COPYRIGHT AND USE OF THIS THESIS

This thesis must be used in accordance with the provisions of the Copyright Act 1968.

Reproduction of material protected by copyright may be an infringement of copyright and copyright owners may be entitled to take legal action against persons who infringe their copyright.

Section 51 (2) of the Copyright Act permits an authorized officer of a university library or archives to provide a copy (by communication or otherwise) of an unpublished thesis kept in the library or archives, to a person who satisfies the authorized officer that he or she requires the reproduction for the purposes of research or study.

The Copyright Act grants the creator of a work a number of moral rights, specifically the right of attribution, the right against false attribution and the right of integrity.

You may infringe the author’s moral rights if you:

- fail to acknowledge the author of this thesis if you quote sections from the work
- attribute this thesis to another author
- subject this thesis to derogatory treatment which may prejudice the author’s reputation

For further information contact the University’s Director of Copyright Services

sydney.edu.au/copyright
The Effect of Age, Sex, and End-Stage Heart Failure on the Human Cardiac Transcriptome

Dr. Maurizio Stefani, MB BS, B.Sc. (Hons I)

A thesis submitted in fulfilment for the award of Doctor of Philosophy

School of Medical Sciences, Sydney Medical School
The University of Sydney, 2015
Statement of Originality

In accordance with the laws of The University of Sydney, the author declares that this thesis describes original work performed under the auspices of the Sydney Medical School at The University of Sydney, Sydney, Australia, as well as at the Victor Chang Cardiac Research Institute, Sydney, Australia and the Université de Nantes, Nantes, France between 2006 and 2012.

All the experimental work and results reported herein are solely by the author except for the following contributions:

1) mRNA microarrays were constructed by Catherine Chevalier of the Université de Nantes, Nantes, France.

2) Blocks of paraffin-embedded formalin-fixed LV samples prepared by the author were cut and arranged into tissue microarrays by Darryl Cameron of The University of Sydney, Sydney, Australia.

This work has not been submitted for any other degree at this or any other institution.

Maurizio Stefani
Abstract

This thesis aimed to describe the transcriptomic changes in the human left ventricle as a result of age and gender, and to relate these changes to the pathophysiology of heart failure, in order to understand why ageing and gender modifies our susceptibility to, and progression of heart failure. Many genes were discovered to be affected by age and gender, in particular, the set of genes that code for the r-proteins, i.e ribosomal subunit coding proteins, are downregulated with age, and are expressed at a lower level in males. I hypothesised that this may compromise the ability of myocardium to engage in necessary hypertrophy/hyperplasia to recover from a myocardial insult such as a myocardial infarction, and thus increases the risk of heart failure developing. However, no evidence of changes in nucleolar abundance, the site of ribosome synthesis, was found with age and gender. In heart failure of a variety of aetiologies, the expression of r-protein mRNA and 45s rRNA was found to be reduced, and the reduction of r-protein mRNA expression was ameliorated by left ventricular assist device support. A reduction in nucleolar abundance was also demonstrated in certain aetiologies of heart failure. I hypothesised that this may be due to an energy shortage in heart failure, partly ameliorated by left ventricular assist device support, leading to a reduction in energetically costly ribosome synthesis which is pathological in the long term.
Acknowledgements

Thank you first and foremost to my supervisor Professor Cris Dos Remedios who gave me opportunity and encouragement in the world of medical research. Thank you to Associate Professor Jean Yee Hwa Yang who assisted with the analysis of my microarray experiments. Thank you to Doctor Marja Steenman, my supervisor in Nantes, France. Many thanks to the other people who have come and gone through the MRU lab, as well as to my colleagues in the VCCRI and at the Université de Nantes. And thank you to my friends and family for all the good times. Finally, I would like to acknowledge the many patients I have met as a junior doctor who are suffering with heart failure, meeting them has made this research project much more meaningful to me.
Publications and Presentations During My PhD Candidature

Publications


Abstracts


# Abbreviations

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>ADP</td>
<td>Adenosine diphosphate</td>
</tr>
<tr>
<td>AMP</td>
<td>Adenosine monophosphate</td>
</tr>
<tr>
<td>AMPK</td>
<td>AMP activated protein kinase</td>
</tr>
<tr>
<td>ATP</td>
<td>Adenosine triphosphate</td>
</tr>
<tr>
<td>AV</td>
<td>Atrioventricular</td>
</tr>
<tr>
<td>BSA</td>
<td>Body surface area</td>
</tr>
<tr>
<td>cAMP</td>
<td>Cyclic adenosine monophosphate</td>
</tr>
<tr>
<td>CDK2</td>
<td>Cyclin-dependent kinase 2</td>
</tr>
<tr>
<td>cDNA</td>
<td>Copy deoxyribonucleic acid</td>
</tr>
<tr>
<td>CI</td>
<td>Cardiac index</td>
</tr>
<tr>
<td>CICR</td>
<td>Ca(^{2+})-induced Ca(^{2+}) release</td>
</tr>
<tr>
<td>CO</td>
<td>Cardiac output</td>
</tr>
<tr>
<td>DE</td>
<td>Differentially expressed</td>
</tr>
<tr>
<td>DFC</td>
<td>Dense fibrillar component</td>
</tr>
<tr>
<td>DNA</td>
<td>Deoxyribonucleic acid</td>
</tr>
<tr>
<td>EC</td>
<td>Excitation-contraction</td>
</tr>
<tr>
<td>EDV</td>
<td>End-diastolic volume</td>
</tr>
<tr>
<td>EDVI</td>
<td>End-diastolic volume index</td>
</tr>
<tr>
<td>EF</td>
<td>Ejection fraction</td>
</tr>
<tr>
<td>eIF</td>
<td>Eukaryotic initiation factor</td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Description</td>
</tr>
<tr>
<td>--------------</td>
<td>-------------</td>
</tr>
<tr>
<td>ES</td>
<td>Enrichment score</td>
</tr>
<tr>
<td>ESV</td>
<td>End-systolic volume</td>
</tr>
<tr>
<td>ESVI</td>
<td>End-systolic volume index</td>
</tr>
<tr>
<td>EtOH</td>
<td>Ethanol</td>
</tr>
<tr>
<td>FC</td>
<td>Fibrillar centre</td>
</tr>
<tr>
<td>FDR</td>
<td>False discovery rate</td>
</tr>
<tr>
<td>GC</td>
<td>Granular component</td>
</tr>
<tr>
<td>GSEA</td>
<td>Gene set enrichment analysis</td>
</tr>
<tr>
<td>GTP</td>
<td>Guanosine triphosphate</td>
</tr>
<tr>
<td>HCM</td>
<td>Hypertrophic cardiomyopathy</td>
</tr>
<tr>
<td>HF</td>
<td>Heart failure</td>
</tr>
<tr>
<td>HR</td>
<td>Heart rate</td>
</tr>
<tr>
<td>IDCM</td>
<td>Idiopathic dilated cardiomyopathy</td>
</tr>
<tr>
<td>IIInt</td>
<td>Integrated intensity</td>
</tr>
<tr>
<td>ISCM</td>
<td>Ischaemic cardiomyopathy</td>
</tr>
<tr>
<td>KCNMA1</td>
<td>Potassium large conductance calcium-activated channel, subfamily M, alpha member 1 (KCNMA1)</td>
</tr>
<tr>
<td>LA</td>
<td>Left atrium</td>
</tr>
<tr>
<td>LIMK1</td>
<td>Lim domain kinase 1</td>
</tr>
<tr>
<td>LV</td>
<td>Left ventricle</td>
</tr>
<tr>
<td>LVAD</td>
<td>Left ventricular assist device</td>
</tr>
<tr>
<td>LVH</td>
<td>Left ventricular hypertrophy</td>
</tr>
<tr>
<td>MI</td>
<td>Myocardial infarction</td>
</tr>
<tr>
<td>mRNAb</td>
<td>Messenger ribonucleic acid</td>
</tr>
<tr>
<td>mtDNAb</td>
<td>Mitochondrial deoxyribonucleic acid</td>
</tr>
<tr>
<td>MTORB</td>
<td>Mammalian target of rapamycin</td>
</tr>
<tr>
<td>NF</td>
<td>Non-failing</td>
</tr>
<tr>
<td>NF-AT</td>
<td>Nuclear factor of activated T-cells</td>
</tr>
<tr>
<td>NFAT3C</td>
<td>Nuclear factor of activated T-cells, cytoplasmic, calcineurin-dependent 3</td>
</tr>
<tr>
<td>NOR</td>
<td>Nucleolar organising region</td>
</tr>
<tr>
<td>PAGE</td>
<td>Polyacrylamide gel electrophoresis</td>
</tr>
<tr>
<td>PBS</td>
<td>Phosphate buffered saline</td>
</tr>
<tr>
<td>PCR</td>
<td>Polymerase chain reaction</td>
</tr>
<tr>
<td>Pi</td>
<td>Inorganic phosphate</td>
</tr>
<tr>
<td>RA</td>
<td>Right atrium</td>
</tr>
<tr>
<td>rDNAb</td>
<td>Ribosomal deoxyribonucleic acid</td>
</tr>
<tr>
<td>RHEB</td>
<td>Ras homolog enriched in brain</td>
</tr>
<tr>
<td>RNA</td>
<td>Ribonucleic acid</td>
</tr>
<tr>
<td>RO</td>
<td>Reverse osmosis</td>
</tr>
<tr>
<td>ROS</td>
<td>Reactive oxygen species</td>
</tr>
<tr>
<td>rRNA</td>
<td>Ribosomal ribonucleic acid</td>
</tr>
<tr>
<td>RT</td>
<td>Reverse transcription</td>
</tr>
<tr>
<td>RV</td>
<td>Right ventricle</td>
</tr>
<tr>
<td>RYR2</td>
<td>Ryanodine receptor 2</td>
</tr>
<tr>
<td>SA</td>
<td>Sinoatrial</td>
</tr>
<tr>
<td>SD</td>
<td>Standard deviation</td>
</tr>
<tr>
<td>Acronym</td>
<td>Term</td>
</tr>
<tr>
<td>---------</td>
<td>------</td>
</tr>
<tr>
<td>SDS</td>
<td>Sodium dodecyl sulfate</td>
</tr>
<tr>
<td>SERCA</td>
<td>Sarco/endoplasmic reticulum Ca(^{2+})-ATPase</td>
</tr>
<tr>
<td>SR</td>
<td>Sarcoplasmic reticulum</td>
</tr>
<tr>
<td>SSC</td>
<td>Saline-sodium citrate buffer</td>
</tr>
<tr>
<td>SV</td>
<td>Stroke volume</td>
</tr>
<tr>
<td>SVI</td>
<td>Stroke volume index</td>
</tr>
<tr>
<td>TLR6</td>
<td>Toll-like receptor 6</td>
</tr>
<tr>
<td>TMA</td>
<td>Tissue microarray</td>
</tr>
<tr>
<td>tRNA</td>
<td>Transfer ribonucleic acid</td>
</tr>
<tr>
<td>TSC1</td>
<td>Tuberous sclerosis complex-1</td>
</tr>
</tbody>
</table>
# Table of Contents

Statement of Originality ........................................................................................................ ii  
Abstract ............................................................................................................................. iii  
Acknowledgements .......................................................................................................... iv  
Publications and Presentations During My PhD Candidature ....................................... v  
  
  *Publications* .................................................................................................................. v  
  *Abstracts* ...................................................................................................................... vi  
Abbreviations .................................................................................................................. vii  

## Chapter 1: General Introduction..................................................................................... 1  

1.1 The Heart and My Thesis: An overview ................................................................... 1  
1.2 The Structure of the Heart ......................................................................................... 2  
  1.2.1 Gross Anatomy ................................................................................................... 2  
  1.2.2 Layers of the Heart ............................................................................................. 3  
  1.2.3 Coronary Vessels ................................................................................................ 4  
  1.2.4 Conduction System of the Heart ....................................................................... 7  
1.2.5 Ultrastructure of the Heart ................................................................................. 8  
  1.2.5.1 Sarcomere ................................................................................................... 10  
  1.2.5.2 Mitochondria ............................................................................................. 13  
  1.2.5.3 Sarcolemma and T-Tubules ....................................................................... 14  
  1.2.5.4 Sarcoplasmic Reticulum ........................................................................... 14  
  1.2.5.5 Ribosomes ................................................................................................. 14  
1.3 Functions of the Heart ............................................................................................ 17
1.3.1 Introduction to Cardiac Contraction ................................................................. 17
1.3.2 Cardiac Contraction ......................................................................................... 19
1.3.3 Regulation of Contraction .............................................................................. 23
  1.3.3.1 Intrinsic Regulation .................................................................................. 23
  1.3.3.2 Regulation by the Autonomic Nervous System .......................................... 24
1.4 Heart Failure ....................................................................................................... 25
  1.4.1 Pathophysiology ........................................................................................... 25
  1.4.2 Ischaemic Cardiomyopathy .......................................................................... 29
  1.4.3 Idiopathic Dilated Cardiomyopathy .............................................................. 29
  1.4.4 Hypertrophic Cardiomyopathy .................................................................... 29
  1.4.5 Recovery with LV Assist Device Implantation .............................................. 30
1.5 The Effects of Age and Gender on the Human Heart .......................................... 31
  1.5.1 Introduction .................................................................................................. 31
  1.5.2 Structural Changes with Age and Gender ...................................................... 31
    1.5.2.1 Macroscopic Changes with Age and Gender ........................................... 31
    1.5.2.2 Cardiomyocyte Compartment Changes with Age and Gender .............. 32
    1.5.2.3 Non-Cardiomyocyte Compartment Changes with Age ......................... 34
  1.5.3 Functional Changes with Age and Gender .................................................... 35
  1.5.4 The Ageing Cell ............................................................................................ 37
  1.5.5 The Risk of HF as Modified by Age and Gender .......................................... 40
1.6 Focus of Thesis .................................................................................................... 42

Chapter 2: General Methods .................................................................................... 43
  2.1 Collection of Human Heart Tissue ................................................................. 43
2.1.1 Human Research Ethics Committee Approval ................................................ 43
2.1.2 Collection of Human Heart Tissue ................................................................. 43

2.2 mRNA Expression Analysis .............................................................................. 44
  2.2.1 RNA Extraction .......................................................................................... 44
  2.2.2 Microarray Analyses .................................................................................. 47
  2.2.3 Real-Time Polymerase Chain Reaction Analysis ........................................ 47

2.3 Protein Expression Analysis ............................................................................ 52
  2.3.1 Protein Extraction ...................................................................................... 52
  2.3.2 Determination of Protein Extract Concentration ......................................... 52
  2.3.3 SDS-PAGE and Western Blotting ............................................................... 53

Chapter 3: Age and Sex-related Changes in Human NF LV ............................... 55
  3.1 Introduction ..................................................................................................... 55
  3.2 Methods .......................................................................................................... 56
    3.2.1 RNA extraction for Microarray and Real-time PCR Analysis .................... 56
    3.2.2 Microarray Analysis of NF LV to Discover Age- and Sex-related Changes ... 56
      3.2.2.1 RNA labelling for microarray analysis, hybridisation and scanning .... 56
      3.2.2.2 Analysis of Microarray Data ................................................................. 59
    3.2.3 Real-Time PCR Analysis of Ribosome-related Genes in NF LV ............... 62
    3.2.4 Western Blotting Analysis of RPS18 in NF LV ........................................ 63
  3.3 Results ............................................................................................................ 63
    3.3.1 Microarray Analysis of NF LV to Discover Age- and Sex-related Changes ... 63
    3.3.2 Real-Time PCR Analysis of Ribosome-related Genes in NF LV ............... 70
    3.3.3 Western Blotting Analysis of RPS18 in NF LV ........................................ 72
3.4 Discussion ............................................................................................................... 73

3.4.1 Technical Interpretation of Results ................................................................. 73
3.4.2 Biological Interpretation of Results ................................................................. 77

Chapter 4: Ribosomal Gene Expression in End-Stage HF and LVAD Support ........ 90

4.1 Introduction ............................................................................................................. 90

4.2 Methods ................................................................................................................... 91

4.2.1 RNA extraction for Real-time PCR Analysis ................................................... 91
4.2.2 Ribosome Gene Set Expression in End-Stage HF and LVAD Support .......... 91
4.2.3 Real-time PCR Analysis of Ribosome-related Genes in End-Stage HF ......... 94
4.2.4 Western Blotting Analysis of RPS18 in End-Stage HF ................................. 95

4.3 Results ..................................................................................................................... 95

4.3.1 Ribosome Gene Set Expression in End-Stage HF and LVAD Support .......... 95
4.3.2 Real-time PCR Analysis of Ribosome-related Genes in End-Stage HF ........ 108
4.3.3 Western Blotting Analysis of RPS18 in End-Stage HF ................................. 109

4.4 Discussion .............................................................................................................. 110

4.4.1 Technical Interpretation of Results ................................................................. 110
4.4.2 Biological Interpretation of Results ................................................................. 112

Chapter 5: The Effect of Age, Sex and End-Stage HF on Nucleolar Remodelling... 119

5.1 Introduction ............................................................................................................. 119

5.2 Methods ................................................................................................................... 123

5.2.1 Patient Demographics .................................................................................... 123
5.2.2 Tissue Microarray Construction ..................................................................... 124
5.2.3 AgNO3 Staining .............................................................................................. 124
Chapter 1: General Introduction

1.1 The Heart and My Thesis: An overview

In this thesis, I would like to address the issue of why people are more susceptible to developing heart failure (HF) after any particular cardiac insult, at greater ages. I would also like to explore the influence of gender on the development of HF. In order to come to grips with these questions, I will first begin by describing the heart, from an anatomical and then functional point of view. I will then discuss the changes that occur within the human heart throughout the adult lifespan that might underlie this increased risk, as well as the differences in cardiac structure and function according to gender. I will also introduce the pathophysiology of HF, and the various clinical entities associated with it that are investigated in this thesis.

The heart is the organ that generates the pressure that drives the flow of blood to all the organs and tissues of the body, including to itself. It must generate sufficient pressure such that adequate flow occurs, and be able to alter the pressure it generates in response to dynamic changes in the requirements of the various tissues of the body. The heart begins to contract at 4 weeks post fertilisation in the embryo, and must never stop.

The function of the heart is dependent on its anatomical architecture, and a myriad of hormonal, nervous and biochemical signalling pathways. These elements are altered during ageing and with the onset of pathology, and affect the ability of the heart to generate sufficient pressure to ensure adequate blood flow to the body. Understanding
how the ageing process and gender and pathology interact to affect the heart, is the 
purpose and scope of this thesis.

1.2 The Structure of the Heart

1.2.1 Gross Anatomy

The human heart is located in the thorax within the middle mediastinum, an anatomical 
space that includes all structures enclosed by the pericardial sac. It is a four-chambered 
organ: left (LA) and right atria (RA), left (LV) and right ventricles (RV) (Fig. 1.1). 
Separating the atria and ventricles are the atroventricular valves: the mitral valve 
separates the left side chambers, and the tricuspid, the right. The aortic and pulmonary 
semi-lunar valves separate the LV and RV from the aorta and pulmonary trunk, 
respectively. Deoxygenated blood returning from the systemic circulation passes into the 
RA from the superior and inferior venae cavae. Oxygenated blood from the pulmonary 
circulation passes into the left atrium from four pulmonary veins (Fig 1.2(b)).
1.2.2 Layers of the Heart

The pericardial sac is composed of two layers. The outer fibrous pericardium has a serous parietal lining that is continuous at the base of the heart with the serous epicardium (or visceral pericardium). The space enclosed by these two layers is the pericardial space, which contains a small amount of serous fluid that allows the two layers to slide over each other as the heart beats. Within the epicardium, are located the major coronary vessels, and variable amounts of adipose tissue. Beneath the pericardium, the next layer is
myocardium, which forms the bulk of the mass of the heart, and is comprised principally of cardiomyocytes, which are the contractile cells of the heart. The myocardium of the LV is typically 3 times thicker than that of the RV, which reflects the differing amounts of pressure that must be generated to drive blood through the systemic and pulmonary circulations, respectively. The inner surface of the atria and ventricles are lined with endocardium, which is continuous with the endothelium of the blood vessels that take blood to and away from the heart. It is comprised of an epithelial cell layer with underlying connective tissue containing fibroblasts and smooth muscle cells.

1.2.3 Coronary Vessels

The heart, like every other tissue in the body, needs an adequate blood supply. The endocardium is nourished sufficiently by the blood that is continuously flowing through the atria and ventricles, but the remainder of the tissue of the heart has a dedicated blood vessel network (Fig. 1.2). The wall of the proximal aorta contains three small dilations: the left, right, and posterior aortic sinuses. From the left aortic sinus, arises the left coronary artery, which divides into the anterior interventricular artery (also known as the left anterior descending), and the left circumflex artery. These arteries usually supply the anterior, left lateral, apical, and sometimes the posterior surfaces of the heart, as well as the anterior portion of the interventricular septum. From the right aortic sinus, arises the right coronary artery, which supplies the right lateral, as well as the posterior surface of the heart in the majority of people, in which it branches to form the posterior interventricular artery. Branches from these epicardial arteries penetrate the myocardium.
and supply the cardiomyocytes (and accompanying fibroblasts), as well as the electrical conduction system that regulates the activity of these contractile cells.
Deoxygenated blood flows from heart tissue through veins that follow the coronary arteries and eventually drain into the coronary sinus (located in the posterior wall of the heart) and then into the RA. A small portion of deoxygenated blood flows directly into the chambers of the heart through small venous channels called Thebesian veins (or venae cordis minima). Blood flow to the heart tissue occurs mainly during diastole, the period during which the ventricles are not contracting.
1.2.4 Conduction System of the Heart

The sinoatrial (SA) node is located in the RA, adjacent to the entry point of the superior vena cava. This group of specialised SA nodal cells generate the wave of depolarisation that initiates heart contraction (Fig. 1.3). From this node, action potentials spread across the two atria, to the atrioventricular (AV) node, a group of conducting cells, located at the base of the interatrial septum. From the AV node, a bundle of conductive cells known as the Bundle of His traverses the fibrous tissue separating the atria and the ventricles. The Bundle of His divides early in the interventricular septum into left (that splits into the anterior and posterior hemi-bundles or fascicles) and right bundle branches. From these bundles, branches split off to form what are known as the Purkinje fibres, and these conduct the action potential to all the cardiomyocytes within the ventricular myocardium.

Figure 1.3: A diagram of conduction system in the heart. The image was taken from http://www.bioen.utah.edu/faculty/sri/Lab8_humancardiovascular.htm.
1.2.5 Ultrastructure of the Heart

The myocardium is by volume, ~ 80% cardiomyocytes, with the remainder made up of blood vessels and connective tissue. Human ventricular cardiomyocytes are approximately cylindrical, measuring 10 to 25 μm in width, and 50 to 100 μm in length. They branch and connect with adjacent cardiomyocytes via specialised junctions known as intercalated discs (Fig. 1.4). These are sites of low electrical resistance that allow the wave of depolarisation to traverse between the cardiomyocytes. The atrial cardiomyocytes are more elliptical, and smaller, measuring approximately 20 μm in width and 50 to 60 μm in length.

Cardiomyocytes contain a centrally located nucleus that comprises about ~ 5% of cell volume, held in place by microtubules supported by the surrounding myofibrils. Between 50 and 60 % of the volume of the cardiomyocyte is occupied by myofibrils, which are the structural units responsible for tension generation. The mitochondria, which are the site of adenosine triphosphate (ATP) generation, make up 25% of the cell volume. Most of the rest of the volume of cardiomyocytes comprises the sarcoplasmic reticulum (the cardiomyocyte equivalent of the endoplasmic reticulum, SR), as well as the T-tubules which are invaginations of the sarcolemma (plasma membrane of cardiomyocytes), and the Golgi apparatus. Numerous lysosomes and ribosomes occupy only a small fraction of the cell volume, but as will be discussed, play an important role in ageing.
Figure 1.4: A diagram demonstrating the spatial relationships of ventricular cardiomyocytes connected by intercalated discs (above), the ultrastructure of the ventricular cardiomyocytes (middle) with myofibrils, mitochondria (purple), SR (pink), sarcolemma and T-tubules (blue), and the structure of the sarcomeres located within the myofibrils (below). Image taken from Zipes, D. P. and Braunwald, E. (2005).
1.2.5.1 Sarcomere

Myofibrils consist of a chain of sarcomeres, defined as the structures between successive Z-disks (or Z-bands). These are the basic functional units of the cardiomyocyte and are responsible for force generation (Figs. 1.4 and 1.5). Sarcomeres are made up of thick filaments containing myosin, thin filaments composed mainly of actin, plus a third filament, titin, that interact during the cross-bridge cycle converting chemical energy (Mg\(^{2+}\)-ATP) into force. Actin is a globular protein that polymerises to form filaments that wind around each other to form a double helix structure. The thin filaments also contain the Ca\(^{2+}\) regulatory molecules, tropomyosin and troponins I, T, and C (Fig. 1.9). Myosin consists of two myosin heavy chains and two pairs of myosin light chains that form a complex resembling a golf club with two heads. The “shafts” of the “golf club” self-assemble to form the bipolar thick filaments at the M-line with the “heads” oriented towards the Z-disks. In the A-bands, the myosin and actin filaments interdigitate and their interaction produces changes in sarcomere length. Titin (nearly 4 mDa) extends as a single molecule from the Z-disk all the way to the M-line.
Figure 1.5: A transmission electron micrograph (EM) of a cardiomyocyte demonstrating the striated appearance and various structures of the myofibrils including the A-band (A), I-band (I), M-line (M), Z-disc (Z), H-band (H), as well as surrounding mitochondria (mit), glycogen granules (g), and a T-tubule (T). Image taken from Zipes, D. P. and Braunwald, E. (2005).

Sarcomeres (and the myofibrils they make up) contain many other proteins that maintain their longitudinal orientation and stability. These include microtubules, γ-actin filaments, titin, α-actinin and intermediate filament proteins, such as desmin and vimentin, as well as many others (Fig. 1.6). These proteins also anchor the myofibrils to the protein complexes of the sarcolemma and extracellular matrix.
Figure 1.6: Major cytoskeletal proteins that connect myofibrils to the other parts of the cardiomyocyte (Dalakas, M. C. et al. 2000).
1.2.5.2 Mitochondria

The mitochondria are cocoon-shaped organelles that lie between myofibrils (myofibrillar mitochondria) and adjacent to the sarcolemma (subsarcolemmal mitochondria) and supply ATP for muscle contraction and cellular metabolism in cardiomyocytes (Fig. 1.7). The mitochondria have a double membrane structure, with the inner mitochondrial membrane being greatly folded to maximise surface area. There are button-shaped outpouchings on the inner mitochondrial membrane called cristae, which contain the enzymes for oxidative phosphorylation and ATP synthesis. The predominant fuels used for cardiac metabolism under physiological conditions are fatty acids, with lactate playing a minor role from a quantitative perspective.

Figure 1.7: The structure of the mitochondrion with demonstration of the various cellular substrates it can take up and use as fuel for ATP synthesis and the output molecules generated. Image taken from Opie, L. H. (2004).
1.2.5.3 Sarcolemma and T-Tubules

The cardiomyocyte sarcolemma contains many channels, pumps, and receptors essential to cell functioning. In particular, it contains the voltage-gated Na⁺ channels that open in response to a wave of depolarisation, which leads to the influx of Ca²⁺ through sarcolemmal L-type Ca²⁺ channels that is ultimately responsible for triggering contraction. The T-tubules (Fig. 1.4) are invaginations of sarcolemma that allow Ca²⁺ entry to occur directly across from the SR and the nearby myofibrils. This allows synchronous activation of all myofibrils within the cardiomyocyte (explained in more detail later).

