pp. 2-9.

DOEHNER, W., 2012. Diagnostic biomarkers in cardiovascular disease: the proteomics approach. *European heart journal*, 33(18), pp. 2249.

DOMANSKI, D., PERCY, A.J., YANG, J., CHAMBERS, A.G., HILL, J.S., FREUE, G.V.C. and BORCHERS, C.H., 2012. MRM-based multiplexed quantitation of 67 putative cardiovascular disease biomarkers in human plasma. *Proteomics*, 12(8), pp. 1222-1243.

DOMON, B., LANGE, V., PICOTTI, P. and AEBERSOLD, R., 2008. Selected reaction monitoring for quantitative proteomics: a tutorial. *Molecular Systems Biology*, 4(1), pp. 222-n/a.

DRAZNER, M.H., 2011. The progression of hypertensive heart disease. *Circulation*, 123(3), pp. 327-334.

DWEIK, R.A., 2002. Pulmonary hypertension and the search for the selective pulmonary vasodilator. *The Lancet*, 360(9337), pp. 886-887.

DWEIK, R.A., LASKOWSKI, D., OZKAN, M., FARVER, C. and ERZURUM, S.C., 2001. High Levels of Exhaled Nitric Oxide (NO) and NO Synthase III Expression in Lesional Smooth Muscle in Lymphangioleiomyomatosis. *American Journal of Respiratory Cell and Molecular Biology*, 24(4), pp. 414.

ECKSTEIN D., KORABATHINA R., 2016. Heart Failure Update: Diagnosis and Classification. FP Essentials, 442:11-7.

ELLIOTT, M.H., SMITH, D.S., PARKER, C.E. and BORCHERS, C., 2009. Current trends in quantitative proteomics. *Journal of mass spectrometry*, 44(12), pp. 1637-n/a.

ELLIS, J.M., BOWMAN, C.E. and WOLFGANG, M.J., 2015. Metabolic and Tissue-Specific Regulation of Acyl-CoA Metabolism: e0116587. *PLoS One*, 10(3).

ENGLARD, S. and SEIFTER, S., 1990. [22] Precipitation techniques. UNITED STATES: *Elsevier Science & Technology*, pp. 285-300.

FANG, Q., LI, Y., LI, N., JIANG, T., LIU, Q. and WANG, R., 2010. Pressor and tachycardic responses to intrathecal administration of neuropeptide FF in anesthetized rats. *Peptides*, 31(4), pp. 683-688.

FARRAH, T., DEUTSCH, E.W., OMENN, G.S., CAMPBELL, D.S., SUN, Z., BLETZ, J.A., MALLICK, P., KATZ, J.E., MALMSTRÖM, J., OSSOLA, R., WATTS, J.D., LIN, B., ZHANG, H., MORITZ, R.L. and AEBERSOLD, R., 2011. A high-confidence human plasma proteome reference set with estimated concentrations in PeptideAtlas. *Molecular & cellular proteomics*, 10(9), pp. M110.006353.

FARRAH, T., DEUTSCH, E.W., OMENN, G.S., SUN, Z., WATTS, J.D., YAMAMOTO, T., SHTEYNBERG, D., HARRIS, M.M. and MORITZ, R.L., 2014. State of the human proteome in 2013 as viewed through PeptideAtlas: comparing the kidney, urine, and plasma proteomes for the biology- and disease-driven Human Proteome Project. *Journal of proteome research*, 13(1), pp. 60-75.

FORGIONE, M.A., CAP, A., LIAO, R., MOLDOVAN, N.I., EBERHARDT, R.T., LIM, C.C., JONES, J., GOLDSCHMIDT-CLERMONT, P.J. and LOSCALZO, J., 2002. Heterozygous cellular glutathione peroxidase deficiency in the mouse: abnormalities in vascular and cardiac function and structure. *Circulation*, 106(9), pp. 1154-1158.

GAASCH WH, L.W., 2007. Assessment of left ventricular diastolic function and recognition of diastolic heart failure. *Circulation*, 116(6):591-593.

GAGGIN, H.K. and JANUZZI, J.L., 2013. Biomarkers and diagnostics in heart failure. Biochimica et Biophysica Acta (BBA) - *Molecular Basis of Disease*, 1832(12), pp. 2442-2450.

GAGNON, P., NG, P., ZHEN, J., ABERIN, C., HE, J. 2006. A Ceramic Hydroxyapatite–Based Purification Platform

GAO, G., XIE, A., HUANG, S., ZHOU, A., ZHANG, J., HERMAN, A.M., GHASSEMZADEH, S., JEONG, E., KASTURIRANGAN, S., RAICU, M., SOBIESKI, 2.,MICHAEL A., BHAT, G., TATOOLES, A., BENZ, J.,EDWARD J., KAMP, T.J. and DUDLEY, J.,SAMUEL C., 2011. Role of RBM25/LUC7L3 in abnormal cardiac sodium channel splicing regulation in human heart failure. *Circulation*, 124(10), pp. 1124-1131.

GAO, G., XIE, A., ZHANG, J., HERMAN, A.M., JEONG, E., GU, L., LIU, M., YANG, K., KAMP, T.J. and DUDLEY, S.C., 2013. Unfolded Protein Response Regulates Cardiac Sodium Current in Systolic Human Heart Failure. *Circulation: Arrhythmia and Electrophysiology*, 6(5), pp. 1018-1024.

GARG, U.C. and HASSID, A., 1989. Nitric oxide-generating vasodilators and 8-bromo-cyclic guanosine monophosphate inhibit mitogenesis and proliferation of cultured rat vascular smooth muscle cells. *The Journal of clinical investigation*, 83(5), pp. 1774-1777.

GASTON, B., DRAZEN, J.M., LOSCALZO, J. and STAMLER, J.S., 1994. The biology of nitrogen oxides in the airways. *American journal of respiratory and critical care medicine*, 149(2 Pt 1), pp. 538.

GERSZTEN, R.E., WANG, T.J., ACCURSO, F., LAGUNA, T.A., BERNARD, G.R., WARE, L.B., CAPRIOLI, R.M., KLEE, E.W., KLEE, G.G., KULLO, I., ROTH, F.P., SABATINE, M. and SRINIVAS, P., 2008. Challenges in translating plasma proteomics from bench to bedside: Update from the NHLBI Clinical Proteomics Programs. *American Journal of Physiology - Lung Cellular and Molecular Physiology*, 295(1), pp. L16-L22.

GIALLOURAKIS, C.C., 2013. Liver complications in patients with congestive heart failure. Gastroenterology & hepatology, 9(4), pp. 244.

GILES, K., PRINGLE, S.D., WORTHINGTON, K.R., LITTLE, D., WILDGOOSE, J.L. and BATEMAN, R.H., 2004. Applications of a travelling wave- based radio- frequency- only stacked ring ion guide. *Rapid Communications in Mass Spectrometry*, **18**(20), pp. 2401-2414.