1.2.5.4 Sarcoplasmic Reticulum

The SR (Fig. 1.4) is a specialised form of endoplasmic reticulum which plays a major role in the Ca²⁺ handling that is integral to cardiomyocyte contraction. It is composed of a fine network of tubing throughout the cytoplasm with bulbous swellings juxtaposed to the T-tubules. Ca²⁺ is stored within the SR bound to Calsequestrin, and its release triggers the contraction of atrial and ventricular cardiomyocytes.

1.2.5.5 Ribosomes

Ribosomes are massive ribonucleoprotein complexes involved in translating the triplet code within messenger ribonucleic acid (mRNA) molecules into the corresponding amino acid sequence to produce all the proteins in the heart. The human ribosome has 2 subunits, a smaller 40S and a larger 60S subunit. The 40S subunit is composed of 33
different r-proteins plus the 18s ribosomal ribonucleic acid (rRNA), and is responsible for decoding the mRNA strand. The 60s subunit is composed of 46 r-proteins plus the 5S, the 28S, and the 5.8S rRNAs and is responsible for forming peptidyl bonds during peptide chain elongation. Ribosomes are found both in the cytosol as well as on the surface of the SR. They are also abundant in the non-cardiomyocyte cells of the heart.

Synthesis of ribosomes is reviewed in Lafontaine, D. L. and Tollervey, D. (2001). Briefly, a large polycistronic 45s rRNA is transcribed by RNA polymerase I, covalently modified and cleaved into the mature 18s, 28s and 5.8s rRNA strands. The 5s rRNA is transcribed from a different region in the genome by RNA polymerase III. R-proteins are translated in the cytoplasm by pre-formed ribosomes and are imported into the nucleus and then into the nucleolus where they are assembled with the rRNA into pre-ribosome subunits that are exported into the cytoplasm and undergo final structural rearrangements to become mature ribosome subunits.

Protein synthesis, i.e. translation of mRNA, is a complex process and is reviewed in Lafontaine, D. L. and Tollervey, D. (2001) and Hannan, R. D. et al. (2003). Briefly, the first step is that a methionyl-transfer ribonucleic acid (tRNA) binds to several eukaryotic initiation factors (eIF) and to the 40s ribosome subunit, this quaternary structure is termed the 43s pre-initiation complex. Then, 43s binds to the eIF4F complex (composed of proteins eIF4A, eIF4E, and EIF4G) which then imbues the complex with the ability to search for mRNA bearing the 5' 7-methylguanosine cap. An mRNA is bound and the complex moves along the 5' untranslated region until it encounters the codon for methionine (AUG). At this point, the eIFs are displaced, and the 60s subunit binds to form the functional 80s ribosome.
There are three key sites on the ribosome that are important in mRNA translation, the A, P and E sites (Fig. 1.8). The methionyl-tRNA is bound at the P site, and the A site is occupied by a complex containing an aminoacyl-tRNA with anticodon complementary to the codon on the mRNA (with the corresponding amino acid), guanosine triphosphate (GTP), GTPase elongation factor 1-α. When correct tRNA-mRNA binding occurs, the GTP is hydrolysed, and the peptide chain (or in the beginning just the methionine) is cleaved from the bound tRNA in the P site and attached to the amino acid via a peptide bond on the aminoacyl-tRNA in the A site. After peptidyl transfer, GTPase elongation factor-G attaches and hydrolyses GTP which leads to translocation of the mRNA along the interface of the two ribosome subunits with the tRNA in the P site moving to the E site, and the tRNA with the growing polypeptide chain in the A site moving to the P site. The process continues until the ribosome reaches a stop codon on the mRNA and the nascent polypeptide is released.
Figure 1.8: Schematic of the eukaryotic 80s ribosome with the 40s and 60s subunits labelled, as well as the mRNA being translated, the E site, the growing polypeptide chain attached by a tRNA (not shown) to the P site, and the incoming aminoacyl-tRNA bound to the A site. Image adapted from Lafontaine, D. L. and Tollervey, D. (2001).

1.3 Functions of the Heart

1.3.1 Introduction to Cardiac Contraction

Cardiomyocytes contract and shorten in response to electrical signals generated originally by the conduction system of the heart, and this results in pressure being generated by the chambers of the heart, which drives blood flow. Blood flow from the heart is usually measured as the cardiac output (CO), defined as the total volume of blood ejected by the ventricles into the pulmonary or systemic circulation, per minute. Cardiac index (CI), a related measure, is CO corrected for body surface area (BSA).
**Figure 1.8:** Depiction on the various phases of the cardiac cycle (bottom), along with the associated pressures in the LA, LV, and aorta, heart sounds, jugular venous pressure (JVP), and electrical events on the electrocardiogram. 'Iso' refers to isovolumic contraction and relaxation, respectively, in the bottom sequence of LA and LV depictions. Image taken from Zipes, D. P. and Braunwald, E. (2005).
The atria and ventricles contract during different phases of the cardiac cycle. The two phases are systole (when the ventricles contract and propel blood to the pulmonary and systemic circulations) and diastole (when the ventricles are relaxing and filling with the aid of atrial contraction in readiness for the next phase of systole). For more detail, see Fig. 1.8.

The heart can vary the force of contraction with each beat, to match the CO to the dynamically changing needs of the tissues of the body. There are multiple mechanisms by which this is achieved including changes in contractility due to changes in ventricular filling, i.e. the Frank-Starling relationship, as well as the varying input of the autonomic nervous system and circulating hormones (elaborated upon in later chapters).

1.3.2 Cardiac Contraction

The cardiomyocyte contracts when the sarcomeres shorten through the interaction of myosin and actin, in response to membrane depolarisation and subsequent elevated cellular Ca\(^{2+}\) concentration. This process is known as excitation-contraction (EC) coupling. When the membrane of a cardiomyocyte is depolarised by changes in ion gradients in the adjacent sarcolemma, the voltage-gated Na\(^+\) channels open, depolarising the sarcolemma further, leading to the opening of L-type voltage-gated Ca\(^{2+}\) channels. This occurs both in the sarcolemma proper, as well as in the T-tubules. The entering Ca\(^{2+}\) binds to ryanodine receptors (RYR2 is the cardiac isoform) present on the adjacent SR, which then triggers the opening of their associated channel. This leads to Ca\(^{2+}\) release from the SR thus elevating the myofibrillar Ca\(^{2+}\) concentration, so-called Ca\(^{2+}\)-induced Ca\(^{2+}\) release (CICR). Myofibrillar Ca\(^{2+}\) then binds to troponin C, which forms part of a
heterotrimer with troponins T and I (Fig. 1.9). This troponin complex is attached to the polymerised actin within the sarcomere, and to the associated tropomyosin. When Ca\(^{2+}\) is absent, tropomyosin covers the myosin-binding site on actin, preventing cross-bridge formation and force generation. However, when Ca\(^{2+}\) binds to troponin C, a conformational change occurs in the troponin complex that causes tropomyosin to move away from the myosin-binding site, allowing myosin to bind, leading to force production via the cross-bridge cycle.

**Figure 1.9:** Diagram illustrating the components of the thick and thin filaments involved in EC coupling and cross-bridge cycling, and the direction of movement of the two filaments relative to each other. Image taken from Kamisago, M. *et al.* (2000).

Myosin, in its non-contractile state, has adenosine diphosphate (ADP) and inorganic phosphate (Pi) bound to it. This is stage 2 of the 4 stage cross-bridge cycle (Fig. 1.10). At this stage, the myosin head is primed, and if the myosin-binding site on the adjacent actin polymer is exposed, it will bind (stage 3) and Pi will be released. Subsequently, myosin
changes conformation such that the head is flexed at the neck of the complex (stage 4), resulting in the pulling of the thin filament towards the M-line, shortening of the sarcomere, and force generation. In the transition to stage 1 from stage 4, ATP binds to the myosin head and the myosin complex detaches from the thin filament. The ATP is then hydrolysed resulting in myosin resuming its primed conformation with ADP and Pi attached (stage 2), and its availability to bind any available exposed myosin-binding sites on actin. This cross-bridge cycle is repeated by the thick and thin filaments as long as Ca\(^{2+}\) is available to bind to troponin C, and the force generated at each cross-bridge summates to generate the total force that results in atrial and ventricular contraction during the cardiac cycle.

Cross-bridge cycling and force generation does not continue indefinitely. Once the atria or ventricles have ejected the requisite amount of blood necessary, they must relax and allow filling to occur. This occurs via a reduction of the myofibrillar Ca\(^{2+}\) concentration by the reuptake of Ca\(^{2+}\) by the SR through sarco/endoplasmic reticulum Ca\(^{2+}\)-ATPase (SERCA). Myofibrillar Ca\(^{2+}\) concentration is also decreased to a lesser extent by efflux through the sarcolemmal Na\(^+-\)Ca\(^{2+}\) exchanger (which utilises the favourable Na\(^+\) electrochemical gradient to drive Ca\(^{2+}\) out of the cell), as well as by the Ca\(^{2+}\)-ATPase present on the sarcolemma, and uptake into the mitochondria.
Figure 1.10: Schematic diagram showing the 4 stages of the cross-bridge cycle. Image taken from Cooke, R. (2004).
1.3.3 Regulation of Contraction

1.3.3.1 Intrinsic Regulation

A ventricle can modify the magnitude of contraction, i.e. the total amount of force generated to decrease the volume of the chamber, depending on the extent to which it is filled at the end of diastole. Increasing end-diastolic volumes (EDV) lead to a more rapid increase in ventricular pressure, a higher peak pressure, and a more rapid rate of relaxation. Thus, a ventricle filled with more blood at the end of diastole, will pump more blood out during systole. This is known as the Frank-Starling relationship, also referred to as length-dependent activation. There are a number of proposed mechanisms (reviewed in Opie, L. H. 2004). One is that stretching of the cardiomyocyte with constant volume, results in decreased distance between thick and thin filaments (with titin and myosin binding protein C possibly playing a role), increasing the probability of cross-bridge cycling at any particular Ca\textsuperscript{2+} concentration. The second is that cross-bridge formation may result in neighbouring actin and myosin molecules being brought into greater vicinity, increasing the likelihood of cross-bridge cycling at those sites. Additionally, myosin ATPase activity increases in response to stretch, and through as yet uncertain mechanisms, stretch seems to increase the opening of Ca\textsuperscript{2+} channels on the sarcolemma leading to greater CICR. Finally, any increase in heart rate (HR), whether mediated by enhanced sympathetic stimulation (see below) or otherwise, leads to an increase in contractility via accumulation of Ca\textsuperscript{2+} in the SR, with greater Ca\textsuperscript{2+} release and cross-bridge formation in successive heart beats. This is known as the Bowditch phenomenon or Treppe effect (Opie, L. H. 2004).
1.3.3.2 Regulation by the Autonomic Nervous System

Both the sympathetic and parasympathetic divisions innervate the heart, and branches from both of these systems form cardiac plexuses. The deep cardiac plexus is located anterior to the carina of the trachea, and the superficial cardiac plexus is located inferior to the aortic arch. Pre-ganglionic parasympathetic fibres arising from the vagus nerve synapse onto post-ganglionic fibres either in these plexuses, or pass through the plexuses to synapse in the walls of the atria. Activation of these fibres results in a decrease in SA node pacemaker activity and thus, a decrease in HR.

Pre-ganglionic sympathetic nerves arising from T1 to T5, pass into the sympathetic trunk and synapse in cervical and upper thoracic sympathetic ganglia. The post-ganglionic fibres pass to the plexuses and mix with the parasympathetic fibres to form nerves that innervate the SA node, as well as the entire atrial and ventricular musculature. Activation of these fibres leads to increases in HR (positive chronotropic effect), contractility (positive inotropic effect), as well as an increase in the rate of relaxation (positive lusitropic effect). This positive lusitropic effect ensures that ventricular filling during diastole does not become limiting at elevated heart rates.

The sympathetic nervous system primarily acts through β1-adrenergic receptors located on the sarcolemma (reviewed in Opie, L. H. 2004). They are activated by norepinephrine released at sympathetic nerve terminals, as well as epinephrine (released from the adrenal medulla in response to sympathetic stimulation). Activation of this G-protein coupled receptor leads to activation of adenyl cyclase which increases cyclic adenosine monophosphate (cAMP) levels within the cardiomyocyte. cAMP binds to and activates protein kinase A, which then phosphorylates multiple cellular targets. These
include the L-type Ca\(^{2+}\) channel that leads to greater Ca\(^{2+}\) entry and the RYR2, leading to increased CICR. These changes are responsible for the positive inotropic effect of adrenergic signalling. The positive lusitropic effect is mediated partly by phosphorylation of proteins of the thin filament that leads to reduced Ca\(^{2+}\) sensitivity (i.e. probability of cross-bridge formation at any particular Ca\(^{2+}\) concentration), as well as by phosphorylation (causing inhibition) of phospholamban, a protein that normally inhibits SERCA. The net result is faster Ca\(^{2+}\) reuptake and thus enhanced relaxation.

1.4 Heart Failure

1.4.1 Pathophysiology

Heart failure (HF) can be defined as the inability of the heart to generate adequate blood flow to supply the needs of the tissues of the body (Opie, L. H. 2004). Depending on the degree of HF, CO may be adequate at rest, but under conditions of increased demand, such as exercise, the heart cannot increase its CO to match, i.e. there exists decreased cardiovascular reserve with HF. In patients with end-stage HF, CO is inadequate even at rest, and the patient has intractable symptoms of dyspnoea, fatigue and drowsiness.

The defects of cardiac contraction can be separated into two interrelated concepts: 1) systolic dysfunction, i.e. inadequate force generation, and 2) diastolic dysfunction, i.e. inability of the ventricles to relax to allow adequate ventricular filling. The various clinical entities that patients are diagnosed with have features of both systolic and diastolic dysfunction.
With systolic dysfunction, there is a decrease in the rate of increase of ventricular pressure and a decrease in peak ventricular pressure. Stroke volume (SV) is defined as the volume of blood pumped out of the ventricle during systole. One measure used to quantify the extent of systolic dysfunction is the ejection fraction (EF), defined as the SV/EDV. With a decreased EF, the LV EDV increases, and this results in greater contractility via Frank-Starling mechanisms. This somewhat compensates for the otherwise decreased contractility, maintaining the stroke volume, which assists in maintaining CO. However, the consequence is a reduction in the scope for further increases in contractility in response to increased venous return to the LV in times of increased CO requirement. In addition, a more dilated LV has increased wall stress during systole as defined by the Laplace Law: Wall Stress = (Pressure * Ventricle Radius) / (2 * Ventricle wall thickness) (Opie, L. H. 2004). This is a measure of the tension within the ventricle wall that must be overcome by the opposing tension generated by cycling cross-bridges in order to pump blood out of the ventricle. Pressure in the wall stress equation refers to the blood pressure in the ventricular cavity and aorta during systole, when the aortic semi-lunar valve is open (also referred to as afterload). Wall stress is a major determinant of oxygen and nutrient consumption per unit mass of the ventricle wall. The response to increased wall stress is LV hypertrophy (LVH), in an attempt to normalise wall stress. LVH as a reaction to systolic dysfunction occurs via an increase in the size of individual cardiomyocytes (cell hypertrophy), as well as by an increase in the number of cardiomyocytes (hyperplasia) (Quaini, F. et al. 1994; Kajstura, J. et al. 1998). However, such hypertrophy is not perfectly compensatory, with inadequate angiogenesis having been suggested to contribute to mild ischaemia and dysfunction of the greater
cardiomyocyte population (Rakusan, K. et al. 1992). Regardless of the underlying mechanisms, there are absolute reductions in phospho-creatine, ATP and the ratio of phospho-creatine/ATP, as well as flux through the creatine kinase pathway in myocardium from patients with end-stage HF, demonstrating a state of energy deficiency (Ten Hove, M. and Neubauer, S. 2007). Note that the levels of phospho-creatine, ATP and the ratio of phospho-creatine/ATP, as well as flux through the creatine kinase pathway in non-failing (NF) LV myocardium do not change with age (Bottomley, P. A. et al. 2009). Additionally, the rate of apoptosis is increased in the failing heart (Olivetti, G. et al. 1997), with higher rates in males than females (Guerra, S. et al. 1999). Furthermore, apoptotic rate has been demonstrated to correlate with the rate of disease progression in HF (Saraste, A. et al., 1999). Necrosis is also increased in HF (Guerra, S. et al. 1999), also with higher rates in males than females. Loss of cardiomyocytes by apoptosis and necrosis increases the load on the remainder, promoting a vicious cycle of haemodynamic overload.

Diastolic dysfunction is promoted by LVH via several mechanisms, including increased interstitial tissue deposition (mainly collagen), as well as decreased energy availability which impairs ATP-dependent reuptake of Ca²⁺ back into the SR by SERCA. This in turn, decreases EDV, which as a result, decreases stroke volume and hence, CO.

Taken together, these defects in cardiac function lead to impairment of LV emptying, leading to increased hydrostatic pressure within the pulmonary circulation and pulmonary oedema. The result is decreased oxygen diffusion, as well as increased lung stiffness leading to the symptoms of dyspnoea, and at end-stage, cyanosis. Back pressure from the failure of the LV and RV, leads to increased pressure in the venae cavae,
resulting in hepatic congestion and dysfunction. Decreased CO (and hypoxia in end-stage HF) leads to abnormalities of cerebral, renal and skeletal muscle perfusion, leading to encephalopathy, renal impairment, and increased fatiguability, respectively.

Compensatory neurohumoral mechanisms that attempt to maintain adequate perfusion of the heart and brain act by increasing systemic arterial pressure through actions on the heart and vasculature. These include increased sympathetic nervous system stimulation of the heart itself via the cardiac nerves to increase contractility, HR, and rate of relaxation, thus increasing CO. Sympathetic stimulation also leads to peripheral arteriolar vasoconstriction in viscera, skeletal muscle, and skin, thereby increasing systemic vascular resistance, and hence, systemic arterial pressure. Sympathetic stimulation of the kidneys leads to increased Na\(^+\) and water reabsorption and decreased urine formation via afferent arteriolar vasoconstriction and activation of the renin-angiotensin-aldosterone system. This results in an increase in total blood volume and venous return to the LV to increase CO, albeit at the risk of exacerbating pulmonary oedema. However, these compensatory mechanisms also increase afterload, and thus increase wall stress. This raises the energetic demands on the cardiomyocytes, promoting dysfunction and cell death, with progressive worsening of HF chronically.

End-stage HF eventually results in death via either progressive HF and circulatory shock with multi-organ failure, sudden death from cardiac arrhythmia, or stroke due to embolism of thrombi formed on the wall of the dilated LA or LV.
1.4.2 Ischaemic Cardiomyopathy

Myocardial infarction (MI) is the most common cause of HF in the Western world. The aetiology of HF due to ischaemic cardiomyopathy (ISCM) is the loss of contractile ventricular tissue from the infarcted site itself, as well as more generalised loss of cardiomyocytes with replacement fibrosis, as a result of atherosclerotic arterial occlusion across the LV (Beltrami, C. A. et al. 1994). The remaining cardiomyocytes display compensatory hypertrophy. This results in systolic dysfunction, with superadded impairment of energy-dependent relaxation due to ischaemia (diastolic dysfunction), playing an additional perhaps more minor role. The result is an increase in LV EDV, a dilated heart, with increased wall stress partially compensated by LVH.

1.4.3 Idiopathic Dilated Cardiomyopathy

Idiopathic Dilated Cardiomyopathy (IDCM) can be considered as a phenotypic group that encompasses patients with HF from a variety of different aetiologies including mutations in cytoskeletal and sarcomeric genes as well as previous subclinical myocarditis and other occult causes (Opie, L. H. 2004). All these various pathologies lead to a similar phenotype characterised by predominantly systolic dysfunction, with a dilated LV and varying levels of LVH.

1.4.4 Hypertrophic Cardiomyopathy

Hypertrophic cardiomyopathy (HCM) is another LV myocardial pathology that predisposes to the development of HF. It is characterised by gross LVH with markedly
reduced LV chamber volume, diastolic dysfunction, and relatively preserved systolic function. For the majority of HCM patients, an underlying genetic cause is found, and these include mutations in a variety of sarcomeric proteins (Opie, L. H. 2004). The common underlying pathophysiology behind HCM may be a deficit in force generation and increased Ca$^{2+}$ sensitivity (Hoskins, A. C. et al. 2010), leading to systolic and diastolic dysfunction, respectively. This initial defect then leads to compensatory, but ultimately excessive LVH (and associated ischaemia), that exacerbates the diastolic dysfunction, and finally results in the development of end-stage HF.

1.4.5 Recovery with LV Assist Device Implantation

A Left Ventricular Assist Device (LVAD) is a mechanical device implanted into patients with LV failure that receives blood from the LV, and then pumps the blood into the ascending aorta, thereby unloading the LV and maintaining adequate CO within the patient. In several trials, LVAD insertion has been shown to improve survival and symptoms in patients with end-stage HF, whether it is being used as a bridge to transplantation or as long-term therapy (reviewed in Birks, E. J. 2010). In a minority of patients, recovery of LV function has been such that the patient no longer required LVAD assistance, nor heart transplantation. In these instances, the patient's hearts seem to have undergone a partial, or almost complete reversal of the defects associated with their HF phenotype.
1.5 The Effects of Age and Gender on the Human Heart

1.5.1 Introduction

The changes in the human heart due to the ageing process are often accompanied by changes recognised as due to the development of pathologies, such as narrowing of the coronary arteries and microcirculation due to atherosclerosis, superimposition of valvular defects associated with older ages, such as calcific aortic stenosis, and progressive systemic arterial hypertension. Different investigators have attempted to separate the effects of the development of pathology from the unremitting ageing process by analysing patients without clinically apparent pathology, however this does not eliminate the effects of sub-clinical pathology. On the other hand, one could argue that these sub-clinical pathologies form part of the ageing process of the heart. I think that the study of the hearts of ageing patients, both with and without overt cardiovascular pathology, have merit. An analysis of the latter population may more properly represent the ageing phenotype, and an analysis of the former population, is more representative of the average patient found in the clinic, who will have an aged heart that has also been subjected to the myriad pathologies associated with older ages.

1.5.2 Structural Changes with Age and Gender

1.5.2.1 Macroscopic Changes with Age and Gender

In the Baltimore Longitudinal Study of Ageing cohort, male and female participants were selected who had no chest wall abnormalities, coronary heart disease (asymptomatic and
with negative stress tests), valve dysfunction, or hypertension. Assessed by echocardiography, it was found that from the ages of 25 to 85 there was an increase in posterior LV wall thickness by ~ 50% (Gerstenblith, G. et al. 1977). In the Framingham cohort, a similar increase with age of LV mass was observed in both men and women (Savage, D. D. et al. 1990). However, in a subset of patients without cardiopulmonary disease, hypertension or obesity, LV mass did not increase in men, and the increase in women was highly attenuated (Dannenberg, A. L. et al. 1989). Echoing these latter findings, Olivetti, G. et al. (1995) showed that in a similar population of healthy individuals that there was no change in wall thickness with age in both the LV and RV of both genders. In fact, it was found that the weight of the myocardium decreases in males with age (0.78 g per year in LV, 0.16 g per year in RV) while remaining constant in females.

Thus, in the absence of pathology, myocardial mass may remain the same with age, but in response to the superimposition of age-related pathology, there may be varying amounts of compensatory LVH.

1.5.2.2 Cardiomyocyte Compartment Changes with Age and Gender

In studies of human hearts that were grossly normal, as well as hearts with eccentric and concentric hypertrophy due to valvular defects, hypertension, chronic pulmonary diseases and congenital abnormalities, no change in cardiomyocyte numbers in either the RV or LV were detected with age (Arai, S. and Machida, A. 1972; Tadokoro, M. and Arai, S. 1972). In contrast, in a group of hearts from patients without significant cardiac related pathology, cardiomyocyte loss has been shown to occur with age (Olivetti, G. et al. 1995;
Olivetti, G. et al. 1991). According to Olivetti, G. et al. (1995), in males, there is a loss of approximately 45 million cardiomyocytes per year in the LV and 19 million cardiomyocytes per year in the RV, corresponding to a decrease of ~ 1 % of the total number of cardiomyocytes per year. In 20 year old males, the LV contains approximately $5.8 \times 10^9$ myocytes of which $5.4 \times 10^9$ are mononucleated and $0.4 \times 10^9$ are binucleated. Due to the decrease in cardiomyocytes with age, in 70 year old males, the LV contains approximately $3.6 \times 10^9$ cardiomyocytes of which $2.9 \times 10^9$ are mononucleated and $0.7 \times 10^9$ are binucleated. Note that in addition to the decrease in overall cardiomyocyte numbers in the LV and RV, there is an increase in the proportion of binucleated cardiomyocytes. This loss of LV cardiomyocytes with age in males may account for the decrease in myocardial weight observed. The female LV, on the other hand, contains on average $4.31 \times 10^9$ cardiomyocytes of which $3.6 \times 10^9$ are mononucleated and $0.67 \times 10^9$ are binucleated, and there was no change with age (Olivetti, G. et al. 1995). Similarly, there were no changes in cardiomyocyte numbers in the RV. The mean cardiomyocyte volume per nucleus was ~ $20,000 \ \mu m^3$ in both men and women at 20 years, and simultaneous with the loss of cardiomyocytes with age in the male heart, mean cardiomyocyte cell volume increased by $158 \ \mu m^3$ per year in the LV and $167 \ \mu m^3$ per year in the RV (Olivetti, G. et al. 1995). There was no change in mean cardiomyocyte cell volume in the female heart. Taken together, it seems that loss of cardiomyocytes in the ageing male heart is associated with hypertrophy of the remaining to cope with an unchanged haemodynamic load, whereas this does not seem to occur in women. In contrast, in the presence of pathology, the number of cardiomyocytes does not tend to
decrease implying a compensatory hyperplastic response to cope with changes in load as a result of superimposition of pathology.

1.5.2.3 Non-Cardiomyocyte Compartment Changes with Age

Collagen content can be used as a surrogate marker for fibroblast abundance and activity. Wegelius, O. and Von, K. (1964) analysed macroscopically normal hearts and did not find a change in collagen content with ageing in either ventricle. Caspari, P. G. et al. (1977) analysed hearts that did not show valvular abnormalities or evidence of ischaemic heart disease at autopsy and also found no relationship between collagen content and age. Furthermore, in a subset of hearts exhibiting hypertrophy, collagen concentration in the LV remained the same. Similarly, Mollova, M. et al. (2013) did not observe a change in fibrotic tissue with age. Olivetti, G. et al. (1991) however, found that rarely, areas of replacement fibrosis and interstitial and perivascular fibrosis are observed in the aged heart.

Little is known about the changes in the microvascular compartment that are associated with age in the human heart. In a study by Rakusan, K. et al. (1992) of hearts from patients who died of non-cardiac causes, with no evidence of coronary disease, mitral regurgitation or aortic valve dysfunction, capillary density decreased from childhood (9 to 14 years) to adolescence and early adulthood (15 to 30 years). However, changes with more advanced age are currently unknown.

It is clear that more work needs to be done to describe adequately the remodelling of the non-cardiomyocyte compartment that occurs during the ageing process in humans.
1.5.3 Functional Changes with Age and Gender

There are a number of age-related changes in human cardiac function that have been described, both at rest, and during times of increased haemodynamic requirement.

CI at rest declines slightly in women but not in men, with age (Fleg, J. L. et al. 1995; Rodeheffer, R. J. et al. 1986). There is a small increase in LV EDV index (LV EDV normalised for body surface area, LV EDVI), LV end systolic volume index (LV ESVI), and stroke volume index (SVI) in men, but not in women (Fleg, J. L. et al. 1995).

However, Rodeheffer, R. J. et al. (1986) did not confirm increases in cardiac volumes at rest in men at older ages. These possible increases in cardiac volumes in the male along with an increased SVI, are opposed by a larger decline in HR at rest with age, as compared to women who experience a smaller decline in resting HR with age. With higher EDVI in males, and higher systolic blood pressure in both sexes, there is an increase in afterload. LV EF at rest, which is consistently higher in females, in accordance with the above findings, does not decline with age (Fleg, J. L. et al. 1995). LV early diastolic filling rate however, does decrease with age in both males and females (Schulman, S. P. et al. 1992; Benjamin, E. J. et al. 1992; Swinne, C. J. et al. 1992). This decrease is countered by an enhanced late filling rate due to more vigorous atrial contraction that maintains diastolic filling (Swinne, C. J. et al. 1992). This decrease in early diastolic filling may represent an increase in stiffness due to changes in the interstitial tissues of the heart, or defects in relaxation at the level of the myofibrils, due to residual Ca²⁺ from the previous systole. In any case, these results imply a maintenance of almost normal cardiac function at rest at older ages, in both males and females.
At maximal exercise, peak CI is reduced in both males and females at older ages (Fleg, J. L. et al. 1995). Maximal HR declines with age in both men and women, and is thought to be primarily responsible for the decrease in maximal CI. EDVI increases in older men during maximal exercise compared to younger men, who do not experience an increase (Fleg, J. L. et al. 1995). This age-related change in EDVI during peak exercise occurs to a lesser extent in women. EF at maximal exercise decreases with age in both sexes, but to a lesser extent in women (Fleg, J. L. et al. 1995). However, coupled with a greater EDVI during exercise, especially in older men, SVI is preserved in both sexes at peak exercise compared to younger people. These results imply that there is a greater reliance on Frank-Starling mechanisms to maintain SVI at older ages, and also, that during maximal exercise, to achieve the same SVI, a higher EDVI is required in men, implying decreased contractility in male hearts compared to female hearts (Fleg, J. L. et al. 1995). As a further piece of evidence of decreased contractility, the same study found that during maximal exercise, LV end systolic pressure/ESV (a marker of LV contractility) decreases with age in both males and females, with the decrease being possibly steeper in males.

There are a variety of possible mechanisms that may explain the change in human LV function with age, at both the microscopic and macroscopic levels. Given the decreased maximal HR, a decrease in contractility would be expected, due to decreased Ca\textsuperscript{2+} accumulation in the SR secondary to the Bowditch phenomenon (Lakatta, E. G. and Levy, D. 2003). In addition, there is decreased post-synaptic adrenergic responsiveness in the older heart which manifests as lower contractility, impaired diastolic filling, as well as lower HR (Opie, L. H. 2004). It was earlier described that in the absence of cardiovascular pathology, there is an age-related loss of cardiomyocytes (with
compensatory cellular hypertrophy) in males, and this may partially contribute to the
greater decrease in systolic function with age in males, at least during states of increased
haemodynamic requirement. However, in the presence of increasing afterload and/or the
development of other cardiovascular pathology with age, LVH occurs as a response in
both sexes (Savage, D. D. et al. 1990; Gerstenblith, G. et al. 1977). This
hypertrophy/hyperplasia of cardiomyocytes may not be coupled with a corresponding
increase in blood vessels (Rakusan, K. et al. 1992), which may lead to impaired oxygen
and nutrient delivery to cardiomyocytes and a combination of systolic and diastolic
dysfunction (Schulz, R. et al. 2001). This would be exacerbated by the presence of age-
related sub-clinical atherosclerotic narrowing of the macro- and micro-circulation.

1.5.4 The Ageing Cell

Another major factor that determines how the human heart changes with age is the ageing
of the cardiomyocytes themselves, and how that ageing impacts on their ability to
generate force, repair and renew themselves.

All cells are subject to the damaging and cumulative effects of oxidative stress,
and the cardiomyocyte may be particularly vulnerable. Reactive oxygen species (ROS)
are a natural byproduct of cellular respiration, and these in turn can oxidatively modify
deoxyribonucleic acid (DNA), proteins and lipids, and lead to a derangement in their
function. The mitochondrion is the major source of ROS in the cardiomyocyte, and is
also the most affected. Mitochondrial DNA (mtDNA) mutations accumulate with age as a
result of oxidative damage, and these in turn result in accumulating mitochondrial deficits
in functioning, including decreased oxidative phosphorylation, as well as increased ROS
formation and further mtDNA damage (Judge and Leeuwenburgh, 2007). As an example of the ageing related defects in mitochondrial function, decreased activities and/or levels of complexes I, III, IV and V have been demonstrated in older hearts in a variety of animals (Judge and Leeuwenburgh, 2007).

Autophagy is a process whereby cellular components, including mitochondria, are entrapped into double-membrane vacuolar structures called phagosomes. These phagosomes fuse with lysosomes, and the cellular components within are digested. This is a mechanism for eliminating damaged organelles and other intracellular structures. Autophagy decreases with age, and this is at least partly due to accumulation of indigestible material (lipofuscin, an aggregate of protein and lipid) within lysosomes that disrupts their function (Terman and Brunk, 2005). This then causes an age-related accumulation of damaged mitochondria which would otherwise be targeted for destruction, and thus further increased ROS production. In addition, one would expect the accumulation of other damaged organelles and cellular structures to decrease cellular functioning with age.

Despite these mechanisms for decreased cardiomyocyte function, there is no increase in the rate of apoptosis with age in humans without cardiovascular disease (Mallat, Z. et al. 2001). However, this does not preclude a difference in susceptibility to apoptosis with the imposition of pathology. In addition, rates of apoptosis are higher in men than women, which may partly account for the maintenance of cardiomyocyte numbers in aged female hearts described above.

Alongside the degradation in the structure and function of cardiomyocytes with age, and their elimination by apoptosis (and necrosis), there also occurs a beneficial
replacement of such cells from either resident stem cells and/or division of differentiated cardiomyocytes. Bergmann, O. et al. (2009) demonstrated that cardiomyocyte renewal occurs in the human heart, and that it is negatively correlated with age. At age 25, the rate of renewal is ~1 % per year, compared to at 75 years, when the rate of renewal is ~ 0.45 %. The data also suggest that the probability of turnover is similar among all the cardiomyocytes. Kajstura, J. et al. (2012) report complementary findings, demonstrating that the number of phosphorylated histone H3 (marker of mitosis) positive cardiomyocytes declines with age (from 0.01% at 1-10 years, to 0.006% above 40 years of age). Similar findings are presented by Mollova, M. et al. (2013), i.e. decreased phosphorylated histone H3 positive cardiomyocytes with age, however, they could not demonstrate contractile rings (required for cytokinesis) in cardiomyocytes from patients > 20 years of age. Therefore, mitosis among cardiomyocytes may be decreasing with age, and the probability that each mitotic event is accompanied by cytokinesis decreases as well, which could account for the increase in the proportion of polyploid, polynucleated cardiomyocytes with age, as was mentioned earlier.

Additionally, Bergmann O. et al. (2009) found the rate of turnover of non-cardiomyocyte cells to be ~ 18 % per year, and this rate did not appear to change with age.

Therefore, there appears to be replacement of damaged and dying cardiomyocytes in the human heart, however this replacement proceeds at a reduced rate at older ages, which would result in the accumulation of less functional cardiomyocytes in the older heart. Decreased renewal can partly explain why the number of cardiomyocytes can decrease with age in males, even if rates of apoptosis remain steady. This decreased
renewal may be due to the accumulation of damage within cardiac stem cells themselves decreasing their proliferative capacities, or to decreased division of differentiated cardiomyocytes. Furthermore, from the cells that continue to proliferate (differentiated cardiomyocyte or stem cell), will arise cardiomyocytes that inherit the accumulated damage of their parent cells. Thus, with age, there are changes in the structure of the heart, at the tissue, cellular, and subcellular levels, and these changes in turn, lead to changes in the function of the heart.

1.5.5 The Risk of HF as Modified by Age and Gender

There is an increase in coronary heart disease (angina pectoris and MI) with age, as well as an increase in HF with age (National Heart, Lung and Blood Institute, 2006). In a large post MI cohort, ~ 1% developed HF per year (risk of HF plus all-cause mortality was ~ 2% per year) (Lewis, E. F. et al. 2003). In the above study, age was the most powerful predictor of HF, with each 1 year increase in age at time of MI associated with a 7 % increase in risk of HF. The second most powerful predictor of development of HF was baseline LV EF post MI. With respect to MI treated with emergent revascularisation, Brodie, B. R. et al. (2006) found that patients over 70 years of age had a greater risk of late mortality with a median follow up of 7 years. Similarly, Parodi, G. et al. (2007) found that age was a significant independent predictor of 5 year mortality. Note that HF is one of several causes of late mortality. With respect to HF only, Terkelsen, C. J. et al. (2011) found that each additional decade in age increased the risk of presenting with HF, with a follow up of up to 7 years, post MI treated with emergent revascularisation. Thus it seems that the heart of an older person does not cope as well with the superimposition of
pathology such as a MI. This decreased coping ability may extend to other cardiac insults as well. In the presence of diagnosed HF, the risk of death is increased with age, and interestingly, the survival advantage of HF with preserved EF is diminished as well with increasing age (MAGGIC et al. 2012).

With respect to gender, all-cause mortality in patients with HF is greater in men than in women, adjusted for age (Martinez-Selles, M. et al. 2012). There are many proposed mechanisms for such a difference in risk, some of which derive from the data in this thesis, and will be presented later. As mentioned earlier, female LV maintain cardiomyocyte numbers with age, thus with the onset of HF there are a greater number of cardiomyocytes to share the load, and they are less likely to undergo apoptosis/necrosis. The cardiomyocytes in ageing males will be larger, which may affect diffusion of nutrients and waste products. In addition, for as yet unclear reasons, females show lower rates of RV dysfunction (measured as low RV EF or large RV EDV) for the same degree of LV systolic dysfunction (Martinez-Selles, M. et al. 2006). Looking at a different aspect of the gender divide in terms of cardiovascular disease, Brodie, B. R. et al. (2006) found no difference in the risk of late mortality during 10 years of follow up post MI treated with emergent revascularisation. However, Parodi, G. et al. (2007) found that female gender was independently associated with recovery of LV EF post MI treated with emergent revascularisation, though this was not observed in a similar study by Van 'T Hof, A. W. et al. (1998). In the same context, Lewis, E. F. et al. (2003) found that gender was not related to HF risk post MI, despite what one might expect from the increased likelihood of LV EF recovery observed by Parodi, G. et al. (2007). However, there are many factors that affect likelihood of HF development including recurrent MI,
differences in treatment among men and women, and others that would obscure the effect of greater LV EF recovery.

1.6 Focus of Thesis

The major purpose of this thesis is to describe the changes that occur in the ageing human heart with respect to gene expression, and then to relate these changes to what occurs in the patient with end-stage HF. I surmise that this approach will shed light on why an older heart will succumb more easily to HF as a consequence of the superimposition of pathology such as a MI or other cardiac insults. With this knowledge we may discover pathways to target for intervention, to reduce the probability and/or rate of development of HF in older patients. Similarly, an understanding of how females seem to have a better prognosis in HF may also assist in the discovery of future therapeutic interventions that can be applied to both genders.
Chapter 2: General Methods

2.1 Collection of Human Heart Tissue

2.1.1 Human Research Ethics Committee Approval

Prior to collection of human heart samples, the Human Research Ethics Committee of the University of Sydney (#09-2009-12146), St. Vincent's Hospital (H91/048/1a), and the Australian Red Cross Blood Service approved the project. Patients undergoing transplantation gave informed consent prior to the collection of end-stage HF heart samples, and the families of heart donors gave informed consent prior to the collection of NF heart samples for research.

2.1.2 Collection of Human Heart Tissue

Collection of all heart tissue was undertaken at St. Vincent's Hospital. NF hearts were obtained from donors who had died from non-cardiac causes but whose organs were maintained by life support until permission was obtained from donor’s next-of-kin. Only donor hearts that were not suitable for cardiac transplantation were used for research. Donor hearts were rejected for transplantation for a range of reasons, but a common cause for rejection was the unavailability of a tissue-compatible recipient. All donor hearts were excised from patients while the donors were on life support, immediately flushed with chilled cardioplegic solution, and transported to Sydney on ice within four
hours of loss of coronary blood flow. Transmural blocks of LV (weighing approximately 1 g) were frozen in de-identified coded cryovials, stored and maintained in liquid nitrogen. With respect to end-stage HF hearts, these were excised from patients undergoing transplantation within 30-40 minutes of aortic cross-clamping, and samples were immediately dissected and stored as described for NF hearts.

2.2 mRNA Expression Analysis

2.2.1 RNA Extraction

Extraction of total RNA was carried out using Trizol reagent (Invitrogen) essentially as described in the manufacturer's protocol. All the following steps were carried out at room temperature unless otherwise stated. Approximately 250 mg of LV tissue was homogenised using a rotor stator homogeniser (PRO200, PRO Scientific, Oxford, CI, UK), in 1 mL of ice-cold Trizol. Specifically, samples were homogenised for 30 seconds at maximum speed and cooled on ice for 1 minute, and this was repeated 2 further times to ensure complete homogenisation while maintaining the temperature of the sample near zero. This was done to ensure minimal RNase-mediated destruction of sample RNA prior to full homogenisation and denaturation of RNase enzymes present in sample. The homogenates were then stored at –80° C. The homogenates were thawed and centrifuged at 12,000 x g at 4° C for 10 minutes and the pellet containing insoluble material (i.e. high molecular weight DNA, extracellular matrix, polysaccharides) was discarded. 0.2 mL of chloroform was added to each supernatant and the mixture was vigorously shaken for 30 seconds. The samples were left to incubate for 2 to 3 minutes and then centrifuged at
12,000 x g at 4°C for 15 minutes. During this step, the mixture separated out into an upper aqueous layer (approximately 60% of volume, containing the total RNA), a milky white interphase (containing DNA), and a lower, pink organic layer (containing protein). 400 µL (approximately 2/3 of the volume) of the aqueous layer was carefully removed and mixed with 0.1 mL of ice-cold isopropanol, and the RNA was left to precipitate for 10 minutes on ice. To pellet the RNA, the mixture was then centrifuged at 12,000 x g at 4°C for 10 minutes. The supernatant was discarded and the pellet washed with excess ice-cold 70% ethanol (EtOH) twice, vortexing for 10 seconds, and centrifugation at 7,500 x g at 4°C for 5 minutes. Finally, the supernatant was completely removed and the pellet left to air-dry. The RNA was resuspended in nuclease-free water, incubating at 55°C for 10 minutes and the resuspended RNA was then treated with Amplification-grade DNase I (Invitrogen) to remove residual genomic DNA contamination. For each LV sample, 20 µg total RNA was incubated with 20 U of DNase I in DNase I reaction buffer (20 mM Tris-HCl, 2 mM MgCl₂, 50 mM KCl) for 15 minutes. To halt the reaction, 20 µL of 25 mM EDTA was added and the mixture was heated to 65°C for 10 minutes. Post DNase treatment, the mixture was purified using a GenElute Mammalian Total RNA Isolation Kit (Sigma) according to manufacturer's instructions. Total RNA was quantified and assayed for purity (260/280 nm and 260/230 nm ratios) using the NanoDrop 1000 (Thermo Scientific) and all RNA samples used for further experiments exhibited A260/A280 ratios > 1.7 and < 2, and A260/A230 ratios > 2. RNA integrity was analysed using the 2100 BioAnalyser (Agilent) and only samples with an RNA integrity number > 6 were used for gene expression experiments (Fig. 2.1). The RNA integrity number is a ranking from 0 (completely degraded mRNA) to 10 (completely intact mRNA) generated
by an algorithm that analyses the electropherogram from the 2100 Bioanalyser for specific features including the sizes of the 28s and 18s peaks and their relation to the total signal, plus the sizes of the downstream regions reflecting the presence of smaller fragments of degraded RNA (Schroeder, A. et al. 2006).

![Bioanalyser electropherogram](image)

**Figure 2.1:** Bioanalyser 2100 (Agilent) electropherogram demonstrating RNA integrity. ‘FU’ stands for fluorescent units, and ‘s’ for seconds.

Note, for RNA samples that were to be used for real-time PCR, DNase treatment and column purification were not performed. Given there is no PCR-like amplification step in the preparation of RNA for microarray analysis, contaminating genomic DNA might make up a significant proportion of total nucleic acid, that might interfere with hybridisation of labelled cDNA to targets on the glass slides, hence the desirability of prior DNase treatment. This was followed by column purification to remove residual...
DNA fragments, salts and DNase enzyme. With respect to real-time PCR, given the primers were designed to cross exon-exon boundaries, there would be no amplification of corresponding r-protein DNA. Pseudogenes present in contaminating genomic DNA may be amplified, however, the number of pseudogene copies to mRNA copies for a particular r-protein is likely at least several orders of magnitude lower, therefore, the effect on Ct values would be insignificant. Instead, an additional RNA precipitation step was used to further purify the RNA samples to maximise the efficiency of downstream enzymatic reactions. To perform this RNA precipitation step, Na acetate, pH 5.2 was added to a final concentration of 0.3 M plus 3 volumes of absolute EtOH, and the samples were incubated overnight at -20°C. The samples were centrifuged at 12,000 x g for 10 minutes at 4°C, and then washed with 70% EtOH, recentrifuged at 7,500 x g for 5 minutes, and then allowed to dry. RNA was resuspended as above and then analysed for purity and integrity as above.

2.2.2 Microarray Analyses

See Sections 3.2 and 4.2

2.2.3 Real-Time Polymerase Chain Reaction Analysis

950 ng of total RNA extracted from each human LV sample was mixed with 50 ng of yeast total RNA (kindly supplied by Dr. Grace Wei, Victor Chang Cardiac Research Institute). The addition of yeast total RNA was carried out so that Scr1, a highly abundant yeast transcript, could be used as a spike-in control to normalise the real-time PCR results
to control for reverse transcription (RT) and polymerase chain reaction (PCR) efficiency differences between RNA samples (Kubista, M. et al. 2006). RT was carried out using the SuperScript® III First-Strand Synthesis System for RT-PCR (Invitrogen). Each RNA sample was made up to 11 µL with nuclease-free H₂O, and 2 µL of a master mix (containing 1 µL of dNTP (10 mM) and 1 µL of random hexamers (50 ng/µL) per RT reaction) was added. This mixture was incubated at 65° C for 5 minutes, and the samples were returned to ice. 7 µL of a second mastermix (containing 1 µL DTT (0.1M), 1 µL RNAseOUT (40 Units/µL), 1 µL Superscript III (50 Units/µL), and 4 µL 5 x first strand buffer) was then added. The mixture was incubated at 50 degrees C for 2 hours, then 75°C for 15 minutes. RT products were stored at -80° C until PCR was carried out.

Table 2.1: Primer sequences for all genes analysed with real-time PCR. Note: primer sequences for Scr1 were taken from Miller, A. et al. (2008), and primers for 45s rRNA were taken from Uemura, M et al. (2012). RHEB, Ras homolog enriched in brain and MTOR, Mammalian target of rapamycin.

<table>
<thead>
<tr>
<th>Gene</th>
<th>Forward Primer (5' to 3')</th>
<th>Backward Primer (5' to 3')</th>
</tr>
</thead>
<tbody>
<tr>
<td>SCR1</td>
<td>CGCGGCTAGACACGGATT</td>
<td>GCACGCGTGCAGAATAGAGAA</td>
</tr>
<tr>
<td>45s rRNA</td>
<td>GAACGGTGTTGTTGTCGTT</td>
<td>GCGTCTCGTCTCGTCTCAG</td>
</tr>
<tr>
<td>MTOR</td>
<td>AGCTGTGCTACACTACAAAC</td>
<td>GGTCTTGACAGATCCTCAG</td>
</tr>
<tr>
<td>RHEB</td>
<td>TTCAACCTTGAGACACAGCC</td>
<td>TCCAAACTTTGGCCATGGG</td>
</tr>
<tr>
<td>RPL23A</td>
<td>CCAAATATTCTCGGAAGAGC</td>
<td>ATGCTTTCTCTCTCCATCA</td>
</tr>
<tr>
<td>RPL26</td>
<td>TAAAGGTCAGCAAATTGGCA</td>
<td>AGTTCTGCATTAGCCCTT</td>
</tr>
<tr>
<td>RPL34</td>
<td>GTGTGCTAAATGTGTCGTT</td>
<td>CTTCTGCCTCTCTGCTTG</td>
</tr>
<tr>
<td>Gene</td>
<td>Forward Primer (5' to 3')</td>
<td>Backward Primer (5' to 3')</td>
</tr>
<tr>
<td>-------</td>
<td>--------------------------</td>
<td>---------------------------</td>
</tr>
<tr>
<td>RPL37</td>
<td>GAACGTCATCGTTTGGAAAG</td>
<td>TCGACTTCTGAAGGTGGTAG</td>
</tr>
<tr>
<td>RPS7</td>
<td>CATGTCGTCTTTATCGCTCA</td>
<td>ATCTAGTTTGACGCGGATTC</td>
</tr>
<tr>
<td>RPS10</td>
<td>TGCCTGGAGACATTCTACT</td>
<td>GGGCAGATGAAGGTAATCAC</td>
</tr>
<tr>
<td>RPS18</td>
<td>AGTTCCAGCATATTTTGCAGA</td>
<td>GTCAATGTCTGCTTTTCCTCA</td>
</tr>
<tr>
<td>RPS21</td>
<td>CGAGTTGCTGGGACCTGTA</td>
<td>CACGTTTCATCTGGATGGATG</td>
</tr>
</tbody>
</table>

Primers for the real-time PCR were designed using Primer-BLAST (Ye, J. et al. 2012) (http://www.ncbi.nlm.nih.gov/tools/primer-blast/) which uses the Primer3 engine (Rozen, S. and Skaletsky, H. 2000). RNA nucleotide sequences transcribed from the various target genes that were used for primer design were derived from the ENSEMBL Genome Browser (http://asia.ensembl.org/index.html). The nucleotide sequences used were selected such that they fulfilled the criteria of being exonic sequences common to all the RNA transcripts deriving from the particular gene. Given that it is was unclear in many cases which was the predominant transcript transcribed from a particular gene (in terms of abundance and in terms of the transcript that would be translated), I decided on a strategy whereby all transcripts would be detected and quantified to ensure that biologically significant regulation would not be missed. This strategy would however, also result in contaminating signal generated by the amplification of non-translated RNA transcripts of unknown biological importance that might make it more difficult to detect the changes in mRNA expression that would result in changes in protein expression that I was interested in quantifying. In cases where there were no overlapping regions common
to all RNA transcripts, the RNA transcripts designated within ENSEMBL to be non-protein coding were discarded. Additional primer design criteria: amplicon to be between 70 and 150 base pairs (bp) (RNA integrity has been shown to have negligible effects on PCR efficiency if amplicons are < 250 bp (Fleige, S. and Pfaffl, M. W. 2006), Tm 57 to 62° C, species is homo sapiens, and the algorithm must use the 'nr' database to look for non-specific amplification. Primers generated by Primer-BLAST which were deemed to not amplify non-specific sequences (except for pseudogenes) were then entered into DNACalc (Sigma Genosys) (http://www.sigma-genosys.com/calc/DNACalc.asp) to characterise their physico-chemical properties in terms of secondary structure (only primers with 'moderate' or less tendency were used), primer dimer (only primers that were deemed to not form primer dimers), and a requirement was set that Tm for each primer within a pair did not differ by more than 4° C. See Table 2.1 for primer sequences used.

For the real-time PCR, 3 µL of RT products (diluted 1:60 in nuclease-free water) were mixed with 1 µl of forward and reverse primers (250 µM) and 5 µL of Lightcycler 480 SYBR Green I Master Mix (Roche). This was done using an epMotion 5075 LH (Eppendorf) liquid handling robot in 384 well plates (Roche). After pipetting, the plates were centrifuged at 800 x g for 2 minutes. Real-time PCR running conditions were as follows unless otherwise specified: 10 minutes at 95° C to heat-activate the Taq polymerase, and then 35 cycles of 1) Binding: 30 seconds at 60°C for Taq polymerase and primer binding to DNA template, 2) Elongation: 72°C for 30 seconds for elongation of DNA strand and acquisition of fluorescence, and 3) Denaturing: 5 seconds at 95°C to denature the double-stranded DNA products. For certain primer pairs that formed non-
specific PCR products as determined from melting curves run post PCR, a modified 'Touchdown' PCR strategy was used: 35 cycles were run as above, but the binding step was modified such that the temperature at which it was carried out was 64°C for the first cycle, and then with each subsequent cycle, the temperature was lowered by 0.25°C. This was done to minimise non-specific binding of PCR primers to non-target template, under the assumption that binding to non-target template would be with a lower affinity than binding to target template, thus, a higher temperature would promote the latter relative to the former. However, there would be an overall decreased amount of binding at a higher temperature resulting in a lower PCR efficiency. With each subsequent cycle, the amount of target template (original copy deoxyribonucleic acid (cDNA) plus newly synthesised amplicons) increases relative to non-target template which promotes primer binding to target template and reduces the requirement for the binding step to be carried out at as high a temperature to ensure intended target amplification. 'Touchdown' PCR was carried out for the genes RPS21 and RPL26.

Melting curves were analysed post real-time PCR run and all primers led to amplification of a single PCR product in the runs used for analysis. This was further confirmed with DNA electrophoresis of the PCR products using a 1% agarose gel.

Pooled cDNA was used at dilutions 1:20, 1:60, and 1:180 to create a relative standard curve. Using these standard curves, Ct values from each PCR reaction were converted into values representing a relative measure of mRNA expression. Each cDNA sample representing a particular LV samples was PCR amplified in triplicate, and the mean of the 3 values obtained was used for statistical analysis. For details of statistical analysis of real-time PCR results, see Sections 3.2 and 4.2.
2.3 Protein Expression Analysis

2.3.1 Protein Extraction

From transmural sections of human LV tissue frozen and stored in liquid nitrogen, 50 to 100 mg was homogenised using a rotor stator homogeniser in 1 ml of ice-cold RIPA buffer (50 mM Tris-HCl, pH 8.0, with 150 mM sodium chloride, 1.0% Igepal CA-630 (NP-40), 0.5% sodium deoxycholate, and 0.1% sodium dodecyl sulfate) (Sigma) plus Protease Inhibitor Cocktail (Sigma) (AEBSF 1.04 mM, aprotinin, 0.8 μM, bestatin 40 μM, E-64 14 μM, leupeptin 20 μM, pepstatin A 15 μM). As was done for the RNA extractions (see above), samples were homogenised for 30 sec at maximum speed and cooled on ice for 1 minute, and this was repeated 2 further times to ensure complete homogenisation. The homogenates were immediately centrifuged at 14,000 x g at 4°C for 10 minutes. The supernatant was collected and stored at -80°C.

2.3.2 Determination of Protein Extract Concentration

The protein concentrations of the LV protein samples were determined using the BCA method described in Smith, P. K. et al. (1985) with the Micro BCA Protein Assay Kit (Thermo Scientific) essentially according to the manufacturer’s instructions. The samples were diluted 1:500 to be within the linear detection range of the assay (i.e. 2 to 40 μg/mL). A Bovine serum albumin standard curve was constructed, with protein concentrations 0.5, 1, 2.5, 5, 10, 20, 40, and 200 μg/mL. 50 μL of freshly made-up
working reagent was mixed with 50 μL of sample in a 96-well polypropylene microtitre plate. This mixture was incubated at 37°C for 2 hours. The plate was then cooled to room temperature, and the PolarStar Galaxy Plate Reader (BMG) was used to quantify absorbance at 562 nm of the standards and the unknown samples. A standard curve of concentration versus absorbance was constructed in Excel (Microsoft) to calculate the concentration of protein in the LV protein samples.