GILES, K., WILLIAMS, J.P. and CAMPUZANO, I., 2011. Enhancements in travelling wave ion mobility resolution. *Rapid Communications in Mass Spectrometry*, 25(11), pp. 1559-1566.

GORDON, S.M., DENG, J., LU, L.J. and DAVIDSON, W.S., 2010. Proteomic characterization of human plasma high-density lipoprotein fractionated by gel filtration chromatography. *Journal of proteome research*, 9(10), pp. 5239.

GÖRG, A., WEISS, W. and DUNN, M.J., 2004. Current two-dimensional electrophoresis technology for proteomics. *Proteomics*, 4(12), pp. 3665-3685.

GREENING, D.W. and SIMPSON, R.J., 2011. Low-molecular weight plasma proteome analysis using centrifugal ultrafiltration. *Methods in molecular biology* (Clifton, N.J.), 728, pp. 109.

GU, H., REN, J.M., JIA, X., LEVY, T., RIKOVA, K., YANG, V., LEE, K.A., STOKES, M.P. and SILVA, J.C., 2016. Quantitative Profiling of Post-translational Modifications by Immunoaffinity Enrichment and LC-MS/MS in Cancer Serum without Immunodepletion. *Molecular & cellular proteomics*, 15(2), pp. 692.

GUTIERREZ, C. and BLANCHARD, D.G., 2004. Diastolic heart failure: challenges of diagnosis and treatment. *American Family Physician*, 69(11), pp. 2609.

HAAB, B.B., PAULOVICH, A.G., ANDERSON, N.L., CLARK, A.M., DOWNING, G.J., HERMJAKOB, H., LABAER, J. and UHLEN, M., 2006. A reagent resource to identify proteins and peptides of interest for the cancer community: a workshop report. *Molecular & cellular proteomics*, 5(10), pp. 1996-2007.

HAKIMI, A., AULUCK, J., JONES, G.D.D., NG, L.L. and JONES, D.J.L., 2014. Assessment of reproducibility in depletion and enrichment workflows for plasma proteomics using label-free quantitative data-independent LC-MS. *Proteomics*, 14(1), pp. 4-13.

HAMDANI, N. and PAULUS, W.J., 2013. Myocardial titin and collagen in cardiac diastolic dysfunction: partners in crime. *Circulation*, 128(1), pp. 5-8.

HARTOG, J.W., VOORS, A.A., BAKKER, S.J., SMIT, A.J. and VAN VELDHUISEN, D.J., 2007. Advanced glycation end products (AGEs) and heart failure: pathophysiology and clinical implications. *European journal of heart failure*, 9(12), pp. 1146-1155.

HEINK, A., DAVIDSON, W.S., SWERTFEGER, D.K., LU, L.J. and SHAH, A.S., 2015. A comparison of methods to enhance protein detection of lipoproteins by mass spectrometry. *Journal of proteome research*, 14(7), pp. 2943-2950.

HELM, D., VISSERS, J.P.C., HUGHES, C.J., HAHNE, H., RUPRECHT, B., PACHL, F., GRZYB, A., RICHARDSON, K., WILDGOOSE, J., MAIER, S.K., MARX, H., WILHELM, M., BECHER, I., LEMEER, S., BANTSCHEFF, M., LANGRIDGE, J.I. and KUSTER, B., 2014. Ion mobility tandem mass spectrometry enhances performance of bottom- up proteomics. *Molecular and Cellular Proteomics*, 13(12), pp. 3709-3715.

HENZEL, W.J., BILLECI, T.M., STULTS, J.T., WONG, S.C., GRIMLEY, C. and WATANABE, C., 1993. Identifying Proteins from Two-Dimensional Gels by Molecular Mass Searching of Peptide Fragments in Protein Sequence Databases. Proceedings of the National Academy of Sciences of the United States of America, 90(11), pp. 5011-5015.

HERRING, S. and OKEN, E., 2011. Obesity and Diabetes in Mothers and Their Children: Can We Stop the Intergenerational Cycle? Current Diabetes Reports, 11(1), pp. 20-27.

HO, C.S., LAM, C.W.K., CHAN, M.H.M., CHEUNG, R.C.K., LAW, L.K., LIT, L.C.W., NG, K.F., SUEN, M.W.M. and TAI, H.L., 2003. Electrospray ionisation mass spectrometry: principles and clinical applications. *The Clinical biochemist. Reviews / Australian Association of Clinical Biochemists*, 24(1), pp. 3-12.

HOFFMANN, U., ESPETER, F., WEIß, C., AHMAD-NEJAD, P., LANG, S., BRUECKMANN, M., AKIN, I., NEUMAIER, M., BORGGREFE, M. and BEHNES, M., 2015. Ischemic biomarker heart-type fatty acid binding protein (hFABP) in acute heart failure - diagnostic and prognostic insights compared to NT-proBNP and troponin I. *BMC cardiovascular disorders*, 15, pp. 50.

HOUGHTON, A.R., 2013. Making sense of echocardiography: a hands-on guide. Second edition. edn. Boca Raton: CRC Press.

HU, S., LOO, J.A. and WONG, D.T., 2006. Human body fluid proteome analysis. Proteomics, 6(23), pp. 6326-6353.

IWAI, N., MANNAMI, T., TOMOIKE, H., ONO, K. and IWANAGA, Y., 2003. An Acyl-CoA Synthetase Gene Family in Chromosome 16p12 May Contribute to Multiple Risk Factors. Hypertension: *Journal of the American Heart Association*, 41(5), pp. 1041-1046.

IYER, S., VISSE, R., NAGASE, H. and ACHARYA, K.R., 2006. Crystal Structure of an Active Form of Human MMP-1. *Journal of Molecular Biology*, 362(1), pp. 78-88.

JELLIS, C., SACRE, J., HALUSKA, B., JENKINS, C., MARWICK, T. and MARTIN, J., 2011. Myocardial Dysfunction and Metabolic Derangement in Type 2 Diabetes: Relationship with Procollagen Biomarkers of Myocardial Fibrosis. *Heart, Lung and Circulation*, 20, pp. S183-S183.

JESSUP, M., ABRAHAM, W.T., CASEY, D.E., FELDMAN, A.M., FRANCIS, G.S., GANIATS, T.G., KONSTAM, M.A., MANCINI, D.M., RAHKO, P.S., SILVER, M.A., STEVENSON, L.W. and YANCY, C.W., 2009. 2009 Focused Update: ACCF/AHA Guidelines for the Diagnosis and Management of Heart Failure in Adults: A Report of the American College of Cardiology Foundation/American Heart Association Task Force on Practice Guidelines Developed in Collaboration With the International Society for Heart and Lung Transplantation. *Journal of the American College of Cardiology*, 53(15), pp. 1343.