2.3.3 SDS-PAGE and Western Blotting

LV protein samples (40 ng) were prepared for sodium dodecyl sulfate – polyacrylamide gel electrophoresis (SDS-PAGE) by mixing with DTT (0.05 M) and 2 x Laemmli sample buffer (4% SDS, 20% glycerol, 10% 2-mercaptoethanol, 0.004% bromophenol blue, 0.125 M Tris HCl, pH 6.8) and then heating to 99°C to ensure protein denaturation. Protein mixtures were kept on ice post denaturation. Protein samples were loaded onto nUView 4-20% gradient Tris-glycine gels (Nusep) and run at 250V for 35 minutes in Mini-PROTEAN II electrophoresis cells (Bio-Rad) with SDS running buffer (25 mM Tris, 192 mM glycine, 0.1% SDS). Dual color Precision Plus Protein Prestained Standards (Bio-Rad) were used to monitor electrophoresis progression and blotting efficiency. At the completion of SDS-PAGE, gels were visualised using the ChemiDoc MP Imaging System (Bio-Rad). As per the 'Stain-free gel' imaging protocol, the gels were UV irradiated for 1 minute prior to UV-induced fluorescence image acquisition. The nUView gels used contain a trihalo compound that covalently binds to tryptophan residues upon UV irradiation. This compound fluoresces under UV light allowing one to visualise the proteins loaded in each well, and thus confirm equal loading per well, similar to using a housekeeping gene as a loading control, except that the former does not
require assumptions regarding the stability of expression of the housekeeping gene used. A drawback is that given the covalent modification that occurs at tryptophan residues, there is a risk of altering the epitopes to which antibodies will subsequently bind, resulting in a lowering of affinity or even failure to bind, thus affecting the sensitivity of the assay.

Immediately after imaging, proteins were transferred to nitrocellulose using the Trans-Blot Turbo Transfer System (Bio-Rad) using the “Low MW” protocol (2.5 A for 5 minutes), a protocol optimised for the transfer of proteins less than 30 kD. Nitrocellulose membranes were then air-dried overnight to maximise protein binding given the low molecular weight of the proteins blotted.

For antibody staining and visualisation, dried nitrocellulose blots were first blocked with Odyssey Blocking Buffer (Li-Cor Biosciences) for 1 hour at room temperature. All subsequent steps were also carried out at room temperature. Blots were incubated in primary antibody (1:200 for rabbit anti-human RPS18) for 1 hour. Primary antibodies were kindly provided by Dr. Jacob Odeberg, KTH Royal Institute of Technology, Sweden. Blots were agitated in phosphate buffered saline (PBS) plus 0.1% Tween 20 for 5 minutes three times to remove any unbound primary antibody. Blots were then incubated in secondary antibody (1:10000 IRDye800 donkey anti-rabbit IgG (Li-Cor Biosciences)) for 30 minutes in the dark. Blots were washed as above and then stored in PBS for immediate imaging.

Blots were imaged and quantified using the Odyssey Clx Imaging System (Li-Cor Biosciences). Integrated intensity (IInt) was calculated as a measure of protein expression, $I_{\text{Int}} = a(\Sigma I_i - \text{Pixels} \ast B)$ where $a = (\text{resolution} \ast 10^{-3})^2$, $\Sigma I_i$ is the sum of the
intensity of all pixels within the rectangular area drawn manually around the protein bands, P is the number of pixels within the rectangular area, and B is the mean background intensity. For details of statistical analysis, see Sections 3.2.4 and 4.2.4.

Chapter 3: Age and Sex-related Changes in Human NF LV

3.1 Introduction

Age is the principal risk factor for cardiovascular disease, as has been discussed above. The human heart undergoes many changes with age, both structural and functional, and they underlie the increased susceptibility to cardiac pathology with age. These changes were described in Section 1.4. The changes that the human heart undergoes must at some level be represented at the level of gene expression. Gene expression changes can be the causal mechanisms behind the age-related changes, or be compensatory responses that attempt to limit the decline in function from age-related derangements in cardiac physiology and/or anatomic structure. Therefore, an analysis of the changes in the transcriptome (and proteome) of the ageing human LV was undertaken.

Gender has an influence on the age-related changes in the structure and function of the human heart as discussed in Sections 1.4.2 and 1.4.3, as well as susceptibility to
the development of cardiac pathology as discussed in Section 1.4.5. Therefore, an analysis was also carried out to discover sex-related differences in the transcriptome (and proteome) of the human LV.

3.2 Methods

3.2.1 RNA extraction for Microarray and Real-time PCR Analysis

For details regarding RNA extraction for microarray analysis and real-time PCR see Section 2.2.1.

3.2.2 Microarray Analysis of NF LV to Discover Age- and Sex-related Changes

3.2.2.1 RNA labelling for microarray analysis, hybridisation and scanning

Microarray analysis was carried out on 57 NF LV samples (age range: 4 to 65 years, 27 females: mean age: 44.6 years and 30 males: mean age: 38.3 years) to determine changes in gene expression associated with age or sex in the cardiac transcriptome.

Microarrays were created in-house at INSERM (Université de Nantes, Nantes, France). The Ocimum Biosciences 30K Oligo Set A (19968 unique probes) was used for construction, and these probes were spotted onto epoxysilane-coated glass slides using the Lucidea Array Spotter (Amersham). These microarrays were constructed and kindly provided to me by Catherine Chevalier of INSERM.

Total RNA samples were transcribed into cDNA and labelled with dye using the CyScribe First-Strand cDNA Labeling Kit (Amersham) as described in the manufacturer's
protocol. RT of total RNA to aminoacyl-cDNA was verified by analysis with the Bioanalyser 2100 (Agilent) (Fig. 3.1). Total RNA from each NF LV sample analysed was labelled with Cy3 and hybridised against a Cy5-labelled reference RNA pool consisting of a mixture of RNA from all NF LV RNA samples. Utilising a competitive hybridisation strategy against a common reference pool is useful as it controls for heterogeneity in spotting. Specifically, some spots on a microarray may have more or less oligonucleotide bound due to imprecision in the microarray creation process, and this will be reflected in the amount of both Cy3-labelled sample cDNA and Cy5 labelled reference pool cDNA bound. Taking the ratio of Cy3 and Cy5 signal intensity for each spot (which corresponds to a different mRNA and therefore, gene) will control for this heterogeneity in spotting.
Figure 3.1: Bioanalyser 2100 (Agilent) electropherogram showing the population of aminoacyl-cDNA derived from total RNA using CyScribe First-Strand cDNA Labeling Kit (Amersham).

Microarrays were pre-hybridized for 1 hour at 42°C (in 10% BSA, 20 x saline-sodium citrate buffer (SSC), 10% SDS). 0.5 μg of labelled sample cDNA was mixed with 0.5 μg of labelled reference cDNA in 40 μL of hybridization mix (50 x Denhardt's solution, 20 x SSC, 10 μg/μL polyA RNA, 10 μg/μL yeast tRNA, 10% SDS, 50% formamide) and hybridized to the microarrays overnight at 42°C in a humidified chamber. Finally, slides were washed in 1 x SSC + 0.05% SDS, then twice in 0.06 SSC, and then dried by
centrifugation. Microarrays were scanned by fluorescence confocal microscopy (Scanarray 4000XL, GSI-Lumonics) (Fig. 3.2). Fluorescence measurements were obtained separately for Cy3- and Cy5-labelled cDNA at 10 µm/pixel resolution.

Figure 3.2: An example of a microarray captured using fluorescence confocal microscopy (Scanarray 4000XL, GSI-Lumonics) and processed in GenePix Pro 5.0 (Axon). Two images, for Cy3-labelled cDNA acquired at 570nm and Cy5-labelled cDNA acquired at 650nm were superimposed. Each spot contains a different oligonucleotide attached to the surface bearing the sequence of a different gene, with variable amounts of bound complementary Cy3- and Cy5-labelled cDNA representing the level of expression in the LV of individual patients and the common reference RNA pool, respectively. Yellow spots indicate the abundance of a particular RNA species in the LV patient samples and the reference RNA pool are equal, green spots indicate the abundance of RNA in the LV patient sample is higher, and conversely for red spots.

3.2.2.2 Analysis of Microarray Data

Microarray images were analyzed with GenePix Pro 5.0 (Axon). The spots to which GenePix assigned a value less than zero to the 'flag' variable (indicating the spots were missing) were given a weight of zero for subsequent analysis and particular mRNA probes weighted zero on more than 75% of the microarrays were removed from
subsequent analysis. The raw Cy3 and Cy5 expression data were analysed in R using the *LIMMA* statistical module (Smyth, G. K. 2005) (Appendix B). Print-tip LOESS and scale normalisation was used for intra-array and inter-array normalisation, respectively (Smyth, G. K. and Speed, T. 2003). Genes that exhibit differential expression (DE) with respect to age were determined using a linear model, where we regressed the normalized log-ratio data (log2 Cy3/Cy5) against covariates age and sex (Smyth, G. K. 2004). The corresponding False Discovery Rate (FDR) was also computed for each gene (Benjamini, Y. and Hochberg, Y. 1995). The FDR refers to the proportion of genes called as DE that will be false-positives in the list of genes with an FDR less than or equal to the FDR cut-off used. More explicitly, if one has a list of genes ordered by likelihood of being truly DE with age in the human heart (i.e. ordered by p-value), and if one selects an FDR of 0.05 as a cut-off, as was done here, then out of all the genes with a FDR < 0.05, 5% will be false-positives.

To gain further insight into biological processes affected by age and gender, we also carried out a separate analysis looking for coordinated DE of entire sets of genes, so-called gene set analysis. There are several reasons for such an analysis. The first is biological. Many genes are coordinately regulated, i.e. all the genes that subserve a particular pathway such as oxidative phosphorylation, are switched on or off at the same time by master regulators in response to cellular conditions. Therefore, it is physiologically apt to consider genes in sets. The second is pragmatic. A small change in mRNA expression, perhaps 20%, in a particular direction of all the genes in a particular pathway, whilst being physiologically significant, may be missed by single gene approaches such as the one used above, as the difference in expression is small relative to
the noise of the microarray platform (Subramanian, A. et al. 2005). Therefore, by combining genes into sets, one can discover physiologically significant gene regulation that would be otherwise missed. We utilized the gene set enrichment analysis (GSEA) method as detailed in Subramanian, A. et al. (2005). The summary statistics used for the GSEA procedure were the t-statistics derived from the linear model measuring the association between gene expression and age or sex. These t-statistics were inputted into the GSEA v2.0 applet (http://www.broadinstitute.org/gsea/software/gsea2_1024mb.jnlp) and the analysis was carried out using the 'GseaPreranked' module. The gene sets used were based on the KEGG knowledgebase from the Molecular Signatures Database (MSigDB, http://www.broad.mit.edu/gsea/msigdb/index.jsp). GSEA calculates an enrichment score (ES) that is a measure of the extent to which the genes in the particular gene set are overrepresented at the extreme ends (top or bottom) of the ranked (by t-statistic in this case) list of genes (Subramanian, A. et al. 2005). An FDR q-value < 0.05 was set as a threshold for determining significantly regulated gene sets against the advice of Subramanian, A. et al. (2005) who suggested FDR q-value < 0.25. Unfortunately, a major caveat of this analysis is that the stated FDR q-values may be significant underestimations of the true likelihood of observing these results if the null hypothesis is true. This is due to the underlying assumption behind the methodology that each gene is independent, which cannot be true (Goeman, J. J. and Buhlmann, P. 2007). Therefore, a more stringent FDR q-value was selected. The FDR q-value, in this context, represents the probability that a particular gene set is a false positive, i.e. that the null hypothesis cannot be rejected. In the context of competitive gene set testing of which our analysis is a variant, the null hypothesis is that the ranks of genes within this gene set (ranked
according to strength of association with phenotype, in this case by t-statistic) are not unevenly distributed towards the extreme ends amongst the ranks of all genes on the microarray. That is, that genes within the particular gene set are no more likely to be DE than genes not in the gene set (Goeman, J. J. and Buhlmann, P. 2007).

3.2.3 Real-Time PCR Analysis of Ribosome-related Genes in NF LV

Real-time PCR was carried out on 38 NF LV samples (20 younger NF subjects: mean age: 20.1, standard deviation (SD): 7.5 years, 6 women and 14 men; and 18 older NF subjects: mean age: 57.3, SD: 4.2 years, 9 females and 9 males) to validate the changes in expression of r-protein mRNA observed to be DE with age and sex by microarray analysis. For details of the experimental protocol see Section 2.2.3. As discussed above, a mean relative measure of expression was computed from triplicate PCR reactions for each NF LV sample. Multiple linear regression was performed with age and sex as the independent variables, and mean relative expression for each gene as the dependent variable in order to test for an effect of age or sex on the expression of r-protein mRNA. In addition, Pearson's correlation coefficients were computed from the mean relative expression measure from the real-time PCR and the corresponding measure (log2 Cy3/Cy5) from the microarrays as an alternative method of validating the microarray analysis. For all statistical tests, p-values < 0.05 were regarded as significant.
3.2.4 Western Blotting Analysis of RPS18 in NF LV

Protein expression analysis was carried out on 6 NF LV samples (older group: 65 and 63 year old males and 55 year old female; younger group: 4, 8 and 19 year old males) to determine change in expression of RPS18 associated with age. For details regarding protein extraction, see Section 2.3.1. For details regarding protein quantification, see Section 2.3.2. For details regarding SDS-PAGE and western blotting, see Section 2.3.3. Int measures for LV samples in older and younger NF groups were compared by unpaired two-tailed t-tests to assess for DE of RPS18 with age. A p-value < 0.05 was considered significant.

3.3 Results

3.3.1 Microarray Analysis of NF LV to Discover Age- and Sex-related Changes

Of the 19968 probes represented on the microarray platform based on the Ocimum Biosciences Oligo Set A probe set, 18725 passed the filtering criteria and were considered suitable for assessment for DE with age and sex. With respect to age, 109 probes were DE in NF LV (FDR q-value < 0.05), and these could be mapped to 94 unique genes using Entrez IDs (Table 3.1). Of these 109 DE probes, 39 probes (38 unique genes) were upregulated, and 70 probes (57 unique genes) were downregulated with age. Using the coefficients derived from the statistical model applied to calculate the expected change in expression over a 50 year period, the majority of genes exhibit changes that amount to less than a 2-fold change in expression. Specifically, only 5 probes (out of 18725) exhibit
a greater than 2-fold increase in expression, and only 2 probes exhibit a greater than 2-
fold decrease in expression, over a 50 year period.

Table 3.1: 109 probes were DE with age in NF human LV found using the microarray platform containing
the probes from the Ocimum Biosciences 30K Oligo Set A. 'Log(2) FC' represents the change in mRNA
expression per year on a log2 scale, e.g. log(2) FC = 1 corresponds to a doubling of mRNA expression, and
– 1 corresponds to a halving of mRNA expression. 'Mean Expression' represents an arbitrary linear scale of
mRNA abundance. 'FDR', for the meaning of which, see Section 3.2.2.2.

<table>
<thead>
<tr>
<th>GenBank Accession</th>
<th>Symbol</th>
<th>Entrez ID</th>
<th>Log(2) FC</th>
<th>Mean Expression</th>
<th>FDR</th>
</tr>
</thead>
<tbody>
<tr>
<td>AC099849</td>
<td></td>
<td>-</td>
<td>-0.00797</td>
<td>9.10329</td>
<td>0.001</td>
</tr>
<tr>
<td>D26309</td>
<td>LIMK1</td>
<td>3984</td>
<td>0.00864</td>
<td>7.90425</td>
<td>0.005</td>
</tr>
<tr>
<td>NM_005493</td>
<td>RANBP9</td>
<td>10048</td>
<td>-0.00648</td>
<td>8.35893</td>
<td>0.005</td>
</tr>
<tr>
<td>NM_016258</td>
<td>YTHDF2</td>
<td>51441</td>
<td>-0.0076</td>
<td>10.30782</td>
<td>0.005</td>
</tr>
<tr>
<td>NM_152729</td>
<td>NT5DC1</td>
<td>221294</td>
<td>0.01297</td>
<td>6.80619</td>
<td>0.011</td>
</tr>
<tr>
<td>NM_004861</td>
<td>GAL3ST1</td>
<td>9514</td>
<td>0.00539</td>
<td>9.20474</td>
<td>0.011</td>
</tr>
<tr>
<td>NM_005990</td>
<td>STK10</td>
<td>6793</td>
<td>-0.00773</td>
<td>10.68771</td>
<td>0.011</td>
</tr>
<tr>
<td>AL591049</td>
<td></td>
<td>-</td>
<td>-0.00592</td>
<td>9.82693</td>
<td>0.013</td>
</tr>
<tr>
<td>AJ697971</td>
<td>SYNM</td>
<td>23336</td>
<td>0.01226</td>
<td>7.93921</td>
<td>0.013</td>
</tr>
<tr>
<td>NM_006068</td>
<td>TLR6</td>
<td>10333</td>
<td>-0.00853</td>
<td>7.17515</td>
<td>0.013</td>
</tr>
<tr>
<td>NM_006639</td>
<td>CYSLTR1</td>
<td>10800</td>
<td>-0.00687</td>
<td>8.00557</td>
<td>0.013</td>
</tr>
<tr>
<td>NM_001014</td>
<td>KCNMA1</td>
<td>3778</td>
<td>-0.00814</td>
<td>10.47064</td>
<td>0.013</td>
</tr>
<tr>
<td>797</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NM_002775</td>
<td>HTRA1</td>
<td>5654</td>
<td>0.00708</td>
<td>8.43547</td>
<td>0.013</td>
</tr>
<tr>
<td>BC065563</td>
<td>VPS11</td>
<td>55823</td>
<td>-0.01161</td>
<td>6.89523</td>
<td>0.013</td>
</tr>
<tr>
<td>BC001405</td>
<td>UPP1</td>
<td>7378</td>
<td>0.01036</td>
<td>10.71248</td>
<td>0.013</td>
</tr>
<tr>
<td>NM_015443</td>
<td>KIAA1267</td>
<td>284058</td>
<td>-0.0061</td>
<td>10.90064</td>
<td>0.013</td>
</tr>
<tr>
<td>NM_000086</td>
<td>CLN3</td>
<td>1201</td>
<td>-0.00551</td>
<td>10.60348</td>
<td>0.013</td>
</tr>
<tr>
<td>NM_138778</td>
<td>WDR85</td>
<td>92715</td>
<td>-0.00636</td>
<td>9.40951</td>
<td>0.013</td>
</tr>
<tr>
<td>NM_003802</td>
<td>MYH13</td>
<td>8735</td>
<td>0.01002</td>
<td>11.66331</td>
<td>0.013</td>
</tr>
<tr>
<td>BC003065</td>
<td>CDK2</td>
<td>1017</td>
<td>0.0094</td>
<td>9.70848</td>
<td>0.013</td>
</tr>
<tr>
<td>GenBank Accession</td>
<td>Symbol</td>
<td>Entrez ID</td>
<td>Log(2) FC</td>
<td>Mean Expression</td>
<td>FDR</td>
</tr>
<tr>
<td>------------------</td>
<td>--------</td>
<td>----------</td>
<td>-----------</td>
<td>-----------------</td>
<td>-----</td>
</tr>
<tr>
<td>AC011498</td>
<td>-</td>
<td>-</td>
<td>-0.0079</td>
<td>9.55644</td>
<td>0.014</td>
</tr>
<tr>
<td>NM_022486</td>
<td>SUSD1</td>
<td>64420</td>
<td>-0.00593</td>
<td>9.26826</td>
<td>0.015</td>
</tr>
<tr>
<td>NM_005281</td>
<td>GPR3</td>
<td>2827</td>
<td>-0.00508</td>
<td>11.89351</td>
<td>0.015</td>
</tr>
<tr>
<td>NM_138796</td>
<td>SPATA17</td>
<td>128153</td>
<td>-0.01473</td>
<td>6.75389</td>
<td>0.017</td>
</tr>
<tr>
<td>NM_000284</td>
<td>PDHA1</td>
<td>5160</td>
<td>0.00524</td>
<td>10.24394</td>
<td>0.017</td>
</tr>
<tr>
<td>AB016898</td>
<td>C6orf124</td>
<td>653483</td>
<td>-0.00493</td>
<td>8.44071</td>
<td>0.018</td>
</tr>
<tr>
<td>NM_005622</td>
<td>ACSM3</td>
<td>6296</td>
<td>-0.00962</td>
<td>6.62112</td>
<td>0.02</td>
</tr>
<tr>
<td>NM_014463</td>
<td>LSM3</td>
<td>27258</td>
<td>-0.00897</td>
<td>8.5315</td>
<td>0.02</td>
</tr>
<tr>
<td>NM_025191</td>
<td>EDEM3</td>
<td>80267</td>
<td>0.01053</td>
<td>6.87895</td>
<td>0.02</td>
</tr>
<tr>
<td>BC002427</td>
<td>CASP2</td>
<td>835</td>
<td>-0.00813</td>
<td>7.19251</td>
<td>0.021</td>
</tr>
<tr>
<td>NM_021184</td>
<td>C6orf47</td>
<td>57827</td>
<td>-0.00635</td>
<td>12.04333</td>
<td>0.028</td>
</tr>
<tr>
<td>NM_025170</td>
<td>PREX2</td>
<td>80243</td>
<td>-0.00729</td>
<td>9.06905</td>
<td>0.028</td>
</tr>
<tr>
<td>NM_004870</td>
<td>MPDU1</td>
<td>9526</td>
<td>-0.00741</td>
<td>9.475</td>
<td>0.028</td>
</tr>
<tr>
<td>NM_005197</td>
<td>FOXN3</td>
<td>1112</td>
<td>-0.00675</td>
<td>7.94658</td>
<td>0.028</td>
</tr>
<tr>
<td>NM_015106</td>
<td>RAD54L2</td>
<td>23132</td>
<td>0.0066</td>
<td>8.36483</td>
<td>0.029</td>
</tr>
<tr>
<td>BC004343</td>
<td>C21orf122</td>
<td>728039</td>
<td>0.01788</td>
<td>6.49715</td>
<td>0.029</td>
</tr>
<tr>
<td>BC092476</td>
<td>RBPMS</td>
<td>11030</td>
<td>-0.00507</td>
<td>12.0179</td>
<td>0.033</td>
</tr>
<tr>
<td>NM_080862</td>
<td>SPSB4</td>
<td>92369</td>
<td>0.00565</td>
<td>13.08059</td>
<td>0.033</td>
</tr>
<tr>
<td>NM_002117</td>
<td>HLA-C</td>
<td>3107</td>
<td>-0.00999</td>
<td>11.27382</td>
<td>0.036</td>
</tr>
<tr>
<td>NM_017481</td>
<td>UBQLN3</td>
<td>50613</td>
<td>-0.00649</td>
<td>11.43234</td>
<td>0.036</td>
</tr>
<tr>
<td>NM_020212</td>
<td>WDR93</td>
<td>56964</td>
<td>-0.00689</td>
<td>8.41272</td>
<td>0.036</td>
</tr>
<tr>
<td>NM_052864</td>
<td>TIFA</td>
<td>92610</td>
<td>-0.00765</td>
<td>7.23797</td>
<td>0.036</td>
</tr>
<tr>
<td>BC014144</td>
<td>ELAVL3</td>
<td>1995</td>
<td>0.00918</td>
<td>9.71658</td>
<td>0.036</td>
</tr>
<tr>
<td>NM_017553</td>
<td>INO80</td>
<td>54617</td>
<td>-0.00511</td>
<td>11.19678</td>
<td>0.036</td>
</tr>
<tr>
<td>AK022468</td>
<td>SORBS1</td>
<td>10580</td>
<td>-0.00531</td>
<td>10.78696</td>
<td>0.037</td>
</tr>
<tr>
<td>NM_020659</td>
<td>TTYH1</td>
<td>57348</td>
<td>0.00536</td>
<td>11.40098</td>
<td>0.037</td>
</tr>
<tr>
<td>NM_002215</td>
<td>ITIH1</td>
<td>3697</td>
<td>0.00932</td>
<td>10.29121</td>
<td>0.037</td>
</tr>
<tr>
<td>AY256821</td>
<td>WWOX</td>
<td>51741</td>
<td>0.00493</td>
<td>11.51324</td>
<td>0.037</td>
</tr>
<tr>
<td>AK027437</td>
<td>-</td>
<td>-</td>
<td>-0.01095</td>
<td>11.631</td>
<td>0.037</td>
</tr>
<tr>
<td>NM_058004</td>
<td>PI4KA</td>
<td>5297</td>
<td>-0.00566</td>
<td>8.35115</td>
<td>0.037</td>
</tr>
<tr>
<td>NM_005284</td>
<td>GPR6</td>
<td>2830</td>
<td>-0.00651</td>
<td>11.72877</td>
<td>0.037</td>
</tr>
<tr>
<td>GenBank Accession</td>
<td>Symbol</td>
<td>Entrez ID</td>
<td>Log(2) FC</td>
<td>Mean Expression</td>
<td>FDR</td>
</tr>
<tr>
<td>------------------</td>
<td>--------</td>
<td>-----------</td>
<td>-----------</td>
<td>----------------</td>
<td>-----</td>
</tr>
<tr>
<td>BC000979</td>
<td>DDX49</td>
<td>54555</td>
<td>-0.00654</td>
<td>10.0822</td>
<td>0.037</td>
</tr>
<tr>
<td>NM_022481</td>
<td>ARAP3</td>
<td>64411</td>
<td>-0.00425</td>
<td>12.1131</td>
<td>0.037</td>
</tr>
<tr>
<td>BC022307</td>
<td>ATM</td>
<td>472</td>
<td>-0.008</td>
<td>7.22298</td>
<td>0.037</td>
</tr>
<tr>
<td>XM_374983</td>
<td>-</td>
<td>-</td>
<td>0.00796</td>
<td>8.58884</td>
<td>0.037</td>
</tr>
<tr>
<td>NM_021140</td>
<td>KDM6A</td>
<td>7403</td>
<td>-0.00758</td>
<td>7.21303</td>
<td>0.037</td>
</tr>
<tr>
<td>BC101301</td>
<td>PAX3</td>
<td>5077</td>
<td>0.00743</td>
<td>11.14169</td>
<td>0.037</td>
</tr>
<tr>
<td>NM_022092</td>
<td>CHTF18</td>
<td>63922</td>
<td>-0.00591</td>
<td>10.41938</td>
<td>0.037</td>
</tr>
<tr>
<td>NM_001184</td>
<td>ATR</td>
<td>545</td>
<td>0.01196</td>
<td>6.83698</td>
<td>0.037</td>
</tr>
<tr>
<td>NM_012103</td>
<td>CNOT7</td>
<td>29883</td>
<td>-0.00802</td>
<td>10.66867</td>
<td>0.037</td>
</tr>
<tr>
<td>NM_013354</td>
<td>CNOT7</td>
<td>29883</td>
<td>-0.00706</td>
<td>8.049</td>
<td>0.037</td>
</tr>
<tr>
<td>NM_213636</td>
<td>PDLIM7</td>
<td>9260</td>
<td>0.00871</td>
<td>9.29493</td>
<td>0.037</td>
</tr>
<tr>
<td>NM_006542</td>
<td>SPHAR</td>
<td>10638</td>
<td>0.02515</td>
<td>6.49186</td>
<td>0.037</td>
</tr>
<tr>
<td>AC078860</td>
<td>-</td>
<td>-</td>
<td>-0.00638</td>
<td>8.29991</td>
<td>0.037</td>
</tr>
<tr>
<td>Z95114</td>
<td>-</td>
<td>-</td>
<td>-0.00883</td>
<td>6.8725</td>
<td>0.038</td>
</tr>
<tr>
<td>NM_015139</td>
<td>SLC35D1</td>
<td>23169</td>
<td>0.00787</td>
<td>8.16734</td>
<td>0.038</td>
</tr>
<tr>
<td>NM_006225</td>
<td>PLCD1</td>
<td>5333</td>
<td>0.00546</td>
<td>11.17581</td>
<td>0.038</td>
</tr>
<tr>
<td>NM_006392</td>
<td>NOP56</td>
<td>10528</td>
<td>-0.00707</td>
<td>7.82859</td>
<td>0.038</td>
</tr>
<tr>
<td>AL590106</td>
<td>-</td>
<td>-</td>
<td>-0.01178</td>
<td>12.75166</td>
<td>0.038</td>
</tr>
<tr>
<td>NM_000272</td>
<td>NPHP1</td>
<td>4867</td>
<td>0.008</td>
<td>7.22654</td>
<td>0.038</td>
</tr>
<tr>
<td>NM_020064</td>
<td>BARHL1</td>
<td>56751</td>
<td>-0.00523</td>
<td>8.92242</td>
<td>0.038</td>
</tr>
<tr>
<td>BC012993</td>
<td>RBMS1</td>
<td>5937</td>
<td>0.0067</td>
<td>11.09068</td>
<td>0.038</td>
</tr>
<tr>
<td>AC122129</td>
<td>-</td>
<td>-</td>
<td>-0.0078</td>
<td>9.99833</td>
<td>0.038</td>
</tr>
<tr>
<td>NM_031297</td>
<td>RNF208</td>
<td>727800</td>
<td>-0.00455</td>
<td>11.40265</td>
<td>0.039</td>
</tr>
<tr>
<td>BC108917</td>
<td>KIR2DS2</td>
<td>100132285</td>
<td>-0.01234</td>
<td>6.76435</td>
<td>0.041</td>
</tr>
<tr>
<td>NM_024122</td>
<td>APOO</td>
<td>79135</td>
<td>-0.00581</td>
<td>8.78299</td>
<td>0.042</td>
</tr>
<tr>
<td>NM_014068</td>
<td>PSORS1C1</td>
<td>170679</td>
<td>-0.00575</td>
<td>11.1356</td>
<td>0.042</td>
</tr>
<tr>
<td>NM_005514</td>
<td>HLA-B</td>
<td>3106</td>
<td>-0.00881</td>
<td>7.8735</td>
<td>0.042</td>
</tr>
<tr>
<td>NM_018976</td>
<td>SLC38A2</td>
<td>54407</td>
<td>-0.01155</td>
<td>9.291</td>
<td>0.042</td>
</tr>
<tr>
<td>NM_014701</td>
<td>SECISBP2L</td>
<td>9728</td>
<td>0.00773</td>
<td>7.90408</td>
<td>0.043</td>
</tr>
<tr>
<td>NM_021095</td>
<td>SLC5A6</td>
<td>8884</td>
<td>-0.00638</td>
<td>12.16871</td>
<td>0.043</td>
</tr>
<tr>
<td>AF288738</td>
<td>-</td>
<td>-</td>
<td>0.00549</td>
<td>10.57222</td>
<td>0.043</td>
</tr>
<tr>
<td>GenBank Accession</td>
<td>Symbol</td>
<td>Entrez ID</td>
<td>Log(2) FC</td>
<td>Mean Expression</td>
<td>FDR</td>
</tr>
<tr>
<td>-------------------</td>
<td>--------</td>
<td>-----------</td>
<td>-----------</td>
<td>----------------</td>
<td>------</td>
</tr>
<tr>
<td>J03639</td>
<td>MCF2</td>
<td>4168</td>
<td>-0.00643</td>
<td>11.86421</td>
<td>0.043</td>
</tr>
<tr>
<td>NM_033052</td>
<td>-</td>
<td>-</td>
<td>-0.00551</td>
<td>9.76119</td>
<td>0.043</td>
</tr>
<tr>
<td>BC069810</td>
<td>-</td>
<td>-</td>
<td>-0.00569</td>
<td>10.44001</td>
<td>0.043</td>
</tr>
<tr>
<td>NM_014555</td>
<td>TRPM5</td>
<td>29850</td>
<td>-0.00774</td>
<td>11.61154</td>
<td>0.043</td>
</tr>
<tr>
<td>NM_003095</td>
<td>SNRPF</td>
<td>6636</td>
<td>0.01074</td>
<td>9.39458</td>
<td>0.044</td>
</tr>
<tr>
<td>BC053585</td>
<td>CSF3R</td>
<td>1441</td>
<td>0.00511</td>
<td>11.83018</td>
<td>0.044</td>
</tr>
<tr>
<td>NM_000098</td>
<td>CPT2</td>
<td>1376</td>
<td>0.0078</td>
<td>7.02202</td>
<td>0.044</td>
</tr>
<tr>
<td>NM_015853</td>
<td>UBXN1</td>
<td>51035</td>
<td>0.00756</td>
<td>8.81844</td>
<td>0.044</td>
</tr>
<tr>
<td>AL122043</td>
<td>C20orf112</td>
<td>140688</td>
<td>0.00771</td>
<td>7.32908</td>
<td>0.044</td>
</tr>
<tr>
<td>NM_006332</td>
<td>IFI30</td>
<td>10437</td>
<td>0.00613</td>
<td>10.83373</td>
<td>0.044</td>
</tr>
<tr>
<td>BC040034</td>
<td>CORT</td>
<td>1325</td>
<td>-0.00615</td>
<td>10.87187</td>
<td>0.045</td>
</tr>
<tr>
<td>NM_024293</td>
<td>FAM134A</td>
<td>79137</td>
<td>-0.00698</td>
<td>10.75081</td>
<td>0.045</td>
</tr>
<tr>
<td>NM_003579</td>
<td>RAD54L</td>
<td>8438</td>
<td>-0.00621</td>
<td>10.62869</td>
<td>0.045</td>
</tr>
<tr>
<td>NM_017709</td>
<td>FAM46C</td>
<td>54855</td>
<td>0.00517</td>
<td>12.63892</td>
<td>0.045</td>
</tr>
<tr>
<td>BC106009</td>
<td>RPL34</td>
<td>6164</td>
<td>-0.00969</td>
<td>10.34143</td>
<td>0.045</td>
</tr>
<tr>
<td>AL354898</td>
<td>-</td>
<td>-</td>
<td>-0.00621</td>
<td>11.29974</td>
<td>0.046</td>
</tr>
<tr>
<td>NM_003750</td>
<td>-</td>
<td>-</td>
<td>-0.00562</td>
<td>10.27669</td>
<td>0.046</td>
</tr>
<tr>
<td>NM_017980</td>
<td>LIMS2</td>
<td>55679</td>
<td>-0.00533</td>
<td>12.00008</td>
<td>0.047</td>
</tr>
<tr>
<td>NM_133468</td>
<td>BMPER</td>
<td>168667</td>
<td>-0.00729</td>
<td>7.92808</td>
<td>0.047</td>
</tr>
<tr>
<td>NM_020374</td>
<td>C12orf4</td>
<td>57102</td>
<td>0.00709</td>
<td>10.16234</td>
<td>0.047</td>
</tr>
<tr>
<td>NM_006500</td>
<td>MCAM</td>
<td>4162</td>
<td>0.0053</td>
<td>10.62599</td>
<td>0.047</td>
</tr>
<tr>
<td>AK074957</td>
<td>KCNQ4</td>
<td>9132</td>
<td>0.00384</td>
<td>11.57865</td>
<td>0.047</td>
</tr>
<tr>
<td>NM_012309</td>
<td>SHANK2</td>
<td>22941</td>
<td>0.0075</td>
<td>8.94692</td>
<td>0.047</td>
</tr>
<tr>
<td>AF213465</td>
<td>DUOX1</td>
<td>53905</td>
<td>-0.00646</td>
<td>12.97164</td>
<td>0.047</td>
</tr>
<tr>
<td>NM_014703</td>
<td>VPRBP</td>
<td>9730</td>
<td>0.00709</td>
<td>8.35103</td>
<td>0.048</td>
</tr>
<tr>
<td>NM_002037</td>
<td>FYN</td>
<td>2534</td>
<td>-0.00634</td>
<td>8.97312</td>
<td>0.048</td>
</tr>
<tr>
<td>NM_014341</td>
<td>MTCH1</td>
<td>23787</td>
<td>-0.00544</td>
<td>9.63417</td>
<td>0.049</td>
</tr>
</tbody>
</table>
With respect to gender, 5 probes corresponding to 5 unique genes were DE in NF LV (FDR q-value < 0.05). Three genes were upregulated in male, and 2 in female LV (Table 3.2)