JHAMANDAS, J.H. and GONCHARUK, V., 2013. Role of neuropeptide FF in central cardiovascular and neuroendocrine regulation. *Frontiers in endocrinology*, 4, pp. 8.

JING-JING CAI JUAN WEN WEI-HONG JIANG JIAN LIN YUAN HONG YUAN-SHAN ZHU, 2016. Androgen actions on endothelium functions and cardiovascular diseases. 13(2),

KARBEK, B., ÖZBEK, M., BOZKURT, N.C., GINIS, Z., GÜNGÜNES, A., ÜNSAL, I.Ö., CAKAL, E. and DELIBASI, T., 2011. Heart-type fatty acid binding protein (H-FABP): relationship with arterial intima-media thickness and role as diagnostic marker for atherosclerosis in patients with ımpaired glucose metabolism. *Cardiovascular diabetology*, 10(1), pp. 37-37.

KATZ, A.M. and ZILE, M.R., 2006. New molecular mechanism in diastolic heart failure. *Circulation*, 113(16), pp. 1922-1925.

KAY, R., BARTON, C., RATCLIFFE, L., MATHAROO-BALL, B., BROWN, P., ROBERTS, J., TEALE, P. and CREASER, C., 2008. Enrichment of low molecular weight serum proteins using acetonitrile precipitation for mass spectrometry based proteomic analysis. *Rapid Communications in Mass Spectrometry*, 22(20), pp. 3255-3260.

KESHISHIAN, H., BURGESS, M.W., GILLETTE, M.A., MERTINS, P., CLAUSER, K.R., MANI, D.R., KUHN, E.W., FARRELL, L.A., GERSZTEN, R.E. and CARR, S.A., 2015. Multiplexed, Quantitative Workflow for Sensitive Biomarker Discovery in Plasma Yields Novel Candidates for Early Myocardial Injury. *Molecular & cellular proteomics*, 14(9), pp. 2375-2393.

KESHISHIAN, H., ADDONA, T., BURGESS, M., KUHN, E. and CARR, S.A., 2007. Quantitative, Multiplexed Assays for Low Abundance Proteins in Plasma by Targeted Mass Spectrometry and Stable Isotope Dilution. *Molecular & Cellular Proteomics*, 6(12), pp. 2212-2229.

KIM, N. and KANG, P.M., 2010. Apoptosis in Cardiovascular Diseases: Mechanism and Clinical Implications. *Korean Circulation Journal*, 40(7), pp. 299-305.

KING, M.K., COKER, M.L., GOLDBERG, A., MCELMURRAY, 3.,JAMES H., GUNASINGHE, H.R., MUKHERJEE, R., ZILE, M.R., O'NEILL, T.P. and SPINALE, F.G., 2003. Selective matrix metalloproteinase inhibition with developing heart failure: effects on left ventricular function and structure. *Circulation research*, 92(2), pp. 177-185.

KITO, K. and ITO, T., 2008. Mass Spectrometry-Based Approaches toward Absolute Quantitative Proteomics. *Current Genomics*, 9(4), pp. 263-274.

KOSCHINSKY, T., HE, C.J., MITSUHASHI, T., BUCALA, R., LIU, C., BUENTING, C., HEITMANN, K. and VLASSARA, H., 1997. Orally absorbed reactive glycation products (glycotoxins): an environmental risk factor in diabetic nephropathy. Proceedings of the National Academy of Sciences of the United States of America, 94(12), pp. 6474-6479.

KUNISHIGE, M., KIJIMA, Y., SAKAI, T., AKUTAGAWA, O., MATSUO, A., NISHIBE, A., NAKAGAWA, Y. and HATA, T., 2007. Transient enhancement of oxidant stress and collagen turnover in patients with acute worsening of congestive heart failure. *Circulation journal*, 71(12), pp. 1893.

LADOMERY, M. and DELLAIRE, G., 2002. Multifunctional zinc finger proteins in development and disease. *Annals of Human Genetics*, 66(5-6), pp. 331-342.

LAKATTA, E.G. and LEVY, D., 2003. Arterial and cardiac aging: major shareholders in cardiovascular disease enterprises: Part II: the aging heart in health: links to heart disease. *Circulation*, 107(2), pp. 346-354.

LAPENNA, D., DE GIOIA, S., CIOFANI, G., MEZZETTI, A., UCCHINO, S., CALAFIORE, A.M., NAPOLITANO, A.M., DI ILIO, C. and CUCCURULLO, F., 1998. Glutathione-related antioxidant defenses in human atherosclerotic plaques. *Circulation*, 97(19), pp. 1930-1934.

LEE, W., LOO, C., ZAVGORODNIY, A.V. and ROHANIZADEH, R., 2013. High protein adsorptive capacity of amino acid-functionalized hydroxyapatite. *Journal of Biomedical Materials Research Part A*, 101A (3), pp. 873-883.

LERMYTE, F., KONIJNENBERG, A., WILLIAMS, J., BROWN, J., VALKENBORG, D. and SOBOTT, F., 2014. ETD Allows for Native Surface Mapping of a 150 kDa Noncovalent Complex on a Commercial Q- TWIMS- TOF Instrument. *Journal of the American Society for Mass Spectrometry; The official journal of The American Society for Mass Spectrometry*, 25(3), pp. 343-350.

LERMYTE, F., VERSCHUEREN, T., BROWN, J.M., WILLIAMS, J.P., VALKENBORG, D. and SOBOTT, F., 2015. Characterization of top- down ETD in a travelling- wave ion guide. *Methods*, 89, pp. 22-29.

LEVINE, B., KALMAN, J., MAYER, L., FILLIT, H.M. and PACKER, M., 1990. Elevated circulating levels of tumor necrosis factor in severe chronic heart failure. *The New England journal of medicine*, 323(4), pp. 236-241.

LI, S., PENG, M., LI, H., LIU, B., WANG, C., WU, J., LI, Y. and ZENG, R., 2009. Sys-BodyFluid: a systematical database for human body fluid proteome research. *Nucleic acids research*, 37(suppl_1), pp. D907-D912.

LIM, D.S., LUTUCUTA, S., BACHIREDDY, P., YOUKER, K., EVANS, A., ENTMAN, M., ROBERTS, R. and MARIAN, A.J., 2001. Angiotensin II blockade reverses myocardial fibrosis in a transgenic mouse model of human hypertrophic cardiomyopathy. *Circulation*, 103(6), pp. 789-791.

LINKE, T., DORAISWAMY, S. and HARRISON, E.H., 2007. Rat plasma proteomics: effects of abundant protein depletion on proteomic analysis. *Journal of chromatography.B, Analytical technologies in the biomedical and life sciences*, 849(1-2), pp. 273-281.

LIU, P.Y., DEATH, A.K. and HANDELSMAN, D.J., 2003. Androgens and Cardiovascular Disease. *Endocrine reviews*, 24(3), pp. 313-340.

LIUMBRUNO, G., D'ALESSANDRO, A., GRAZZINI, G. and ZOLLA, L., 2010. Blood-related proteomics. *Journal of Proteomics*, 73(3), pp. 483-507.

LY, L. and WASINGER, V.C., 2011. Protein and peptide fractionation, enrichment and depletion: Tools for the complex proteome. *Proteomics*, 11(4), pp. 513-534.

LY, T., AHMAD, Y., SHLIEN, A., SOROKA, D., MILLS, A., EMANUELE, M.J., STRATTON, M.R. and LAMOND, A.I., 2014. A proteomic chronology of gene expression through the cell cycle in human myeloid leukemia cells. *eLife*, 3, pp. e01630.

MACHADO, R.F., LONDHE NERKAR, M., DWEIK, R.A., HAMMEL, J., JANOCHA, A., PYLE, J., LASKOWSKI, D., JENNINGS, C., ARROLIGA, A.C. and ERZURUM, S.C., 2004. Nitric oxide and pulmonary arterial pressures in pulmonary hypertension. *Free Radical Biology and Medicine*, 37(7), pp. 1010-1017.

MACLEAN, B., TOMAZELA, D.M., SHULMAN, N., CHAMBERS, M., FINNEY, G.L., FREWEN, B., KERN, R., TABB, D.L., LIEBLER, D.C. and MACCOSS, M.J., 2010. Skyline: an open source document editor for creating and analyzing targeted proteomics experiments. *Bioinformatics*, 26(7), pp. 966-968.

MAHN, A. and ISMAIL, M., 2011. Depletion of highly abundant proteins in blood plasma by ammonium sulfate precipitation for 2D-PAGE analysis. *Journal of Chromatography B*, 879(30), pp. 3645.

MAKAROV, A., 2000. Electrostatic Axially Harmonic Orbital Trapping: A High. Performance Technique of Mass Analysis.(Statistical Data Included). *Analytical Chemistry*, 72(6), pp. 1156.

MANN, D.L., 2005. Targeted anticytokine therapy and the failing heart. *The American Journal of Cardiology*, 95(11A), pp. 9C-16.

MARKIEWSKI, M.M. and LAMBRIS, J.D., 2007. The Role of Complement in Inflammatory Diseases From Behind the Scenes into the Spotlight. *The American Journal of Pathology*, 171(3), pp. 715-727.

MARTIN, M., GOTTSÄTER, A., NILSSON, P.M., MOLLNES, T.E., LINDBLAD, B., BLOM, A.M., 2009. Complement activation and plasma levels of C4b-binding protein in critical limb ischemia patients. *Journal of Vascular Surgery*, 50(1), pp. 100-106.

MCLAFFERTY, F. W. and BRYCE, T. A. (1967). Metastable-ion characteristics: characterization of isomeric molecules. *Chemical Communications*. 1215-1217.

MEBAZAA, A., VANPOUCKE, G., THOMAS, G., VERLEYSEN, K., COHEN-SOLAL, A., VANDERHEYDEN, M., BARTUNEK, J., MUELLER, C., LAUNAY, J., VAN LANDUYT, N., D'HONDT, F., VERSCHUERE, E., VANHAUTE, C., TUYTTEN, R., VANNESTE, L., DE CREMER, K., WUYTS, J., DAVIES, H., MOERMAN, P., LOGEART, D., COLLET, C., LORTAT-JACOB, B., TAVARES, M., LAROY, W., JANUZZI, J.L., SAMUEL, J. and KAS, K., 2012. Unbiased plasma proteomics for novel diagnostic biomarkers in cardiovascular disease: identification of quiescin Q6 as a candidate biomarker of acutely decompensated heart failure. *European heart journal*, 33(18), pp. 2317.

MILLIONI, R., TOLIN, S., PURICELLI, L., SBRIGNADELLO, S., FADINI, G.P., TESSARI, P. and ARRIGONI, G., 2011. High abundance proteins depletion vs low abundance proteins enrichment: comparison of methods to reduce the plasma proteome complexity. *PloS one*, 6(5), pp. e19603.

MORAN, A.E., FOROUZANFAR, M.H., ROTH, G.A., MENSAH, G.A., EZZATI, M., FLAXMAN, A., MURRAY, C.J.L. and NAGHAVI, M., 2014. The Global Burden of Ischemic Heart Disease in 1990 and 2010: The Global Burden of Disease 2010 Study. *Circulation*, 129(14), pp. 1493-1501.

MORAN, A.E., FOROUZANFAR, M.H., ROTH, G.A., MENSAH, G.A., EZZATI, M., MURRAY, C.J.L. and NAGHAVI, M., 2014. Temporal Trends in Ischemic Heart Disease Mortality in 21 World Regions, 1980 to 2010: The Global Burden of Disease 2010 Study. *Circulation*, 129(14), pp. 1483-1492.

MORENO-GONZALO, O., VILLARROYA-BELTRI, C. and SÁNCHEZ-MADRID, F., 2014. Post-translational modifications of exosomal proteins. *Frontiers in immunology*, 5, pp. 383.

MROZINSKI P., ZOLOTARJOVA, N., and CHEN, H. Application note: Human serum and plasma protein depletion–novelhigh-capacity affinity column for the removal of the "Top 14" abundant proteins. Available online:https://www.agilent.com/cs/library/applications/5989-7839EN.pdf (accessed online 08 November 2016).

MUKHERJEE, D. and SEN, S., 1990. Collagen phenotypes during development and regression of myocardial hypertrophy in spontaneously hypertensive rats. *Circulation research*, 67(6), pp. 1474-1480.

MYAGMAR, B., UMIKAWA, M., ASATO, T., TAIRA, K., OSHIRO, M., HINO, A., TAKEI, K., UEZATO, H. and KARIYA, K., 2005. PARG1, a protein-tyrosine phosphatase-associated RhoGAP, as a putative Rap2 effector. *Biochemical and biophysical research communications*, 329(3), pp. 1046-1052.

NAGANO, F., SASAKI, T., FUKUI, K., ASAKURA, T., IMAZUMI, K. and TAKAI, Y., 1998. Molecular Cloning and Characterization of the Noncatalytic Subunit of the Rab3 Subfamily-specific GTPase-activating Protein. *Journal of Biological Chemistry*, 273(38), pp. 24781-24785.

NEUMANN, F.J., OTT, I., GAWAZ, M., RICHARDT, G., HOLZAPFEL, H., JOCHUM, M. and SCHÖMIG, A., 1995. Cardiac release of cytokines and inflammatory responses in acute myocardial infarction. *Circulation*, 92(4), pp. 748-755.