Table 3.2: Five probes were DE with gender in NF human LV using the microarray platform containing the probes from the Ocimum Biosciences 30K Oligo Set A. 'Log(2) FC' represents the change in mRNA expression as a function of gender on a log₂ scale, e.g. log(2) FC = 1 corresponds to mRNA expression in a male being twice that of a female, and – 1 corresponds mRNA expression in a male being half that of a female. 'Mean Expression' represents an arbitrary linear scale of mRNA abundance. 'FDR' represents the FDR q-value, the meaning of which is explained in Section 3.2.2.2.

<table>
<thead>
<tr>
<th>Genbank Accession</th>
<th>Symbol</th>
<th>Entrez ID</th>
<th>Log(2) FC</th>
<th>Mean Expression</th>
<th>FDR</th>
</tr>
</thead>
<tbody>
<tr>
<td>NM_004653</td>
<td>KDM5D</td>
<td>8284</td>
<td>0.49637</td>
<td>7.28068</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>NM_032490</td>
<td>C14orf142</td>
<td>84520</td>
<td>0.55457</td>
<td>6.47439</td>
<td>0.007</td>
</tr>
<tr>
<td>NM_004555</td>
<td>NFATC3</td>
<td>4775</td>
<td>0.27737</td>
<td>7.33425</td>
<td>0.011</td>
</tr>
<tr>
<td>NM_001007</td>
<td>RPS4X</td>
<td>6191</td>
<td>-0.40798</td>
<td>10.53845</td>
<td>0.015</td>
</tr>
<tr>
<td>NM_016509</td>
<td>CLEC1B</td>
<td>51266</td>
<td>-0.81863</td>
<td>6.29678</td>
<td>0.039</td>
</tr>
</tbody>
</table>

The Genbank Accession numbers for each of the 18725 filtered probes of the microarray platform based on the Ocimum 30K Oligo Set A were entered into the GSEA software, along with the two t-statistics describing the relationship between gene expression in human LV and age and sex, respectively. Of those 18725 probes, the software collapsed the list to 16112 unique Genbank accession numbers, then collapsed this list to a corresponding list of 9954 unique genes as defined by Entrez ID (taking the mean of the t-statistics for all the probes that corresponded to the same unique gene). This collapsed gene list was used in the analysis. The GSEA analysis found 3 KEGG pathway gene sets to be significantly downregulated with age in human LV (KEGG_RIBOSOME, 54
probes, ES = -0.71, FDR q-value < 0.005 (Fig. 3.3);
KEGG_ANTIGEN_PROCESSING_AND_PRESENTATION, 44 probes, ES = -0.51, FDR q-value = 0.049; and KEGG_DORSO_VENTRAL_AXISFORMATION, 18 probes, ES = -0.63, FDR q-value = 0.037) and zero KEGG pathway gene sets to be significantly upregulated with age. The GSEA analysis found 2 KEGG pathway gene sets was expressed at a higher level in female LV (KEGG_RIBOSOME, 54 probes, ES = -0.80, FDR q-value < 0.0005; KEGG_PROTEIN_EXPORT, 17 probes, ES = -0.64, FDR q-value = 0.05), and zero at lower levels in female LV, compared to male LV.

![Enrichment plot: KEGG_RIBOSOME](image)

**Figure 3.3:** Enrichment plot generated by the GSEA applet (v 2.0), for the KEGG_RIBOSOME gene set. Note that the red side of the bar indicates genes more upregulated with age and the violet side indicates genes more downregulated with age, and that r-protein coding genes, individually represented by vertical black lines, are concentrated towards the violet end. A similar pattern is seen for the comparison between female and male NF LV.
3.3.2 Real-Time PCR Analysis of Ribosome-related Genes in NF LV

For confirmation of changes in mRNA expression discovered using microarray analysis, real-time PCR analysis was carried out on RNA samples extracted from the LV of 38 NF hearts. The GSEA results indicated that the genes contained within the KEGG Ribosome gene set, i.e. genes that code for proteins that comprise the subunits of the ribosome ribonucleoprotein complex, were downregulated with age and were also expressed at a higher abundance in females. To confirm, real-time PCR was used to quantify the expression of RPL23A, RPL26, RPL34, RPL37, RPS7, RPS10, RPS18, RPS21, 45s rRNA, MTOR, and RHEB.

Table 3.3: The results of the multiple linear regression model with age and sex as independent and gene expression as the dependent variable, i.e. gene expression = y-intercept + β1 * age + β2 * sex (n = 38). “% Change / 50 years” was derived from β1 and refers to the difference in expression observed after 50 years of ageing. “% Change in Males” was derived from β2 and refers to the difference in expression in males versus females across all ages.

<table>
<thead>
<tr>
<th>Gene</th>
<th>% Change / 50 years</th>
<th>P-value</th>
<th>% Change in Males</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>45s rRNA</td>
<td>-11.4</td>
<td>0.41</td>
<td>-21.2</td>
<td>0.06</td>
</tr>
<tr>
<td>MTOR</td>
<td>-15.9</td>
<td>0.06</td>
<td>3.1</td>
<td>0.63</td>
</tr>
<tr>
<td>RHEB</td>
<td>-21.5</td>
<td>0.09</td>
<td>-12.2</td>
<td>0.20</td>
</tr>
<tr>
<td>RPL23A</td>
<td>-8.5</td>
<td>0.40</td>
<td>-6.5</td>
<td>0.42</td>
</tr>
<tr>
<td>RPL26</td>
<td>-15.7</td>
<td>0.13</td>
<td>-8.5</td>
<td>0.29</td>
</tr>
<tr>
<td>RPL34</td>
<td>-15.5</td>
<td>0.15</td>
<td>-4.4</td>
<td>0.59</td>
</tr>
<tr>
<td>RPL37</td>
<td>-24.3</td>
<td>0.015</td>
<td>-6.2</td>
<td>0.39</td>
</tr>
<tr>
<td>RPS7</td>
<td>-17.4</td>
<td>0.23</td>
<td>-11.6</td>
<td>0.30</td>
</tr>
<tr>
<td>RPS10</td>
<td>-18.9</td>
<td>0.046</td>
<td>-11.3</td>
<td>0.12</td>
</tr>
<tr>
<td>RPS18</td>
<td>-28.8</td>
<td>0.006</td>
<td>-8.8</td>
<td>0.25</td>
</tr>
<tr>
<td>RPS21</td>
<td>-14.0</td>
<td>0.14</td>
<td>-12.0</td>
<td>0.10</td>
</tr>
</tbody>
</table>
With respect to age, RPS10, RPS18, and RPL37 were statistically significantly downregulated with age (p-values = 0.046, 0.006, and 0.015, respectively) (Table 3.3). All other genes were not shown to be DE with age, although the fold-change estimates for all mRNA coding for r-proteins, as well as for 45s rRNA, were in the direction of downregulation. With respect to gender, no genes were shown to be DE, however, all fold-change estimates for mRNA coding for r-proteins, plus for 45s rRNA, were in the direction of decreased expression in males.

**Table 3.4**: The results of the Pearson's correlation analyses to determine if the gene expression measurements derived from the microarrays are linearly related to the measurements from the real-time PCR experiments (n=18). 'Normalised' refers to calculations done with real-time PCR data normalised to the internal control gene, SCR1, and 'Raw' refers to the use of non-normalised real-time PCR data. Note, RPS10 has an additional probe represented on the array, for this probe: normalised to Scr1: Pearson's r = 0.5006826, p-value = 0.01716; non-normalised: Pearson's r = 0.3152774, p-value = 0.1013. Similarly, RPS21 has an additional probe: normalised to Scr1: Pearson's r = 0.2403324, p-val = 0.1684; non-normalised: Pearson's r = 0.1991245, p-val = 0.2141.

<table>
<thead>
<tr>
<th>Gene</th>
<th>Pearson's r (Normalised)</th>
<th>P-value</th>
<th>Pearson's r (Raw)</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>RPL23A</td>
<td>0.06</td>
<td>0.41</td>
<td>0.09</td>
<td>0.36</td>
</tr>
<tr>
<td>RPL26</td>
<td>0.66</td>
<td>0.001</td>
<td>0.68</td>
<td>0.001</td>
</tr>
<tr>
<td>RPL34</td>
<td>0.64</td>
<td>0.002</td>
<td>0.58</td>
<td>0.01</td>
</tr>
<tr>
<td>RPL37</td>
<td>0.21</td>
<td>0.20</td>
<td>0.03</td>
<td>0.46</td>
</tr>
<tr>
<td>RPS7</td>
<td>0.66</td>
<td>0.001</td>
<td>0.65</td>
<td>0.002</td>
</tr>
<tr>
<td>RPS10</td>
<td>0.71</td>
<td>0.0004</td>
<td>0.60</td>
<td>0.004</td>
</tr>
<tr>
<td>RPS18</td>
<td>0.74</td>
<td>0.0002</td>
<td>0.67</td>
<td>0.001</td>
</tr>
<tr>
<td>RPS21</td>
<td>0.52</td>
<td>0.01</td>
<td>0.48</td>
<td>0.02</td>
</tr>
</tbody>
</table>
With respect to the Pearson's correlation between the microarrays and real-time PCR experiments, the gene expression measurements for 6 out of 8 r-protein mRNA were significantly correlated, irrespective of normalisation of real-time PCR results to the internal control gene, Scr1 (Table 3.4).

### 3.3.3 Western Blotting Analysis of RPS18 in NF LV

The rabbit ant-human RPS18 antibody stained a band in the Western blot experiments at the correct migration point of ~17 kDa (Fig. 3.5 and Appendix A). Note equal loading of protein among all wells is observed (Fig. 3.4). However, no difference in expression between younger and older NF hearts was observed, mean IInt of 0.82 vs 0.79, respectively, p-value = 0.82 (Fig. 3.5).

![Figure 3.4](image)

*Figure 3.4: SDS-PAGE using a Nusep 4-20% gradient gel loaded with 40 μg LV total protein. The gel was UV irradiated for 1 minute prior to imaging to activate the Stain-free compound to visualise tryptophan-containing proteins. Protein bands corresponding to size range 75 (top) to 25 kDa (bottom) are shown. Samples from left: alternating older NF, HCM, and younger NF LV samples.*
3.4 Discussion

3.4.1 Technical Interpretation of Results

The validity of the microarray analysis is supported by the statistically significant correlation of the real-time PCR results with that of the microarray results. Furthermore, 3 mRNA coding for r-proteins were found to be significantly downregulated with age by real-time PCR. The other mRNA coding for r-proteins tested by real-time PCR were not shown to be DE with age or sex, despite the gene set analysis of the microarray data implying that all r-proteins might be downregulated. This was expected, as out of the 54 r-proteins with corresponding mRNA represented on the microarrays, only RPL34 was found to be DE (downregulated) with age with a FDR q-value < 0.05 in the initial gene-by-gene microarray analysis. Only with the gene set analysis was it possible to tease out the existence of an effect of age (and sex) on the coordinated expression of multiple r-proteins. One can consider a gene-by-gene analysis of microarray data to lack sensitivity to discover DE genes, due to the large number of genes that are tested, and thus the need to control for multiplicity of testing. If one uses the raw p-value associated with each gene tested on a microarray to determine DE, one would expect 0.05 * (no. of genes on microarray) false positives. Therefore, one must use a correction such as the FDR.
correction employed here to minimise the false positive rate (Benjamini, Y. and Hochberg, Y. 1995). However, this comes at the cost of significantly reducing the true positive rate of DE gene discovery.

I had hypothesised that a decrease in the expression with age (and male gender) of mRNA coding for r-proteins would be accompanied by a decrease in synthesis of the 45s rRNA that could be detected by real-time PCR. This was not detected, and there are a number of possible reasons for this. Note that the 45s rRNA amplicon in the real-time PCR was located in the 5' upstream region not present within the mature rRNA species. If we assume that there is in fact a true decrease in the rate of transcription of the 45s rRNA within LV tissue with age (and male gender), and it is coupled with an equal decrease in the rate of processing of the 45s rRNA into the mature 18s, 28s, and 5.8s rRNA species, then there would not be a change in abundance of the 45s rRNA. This would be despite there being a decrease in the rate of production of mature rRNA to be incorporated into newly-synthesised ribosomes. Another possibility is that per unit mass of tissue, there is a decrease in 45s rRNA abundance secondary to a decrease in the rate of transcription, and this leads to a corresponding decrease in the levels of mature rRNA per unit mass of tissue. Note that given that rRNA constitutes approximately 80% of total RNA, this would result in a decrease in extracted RNA per unit mass of tissue. Unfortunately, documenting yields of RNA per unit mass of tissue is complicated by various aspects of the RNA extraction process that are difficult to control for. These include: (1) differences in the amount of precipitated moisture present on LV tissue samples when they are removed from liquid nitrogen and being fragmented into a suitable size for extraction; (2) variable amounts of LV tissue per equal amounts of extraction reagent which may alter
the efficiency of RNA extraction; (3) variation in the efficiency with which tissue is homogenised in extraction reagent prior to the first centrifugation step; (4) variable loss of RNA at different stages of the extraction procedure; and (5) potentially other sources of variability. Thus, it was decided not to analyse RNA yield to assess for changes in RNA concentration with age or gender. Then, a constant amount of RNA was added to each RT reaction, the consequence of which was that one could only discover changes in particular RNA species relative to the overall RNA pool. If mature rRNA abundance decreases with age alongside the decrease in 45s rRNA, then the tissue from older ages will yield less RNA, and then to add the same amount of total RNA to each RT reaction, total RNA from a greater mass of tissue will have been added to the RT reactions of the older NF hearts. In this way, any detectable change in 45s rRNA between younger and older NF hearts would be attenuated, which would explain the inability to discover a statistically significant decrease in 45s rRNA with age (or male gender).

Regarding the Western blot results, we did not find that RPS18 protein levels were different in old and young NF hearts. There are several possible explanations. The first is that the change in mRNA expression of r-proteins is not associated with a corresponding change in r-protein abundance, i.e. while there may be a reduced abundance of r-protein mRNA transcripts, they may be more stable (longer half-life within the cytoplasm) and/or are translated at a higher rate at older ages, and thus the expression of r-proteins is maintained. The alternative is that the lack of appreciation of a true downregulation in RPS18 protein abundance is due to sampling error and/or technical variation. The sample size was quite small and there is massive variation among NF donors in terms of lifestyle, including nutrition and physical activity, subclinical
cardiac pathology, genetics, as well as cause of death, all of which may impact upon RPS18 protein expression/abundance. As an example, Luczak, E. D. et al. (2011) have shown that in rats, a casein-rich diet increases the expression of several r-proteins including RPS18. In addition, sampling error notwithstanding, the Western blot is a multi-step technique, each step of which introduces variability. This variability, coupled with the expectation of only a small change in protein abundance from the real-time PCR experiments, which demonstrated an approximate reduction of 30% in mRNA abundance over 50 years of ageing, makes it conceivable that the experiment would not allow appreciation of such a small change.

A significant technical limitation of the microarray analysis was the incomplete coverage of the genome (12087 unique Entrez IDs in the Ocimum 30K oligo set A), which prevented the discovery of all the transcriptomic changes associated with cardiac ageing and gender.

Another important limitation was that the location within the LV from which the biopsies were derived was unknown, thus regional variation in gene expression could not be appreciated. As an example, Heerdt, P. M. et al. (2000) found that SERCA2A mRNA is more highly expressed in LV apex versus free wall. Furthermore, this regional variation would have introduced biological variation not captured in the statistical model applied, therefore reducing the power to detect statistically significant changes in gene expression with age or gender common to all regions of LV myocardium.
3.4.2 Biological Interpretation of Results

The KEGG_RIBOSOME gene set, which was significantly downregulated, contains the genes that code for the r-proteins, of which 79 were represented in the microarray platform. Note that these must be present in equimolar proportions, along with rRNA, in order to form functioning ribosomes (Hannan, R. D. et al. 2003). Therefore a decrease in any r-proteins would be reflected in a decrease in ribosome synthesis. This has been shown by Robledo, S. et al. (2008) who found that siRNA mediated knockout of any r-protein led to a decrease in protein abundance of all other r-proteins comprising the same subunit, a reduction in abundance of the particular subunit, plus a reduction in abundance of the mature 80s ribosome. An increase in cardiomyocyte size, or number, requires increased protein synthesis. This can be achieved in several ways. One way is to increase the number of ribosomes within a cell, which will allow the number of mRNA molecules that can be translated in parallel to increase, assuming that ribosome abundance is limiting. Beltrami, A. P. et al. (2001) have demonstrated myocyte division in the border zone and in distant myocardium in the days following MI within the LV. This implies that the heart is attempting to compensate for the loss of contractile tissue by hyperplasia (and concomitant hypertrophy) of myocardial cells to replace infarcted tissue. With age, if there is a decrease in ribosome abundance or a decrease in the capacity to synthesise additional ribosomes (if decreased r-protein mRNA abundance with age is limiting for increasing the rate of ribosome synthesis), then this may interfere with this compensatory hyperplasia/hypertrophy. Without replacement of infarcted tissue, assuming similar haemodynamic demands post MI, the load on the remaining cardiomyocytes will have increased and thus, the probability of cardiomyocyte apoptosis and/or necrosis will also
increase, with further worsening of haemodynamic load on each remaining cardiomyocyte. As noted before, the likelihood of HF developing post MI is increased with the age at which the MI occurred (Lewis, E. F. et al. 2003; Terkelsen, C. J. et al. 2011), and increased age also increases the risk of death in the context of established HF (MAGGIC et al. 2012). Decreased ribosome synthesis at older ages may contribute. In end-stage HF it has already been mentioned that there is ongoing increased hyperplasia (Quaini, F. et al. 1994; Kajstura, J. et al. 1998) of the cardiomyocytes relative to NF myocardium, reflecting inadequacy of contractile function of the cardiomyocytes with an attempt at compensation. A decreased abundance of ribosomes or inadequate capacity to upregulate the synthesis of ribosomes with age, in the end-stage HF state of increased need for protein synthesis, might represent a contributing factor to the inadequacy of hyperplasia. This inadequacy of hyperplasia in the aged heart is reflected in the findings of Bergmann, O. et al. (2009) of decreased cardiomyocyte renewal. Finally, the increased rate of apoptosis in end-stage HF (Olivetti, G. et al. 1997) may reflect in part the increased haemodynamic load on the available pool of cardiomyocytes that is not attenuated by cardiomyocyte hyperplasia to share the haemodynamic burden.

If the effects of male gender are similar to the effects of age on the capacity to synthesise ribosomes, then one would expect that the male heart would be less able to cope with the imposition of an MI, and would have an increased risk of HF, all other factors being equal. Parodi, G. et al. (2007) found greater recovery of LV EF post MI in females, and this may be related to greater ribosome synthesis affording greater hyperplasia/hypertrophy of the remaining viable myocardium in females. However, protection from development of HF in females was not demonstrated in a number of
studies mentioned above, the reasons for which are unclear. All-cause mortality in patients with HF is greater in men than in women, adjusted for age (Martinez-Selles, M. et al. 2012). This greater risk in males was found to be modified by aetiology of HF and diabetes. Specifically, in subgroup analyses, the association of female gender with greater survival appeared more marked in non-ischaemic aetologies of HF, and separately, diabetes attenuated the protective effect of female gender in HF (Martinez-Selles, M. et al. 2012). It is possible that the advantage of female gender is to be found only prior to the development of ischaemia from coronary obstruction, either at the micro- or macro-vascular level (which are features of diabetes and ISCM). Ribosome synthesis is energetically costly and in conditions of reduced blood flow it might be physiologically rational to reduce the synthesis of ribosomes to conserve ATP for use by the contractile machinery. If however, ischaemia due to atherosclerosis does not supervene in end-stage HF, then we might see that ribosome synthesis is greater in female hearts within the viable myocardium, and this confers an advantage in terms of greater capacity for hyperplasia/hypertrophy to compensate for the contractile deficit. This theory implies that blood flow is not limiting in end-stage HF, or at least during the evolution of HF, except in patients where coronary artery/arteriolar obstruction is the predominant cause of the HF phenotype. Perhaps the discrepancy between the results of Parodi, G. et al. (2007) and Van 'T Hof, A. W. et al. (1998) lies in the higher rate of multi-vessel disease in the latter study, abrogating the advantage of female gender.