NICHOLS, R., BASS, C., DEMERS, L., LARSEN, B., LI, E., BLEWETT, N., CONVERSO-BARAN, K., RUSSELL, M.W. and WESTFALL, M.V., 2012. Structure-activity studies of RFamide-related peptide-1 identify a functional receptor antagonist and novel cardiac myocyte signaling pathway involved in contractile performance. *Journal of medicinal chemistry*, 55(17), pp. 7736-7745.

NUNES, K.P., RIGSBY, C.S. and WEBB, R.C., 2010. RhoA/Rho-kinase and vascular diseases: what is the link? *Cellular and molecular life sciences*, 67(22), pp. 3823-3836.

OKTAY, A.A., RICH, J.D. and SHAH, S.J., 2013. The Emerging Epidemic of Heart Failure with Preserved Ejection Fraction. *Current heart failure reports*.

OLSEN, J.V., ONG, S. and MANN, M., 2004. Trypsin Cleaves Exclusively C-terminal to Arginine and Lysine Residues. *Molecular & Cellular Proteomics*, 3(6), pp. 608-614.

OMENN, G.S., EDDES, J.S., KAPP, E.A., MORITZ, R.L., CHAN, D.W., RAI, A.J., ADMON, A., AEBERSOLD, R., ENG, J., HANCOCK, W.S., HEFTA, S.A., STATES, D.J., MEYER, H., PAIK, Y., YOO, J., PING, P., POUNDS, J., ADKINS, J., QIAN, X., WANG, R., WASINGER, V., WU, C.Y., ADAMSKI, M., ZHAO, X., ZENG, R., ARCHAKOV, A., TSUGITA, A., BEER, I., PANDEY, A., PISANO, M., ANDREWS, P., TAMMEN, H., SPEICHER, D.W., BLACKWELL, T.W., HANASH, S.M., MENON, R., HERMJAKOB, H., APWEILER, R., HAAB, B.B. and SIMPSON, R.J., 2005. Overview of the HUPO Plasma Proteome Project: results from the pilot phase with 35 collaborating laboratories and multiple

analytical groups, generating a core dataset of 3020 proteins and a publicly-available database. *Proteomics*, 5(13), pp. 3226-3245.

OPIE, L.H., 2004. Cellular basis for therapeutic choices in heart failure. *Circulation*, 110(17), pp. 2559-2561.

PAUL, W., and STEINWEDEL H., 1953. Ein neues Massenspektrometer ohne Magnetfeld. *Zeitschrift Naturforschung Teil A*. 8, 448.

PAVLAKIS, E., MAKRYGIANNIS, A.K., CHIOTAKI, R. and CHALEPAKIS, G., 2008. Differential localization profile of Fras1/Frem proteins in epithelial basement membranes of newborn and adult mice. *Histochemistry and cell biology*, 130(4), pp. 785-793.

PEPE, M.S. and THOMPSON, M.L., 2000. Combining diagnostic test results to increase accuracy. *Biostatistics*, **1**(2), pp. 123-140.

.

PFEFFER, M.D., PH.D, MARC A., 1995. LEFT VENTRICULAR REMODELING AFTER ACUTE MYOCARDIAL INFARCTION. *Annual Review of Medicine*, 46(1), pp. 455-466.

PFÜTZNER, A., SCHÖNDORF, T., HANEFELD, M. and FORST, T., 2010. High-sensitivity C-reactive protein predicts cardiovascular risk in diabetic and nondiabetic patients: effects of insulin-sensitizing treatment with pioglitazone. *Journal of diabetes science and technology*, 4(3), pp. 706-716.

PIEPER, R., SU, Q., GATLIN, C.L., HUANG, S., ANDERSON, N.L. and STEINER, S., 2003. Multi-component immunoaffinity subtraction chromatography: an innovative step towards a comprehensive survey of the human plasma proteome. *Proteomics*, 3(4), pp. 422-432.

PLEGER, T., S., BRINKS, J., H., RITTERHOFF, A., J., RAAKE, A., P., KOCH, A., W., KATUS, A., H. and MOST, A., P., 2013. Heart Failure Gene Therapy: The Path to Clinical Practice. *Circulation research*, **113**(6), pp. 792-809.

PREIS, S.R., HWANG, S., COADY, S., PENCINA, M.J., D'AGOSTINO, S., RALPH B., SAVAGE, P.J., LEVY, D. and FOX, C.S., 2009. Trends in all-cause and cardiovascular disease mortality among women and men with and without diabetes mellitus in the Framingham Heart Study, 1950 to 2005. *Circulation*, 119(13), pp. 1728.

PRINGLE, S.D., GILES, K., WILDGOOSE, J.L., WILLIAMS, J.P., SLADE, S.E., THALASSINOS, K., BATEMAN, R.H., BOWERS, M.T. and SCRIVENS, J.H., 2007. An investigation of the mobility separation of some peptide and protein ions using a new hybrid quadrupole/travelling wave IMS/oa-ToF instrument. *International Journal of Mass Spectrometry*, 261(1), pp. 1-12.

RADULOVIC, D., JELVEH, S., RYU, S., HAMILTON, T.G., FOSS, E., MAO, Y. and EMILI, A., 2004. Informatics Platform for Global Proteomic Profiling and Biomarker

Discovery Using Liquid Chromatography-Tandem Mass Spectrometry. *Molecular & Cellular Proteomics*, 3(10), pp. 984-997.

MARTOS, R., BAUGH, J., LEDWIDGE, M., O'LOUGHLIN, C., CONLON, C., PATLE, A., DONNELLY, S.C. and MCDONALD, K., 2007. Diastolic heart failure: evidence of increased myocardial collagen turnover linked to diastolic dysfunction. *Circulation*, 115(7), pp. 888-895.

REDDY, J.V., BURGUETE, A.S., SRIDEVI, K., GANLEY, I.G., NOTTINGHAM, R.M. and PFEFFER, S.R., 2006. A functional role for the GCC185 golgin in mannose 6-phosphate receptor recycling. *Molecular biology of the cell*, 17(10), pp. 4353-4363.

RIGHETTI, P.G. and BOSCHETTI, E., 2007. Sherlock Holmes and the proteome − a detective story. *FEBS Journal*, 274(4), pp. 897-905.

RIGHETTI, P.G., BOSCHETTI, E., ZANELLA, A., FASOLI, E. and CITTERIO, A., 2010. Plucking, pillaging and plundering proteomes with combinatorial peptide ligand libraries. *Journal of Chromatography A*, 1217(6), pp. 893-900.

ROGER, V.L., FOX, C.S., FULLERTON, H.J., GILLESPIE, C., HAILPERN, S.M., HEIT, J.A., HOWARD, V.J., KISSELA, B.M., KITTNER, S.J., LACKLAND, D.T., LICHTMAN, J.H., GO, A.S., LISABETH, L.D., MAKUC, D.M., MARCUS, G.M., MARELLI, A., MATCHAR, D.B., MOY, C.S., MOZAFFARIAN, D., MUSSOLINO, M.E., NICHOL, G., PAYNTER, N.P., LLOYD-JONES, D.M., SOLIMAN, E.Z., SORLIE, P.D., SOTOODEHNIA, N., TURAN, T.N., VIRANI, S.S., WONG, N.D., WOO, D., TURNER, M.B., BENJAMIN, E.J., BERRY, J.D., BORDEN, W.B., BRAVATA, D.M., DAI, S., FORD, E.S., 2012. Heart disease and stroke statistics--2012 update: a report from the American Heart Association. *Circulation*, 125(1), pp. e2.

ROSENBERGER, G., CHING, C.K., GUO, T., HANNES L. RÖST, KOUVONEN, P., COLLINS, B.C., HEUSEL, M., LIU, Y., CARON, E., VICHALKOVSKI, A., FAINI, M., SCHUBERT, O.T., FARIDI, P., ALEXANDER EBHARDT, H., MATONDO, M., LAM, H., BADER, S.L., CAMPBELL, D.S., DEUTSCH, E.W., MORITZ, R.L., TATE, S. and AEBERSOLD, R., 2014. A repository of assays to quantify 10,000 human proteins by SWATH- MS. *Scientific Data*, 1.

ROSENBLAT, M. and AVIRAM, M., 1998. Macrophage glutathione content and glutathione peroxidase activity are inversely related to cell-mediated oxidation of LDL: in vitro and in vivo studies. *Free radical biology & medicine*, 24(2), pp. 305.

ROSSING, K., BOSSELMANN, H.S., GUSTAFSSON, F., ZHANG, Z., GU, Y., KUZNETSOVA, T., NKUIPOU-KENFACK, E., MISCHAK, H., STAESSEN, J.A., KOECK, T. and SCHOU, M., 2016. Urinary Proteomics Pilot Study for Biomarker Discovery and Diagnosis in Heart Failure with Reduced Ejection Fraction. *PLoS One*, 11(6), pp. e0157167.

ROUET-BENZINEB, P., BUHLER, J.M., DREYFUS, P., DELCOURT, A., DORENT, R., PERENNEC, J., CROZATIER, B., HARF, A. and LAFUMA, C., 1999. Altered balance between matrix gelatinases (MMP-2 and MMP-9) and their tissue inhibitors in human dilated cardiomyopathy: potential role of MMP-9 in myosin-heavy chain degradation. *European journal of heart failure*, 1(4), pp. 337.

SAMSTAD, E.O., NIYONZIMA, N., NYMO, S., AUNE, M.H., RYAN, L., BAKKE, S.S., LAPPEGÅRD, K.T., BREKKE, O., LAMBRIS, J.D., DAMÅS, J.K., LATZ, E., MOLLNES, T.E. and ESPEVIK, T., 2014. Cholesterol crystals induce complement-dependent inflammasome activation and cytokine release. *Journal of immunology* (Baltimore, Md.: 1950), 192(6), pp. 2837-2845.

SANDERSON, J.E., 2007; 2005. Heart failure with a normal ejection fraction. Heart (British Cardiac Society), 93(2), pp. 155-158.

SANDERSON, J.E., 2014. HFNEF, HFpEF, HF-PEF, or DHF. JACC: *Heart Failure*, 2(1), pp. 93-94.

SANDERSON, J.E., 2002. New treatments for myocardial fibrosis. Cardiovascular drugs and therapy / sponsored by the *International Society of Cardiovascular Pharmacotherapy*, 16(3), pp. 181-182.

SCHIESS, R., WOLLSCHEID, B. and AEBERSOLD, R., 2009. Targeted proteomic strategy for clinical biomarker discovery. *Molecular Oncology*, 3(1), pp. 33-44.

SCHULZ, R. and HEUSCH, G., 2009. Tumor necrosis factor-alpha and its receptors 1 and 2: Yin and Yang in myocardial infarction? *Circulation*, 119(10), pp. 1355-1357.

SENTE, T., VAN BERENDONCKS, A.M., HOYMANS, V.Y. and VRINTS, C.J., 2016. Adiponectin resistance in skeletal muscle: pathophysiological implications in chronic heart failure: Adiponectin resistance in skeletal muscle. *Journal of Cachexia, Sarcopenia and Muscle*, 7(3), pp. 261-274.

SEO, J., KIM, J. and KIM, M., 2001. Cloning of androgen-inducible gene 1 (AIG1) from human dermal papilla cells. *Molecules and cells*, 11(1), pp. 35.

SERTIÉ, A.L., MOREIRA, M., ZATZ, E.S., PASSOS-BUENO, J., QUIMBY, M., MURRAY, S.E. and ANTONARAKIS, M.R., 1996. A gene which causes severe ocular alterations and occipital encephalocele (Knobloch syndrome) is mapped to 21q22.3. *Human molecular genetics*, 5(6), pp. 843-847.

SERTIÉ, A.L., SOSSI, V., CAMARGO, A.A., ZATZ, M., BRAHE, C. and PASSOS-BUENO, M.R., 2000. Collagen XVIII, containing an endogenous inhibitor of angiogenesis and tumor growth, plays a critical role in the maintenance of retinal structure and in neural tube closure (Knobloch syndrome). *Human molecular genetics*, 9(13), pp. 2051-2058.

SHARPE, N., 1992. Ventricular remodeling following myocardial infarction. *The American Journal of Cardiology*, 70(10), pp. 20-26.

SHEPARD, S.R., BRICKMAN-STONE, C., SCHRIMSHER, J.L. and KOCH, G., 2000. Discoloration of ceramic hydroxyapatite used for protein chromatography. *Journal of Chromatography A*, 891(1), pp. 93-98.

SHI, T., ZHOU, J., GRITSENKO, M.A., HOSSAIN, M., CAMP, D.G., SMITH, R.D. and QIAN, W., 2012. IgY14 and SuperMix immunoaffinity separations coupled with liquid chromatography–mass spectrometry for human plasma proteomics biomarker discovery. *Methods*, 56(2), pp. 246-253.

SHIOMI, T., TSUTSUI, H., MATSUSAKA, H., MURAKAMI, K., HAYASHIDANI, S., IKEUCHI, M., WEN, J., KUBOTA, T., UTSUMI, H. and TAKESHITA, A., 2004. Overexpression of glutathione peroxidase prevents left ventricular remodeling and failure after myocardial infarction in mice. *Circulation*, 109(4), pp. 544-549.

SHIPMAN, K.L., ROBINSON, P.J., KING, B.R., SMITH, R. and NICHOLSON, R.C., 2006. Identification of a family of DNA-binding proteins with homology to RNA splicing factors. *Biochemistry and Cell Biology*, 84(1), pp. 9-19.