The effects of a proposed reduction in the capacity to synthesise ribosomes with age (and male gender) has been discussed, but what has so far not been addressed, is the underlying basis for this reduction. The expression of RHEB and MTOR were examined
by real-time PCR to discern whether they might be responsible for the decrease in expression of r-protein mRNA. RHEB is a GTP-binding protein with intrinsic GTPase activity, which in its active GTP-bound form, activates mTOR protein (Bai, X. et al. 2007). MTOR in turn, is a well-established promoter of cardiac hypertrophy through a number of mechanisms including the stimulation of ribosome biogenesis (Mayer, C. and Grummt, I. 2006). However, neither RHEB nor MTOR mRNA expression were shown to be differentially expressed with age, though the direction of change was of possible downregulation. This does not preclude changes in protein abundance and activity as potential explanations for decreased r-protein mRNA expression with age (and in males), and further studies to investigate this possibility would be useful. Another potential mechanism for decreased ribosome synthesis with age was demonstrated by Strehler, B. L. et al. (1979) who reported decreased hybridisation of labelled rat rRNA to human myocardial DNA in older versus younger males. This may represent either loss via deletion of rRNA gene segments with age, or some other modification such as accumulation of point mutations or covalent cross-linking which would interfere with such complementary binding. A reduced number of structurally intact rRNA genes which can be transcribed simultaneously may lead to a decreased rate of production of 45s rRNA. It has been proposed that this may feedback negatively to pathways responsible for r-protein synthesis (Hannan, R. D. et al. 2003) resulting in the reduction of r-protein mRNA abundance found with age. A future analysis to see if such a modification of 45s rRNA genes also occurs in female LV with age, and at what rate, should be undertaken, which may or may not be able to explain the increased r-protein mRNA expression in females. Finally, depletion of ATP and decreased flux through the creatine kinase
pathway has not been shown to occur with age (Bottomley, P. A. et al. 2008), therefore energy depletion does not seem to be the explanation for decreased ribosome synthesis in aged NF LV myocardium.

Another DE gene set with age was the KEGG_ANTIGEN_PROCESSING_AND_PRESENTATION gene set, which was downregulated with age. The most highly ranked DE genes within the gene set include the various haplotypes of the Major histocompatibility complex Class I and II genes, beta-2-microglobulin, cathepsin S, interferon alpha 6 and 10, and transporter 1, ATP-binding cassette, sub-family B. These proteins are involved in cell-mediated immunity against intracellular microbes. Decrease in expression of the various components of the immune response to intracellular microbes may lead to increased susceptibility to myocarditis with age. Future studies looking at changes in the protein expression of the above genes should be undertaken.

The genes discovered by the initial, gene-by-gene analysis include several with known cardiovascular roles in humans. The mRNA coding for Potassium (K) large conductance calcium-activated channel, subfamily M, alpha member 1 (KCNMA1), a K⁺ channel responsive to voltage and Ca^{2+} in human coronary artery smooth muscle cells, was shown to decrease in expression with age (Fig. 3.6). Activation of this channel is associated with relaxation of coronary arterioles (Lu, R. et al. 2006), and thus, a reduction in expression with age may increase susceptibility to myocardial ischaemia. Incidentally, expression of KCNMA1 has also been detected in human cardiac fibroblasts (Wang, Y. J. et al. 2006), the functional implications of which are currently unknown.
LIM domain kinase 1 (LIMK1) was the most highly ranked gene shown to be upregulated with age in our dataset (Fig. 3.7). It is a kinase that phosphorylates and thereby inactivates cofilin proteins (Scott, R. W. and Olson, M. F. 2007). Cofilin proteins are F-actin depolymerising enzymes, suggesting that an increased abundance of LIMK1 might lead to a decrease in F-actin disassembly. Although this is highly speculative, decreased F-actin disassembly might have several effects on ageing cardiomyocytes. The first is to prevent the disassembly of sarcomeres which would be required in order to allow cardiomyocytes to undergo cytokinesis and hence, hyperplasia, in response to increased haemodynamic demands. Note, as mentioned earlier, that cardiomyocyte renewal decreases with age in the human heart (Bergmann, O. et al. 2009), though it is at present unknown whether this renewal is from cardiomyocyte division or from cardiac stem cell differentiation. The second is a reduction in turnover of F-actin, leading to the accumulation within sarcomeres of post-translationally modified actin (e.g oxidised) which would decrease the sarcomere's ability to generate force. An alternate effect might be to reduce the breakdown of F-actin and thus preserve myofibril mass in light of increasing haemodynamic demands with age.

Another gene that was shown to be upregulated with age was cyclin-dependent kinase 2 (CDK2), a kinase that complexes with cyclin proteins to control the cell cycle (Fig. 3.8). In rat cardiomyocyte cultures, activation of cyclin E-CDK2 is involved in angiotensin II-induced hypertrophy as well as serum- and angiotensin II-induced nuclear replication (Hinrichsen, R. et al. 2008). Given the decreased cardiomyocyte turnover that seems to occur with age as described earlier, this increase in CDK2 mRNA might be the opposite of what one would expect. However, cardiomyocyte cell turnover in the aged
heart is possibly frustrated by a bottleneck at another stage of cell division, and this increased CDK2 expression reflects the increasing need for hypertrophy/hyperplasia with age (particularly in the male heart with the observed loss of cardiomyocytes with age).

**Figure 3.6:** mRNA expression of KCNMA1 from the microarray analysis. The line represents the least-squares regression line for the effect of age on the expression of KCNMA1 in LV NF myocardium (FDR q-value = 0.01). Red spots denote female and blue spots denote male subjects.
Figure 3.7: mRNA expression of LIMK1 from the microarray analysis. The line represents the least-squares regression line for the effect of age on the expression of LIMK1 in LV NF myocardium (FDR q-value = 0.005). Red spots denote female and blue spots denote male subjects.

Toll-like receptors, found both on the plasma membrane and on endosomes intracellularly, can bind to conserved structures present on microbes and are involved in activation of the innate immune system. Toll-like receptor 6 (TLR6), as well as others of this class, is expressed on cardiomyocytes (Frantz, S. et al. 2007) and polymorphisms of
TLR6 have been linked to propensity to develop LV wall thickening in response to systolic hypertension in women (Sales, M. L. et al. 2010). Specifically, the authors found that the TLR6 polymorphism associated with reduced activation in response to the agonist zymosan, was associated with reduced LV wall thickness in the context of systolic hypertension. TLR6 mRNA was found to be downregulated with age in the microarray experiments (Fig. 3.9), and this downregulation might inhibit LV hypertrophy/hyperplasia in response to the increased haemodynamic loads in the LV, adversely affecting CO at older ages. As to why TLR6 polymorphisms only affect LV wall thickness in women, and the relation to the functioning of the innate immune system in the heart, is unclear.

Nuclear factor of activated T-cells (NF-AT) is a transcription factor that is involved in promotion of hypertrophy in response to increased haemodynamic load in cardiomyocytes, as well as being involved in the inhibition of apoptosis (Vega, R. B. et al. 2003). NF-AT translocates to the nucleus after it is dephosphorylated by calcineurin, which is a phosphatase activated by binding to the Ca\(^{2+}\)-calmodulin complex (Vega, R. B. et al. 2003). In this way, NF-AT transcriptional activity is stimulated by increased cytoplasmic Ca\(^{2+}\), which reflects contractile activity and is chronically elevated in states of increased haemodynamic load. The mRNA coding for nuclear factor of activated T-cells, cytoplasmic, calcineurin-dependent 3 (NFATC3), the isoform of NF-AT found in cardiomyocytes, was found to be more highly expressed in males than in female LV in our study (Fig. 3.10). This may represent an upstream actor in the mechanism of hypertrophy of cardiomyocytes in the male LV with age, to compensate for the
simultaneous decrease in abundance of cardiomyocytes with age, which does not occur (or at least occurs to a lesser extent) in female LV.

Figure 3.8: mRNA expression of CDK2 from the microarray analysis. The line represents the least-squares regression line for the effect of age on the expression of CDK2 in LV NF myocardium (FDR q-value = 0.01). Red spots denote female and blue spots denote male subjects.

In conclusion, the transcriptome of human LV undergoes substantial change with age (and is modified to a lesser extent by gender) and close examination of such changes will
allow discovery of age- and sex-related mechanisms that are related to the anatomical and functional changes that the heart undergoes as described above. More importantly, understanding the changes in the LV transcriptome will shed light on the effect of age and gender on the evolution of various cardiovascular pathologies, including MI, valvular disorders, and HF of a variety of aetiologies.
Figure 3.9: mRNA expression of TLR6 from the microarray analysis. The line represents the least-squares regression line for the effect of age on the expression of TLR6 in LV NF myocardium (FDR q-value = 0.01). Red spots denote female and blue spots denote male subjects.
Figure 3.10: mRNA expression of NFATC3 from the microarray analysis indicating the difference in expression between female and male LV NF myocardium (FDR q-value = 0.01). The limits of the boxes represent the interquartile range for the values, whiskers are 1.5 * inter-quartile range distant from the median (middle line in the boxes).
Chapter 4: Ribosomal Gene Expression in End-Stage HF and LVAD Support

4.1 Introduction

As described in Chapter 3, with ageing of the LV myocardium, ribosome synthesis seems to decrease. In addition, it was also shown that ribosome synthesis may be lower in male LV compared to female LV. This may affect the susceptibility of the aged (and male) heart to the development of HF in response to the superimposition of pathologies like MI and valvular dysfunction. To extend this work, I examined the expression of ribosome coding genes in end-stage HF, and what effect LVAD support may have. It was postulated earlier that the reduced energy status of the failing heart (Ten Hove, M. and Neubauer, S. 2007) may lead to a rational decrease in the synthesis of ribosomes to divert scarce ATP to the myofibrils to maintain contractility. One might expect that mechanical unloading of the LV via LVAD implantation might lead to normalisation of the energy status of the cardiomyocyte, and thus, the expression of ribosome genes was also examined in failing hearts pre- and post-LVAD implantation to further test the above postulate.
4.2 Methods

4.2.1 RNA extraction for Real-time PCR Analysis

For details regarding RNA extraction for real-time PCR see Section 2.2.1.

4.2.2 Ribosome Gene Set Expression in End-Stage HF and LVAD Support

To analyse gene expression changes related to r-protein mRNA associated with HF, two different types of microarray datasets were downloaded from the microarray dataset repository, Gene Expression Omnibus (www.ncbi.nlm.nih.gov/geo/). The first kind were of comparisons of LV samples from patients with end-stage HF versus NF LV samples (GSE1145, GSE1869, GSE3585, GSE3586, GSE5406, and GSE21610), and the second were of comparisons of LV samples from patients with end-stage HF at LVAD implantation versus LV samples from the same patients at LVAD explantation (GSE21610, GSE974, and GSE430). For patient demographics see Table 4.1.

Table 4.1: Patient demographics for the various end-stage HF versus NF LV and end-stage HF pre- and post-LVAD support datasets. *Summary statistics only were available, 'na' means data was not available.

<table>
<thead>
<tr>
<th>GEO Dataset</th>
<th>Reference</th>
<th>Phenotype</th>
<th>Mean Age (SD)</th>
<th>Gender</th>
</tr>
</thead>
<tbody>
<tr>
<td>GSE430</td>
<td>Chen, Y. et al.</td>
<td>IDC (n=7)</td>
<td>47.4 (8.6)</td>
<td>2 females, 5 males</td>
</tr>
<tr>
<td></td>
<td>(2003)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GSE974</td>
<td>Hall, J. L. et al.</td>
<td>IDC (n=8)</td>
<td>44.0 (7.0)</td>
<td>4 females, 4 males</td>
</tr>
<tr>
<td></td>
<td>(2004)</td>
<td>ISCM (n=5)</td>
<td>58.8 (6.1)</td>
<td>0 females, 5 males</td>
</tr>
<tr>
<td></td>
<td></td>
<td>MI (n=6)</td>
<td>53.2 (6.4)</td>
<td>0 females, 6 males</td>
</tr>
<tr>
<td>GEO Dataset</td>
<td>Reference</td>
<td>Phenotype</td>
<td>Mean Age (SD)</td>
<td>Gender</td>
</tr>
<tr>
<td>-------------</td>
<td>-----------</td>
<td>-----------</td>
<td>---------------</td>
<td>--------</td>
</tr>
<tr>
<td>GSE1145</td>
<td>Tsubakihara, M. et al. (2006)</td>
<td>NF (n=11)</td>
<td>61.4 (6.7)</td>
<td>2 females, 9 males</td>
</tr>
<tr>
<td></td>
<td></td>
<td>IDCMM (n=15)</td>
<td>53.7 (18.1)</td>
<td>7 females, 8 males</td>
</tr>
<tr>
<td></td>
<td></td>
<td>ISCM (n=11)</td>
<td>54.1 (13.8)</td>
<td>4 females, 7 males</td>
</tr>
<tr>
<td>GSE1869</td>
<td>Kittleson, M. M. et al. (2005)</td>
<td>NF (n=6)</td>
<td>43.0 (13.1)</td>
<td>1 female, 5 males</td>
</tr>
<tr>
<td></td>
<td></td>
<td>IDCMM (n=21)</td>
<td>45.6 (11.0)</td>
<td>6 females, 15 males</td>
</tr>
<tr>
<td></td>
<td></td>
<td>ISCM (n=10)</td>
<td>55.8 (na) *</td>
<td>0 females, 10 males</td>
</tr>
<tr>
<td>GSE3585</td>
<td>Barth, A. S. et al. (2006)</td>
<td>NF (n=5)</td>
<td>48.8 (9.4)</td>
<td>4 females, 1 male</td>
</tr>
<tr>
<td></td>
<td></td>
<td>IDCMM (n=7)</td>
<td>37.0 (14.1)</td>
<td>3 females, 4 males</td>
</tr>
<tr>
<td>GSE3586</td>
<td>Barth, A. S. et al. (2006)</td>
<td>NF (n=15)</td>
<td>53.7 (4.1)</td>
<td>4 females, 11 males</td>
</tr>
<tr>
<td></td>
<td></td>
<td>IDCMM (n=13)</td>
<td>44.3 (17.5)</td>
<td>5 females, 8 males</td>
</tr>
<tr>
<td>GSE5406</td>
<td>Hannenhalli, S. et al. (2006)</td>
<td>NF (n=16)</td>
<td>54 (12) *</td>
<td>na</td>
</tr>
<tr>
<td></td>
<td></td>
<td>HF (n=194)</td>
<td>57 (12) *</td>
<td>na</td>
</tr>
<tr>
<td>GSE21610</td>
<td>Schwientek, P. et al. (2010)</td>
<td>NF (n=8)</td>
<td>29.0 (17.4)</td>
<td>2 females, 6 males</td>
</tr>
<tr>
<td></td>
<td></td>
<td>IDCMM (n=21)</td>
<td>48.6 (13.4)</td>
<td>2 females, 19 males</td>
</tr>
<tr>
<td></td>
<td></td>
<td>ISCM (n=10)</td>
<td>57.6 (8.9)</td>
<td>0 females, 9 males</td>
</tr>
</tbody>
</table>

Raw CEL file image data was used if available, otherwise pre-normalised data were analysed. Normalisation of Affymetrix CEL data was done using Robust multi-array average (Irizarry, R. A. et al. 2003) implemented in Bioconductor's affy package.

Analyses carried out are detailed further in Appendix C. For the end-stage HF versus NF
datasets, t-statistics for each gene were first generated in R using a linear model for each probe set, where we regressed the normalized expression values against phenotype (Smyth, G. K. 2004). T-statistics were similarly generated for each gene in the comparisons of pre- and post-LVAD implantation HF samples. These t-statistics were then inputted into GSEA to calculate the ES and statistical significance of the association of ribosome gene set expression with phenotype, essentially as described in Section 3.2.2.2. However, given that we were only testing one gene set, we relied on the p-value to decide significance, with a p-value < 0.05 regarded as significant.

To assess the similarity of the datasets analysed, specifically, of the Affymetrix datasets comparing end-stage IDCM (GSE1145, GSE1869, GSE3585, GSE5406, and GSE21610) and ISCM (GSE1145, GSE1869, GSE5406, and GSE21610) with NF LV, we undertook two different, complementary analyses. The first was an analysis of the overlap of the lists of ranked DE genes between the various datasets comparing the same phenotype of HF. This was done by computing the number of genes in common between the top 2000 most highly DE genes in each dataset. To estimate the likelihood of the particular number of overlapping genes occurring by chance, i.e. the p-value, for the null hypothesis of no relation between the various datasets comparing HF LV with NF LV of a particular aetiology, a frequency distribution of number of overlapping genes was computed by randomly sampling 2000 genes from each dataset and computing the overlap for each random sample. A p-value < 0.05 was considered to indicate a statistically significant relationship between the datasets. Note that the above analysis does not take into account that genes show correlated expression patterns. Thus, the frequency distribution generated would be narrower than it otherwise would be, and thus,
the resultant p-values may be underestimations of the probability of observing a particular level of overlap or greater, were the null hypothesis true. The second analysis was more descriptive in nature, and involved the construction of intersection plots illustrating the level of overlap of DE genes among the various datasets comparing end-stage HF LV with NF LV of a particular aetiology. The GSE3586 dataset was not included in these two analyses as it was carried out on a different platform to the others, with probes against different sections of the nucleic acid strands of each particular gene, as well as coverage of fewer genes, making comparison difficult. The GSE21610 and GSE1145 datasets derive from the Affymetrix H133 plus 2 platform which carries 54675 probes, which includes the 22283 probes represented on the Affymetrix H133A platform from which the other datasets derive, and only the subset of probes common to both platforms was analysed.

4.2.3 Real-time PCR Analysis of Ribosome-related Genes in End-Stage HF

Real-time PCR was carried out on LV from 18 NF patients (mean age: 57.3, SD: 4.2 years, 9 females and 9 males), 19 ISCM (mean age: 58.8, SD: 4.1 years, 2 females and 17 males), 10 IDCM (mean age: 57.1, SD: 4.0 years, 2 females and 8 males), and 4 HCM (mean age: 61.5, SD: 3.4, 2 females and 2 males) patients with end-stage HF. This was to validate the changes in expression of r-protein mRNA discovered to be DE with end-stage HF by the microarray analysis. For details of the experimental protocol see Section 2.2.3. As discussed above, a mean relative measure of expression was computed from triplicate PCR reactions for each LV sample. Unpaired two-tailed t-tests were used to compare mRNA expression between NF LV samples and LV samples from patients with
end-stage HF due to the above aetiologies. For all statistical tests, p-values < 0.05 were regarded as significant.

4.2.4 Western Blotting Analysis of RPS18 in End-Stage HF

Protein expression analysis was carried out on LV from 5 NF subjects (55 year old females and 65, 63, 61 and 55 year old males), 4 IDCM (59, 56, and two 52 year old males) and 4 ISCM (65, 59, 58, and 56 year old males) patients with end-stage HF, and separately on 3 NF subjects (55 year old female, 65 and 63 year old males) and 3 HCM (57 year old female, 61 and 26 year old males) patients with end-stage HF to determine changes in expression of RPS18. For details regarding protein extraction, see Section 2.3.1. For details regarding protein quantification, see Section 2.3.2. For details regarding SDS-PAGE and Western blotting, see Section 2.3.3. Mean IInt for LV samples from each phenotype group were compared by unpaired two-tailed t-tests to assess for DE. For all statistical tests, p-values < 0.05 were considered significant.

4.3 Results

4.3.1 Ribosome Gene Set Expression in End-Stage HF and LVAD Support

In two microarray studies comparing end-stage IDCM with NF LV, the ribosome gene set was highly significantly downregulated (GSE5406 and GSE1145), with an additional study showing borderline significance for downregulation with IDCM (GSE21610), and another showing insignificant downregulation (GSE3586) (Table 4.2). Note that the final study had the lowest representation of probes targeting r-protein mRNA species, and thus
the sensitivity to detect significant downregulation would have been reduced as compared
with the other studies. However, in two additional studies (GSE1869 and GSE3585), the
ribosome gene set was significantly upregulated in end-stage IDCM.

In three microarray studies comparing end-stage ISCM with NF LV, the ribosome gene
set was highly significantly downregulated (GSE5406, GSE1145, and GSE21610), and in
one dataset it was found to be significantly upregulated (GSE1869). Interestingly, this last
dataset was a subset of the same microarray dataset in which was discovered an
upregulation in ribosome gene set expression with end-stage IDCM.

Table 4.2: Ribosome gene set (HSA03010_RIBOSOME) expression in various datasets comparing end-
stage HF due to IDCM and ISCM with NF LV. Within the 'Study design' column one finds the phenotype
comparisons carried out with the sample sizes for each group in parentheses. 'Size' refers to how many
nucleic acid targets are present on each platform that bind to mRNA coding for r-proteins. 'ES' is
enrichment score, explained in Section 3.2.2.2, however, in this context, a negative 'ES' means that the
ribosome gene set is downregulated in the end-stage HF or pre-LVAD groups. The studies are ordered in
descending sample size.

<table>
<thead>
<tr>
<th>GEO Dataset</th>
<th>Reference</th>
<th>Study design</th>
<th>Platform</th>
<th>Size</th>
<th>ES</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>GSE5406</td>
<td>Hannenhalli, S. et al. (2006)</td>
<td>IDCM (n=108) vs NF (n=16)</td>
<td>HG-U133A</td>
<td>78</td>
<td>-0.4</td>
<td>0.004</td>
</tr>
<tr>
<td>GSE5406</td>
<td>Hannenhalli, S. et al. (2006)</td>
<td>ISCM (n=86) vs NF (n=16)</td>
<td>HG-U133A</td>
<td>78</td>
<td>-0.43</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td>GSE21610</td>
<td>Schwientek, P. et al. (2010)</td>
<td>IDCM (n=21) vs NF (n=8)</td>
<td>HG-U133_Plus_2</td>
<td>79</td>
<td>-0.29</td>
<td>0.057</td>
</tr>
<tr>
<td>GSE3586</td>
<td>Barth, A. S. et al. (2006)</td>
<td>IDCM (n=13) vs NF (n=15)</td>
<td>Unigene3.1 cDNA Array 37.5K v1.0</td>
<td>53</td>
<td>-0.21</td>
<td>0.78</td>
</tr>
<tr>
<td>GEO Dataset</td>
<td>Reference</td>
<td>Study design</td>
<td>Platform</td>
<td>Size</td>
<td>ES</td>
<td>P-value</td>
</tr>
<tr>
<td>-------------</td>
<td>-----------</td>
<td>--------------</td>
<td>----------</td>
<td>------</td>
<td>----</td>
<td>---------</td>
</tr>
<tr>
<td>GSE1869</td>
<td>Kittleson, M. M. et al. (2005)</td>
<td>IDC (n=21) vs NF (n=6)</td>
<td>HG-U133A</td>
<td>78</td>
<td>0.66</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td>GSE1145</td>
<td>Tsubakihara, M. et al. (2005)</td>
<td>IDC (n=15) vs NF (n=11)</td>
<td>HG-U133_Plus_2</td>
<td>79</td>
<td>-0.54</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td>GSE1145</td>
<td>Tsubakihara, M. et al. (2005)</td>
<td>ISCM (n=11) vs NF (n=11)</td>
<td>HG-U133_Plus_2</td>
<td>79</td>
<td>-0.45</td>
<td>0.001</td>
</tr>
<tr>
<td>GSE21610</td>
<td>Schwientek, P. et al. (2010)</td>
<td>ISCM (n=9) vs NF (n=8)</td>
<td>HG-U133_Plus_2</td>
<td>79</td>
<td>-0.71</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td>GSE1869</td>
<td>Kittleson, M. M. et al. (2005)</td>
<td>ISCM (n=10) vs NF (n=6)</td>
<td>HG-U133A</td>
<td>78</td>
<td>0.61</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td>GSE3585</td>
<td>Barth, A. S. et al. (2006)</td>
<td>IDC (n=7) vs NF (n=5)</td>
<td>HG-U133A</td>
<td>78</td>
<td>0.77</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td>GSE21610</td>
<td>Schwientek, P. et al. (2010)</td>
<td>Paired IDC post- vs pre-LVAD (n=21)</td>
<td>HG-U133_Plus_2</td>
<td>79</td>
<td>0.32</td>
<td>0.05</td>
</tr>
<tr>
<td>GSE21610</td>
<td>Schwientek, P. et al. (2010)</td>
<td>Paired ISCM post- vs pre-LVAD (n=9)</td>
<td>HG-U133_Plus_2</td>
<td>79</td>
<td>0.81</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td>GSE974</td>
<td>Hall, J. L. et al. (2004)</td>
<td>Paired IDC post- vs pre-LVAD (n=8)</td>
<td>HG-U133A</td>
<td>78</td>
<td>0.49</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td>GSE430</td>
<td>Chen, Y. et al. (2003)</td>
<td>Paired IDC post- vs pre-LVAD (n=7)</td>
<td>HG-U133A</td>
<td>78</td>
<td>0.72</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td>GSE974</td>
<td>Hall, J. L. et al. (2004)</td>
<td>Paired ISCM post- vs pre-LVAD (n=5)</td>
<td>HG-U133A</td>
<td>78</td>
<td>0.5</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td>GSE974</td>
<td>Hall, J. L. et al. (2004)</td>
<td>Paired AMI post- vs pre-LVAD (n=6)</td>
<td>HG-U133A</td>
<td>78</td>
<td>-0.62</td>
<td>&lt; 0.001</td>
</tr>
</tbody>
</table>
In order to understand the above discrepancies in results, I undertook to explore how similar the various datasets were to each other. Almost all the Affymetrix datasets showed more overlap of genes ranked as DE than would be expected by chance, however certain studies showed little overlap with each other. For studies comparing end-stage IDCM with NF LV, out of the top 2000 genes in order of strength of DE by p-value, only 141 genes were found in common to GSE1145 and GSE21610, an amount of overlap that would be expected by chance (p-value = 0.99), and similarly, only 77 in common between GSE1145 and GSE3585 (p-value = 0.59). For studies comparing end-stage ISCM LV with NF LV, all studies showed significant overlap in terms of the number of genes in common ranked in the top 2000 most highly DE genes. To illustrate the level of overlap between datasets in a more intuitive way, see Figs. 4.1 to 4.8. Note firstly, the relatively low proportion of highly ranked DE Affymetrix probes that overlap amongst the various datasets for both IDCM and ISCM comparisons (Figs. 4.1 to 4.4). Despite this lack of substantial overlap, a few key observations can be made. The first is that it appears that the datasets most concordant with each other with respect to IDCM are GSE21610 and GSE5406, which happen to also be the two largest datasets in terms of sample size. Note that the analysis of both these datasets found a downregulation of the ribosome gene set with end-stage IDCM. With respect to ISCM, GSE5406 and GSE1145 are the most concordant with each other, and these two datasets also have the largest sample sizes for the ISCM LV versus NF LV comparison, and in both was a significant downregulation of the ribosome gene set observed in end-stage ISCM.
Figure 4.1: Intersection plot showing the proportion of overlapping Affymetrix probes ranked by strength of evidence of DE (i.e. p-value) among the datasets comparing end-stage IDCM LV and NF LV. Black represents the overlap between GSE21610 and GSE3585, blue represents the overlap between GSE5406 and GSE1869, brown represents the overlap between GSE21610 and GSE5406, cyan represents the overlap between GSE21610 and GSE1145, green represents the overlap between GSE1145 and GSE5406, grey represents the overlap between GSE1145 and GSE3585, orange represents the overlap between GSE5406 and GSE3585, purple represents the overlap between GSE21610 and GSE1869, red represents the overlap between GSE1145 and GSE1869, and yellow represents the overlap between GSE3585 and GSE1869.
Figure 4.2: Intersection plot showing the proportion of overlapping Affymetrix probes ranked by strength of evidence of DE (i.e. p-value) among the datasets comparing end-stage IDC M LV and NF LV for the top 2000 most highly ranked Affymetrix probes in each dataset. Black represents the overlap between GSE21610 and GSE3585, blue represents the overlap between GSE5406 and GSE1869, brown represents the overlap between GSE21610 and GSE5406, cyan represents the overlap between GSE21610 and GSE1145, green represents the overlap between GSE1145 and GSE5406, grey represents the overlap between GSE1145 and GSE3585, orange represents the overlap between GSE5406 and GSE3585, purple represents the overlap between GSE21610 and GSE1869, red represents the overlap between GSE1145 and GSE1869, and yellow represents the overlap between GSE3585 and GSE1869.
Figure 4.3: Intersection plot showing the proportion of overlapping Affymetrix probes ranked by strength of evidence of DE (i.e. p-value) among the datasets comparing end-stage ISCM LV and NF LV. Blue represents the overlap between GSE5406 and GSE1869, brown represents the overlap between GSE21610 and GSE5406, cyan represents the overlap between GSE21610 and GSE1145, green represents the overlap between GSE1145 and GSE5406, purple represents the overlap between GSE21610 and GSE1869, red represents the overlap between GSE1145 and GSE1869.
**Figure 4.4:** Intersection plot showing the proportion of overlapping Affymetrix probes ranked by strength of evidence of DE (i.e. p-value) among the datasets comparing end-stage ISCM LV and NF LV for the top 2000 most highly ranked Affymetrix probes in each dataset. Blue represents the overlap between GSE5406 and GSE1869, brown represents the overlap between GSE21610 and GSE5406, cyan represents the overlap between GSE21610 and GSE1145, green represents the overlap between GSE1145 and GSE5406, purple represents the overlap between GSE21610 and GSE1869, red represents the overlap between GSE1145 and GSE1869.
If we separate the most highly ranked DE Affymetrix probes according to direction of expression change with HF, we can observe more clearly the similarities between the various datasets (Figs. 4.5 to 4.8). For both IDCM and ISCM comparisons, GSE1145 and GSE5406 are more highly concordant with respect to the other datasets. In addition, the concordance between GSE5406 and GSE21610 for both IDCM and ISCM comparisons is also high with respect to the other datasets. By contrast, the concordance between GSE1145 and GSE1869 is particularly low, especially for the IDCM comparison. To conclude, it appears that the most highly concordant datasets for both IDCM and ISCM, seem to be the datasets with the highest sample sizes, and it is these datasets that also consistently show a downregulation of ribosome gene set expression with end-stage HF related to IDCM and ISCM. The reverse effect is seen when comparing studies with lower sample sizes. Note GSE3583 and GSE1869 analysing end-stage IDCM, where we observe a moderately low concordance, and GSE3585 and GSE1145 also analysing end-stage IDCM, where we observe a similarly low level of concordance.