SINGH, V.P., BALI, A., SINGH, N. and JAGGI, A.S., 2014. Advanced glycation end products and diabetic complications. *The Korean journal of physiology & pharmacology*. 18(1), pp. 1.

SILVA, J.C., DENNY, R., DORSCHEL, C.A., GORENSTEIN, M., KASS, I.J., LI, G., MCKENNA, T., NOLD, M.J., RICHARDSON, K., YOUNG, P. and GEROMANOS, S., 2005. Quantitative proteomic analysis by accurate mass retention time pairs. *Analytical Chemistry*, 77(7), pp. 2187-2200.

SILVA, J.C., GORENSTEIN, M.V., LI, G., VISSERS, J.P.C. and GEROMANOS, S.J., 2006. Absolute quantification of proteins by LCMSE: a virtue of parallel MS acquisition. *Molecular & cellular proteomics* : MCP, 5(1), pp. 144.

SMITH, M.P.W., WOOD, S.L., ZOUGMAN, A., HO, J.T.C., PENG, J., JACKSON, D., CAIRNS, D.A., LEWINGTON, A.J.P., SELBY, P.J. and BANKS, R.E., 2011. A systematic analysis of the effects of increasing degrees of serum immunodepletion in terms of depth of coverage and other key aspects in top-down and bottom-up proteomic analyses. *Proteomics*, 11(11), pp. 2222-2235.

SPEIDL, W.S., KASTL, S.P., HUBER, K. and WOJTA, J., 2011. Complement in atherosclerosis: friend or foe? *Journal of Thrombosis and Haemostasis*, 9(3), pp. 428-440.

SPINALE, F.G., 2002. Matrix metalloproteinases: regulation and dysregulation in the failing heart. *Circulation research*, 90(5), pp. 520-530.

STATES D.S., GILBERT, S.O., THOMAS, W.B., FERMIN, D., ENG, J., DAVID, W.S. and SAMIR, M.H., 2006. Challenges in deriving high- confidence protein identifications from data gathered by a HUPO plasma proteome collaborative study. *Nature biotechnology*, 24(3), pp. 333.

SUCH-SANMARTÍN, G., VENTURA-ESPEJO, E. and JENSEN, O.N., 2014. Depletion of abundant plasma proteins by poly(N-isopropylacrylamide-acrylic acid) hydrogel particles. *Analytical Chemistry*, 86(3), pp. 1543.

SUN, M., CHEN, M., DAWOOD, F., ZURAWSKA, U., LI, J.Y., PARKER, T., KASSIRI, Z., KIRSHENBAUM, L.A., ARNOLD, M., KHOKHA, R. and LIU, P.P., 2007. Tumor necrosis factor-alpha mediates cardiac remodeling and ventricular dysfunction after pressure overload state. *Circulation*, 115(11), pp. 1398-1407.

SUN, M., DAWOOD, F., WEN, W.H., CHEN, M., DIXON, I., KIRSHENBAUM, L.A. and LIU, P.P., 2004. Excessive tumor necrosis factor activation after infarction contributes to susceptibility of myocardial rupture and left ventricular dysfunction. *Circulation*, 110(20), pp. 3221-3228.

SURINOVA, S., SCHIESS, R., HÜTTENHAIN, R., CERCIELLO, F., WOLLSCHEID, B. and AEBERSOLD, R., 2011. On the development of plasma protein biomarkers. *Journal of proteome research*, 10(1), pp. 5.

SYKA, J. E., COON, J. J.,SCHROEDER, M. J.,SHABANOWITZ, J. and HUNT. D. F. (2004). Peptide and protein sequence analysis by electron transfer dissociation mass spectrometry. *Proc Natl Acad Sci U S A*. 101, 9528-33

TAKEMOTO, M., SUN, J., HIROKI, J., SHIMOKAWA, H. and LIAO, J.K., 2002. Rho-kinase mediates hypoxia-induced downregulation of endothelial nitric oxide synthase. *Circulation*, 106(1), pp. 57-62.

TONELLI, A.R., HASERODT, S., AYTEKIN, M. and DWEIK, R.A., 2013. Nitric oxide deficiency in pulmonary hypertension: Pathobiology and implications for therapy. *Pulmonary Circulation*, 3(1), pp. 20-30.

TOWSEND, N., WILLIAMS, J., BHATNAGAR, P., KREMLIN, W., and RAYNER, M., 2014. Cardiovascular disease statistics. *British Heart Foundation*. (19), pp. 6-15

TRIPOSKIADIS, F., KARAYANNIS, G., GIAMOUZIS, G., SKOULARIGIS, J., LOURIDAS, G. and BUTLER, J., 2009. The Sympathetic Nervous System in Heart Failure: Physiology, Pathophysiology, and Clinical Implications. *Journal of the American College of Cardiology*, 54(19), pp. 1747-1762.

TSCHÖPE, C. and LAM, C.S.P., 2012. Diastolic heart failure: What we still don't know: Looking for new concepts, diagnostic approaches, and the role of comorbidities. *Herz*, 37(8), pp. 875-879.

TU, C., RUDNICK, P.A., MARTINEZ, M.Y., CHEEK, K.L., STEIN, S.E., SLEBOS, R.J.C. and LIEBLER, D.C., 2010. Depletion of abundant plasma proteins and limitations of plasma proteomics. *Journal of proteome research*, 9(10), pp. 4982.

VALENTINE, S.J., PLASENCIA, M.D., LIU, X., KRISHNAN, M., NAYLOR, S., UDSETH, H.R., SMITH, R.D. and CLEMMER, D.E., 2006. Toward plasma proteome

profiling with ion mobility-mass spectrometry. *Journal of proteome research*, 5(11), pp. 2977-2984.

VAN EMPEL, V.P.M., BERTRAND, A.T.A., HOFSTRA, L., CRIJNS, H.J., DOEVENDANS, P.A. and DE WINDT, L.J., 2005. Myocyte apoptosis in heart failure. *Cardiovascular research*, 67(1), pp. 21-29.

VAN HEEREBEEK, L., FRANSSEN, C.P.M., HAMDANI, N., VERHEUGT, F.W.A., SOMSEN, G.A. and PAULUS, W.J., 2012. Molecular and cellular basis for diastolic dysfunction. *Current heart failure reports*, 9(4), pp. 293-302.

VAN KIMMENADE, R. and JANUZZI, J., 2012. Emerging Biomarkers in Heart Failure.

WALTHER, T.C. and FARESE, R.V., 2012. Lipid Droplets and Cellular Lipid Metabolism. *Annual Review of Biochemistry*, 81, pp. 687-714.