With regards to the studies examining the effect of LVAD support in patients with end-stage ISCM, both GSE21610 and GSE974 showed a highly significant upregulation of the ribosome gene set (Table 4.2). Similarly for end-stage IDCM, GSE974 and GSE430 showed a highly significant upregulation of the ribosome gene set, with GSE21610 also showing upregulation, but only with borderline statistical significance (Table 4.2). In a subset of the study (GSE21610) by Hall, J. L. et al. (2004), LV tissue sampled at the time of LVAD insertion, within 10 days of an MI significant enough to cause acute HF, was analysed. This tissue was compared to LV tissue from the same
patients at the time of LVAD removal and heart transplantation. In this comparison, there was observed a decrease in ribosome gene set expression post LVAD support.

Figure 4.5: Intersection plot showing the proportion of overlapping Affymetrix probes that are upregulated ranked by strength of evidence of DE (i.e. p-value) among the datasets comparing end-stage IDCM LV and NF LV for the top 2000 most highly ranked Affymetrix probes in each dataset. Black represents the overlap between GSE21610 and GSE3585, blue represents the overlap between GSE5406 and GSE1869, brown represents the overlap between GSE21610 and GSE5406, cyan represents the overlap between GSE21610 and GSE1145, green represents the overlap between GSE1145 and GSE5406, grey represents the overlap between GSE1145 and GSE3585, orange represents the overlap between GSE5406 and GSE3585, purple
represents the overlap between GSE21610 and GSE1869, red represents the overlap between GSE1145 and GSE1869, and yellow represents the overlap between GSE3585 and GSE1869.

Figure 4.6: Intersection plot showing the proportion of overlapping Affymetrix probes that are downregulated ranked by strength of evidence of DE (i.e. p-value) among the datasets comparing end-stage IDCMLV and NF LV for the top 2000 most highly ranked Affymetrix probes in each dataset. Black represents the overlap between GSE21610 and GSE3585, blue represents the overlap between GSE5406 and GSE1869, brown represents the overlap between GSE21610 and GSE5406, cyan represents the overlap between GSE21610 and GSE1145, green represents the overlap between GSE1145 and GSE5406, grey represents the overlap between GSE1145 and GSE3585, orange represents the overlap between GSE5406
and GSE3585, purple represents the overlap between GSE21610 and GSE1869, red represents the overlap between GSE1145 and GSE1869, and yellow represents the overlap between GSE3585 and GSE1869.

**Figure 4.7:** Intersection plot showing the proportion of overlapping Affymetrix probes that are upregulated ranked by strength of evidence of DE (i.e. p-value) among the datasets comparing end-stage ISCM LV and NF LV for the top 2000 most highly ranked Affymetrix probes in each dataset. Blue represents the overlap between GSE5406 and GSE1869, brown represents the overlap between GSE21610 and GSE5406, cyan represents the overlap between GSE21610 and GSE1145, green represents the overlap between GSE1145 and GSE5406, purple represents the overlap between GSE21610 and GSE1869, red represents the overlap between GSE1145 and GSE1869.
**Figure 4.8:** Intersection plot showing the proportion of overlapping Affymetrix probes that are downregulated ranked by strength of evidence of DE (i.e. p-value) among the datasets comparing end-stage ISCM LV and NF LV for the top 2000 most highly ranked Affymetrix probes in each dataset. Blue represents the overlap between GSE5406 and GSE1869, brown represents the overlap between GSE21610 and GSE5406, cyan represents the overlap between GSE21610 and GSE1145, green represents the overlap between GSE1145 and GSE5406, purple represents the overlap between GSE21610 and GSE1869, red represents the overlap between GSE1145 and GSE1869.
4.3.2 Real-time PCR Analysis of Ribosome-related Genes in End-Stage HF

To confirm the above findings, real-time PCR was carried out on samples of LV from patients with end-stage HF due to IDCM, ISCM, as well as HCM, and compared to NF LV (Table 4.3). With respect to ISCM, 4 out of 8 r-protein mRNA tested were confirmed to be downregulated in end-stage HF. With respect to IDCM, 5 out of 8 r-protein mRNA were confirmed to be downregulated. And finally, in HCM, 4 out of the 8 r-protein mRNA tested were found to be downregulated. Note that even for the r-protein mRNA found not to be statistically significantly DE, the direction of DE for all were towards downregulation in end-stage HF. In end-stage HF from IDCM, ISCM and HCM, 45s rRNA was found to be statistically significantly downregulated (Table 4.3). Finally, RHEB mRNA was found to be downregulated only in end-stage HF due to HCM.

Table 4.3: Real-time PCR results comparing mRNA expression of various genes in end-stage HF versus NF hearts. Results are presented as percentage change from NF LV (n=8) with associated p-values in parentheses. Sample size is in parentheses beside each phenotype group label, with 'End-stage HF' being a composite group of all HF patients. Real-time PCR results normalised to Scr1.

<table>
<thead>
<tr>
<th>Gene</th>
<th>End-stage HF (n=33)</th>
<th>ISCM (n=19)</th>
<th>IDC (n=10)</th>
<th>HCM (n=4)</th>
</tr>
</thead>
<tbody>
<tr>
<td>45s rRNA</td>
<td>-28.8 (0.009)</td>
<td>-24.7 (0.03)</td>
<td>-34.2 (0.005)</td>
<td>-37.7 (0.006)</td>
</tr>
<tr>
<td>RPS10</td>
<td>-16.1 (0.03)</td>
<td>-16.1 (0.04)</td>
<td>-11.0 (0.25)</td>
<td>-34.7 (0.07)</td>
</tr>
<tr>
<td>RPS21</td>
<td>-20.3 (0.007)</td>
<td>-18.7 (0.02)</td>
<td>-19.5 (0.01)</td>
<td>-30.9 (0.006)</td>
</tr>
<tr>
<td>RPL26</td>
<td>-19.4 (0.01)</td>
<td>-15.5 (0.04)</td>
<td>-20.2 (0.02)</td>
<td>-31.8 (0.04)</td>
</tr>
<tr>
<td>RPL37</td>
<td>-14.5 (0.09)</td>
<td>-10.0 (0.30)</td>
<td>-18.2 (0.04)</td>
<td>-29.1 (0.09)</td>
</tr>
<tr>
<td>RPS7</td>
<td>-29.1 (0.01)</td>
<td>-28.4 (0.01)</td>
<td>-27.7 (0.02)</td>
<td>-37.6 (0.046)</td>
</tr>
<tr>
<td>RPS18</td>
<td>-9.4 (0.30)</td>
<td>-11.1 (0.25)</td>
<td>-2.6 (0.82)</td>
<td>-18.8 (0.25)</td>
</tr>
<tr>
<td>RPL23A</td>
<td>-11.6 (0.16)</td>
<td>-10.7 (0.25)</td>
<td>-13.4 (0.11)</td>
<td>-13.4 (0.47)</td>
</tr>
<tr>
<td>RPL34</td>
<td>-4.8 (0.52)</td>
<td>-1.6 (0.81)</td>
<td>-3.2 (0.73)</td>
<td>-20.6 (0.21)</td>
</tr>
</tbody>
</table>
## 4.3.3 Western Blotting Analysis of RPS18 in End-Stage HF

The rabbit anti-human RPS18 antibody stained a band in the Western blot experiments at the correct migration point of ~17 kDa (Fig. 4.10 and Appendix A). Note similar loading of protein among all wells is observed (Fig. 4.9). However, no difference in expression between end-stage IDCM and ISCM LV versus NF LV were observed (mean IInt of 0.73 and 0.75 versus 0.70 with associated p-values = 0.77 and 0.62, respectively) (Fig. 4.10). Similarly, no difference in RPS18 protein expression was observed between end-stage HCM and NF LV (mean IInt 1.03 and 0.79, respectively, p-value = 0.19) (Figs. 3.4 and 3.5).

<table>
<thead>
<tr>
<th></th>
<th>End-stage HF (n=33)</th>
<th>ISCM (n=19)</th>
<th>IDCM (n=10)</th>
<th>HCM (n=4)</th>
</tr>
</thead>
<tbody>
<tr>
<td>RHEB</td>
<td>-17.2 (0.09)</td>
<td>-12.5 (0.26)</td>
<td>-19.5 (0.09)</td>
<td>-37.5 (0.008)</td>
</tr>
<tr>
<td>MTOR</td>
<td>7.5 (0.41)</td>
<td>11.8 (0.32)</td>
<td>3.2 (0.70)</td>
<td>-4.3 (0.04)</td>
</tr>
</tbody>
</table>

**Figure 4.9:** SDS-PAGE using a Nusep 4-20% gradient gel loaded with 40 μg LV total protein. The gel was UV irradiated for 1 minute prior to imaging to activate the Stain-free compound to visualise tryptophan-containing proteins. Protein bands corresponding to size range 75 (top) to 25 kDa (bottom) are shown. Samples from left: alternating NF, IDCM, and ISCM LV samples.
Figure 4.10: Western blot with 1:500 rabbit anti-human RPS18 antibody and 1:10000 IRDye800 donkey anti-rabbit IgG (Li-Cor). Samples from left: alternating NF, IDC, and ISCM LV samples.

4.4 Discussion

4.4.1 Technical Interpretation of Results

It has been demonstrated among several microarray datasets comparing NF LV with LV from patients with end-stage HF due to IDC and ISCM, that there is a decrease in expression of ribosome subunit coding mRNA. More precisely, mRNA coding for r-proteins are more likely to be downregulated than the average mRNA transcript. However, there were discrepancies in the results of the analyses of the microarray studies above, as already mentioned. Not all microarray studies comparing end-stage IDC and ISCM with NF LV found a decrease in r-protein expression, and in fact, analysis of the datasets from Kittleson, M. M. et al. (2005) and Barth, A. S. et al. (2006) demonstrated the reverse. These studies had smaller sample sizes, making them more susceptible to idiosyncrasies in the biology of the particular patients involved, which would be unrelated to HF pathology. This would explain the relatively low concordance of ranked DE genes in end-stage IDC among the lower sample size datasets derived from Kittleson, M. M. et al. (2005), Barth, A. S. et al. (2006), and Tsubakiwa, M. et al. (2005). It is also possible that there may have been confounding by age and gender. As
demonstrated in Chapter 3, there is evidence to suggest that increasing age and male
gender is associated with lower expression of r-protein mRNA. These factors were not
incorporated into the analyses as age and/or gender information was not available for all
the datasets. For the datasets for which age and/or gender data was available, matching
the human subjects in each phenotype group according to said factors would have
dramatically reduced the sample sizes of the groups. The assumption was thus made that
the effect on the transcriptome of end-stage HF would far outweigh the effects of age and
gender. The strength of the findings above rests on the validity of this assumption. As
discussed in Chapter 3, the particular gene set testing strategy used may underestimate
the p-values for the association between the expression of the r-protein gene set and the
phenotype in question (Goeman, v J. J. and Buhlmann, P. 2007). However, given that
multiple microarray datasets have been analysed, which constitutes a form of biological
replication, the results can be accepted with more confidence.

Downregulation of particular r-protein mRNA in end-stage HF due to IDC and
ISCM, and additionally in HCM, has also been confirmed using real-time PCR. This
lends support to the results of the microarray experiments, as well as extending them to
the particular aetiology of HCM. Note that 45s rRNA expression was also shown to be
downregulated in all forms of end-stage HF, and this lends support to the idea of a
coordinated downregulation of mRNA coding for ribosomal subunits. Given the
argument put forth explaining the inability to demonstrate a decrease of 45s rRNA
expression with age or sex in Section 3.4.1, the demonstration of such a downregulation
in end-stage HF due to ISCM, IDC, and HCM increases our confidence that this is a
true feature of end-stage HF.
We did not observe a difference in protein expression of RPS18 in end-stage HF due to ISCM, IDC and HCM, despite the differences in mRNA expression observed for many other ribosomal subunit coding genes. Whilst we did not observe a statistically significant downregulation of RPS18 mRNA in any end-stage HF group in our real-time PCR experiments, we would expect the protein abundance of RPS18 to be downregulated in the context of decreased expression of mRNA coding for other r-proteins that make up the same subunit (40s or 60s) (Robledo, S. et al. 2008). It has been explained above in Section 3.4.1 how such a difference may not be appreciable using conventional protein quantification methods employed here. In addition, the sample size particularly in these experiments was severely limiting, and a larger study should be contemplated to properly address this question.

4.4.2 Biological Interpretation of Results

The above experiments demonstrate a decrease in the synthesis of components that make up the ribosome in end-stage HF of a variety of aetiologies. This would result in a decrease in ribosome abundance within cardiomyocytes in end-stage HF, leading to a decrease in the capacity for protein synthesis, turnover, hypertrophy, and hyperplasia.

Pagani, E. D. et al. (1988) have found that myofibrillar protein content per unit wet weight of LV myocardium is decreased in end-stage IDC and ISCM hearts when compared to NF hearts. Scholz, D. et al. (1994) have demonstrated that the myofibrillar volume density of LV cardiomyocytes from patients with end-stage IDC is reduced. These findings could be explained by a decrease in myofibril content in cardiomyocytes, as well as by the replacement of cardiomyocytes with non-muscle cells. Extending the
above findings, Hein, S. et al. (1994) have demonstrated reduced expression of actin, myosin, troponin T, tropomyosin and titin in end-stage IDCM hearts using immunofluorescence, with marked heterogeneity of loss of expression between and within, individual cardiomyocytes. Furthermore, Hammond, E. H. et al. (1987) found that myofibril loss in IDCM as assessed by EM is predictive of worse prognosis, similar to findings from a study by Figulla, H. R. et al. (1985) in ISCM. Decreased ribosome abundance may lead to decreased production of myofibrillar proteins within cardiomyocytes in end-stage HF. Under conditions of decreased capacity for protein turnover, it is also possible that the loss of appreciable myofibrillar proteins may represent an increase in deleterious post-translational modifications, leading to alterations in tertiary structure and breakdown of myofibrils, with the accumulation of misfolded proteins in the cytosol that are not detected by the various techniques for quantifying myofibrils used in the above studies. Supporting the latter proposal, Pagani, E. D. et al. (1988) found that the Mg-ATPase activity of myofibrils extracted from end-stage IDCM and ISCM LV is reduced compared to NF LV, and this could be due to reduced protein turnover leading to the accumulation of actin and myosin filaments with increased deleterious post-translational modifications resulting in reduced enzymatic activity. Similar decreases in myofibril ATPase activity in LV tissue from patients with end-stage HF was demonstrated by Peters, T. J. et al. (1977). However, in this study myofibril ATPase activity was expressed as a function of total extracted protein, and thus reduced synthesis of actin/myosin or other myofibril proteins that modulate ATPase activity may account for the findings. Pagani, E. D. et al. (1998) also found that Mg-ATPase activity was depressed in LV papillary muscle from patients undergoing mitral valve replacement.
for severe mitral regurgitation. These hearts were not in end-stage HF at the time of surgery, and this group was not found to have reduced myofibril proteins per unit wet weight of myocardial tissue. It is possible that pre-end-stage HF hearts such as these with severe mitral regurgitation are suffering an energy deficit that is not as profound as that found in end-stage HF, such that there is a decrease in protein turnover, but the rate of synthesis is not depressed to such an extent that there is a net loss of myofibrils. Therefore, a decrease in Mg-ATPase activity is found due to accumulation of myofibril proteins with accumulated deleterious post-translational modifications. Evidence for this proposed mechanism comes from a study by Gubdajarnason, S. et al. (1964), who measured protein turnover/synthesis by radioactive glycine accumulation in rabbit LV myocardium. It was found that the induction of aortic stenosis led to increased protein synthesis 2 days post surgery, and protein turnover was mildly increased 1-6 months post operation, as compared with sham operated rabbits. However, for the rabbits that developed acute HF post operation, protein synthesis was increased, but not to the extent of the group who did not develop acute HF, and finally, in the rabbits who developed chronic HF 1-6 months post operation, protein turnover was found to be lower compared to all other groups.

In the microarray datasets analysed above, LVAD implantation into hearts with end-stage HF due to IDC or ISCM was shown to lead to an upregulation of ribosome gene set expression. Furthermore, in their own analysis, Schwientek, P. et al. (2010) found that LVAD implantation corrected the downregulation in end-stage HF of elongation factor, RNA polymerase II, 2, which is a protein that is believed to increase the overall rate of mRNA transcription (Shilatifard, A. et al. 1997). The analysis of
ribosome gene set expression with LVAD support is consistent with the general contention of this thesis that energy deficits drive downregulation of ribosome production, by demonstrating that reducing myocardial energy expenditure by mechanical unloading leads to an increase in r-protein coding mRNA expression. My results also concur with Razeghi, P. et al. (2006) who found increased RPS10 mRNA and 18s rRNA expression post LVAD implantation in end-stage HF due to IDCM and ISCM. Similarly, Matkovich, S. J. et al. (2009) found that RPL17 is downregulated in end-stage HF, and that this decrease was attenuated by LVAD support. At the same time, LVAD support leads to increased mRNA expression of multiple sarcomeric and non-sarcomeric genes including β-myosin heavy chain (Rodrigue-way, A. et al. 2005). LVAD support has also been demonstrated to ameliorate, though incompletely, the distorted myofibrillar structure within cardiomyocytes from patients with end-stage HF (de Jonge, N. et al. 2002). These findings are consistent with a normalising of myocardial energy status leading to a resumption of protein synthesis/repair of cellular structures, assisted by an increase in ribosome synthesis. LVAD support appears to cause reduced cardiomyocyte volume, length and width in conditions with pathological hypertrophy associated with end-stage HF (Ambardekar, A. V. and Buttrick, P. M. 2011), but the effect on the total volume of the myofibrillar compartment is unclear, and prolonged unloading may be leading to cardiomyocyte atrophy which might be harmful (more on this later). LV cardiomyocytes from patients with end-stage HF exhibit decreased rates of shortening, magnitude of shortening and rates of relaxation, alongside decreased peak cytosolic Ca\(^{2+}\) and rate of Ca\(^{2+}\) uptake, and LVAD support has been demonstrated to reverse these deficits (Dipla, K. et al. 1998). Similarly, LVAD support has been shown to reverse, albeit
incompletely, the decreased maximal force able to be generated by myofibrils from patients with end-stage IDC (Ambardekar, A. V. et al. 2001). It is suggested here that an increase in r-protein mRNA represents an increase in ribosome synthesis that then leads to increased protein turnover, which reverses the accumulation of proteins with deleterious post-translational modifications which contributes to the aforementioned defects. Muscle ring finger proteins are ubiquitin ligases that have been demonstrated to target multiple proteins (including myosin and troponin I) in the cardiac sarcomere for proteasome-dependent destruction, and knockout mice display impaired cardiac muscle performance, cardiac hypertrophy and early-onset HF (Willis, M. S. et al. 2009). While it has been argued that net accumulation of particular sarcomeric proteins leads to the myofibrillar disarray and hypertrophy associated with these phenotypes, it is also possible that decreased turnover of damaged sarcomeric proteins leads to impairment of force generation that results in cardiac hypertrophy via enhanced neurohumoral signalling in the context of haemodynamic insufficiency. In humans, bortezomib (a proteasome inhibitor) administration for cancer has been observed to lead to HF (Enrico, O. et al. 2007), and the mechanism may be decreased turnover of cardiomyocyte proteins, echoing the effect of ribosome deficiency. As a potential strategy for the testing of the conjecture of reduced protein turnover in end-stage HF, and its reversal with LVAD implantation, one could use the context of 'bridge to transplant' LVAD implantation/explantation. Specifically, patients with end-stage HF scheduled for LVAD insertion could be injected with radioactive amino acids just prior, and then at LVAD implantation, apical tissue removed could be analysed for incorporated radioactivity, as a measure of protein turnover in end-stage HF. At LVAD explantation and heart transplantation, one could
carry out the same injection just prior, and then analyse the explanted heart tissue for incorporation, as a measure of protein turnover post mechanical unloading.

As mentioned earlier, end-stage HF is associated with increased mitosis and cardiomyocyte division (Quaini, F. et al. 1994; Kajstura, J. et al. 1998; Kajstura, J. et al. 2012), presumably in an attempt to compensate for the deficit in contractility. However, clearly this attempt at compensation is inadequate. Wohlschlaeger, J. et al. (2010) have observed a decrease in mean cardiomyocyte DNA content and an increase and decrease in diploid and polyploid cardiomyocytes, respectively, post LVAD support in end-stage HF due to a variety of aetiologies. They hypothesise that this is due to increased cardiomyocyte division post LVAD support. It is possible that this represents an augmentation of the pre LVAD inadequate hyperplastic response in end-stage HF. An increase in ribosome synthesis post LVAD implantation would be in accordance with the aims of cardiomyocyte hyperplasia (and associated hypertrophy). The findings from Wohlschlaeger, J. et al. (2010) can also explain the observation that LVAD support appears to reverse the increased volume, length, and width of cardiomyocytes in LV from patients with end-stage HF (Ambardekar, A. V. and Buttrick, P. M. 2011). This may be due to an increase in cardiomyocyte division, with smaller daughter cells, rather than due solely to atrophy of cardiomyocytes.

RHEB mRNA expression was shown to be decreased in end-stage HF due to HCM. Notwithstanding that RHEB protein expression has not yet been shown to be decreased in end-stage HCM, I think this result must be accepted cautiously as only 4 patients were present in this group. Also, the fact that such a downregulation was not seen in IDC and ISCM, highlights that even if such an effect is true, it does not represent a
feature of end-stage HF in general, but rather a peculiarity of HCM itself, which is not the focus of my thesis.

Analysing the dataset from Hall, J. L. *et al.* (2004), in patients in whom LVAD insertion was performed who were < 10 days post MI severe enough to cause acute HF, r-protein mRNA expression was lower post LVAD explantation. This can be explained by the observation that hypertrophy/hyperplasia is markedly increased in the remaining viable myocardial tissue immediately after an MI occurs (Beltrami, A. P. *et al.* 2001). This process, which would require increased ribosome synthesis, would diminish over time as the deficit in contractile tissue is reduced. Thus at LVAD explantation, ribosome synthesis would be at a lower level compared to immediately post MI during which a burst of increased protein synthesis was occurring. This is despite ribosome synthesis at LVAD explantation being higher than it otherwise would be, had the LVAD not been inserted.
Chapter 5: The Effect of Age, Sex and End-Stage HF on Nucleolar Remodelling

5.1 Introduction

The nucleolus, the site of ribosome synthesis (Section 1.2.5.5), is a roughly spherical structure of approximately 1 μm in diameter. It is subdivided into three distinct regions: 1) the fibrillar centre (FC) containing ribosomal deoxyribonucleic acid (rDNA) and RNA polymerase I subunits, 2) the dense fibrillar component (DFC), and 3) the granular component (GC) (Fig. 5.1 and 5.2). The nucleolus is the exclusive site of ribosome synthesis, but it performs a number of other functions, including regulation of mitosis, cell cycle arrest, and apoptosis in response to cellular stress via actions of nucleolar proteins on the levels of p53 (reviewed in Boisvert, F. M. et al. 2007). The nucleolus contains nucleolar organising regions (NOR) that are composed of clusters of rDNA repeat units. In man, there are approximately 400 copies of 43 kb repeat units distributed along chromosomes 13, 14, 15, 21, 22). Active NORs are a subset of NORs that are being actively transcribed, and are bound to the RNA pol I complex within the FC and the border zone between the FC and the DFC. RNA pol I, along with nucleolin and a number of other proteins that are localised to the FC and the DFC stain avidly with silver ions,
and the stained nucleoprotein complexes are known as AgNORs (Derenzini, M. 2000). The AgNOR area within nuclei is directly related to the size/abundance of nucleoli and the level of ribosome synthesis (Derenzini, M. 2000).

Figure 5.1: This figure depicts the various sections of the nucleolus and associated molecular species. Within the FC or at the border between the FC and the DFC, rDNA is transcribed by the RNA pol I complex. The 45s rRNA then migrates to the DFC and then to the GC where various accessory proteins and enzymes process the rRNA precursor into the various mature rRNA species, and r-proteins bind, to form the

120
complete 40s and 60s ribosomal subunits. These exit the nucleolus and form the complete ribosome in the cytoplasm where the process of protein translation occurs. Adapted from Boisvert, F. M. et al. (2007).