WANG, Q., HE, K., LI, Z., CHEN, J., LI, W., WEN, Z., SHEN, J., QIANG, Y., JI, J., WANG, Y. and SHI, Y., 2014. The CMYA5 gene confers risk for both schizophrenia and major depressive disorder in the Han Chinese population. *The World Journal of Biological Psychiatry*, 15(7), pp. 553-560.

WANG, T., YAO, S., XIA, Z. and IRWIN, M.G., 2013. Adiponectin: mechanisms and new therapeutic approaches for restoring diabetic heart sensitivity to ischemic post-conditioning. *Frontiers of Medicine*, 7(3), pp. 301-305.

WANG, T.J., GONA, P., LARSON, M.G., TOFLER, G.H., LEVY, D., NEWTON-CHEH, C., JACQUES, P.F., RIFAI, N., SELHUB, J., ROBINS, S.J., BENJAMIN, E.J., D'AGOSTINO, R.B. and VASAN, R.S., 2006. Multiple Biomarkers for the Prediction of First Major Cardiovascular Events and Death. *The New England journal of medicine*, **355**(25), pp. 2631-2639.

WANG, W., ZHOU, H., LIN, H., ROY, S., SHALER, T.A., HILL, L.R., NORTON, S., KUMAR, P., ANDERLE, M. and BECKER, C.H., 2003. Quantification of proteins and metabolites by mass spectrometry without isotopic labeling or spiked standards. *Analytical Chemistry*, 75(18), pp. 4818-4826.

WANG, Y., YANG, F., GRITSENKO, M.A., WANG, Y., CLAUSS, T., LIU, T., SHEN, Y., MONROE, M.E., LOPEZ-FERRER, D., RENO, T., MOORE, R.J., KLEMKE, R.L., CAMP, D.G. and SMITH, R.D., 2011. Reversed-phase chromatography with multiple fraction concatenation strategy for proteome profiling of human MCF10A cells. *Proteomics*, 11(10), pp. 2019-2026.

WASINGER, V.C., ZENG, M. and YAU, Y., 2013. Current status and advances in quantitative proteomic mass spectrometry. *International journal of proteomics*, 2013, pp. 180605-12.

WEAVER, P.J., LAURES, A.M.-. And WOLFF, J., 2007. Investigation of the advanced functionalities of a hybrid quadrupole orthogonal acceleration time-of- flight mass spectrometer. *Rapid communications in mass spectrometry: RCM*, **21**(15), pp. 2415.

WEISS, W. and GÖRG, A., 2009. High-resolution two-dimensional electrophoresis. *Methods in molecular biology* (Clifton, N.J.), 564, pp. 13.

WILEY, W. C., AND MCLAREN, I. H., 1955. Time-of-Flight Mass Spectrometer with improved Resolution. *Review of Scientific Instruments*. 26, 1150-1157.

WILLEMSEN, S., HARTOG, J.W.L., HUMMEL, Y.M., VAN RUIJVEN, M.H.I., VAN DER HORST,IWAN C C., VAN VELDHUISEN, D.J. and VOORS, A.A., 2011. Tissue advanced glycation end products are associated with diastolic function and aerobic exercise capacity in diabetic heart failure patients. *European journal of heart failure*, 13(1), pp. 76-82.

WILM, M., MANN, M., BREIT, S., SCHWEIGERER, L., FOTSIS, T., SHEVCHENKO, A. and HOUTHAEVE, T., 1996. Femtomole sequencing of proteins from polyacrylamide gels by nano-electrospray mass spectrometry. *Nature*, 379(6564), pp. 466-469.

WOOD, P., PIRAN, S. and LIU, P.P., 2011. Diastolic heart failure: progress, treatment challenges, and prevention. *The Canadian journal of cardiology*, 27(3), pp. 302-310.

World Medical Association Declaration of Helsinki: ethical principles for medical research involving human subjects. 2013. *JAMA*, 310(20), pp. 2191.

WU, D., ZHU, B. and WANG, X., 2011. Metabonomics-based omics study and atherosclerosis. *Journal of clinical bioinformatics*, 1(1), pp. 30-30.

XIE, X., ZHANG, C. and TUAN, R.S., 2014. Biology of platelet-rich plasma and its clinical application in cartilage repair. *Arthritis research & therapy*, 16(1), pp. 204-204.

YADAV, A.K., BHARDWAJ, G., BASAK, T., KUMAR, D., AHMAD, S., PRIYADARSHINI, R., SINGH, A.K., DASH, D. and SENGUPTA, S., 2011. A systematic analysis of eluted fraction of plasma post immunoaffinity depletion: implications in biomarker discovery. *PloS one*, 6(9), pp. e24442.

YANG, K., RICKLIN, D., LAMBRIS, J.D. and HAJISHENGALLIS, G., 2010. Complement: a key system for immune surveillance and homeostasis. *Nature immunology*, 11(9), pp. 785-797.

ZECHNER, R., ZIMMERMANN, R., EICHMANN, T., KOHLWEIN, S., HAEMMERLE, G., LASS, A. and MADEO, F., 2012. FAT SIGNALS - Lipases and Lipolysis in Lipid Metabolism and Signaling. *Cell Metabolism*, 15(3), pp. 279-291.

ZILE, M.R. and BAICU, C.F., 2013. Biomarkers of diastolic dysfunction and myocardial fibrosis: application to heart failure with a preserved ejection fraction. *Journal of cardiovascular translational research*, 6(4), pp. 501-515.

ZILE, M.R., DESANTIS, S.M., BAICU, C.F., STROUD, R.E., THOMPSON, S.B., MCCLURE, C.D., MEHURG, S.M. and SPINALE, F.G., 2011. Plasma biomarkers that reflect determinants of matrix composition identify the presence of left ventricular hypertrophy and diastolic heart failure. Circulation. *Heart failure*, 4(3), pp. 246-256.

ZOLOTARJOVA, N., MROZINSKI, P., CHEN, H. and MARTOSELLA, J., 2008. Combination of affinity depletion of abundant proteins and reversed-phase fractionation in proteomic analysis of human plasma/serum. *Journal of Chromatography* A, 1189(1), pp. 332-338.

ZHOU, H. and LEE, J., 2011. Nanoscale hydroxyapatite particles for bone tissue engineering. *Acta Biomaterialia*, **7**(7), pp. 2769-2781.

ZHU, S., ZHANG, X., GAO, M., HONG, G., YAN, G. and ZHANG, X., 2012. Developing a strong anion exchange/ RP (SAX/ RP) 2D LC system for high- abundance proteins depletion in human plasma. *Proteomics*, 12(23-24), pp. 3451-3463.

ZUBAREV, R. A., KELLEHER N.L., AND F. W. MCLAFFERTY F.W., 1998. Electron Capture Dissociation of Multiply Charged Protein Cations. A Nonergodic Process. *Journal of the American Chemical Society*. 120, 3265-3266.