**Figure 5.2:** Images of the nucleolus. 'a' is a differential interference contrast (DIC) image of a HeLa cell showing prominent nucleoli (indicated by arrows) within the nucleus. 'c' is another DIC image showing nucleoli purified from HeLa cells (inset is a scanning electron microscopy (EM) image of same). 'd' is a transmission EM image of a uranyl-acetate-stained cell section showing a nucleus with a nucleolus inside (labelled 'Nu'). 'e' is an electron spectroscopic image with phosphate enrichment showing nucleoplasm with a nucleolus with its distinct sections indicated: FC, DFC and GC. Adapted from Boisvert, F. M. et al. (2007).
As discussed in previous chapters, if the haemodynamic load increases, LV cardiomyocytes engage in hypertrophy/hyperplasia to increase the available contractile tissue, to meet the increased load. A paradigmatic example of this is the hypertrophy/hyperplasia that occurs after the loss of LV mass during an MI (Beltrami, A. P. et al. 2001). However, during end-stage HF of any particular aetiology, given the decrease in 45s rRNA and r-protein expression (at least at the mRNA level) that was observed in this study (Chapter 4), I hypothesised that the LV myocardium makes the compromise of diverting all available ATP to contraction and deprioritises ribosome synthesis which is energetically costly, as a strategy to maximise cardiac output in the short-term. If this is correct, one might expect to see a decrease in abundance of nucleoli in the LV cardiomyocyte nuclei of patients with end-stage HF, regardless of aetiology. This was assessed by quantifying the area of silver-stained AgNORs in LV samples from patients with end-stage HF of various aetiologies and comparing with LV samples from NF hearts.

With age, it has already been demonstrated that r-protein mRNA synthesis is decreased with age, and there is lower r-protein mRNA synthesis in males versus females. However, 45s rRNA synthesis was not found to decrease with age, nor differ by sex. Neither were protein levels of RPS18 shown to decrease with age or differ by sex. Regardless, the effect of age and gender on AgNOR staining area in NF LV was also assessed.
5.2 Methods

5.2.1 Patient Demographics

In the first part of this study, AgNO₃ staining (for AgNOR area) was carried out on LV samples from 34 NF subjects, age range: 2 months to 66 years, to determine if abundance of nucleoli was altered with age or sex. Of these 33, there were 15 females, mean age: 37.3, SD: 14.1 years and 18 males, mean age: 35.5, SD: 17.6 years. In the second part of this study, AgNO₃ staining was carried out on LV samples from 8 NF subjects (mean age: 38.3, SD: 13.4 years, 6 females and 2 males), and 19 IDCM (mean age: 43.6, SD: 14.1 years, 4 females and 15 males), 3 HCM (mean age: 45.7, SD: 4.2 years, 3 females), and 11 ISCM (mean age: 54, SD: 6.6 years, 11 males) patients with end-stage HF to determine if abundance of nucleoli was altered in different forms of end-stage HF. For details regarding the collection of human heart tissue see Section 2.1.

5.2.2 Tissue Microarray Construction

To construct the tissue microarrays (TMA), samples of human LV stored in liquid nitrogen were slowly thawed to reduce damage from ice crystal formation. This was done by removing the samples in cryotubes from liquid nitrogen storage and placing them immediately onto dry ice, making sure that the cryotubes were completely covered. After 45 minutes of dry ice incubation, the cryotubes were transferred to wet ice, and incubated for a further 30 minutes. The tissue was then removed from the cryotubes and cut into 1 cm cubed pieces and incubated in 5% phosphate buffered formaldehyde for 24 to 36 hours. The fixed tissue samples were then embedded in paraffin and 1 mm diameter cores
were taken and embedded in a separate paraffin block in a regular grid pattern. 4 μm thick sections were cut from these blocks of assembled cores and affixed to a glass slide. These sections were stained as described below within 48 hours of sectioning.

5.2.3 AgNO₃ Staining

All the following steps were carried out at room temperature in a chemical fume hood unless otherwise stated. TMA staining with AgNO₃ was carried out as described in Trere, D. (2000) with some modifications. The slides were dewaxed and hydrated as follows: 2 x 5 minutes in histolene, 2 x 2 minutes in 100% EtOH, 1 x 2 minutes in 90% EtOH, 1 x 5 minutes in 70% EtOH, 2 minutes in 50% EtOH, 2 x 2 minutes in reverse osmosis (RO) water. An antigen retrieval step was then performed, with the slides immersed in 0.01 M sodium-citrate monohydrate, pH 6.0 and incubated at 120° C for 20 minutes in a wet autoclave. The slides were allowed to return to room temperature and then rinsed 3 times in RO water for 1 minute. Immediately before use, 2 parts of 50% AgNO₃ solution was mixed with 1 part of 2% gelatin plus 1% formic acid solution. Both solutions were at 25° C. The slides were then incubated in this solution for 15 minutes at 25° C in a dark, humidified chamber. After this incubation, slides were washed three times in RO water, agitating for 1 minute in each wash. To improve stain specificity, the slides were then incubated for 5 minutes in 5% sodium-thiosulfate, after which they were washed three times in RO water as above (Lindner, L. E. 1993). The slides were then dehydrated and cleared as follows: 1 x 5 minutes in 70% EtOH, 1 x 2 min in 90% EtOH, 2 x 2 min in 100% EtOH, and 2 x 10 minutes in histolene. The slides were mounted in DPX and stored at room temperature in the dark and imaged within 48 hours.
5.2.4 Imaging and Quantification of AgNO₃ Staining

The AgNO₃ stained LV samples were imaged using a DMRBE Microscope (Leica) controlled by Microscope and Stereo Investigator V8 (MBF Biosciences). The procedure was as follows: at 5 x magnification, contours were drawn around each LV section present on the TMA and labelled with a serial number that permitted blinding of the observer to the clinical details of the section. Then at 63 x magnification, five randomly located images were taken from each section (Figs. 5.3 and 5.5). If an image contained greater than ~20% empty space due to a split or missing section, or large blood vessel, a new random location was imaged. All sections to be analysed together were imaged on the same day with the same microscope settings to ensure similar overall intensity of images, and thus comparability in quantification and analysis.

Each image, five per tissue section, was analysed in Metamorph V7.6 (Molecular Devices). Given that AgNO₃ is not a specific stain, and that some extracellular staining is common (Fig. 5.3), it was necessary to ensure that only AgNO₃ stained regions within nuclei were included in the quantification. Nuclei were identified by their pale diffuse beige staining (Fig. 5.3), and regions were manually drawn around each nucleus identified. A threshold was set for all images, and pixels darker than said threshold were defined as being occupied by the argyrophilic proteins that stain with AgNO₃ and that make up the AgNOR (Trere, D. 2000).

For each image, for each identified nucleus, the nuclear area (μm²) and AgNOR area (μm²) were recorded. The values for each nucleus were summated to give a total nuclear and total AgNOR area (in μm²) for each image. The mean nuclear area and AgNOR area among the five images for a particular LV section (i.e. a particular LV
sample) were computed. AgNOR area is thus a quantitative measure of the amount of AgNOR present per unit volume of heart tissue. These two values for each particular LV section were also used to compute AgNOR area / nuclear area to give a value that represents the proportion of nuclear area occupied by AgNOR material for each LV section. This was done in order to gain a quantitative measure of AgNOR area that was not affected by the area of nucleus present on each different section thus reducing sampling error introduced by taking random tissue sections from each LV samples with varying amounts of nuclei present. This measure would also represent the amount of AgNOR present per nucleus, rather than per unit volume of tissue.

5.2.5 Statistical Analysis

In order to investigate changes in AgNOR and nuclear area with age and sex, a linear model was applied with the dependent variables: AgNOR area, nuclear area, AgNOR area / nuclear area, and covariates: age and sex. To investigate changes in AgNOR and nuclear area with end-stage HF, the mean and SD of AgNOR area, nuclear area, AgNOR area / nuclear area, were calculated for each group of patients (IDCM, HCM and ISCM) and compared with NF hearts using a two-tailed, unpaired Student's t-test assuming homogeneity of variance between groups. For all statistical tests, a p-value of < 0.05 was considered significant.
5.3 Results

5.3.1 Effects of Age and Gender on AgNOR Abundance

Across all NF subjects, mean AgNOR area was 6672.0 μm² (SD: 8960.2), mean nuclear area was 67487.6 μm² (SD: 34169.6 μm²), and mean AgNOR area / nuclear area was 0.082 (SD: 0.065) (Fig. 5.3).

Regarding the effects of age on AgNOR area, there was a statistically significant decrease in AgNOR area with increasing age (-223.9 μm² per year, p-value = 0.01, Fig. 5.4). However, excluding subjects < 20 years old, there was no significant change in AgNOR area with age (79.7 μm² per year, p-value = 0.10). There was a decrease in nuclear area with age (-767.6 μm² per year, p-value = 0.02) (Appendix D, Fig. D.1). However, again, once NF subjects < 20 years old were excluded, there was no decrease in nuclear area with age (-4.0 μm² per year, p-value = 0.99). Finally, there was no change in AgNOR area / nuclear area with age in NF subjects (-0.001 per year, p-value = 0.10), and in the subset of NF subjects greater than 20 years old (0.001 per year, p-value = 0.12) (Appendix D, Fig. D.2).
Figure 5.3: AgNO₃ stained images of NF LV samples. Clockwise from upper left: 25 year old female, 55 year old female, 65 year old male, and 19 year old male. Note the darkly staining nucleoli present within lightly staining, beige nuclei.
Figure 5.4: Plot illustrating mean AgNOR area (in μm²) versus age (in years) in NF hearts (n = 33). LV samples from females in red, males in blue. The least-squares regression lines for relationship between age and mean AgNOR for all NF hearts (yellow), and for subset of NF subjects >= 20 years old (green) are depicted. Asterisk indicates p-value < 0.05 for linear relationship between age and mean AgNOR area (controlled for gender). Nil statistically significant changes of mean AgNOR area with age (in subset of NF subjects >= 20 years) and gender were observed (p-values > 0.05 for both analyses).

Gender had no effect on AgNOR area (3249 μm² higher in females, p-value = 0.27), controlling for age (Fig. 5.4). This lack of effect was similar in the subset of NF subjects
greater than 20 years old (2032.1 μm² higher in females, p-value = 0.07). Gender had no
effect on nuclear area (18382.3 μm² higher in females, p-value = 0.10), and neither did
gender have an effect in the subset of NF subjects greater than 20 years old (18316.0 μm²
higher in females, p-value = 0.08) (Appendix D, Fig. D.1). Finally, there was no change
in AgNOR area / nuclear area with gender in NF subjects (0.011 higher in females, p-
value = 0.63), nor with subset of NF subjects greater than 20 years old (0.016 higher in
females, p-value = 0.36) (Appendix D, Fig. D.2).

5.3.2 Effect of End-Stage HF on AgNOR Abundance

There was no difference between IDCM samples and NF samples with respect to mean
AgNOR area (4126.3 versus 3436.6 μm², p-value = 0.56, Fig. 5.6), mean nuclear area
(77618.7 versus 75171.7 μm², p-value = 0.86), and mean AgNOR area / nuclear area
(0.051 versus 0.046, p-value = 0.71).

Mean AgNOR area was lower in ISCM samples compared with NF samples
(1487.5 versus 3436.6 μm², p-value = 0.02, Fig. 5.6), and similarly, mean AgNOR area /
nuclear area was lower in ISCM samples compared with NF samples (0.022 versus 0.046,
p-value = 0.04). There was no difference in mean nuclear area between ISCM and NF
samples (62025.1 versus 75171.7 μm², p-value = 0.37).
Figure 5.5: AgNO₃ stained images of LV samples. Clockwise from upper left: 33 year old male NF donor, 29 year old male IDCM, 41 year old female HCM, and 59 year old male ISCM. Note the darkly staining nucleoli present within lightly staining, beige nuclei.
Figure 5.6: Plot illustrating mean AgNOR area (in μm²) in the LV from patients with end-stage HF (ISCM, n = 11; HCM, n = 3; IDCM, n = 19) versus NF subjects (n = 8). Asterisks indicate statistically significant decreases (p-value < 0.05) for comparison with NF patients.

Mean AgNOR area was lower in HCM samples compared with NF samples (1104.267 versus 3436.6 μm², p-value = 0.03, Fig. 5.6), however, mean AgNOR area / nuclear area...
was not different between HCM and NF samples (0.021 versus 0.046, p-value = 0.21). There was no difference in nuclear area in HCM samples compared to NF samples (40338.7 versus 75171.7 μm², p-value = 0.09).

5.4 Discussion

The AgNOR quantification method did not distinguish between the various cell types within the LV myocardium, however, given that cardiomyocytes make up the vast majority of myocardial tissue in terms of volume, it will be henceforth assumed that the changes in AgNOR abundance are occurring in the cardiomyocyte compartment for the purpose of interpretation of results.

We observed an effect of age on the abundance of AgNOR in NF hearts, however, this effect was not observed within the subset of NF patients greater than 20 years old. If AgNOR abundance can be taken as an indicator of nucleolar abundance, and more importantly, an indicator of rate of ribosome synthesis, then no histological evidence was obtained of a change in ribosome synthesis with age in adulthood. Neither was there any effect of gender observed on AgNOR area, nucleolar abundance, and by inference, ribosome synthesis. However, this does not preclude there being a difference in the potential to increase ribosome synthesis via increasing nucleolar capacity at different ages, and in males versus females, when haemodynamic load is increased.

Given the absence of a demonstrable effect of age or gender on the AgNOR area in NF hearts, it was deemed unnecessary to age and gender match the comparisons of NF hearts with end-stage HF hearts of IDCM, HCM, and ISCM aetiologies.
In end-stage ISCM, a statistically significant decrease in AgNOR abundance per unit volume of LV tissue was observed. This finding accords with Mamaev N. N. et al. (1998) who found lower AgNOR numbers in patients with more severe ISCM, as compared with NF hearts. With respect to end-stage HF in HCM patients, there also seems to be a decrease in AgNOR abundance per unit volume of heart tissue. Note that there was no difference in AgNOR volume per volume of nucleus in end-stage HCM LV versus LV from NF hearts. However, given the reduced volume of nuclei per unit volume of LV tissue, there was a reduced amount of nucleoli for the amount of LV contractile tissue that must be maintained. These data accord with the results of Chapter 4 which demonstrated the decreased synthesis of r-protein mRNA and 45s rRNA in end-stage HF from ISCM, IDC, and HCM. However, in these experiments, end-stage IDC does not seem to be associated with a decrease in AgNOR abundance compared to NF LV, the reasons for which are unclear.

There are several limitations to these studies. The first is the AgNOR staining technique, and subsequent quantification, did not distinguish between AgNOR located within cardiomyocytes, fibroblasts, endothelial cells, smooth muscle cells, or other cells present in the myocardium. To demonstrate that reduced AgNOR abundance is occurring in end-stage HF within the cardiomyocyte compartment would require the use of a staining protocol that allows for the positive identification of cardiomyocytes, such as using an antibody to cardiac actin or troponin, linked to a secondary antibody conjugated to a fluorophore or horseradish peroxidase, in order to direct the subsequent quantification to cardiomyocytes only. Another limitation relates to the location of samples of myocardial tissue within the LV in each patient. Within hearts from patients
with end-stage ISCM, there are commonly macroscopic regions of infarcted tissue, sampling of which, will yield mostly fibrotic tissue devoid of cardiomyocytes. A considerable effort was made to avoid sampling from macroscopically fibrotic regions of LV myocardium, which are clearly distinguishable from non-infarcted myocardium. However, it is more difficult to ensure that LV myocardial samples used for preparation of the TMAs do not contain smaller regions of fibrosis, too small to be appreciable by the naked eye, and that these minute fibrotic regions were not incorporated into the tissue sections used for AgNOR quantification. This methodological limitation is shared by the studies in Chapter 4, but is potentially more severe in this study, as the volume of LV myocardium sampled is much smaller, and thus, more likely to be affected by the incorporation of regions of fibrotic tissue which may in the worst case, make up the entire tissue section leading to the analysis being invalid for that sample. The use of a cardiomyocyte-specific stain to positively identify cardiomyocytes as suggested above would be expected to circumvent this, as one could then readily identify a tissue section containing only, or mostly fibrotic tissue, and exclude that section from the analysis.
Chapter 6: General Discussion and Future Work

There remains a lot of additional experimental work that will be required in order to make solid the contention that ribosome synthesis is downregulated with age, and in the presence of end-stage HF, and that there is upregulation with mechanical unloading of the LV in patients with end-stage HF. The mRNA changes in r-protein expression have to be confirmed at the protein level, and this will require further Western blotting with antibodies targeted to multiple r-proteins, not just RPS18, which was focussed on in this thesis. In addition, an analysis of r-protein and 45s rRNA levels in LV samples from patients pre- and post-LVAD insertion will need to be carried out as well. The limitations imposed by protein extraction leading to the necessity of measuring the abundance of r-protein levels relative to total protein, which similarly affects 45s rRNA quantitation, have been discussed previously in Section 3.4.1. Another approach would be to use immunofluorescence/immunohistochemistry with tissue sections from LV to quantify changes in r-protein levels with age, sex, HF, and pre- and post-LVAD support. This would also allow the quantitation of changes in r-proteins in particular cell types (with appropriate cell-specific antibodies), and thus allow us to confirm that changes in ribosome abundance occur in cardiomyocytes, rather than in other cell types. The quantity of fibrotic tissue (both diffuse and focally distributed) is increased in LV from patients with end-stage HF (De Leeuw, N. et al. 2001), and cardiac fibroblasts may have
reduced amounts of ribosome synthesis due to their particular roles in cardiac tissue. If this is the case, then the reduced ribosome synthesis demonstrated in the above experiments may have been partly due to a greater proportion of fibrotic tissue in LV samples from end-stage HF patients versus NF donors. This illustrates the importance of follow up experiments of a nature suggested above. The disadvantage of immunofluorescence and immunohistochemical analyses is however, that issues surrounding the sampling of minute amounts of myocardial tissue would then arise as discussed in Section 5.4. On another note, as r-proteins produced may not necessarily be incorporated into functional ribosomes, direct assessment of intact ribosome abundance per unit volume/weight of LV tissue should be carried out using a method similar to Earl, D. C. and Morgan, H. E. (1968). Nevertheless, this thesis has demonstrated evidence of a decrease in 45s rRNA abundance in ISCM, IDCM, and HCM suggesting that ribosome synthesis is decreased in end-stage HF. And this was supported indirectly by the demonstration of decreased nucleolar abundance, at least in end-stage ISCM and HCM.

The mechanistic basis for the changes in rRNA and r-protein mRNA in the various pathological states examined in this thesis is still unclear. It was mentioned earlier in Chapters 3 and 4 that changes in RHEB and mTOR activity may underlie these changes. The AMP activated protein kinase (AMPK) is an intracellular enzyme activated by energy depletion, manifested as an increase in AMP relative to ATP (Horman, S. et al. 2012). HF, as has been discussed in this thesis, and reviewed in Ten Hove and Neubauer (2007), is a state that is associated with energy depletion within the cardiomyocyte. Specifically, there seems to be reduced concentrations of ATP and phospho-creatine within cardiomyocytes in end-stage HF, however, in less severe HF, ATP can be normal,
but there is decreased phospho-creatine and decreased flux through the creatine phosphate pathway (Weiss, R. G. et al. 2005; Smith, C. S. et al. 2006). Therefore, there is decreased ATP reaching the myofibrils leading to impaired contractility, and also presumably, decreased energy delivery to other cellular structures. AMPK can act to ameliorate cellular energy levels by inhibiting the action of mTOR and the downstream hypertrophic response (Horman, S. et al. 2012). One of the effects of mTOR inhibition is to decrease ribosome synthesis by interfering with mTOR-dependent transcription of rRNA and r-protein mRNA (Mayer, C. and Grummt, I. 2006). A decrease in mTOR mRNA was not detected in ageing nor in end-stage HF, but as mentioned in Section 3.4.2, this does not preclude a change in mTOR activity via changes in protein abundance or activation. mTOR is activated by the binding of RHEB-GTP, a binding which is antagonised by Tuberous Sclerosis Complex-1 (TSC1), a GTPase activating protein that acts on RHEB. Akt/Protein kinase B, a downstream effector of the insulin/Igf-1 receptor/phosphoinositide 3-kinase pro-hypertrophy pathway, inhibits TSC1 GTPase activity, thereby increasing RHEB-GTP abundance, and mTOR activity. AMPK, on the other hand, disinhibits TSC1 GTPase activity, leading to repression of mTOR activity. The analysis of the protein abundance and activities of these various cellular mediators in ageing, between genders, and HF deserves further study.

A major limitation of the work presented in this thesis is the inability to distinguish cause from effect. Many changes at the transcriptomic, proteomic, and nuclear level have been described that are associated with end-stage HF. However, it is difficult to determine whether these changes are further upstream, that is, they were involved in the initiation of the evolution of HF, or further downstream. That is, that they
came after the HF phenotype began to be established, perhaps representing compensatory changes. The distinction is important, as the more upstream a change is, the more likely that intervention to reverse that change will lead to reversal of the HF phenotype. Or perhaps, the greater will be the magnitude of amelioration of the HF phenotype to be achieved. On the other hand, intervention to reverse phenotypic changes associated with HF that are in fact compensatory, would have the unintended effect of worsening the magnitude of the HF condition. It has been hypothesised in this thesis that ribosome synthesis is downregulated in the presence of end-stage HF, as a potentially compensatory mechanism designed to conserve energy, in order to keep available maximal amounts of ATP to be used for force production. And while this compensatory change is beneficial in the short term, over the long term, it might exacerbate the primary defect of decreased force production by leading to the net loss of myofibrils per unit mass of myocardial tissue and/or the accumulation of myofibrillar components with post-translational modifications which renders them less able to efficiently generate force. Decreased turnover of cellular proteins would be expected of course to affect a myriad of cellular activities. Nevertheless, one can consider such a downregulation of ribosome synthesis as compensatory. Thus, if one were to intervene pharmacologically to increase ribosome synthesis, such as by designing a small molecule to decrease RHEB GTPase activity thus leading to increased mTOR activity, it might worsen the energy status of a cardiomyocyte and provoke acute LV failure and death from circulatory insufficiency or ventricular arrhythmia. With respect to female gender, it was hypothesised in Chapter 3 that increased ribosome synthesis might improve HF via increased hypertrophy and hyperplasia, however, this contradicts the ribosome synthesis as detrimental argument put
forth here. It is possible that increased ribosome synthesis and associated hypertrophy/hyperplasia of cardiomyocytes is beneficial at earlier stages of the evolution of the HF phenotype, when the energetic status of the cardiomyocyte is more favourable. On the other side of the energy availability spectrum, viable LV myocardium in patients post MI in the absence of HF seems to exhibit nil deficits in ATP, phospho-creatine, and flux through the creatine kinase pathway (Bottomley, P. A. et al. 2009). Therefore, promotion of ribosome synthesis within cardiomyocytes in this clinical scenario may facilitate increased hypertrophy/hyperplasia, and thus a more complete replenishment of the deficit of contractile tissue, with a decrease in the probability of HF eventually supervening.

What has been missed from the analyses undertaken in this thesis is the changes in the RV that occur with age, gender, end-stage HF, and LVAD implantation. How the RV responds to the contractile dysfunction of the LV determines the clinical course of HF, and a repeat of the above studies in the RV would add much insight into the pathophysiology of HF, and how it is modified by age, sex, and LVAD implantation. As an example, it was noted earlier that female RV is more resistant to the effects of LV systolic dysfunction compared with male RV (Martinez-Selles, M. et al. 2006). Perhaps it is greater ribosome synthesis and associated greater capacity for hypertrophy/hyperplasia of the RV cardiomyocytes in the context of LV failure, that contributes to the better prognosis of HF in women.

To restate the overarching proposal of this thesis, in end-stage HF, ribosome synthesis is not pragmatic, and it is inhibited, whether through the AMPK pathway described above, or through another pathway that awaits description. However, in concert
with mechanical unloading of the LV with an LVAD, the promotion of ribosome synthesis becomes energetically feasible, and turnover of damaged myofibrillar components can be promoted, as well as the addition of more myofibrils. And thus, increasing the likelihood of recovery of LV myocardium and the ability to successfully wean the patient from mechanical support without recourse to transplantation. LVAD implantation however, does not usually lead to sufficient myocardial recovery to allow LVAD removal without subsequent transplantation (Birks, E. J. 2010), and this may be due to insufficient recovery such that the LV myocardium cannot “stand on its own”. It may also be related to the myocardial atrophy that occurs with prolonged unloading (Amebardekar and Buttrick, 2011). Hall, J. L. et al. (2007) describe a strategy of LVAD implantation combined with administration of Clenbuterol, a β2-adrenergic agonist which promotes hypertrophy, and which resulted in the majority of patients achieving sufficient myocardial recovery to be weaned from LVAD support with sustained recovery. They report that among the many transcriptional changes associated with myocardial recovery was an increase in mRNA expression of Ribosomal protein 6 (Hall, J. L. et al. 2007), plus increased protein expression of myofibrillar proteins including sarcomeric actin and myosin heavy chain (Latif, N. et al. 2007). Clenbuterol has been shown to act through activation of mTOR (Kline, W. O. et al. 2007), thus one of its many actions may be to stimulate ribosome synthesis, with increased ribosome synthesis only being beneficial in the context of the favourable energy status of the unloaded cardiomyocyte.

This thesis has presented novel data on the effect of age and gender on the cardiac transcriptome, interpretation of which has only just begun. From my initial attempt at interpreting this trove of data, I was led to examine ribosome gene expression in HF.
From these studies has arisen my suggestion that the promotion of ribosome synthesis in LV myocardium post MI, or subject to other insults likely to lead to HF, may improve the clinical course, or perhaps even prevent the development of HF entirely. This is an avenue of research that I believe is worthy of further investigation, and will hopefully assist in the understanding and future treatment of this terrible illness.
References


Tsubakihara, M. (2005) Transcription Profiling Study of the Human Heart. *Faculty of Medicine, University of Sydney.*


Appendix A: Further Details of Western blot of RPS18

Figure A.1: Full Western blot demonstrating the specificity of the rabbit anti-RPS18 (1:500) and IRDye800 donkey anti-rabbit IgG (1:10000) (Li-Cor) used in the experiments described in Chapters 3 and 4. Note, the bright band at ~ 17kDa which represents the predominant splice form of RPS18 denoted as ENSP00000211372. There is also a faint band at ~ 10 kDa which may correspond to the smaller alternatively spliced forms ENST00000476288 and ENST00000477055 which weigh approximately 9.7 kDa.
Appendix B: R code to Generate the

Results of Chapter 3

# Code to analyse the raw images of the microarrays analysing the effects of age and
# sex on the human LV myocardium, 'morris' is the name of the dataset I used. #
# signifies the beginning of a comment, and > signifies the beginning of output of the
# R environment
library(limma)
morristargets <- readTargets()
RG <- read.maimages(morristargets,source="genepix.median", columns=list(R='F532 Median', G='F635 Median', Rb='B532 Median', Gb='B635 Median'),
wt.fun=wtflags(weight=0,cutoff=-40) )
RG$genes <- readGAL('223272_low.gpr')
RG$printer <- getLayout(RG$genes)
RG <- backgroundCorrect(RG, method="normexp", offset=50)
MA <- normalizeWithinArrays(RG)
remove(RG)
morrisnormaldata <- normalizeBetweenArrays(MA, method='scale')
remove(MA)
# To eliminate genes/probes with fewer than 13 arrays that were analysable
vectorofweightsums<-c()
for (count in 1:20160) {
    vectorofweightsums[count]<-sum(morrisnormaldata[count,]$weights);
}
morrisnormaldata<-morrisnormaldata[vectorofweightsums>13,]
ageofmorris <- c( 19, 48, 37, 8, 37, 31, 39, 55, 50, 19, 41, 39, 23, 39, 47, 52, 40, 8, 55,
37, 53, 42, 23, 65, 44, 17, 4, 25, 26, 21, 36, 45, 50, 49, 49, 56, 24, 51, 52, 49, 52, 48, 53,
50, 60, 49, 65, 52, 65, 34, 27, 23, 56, 52, 48, 54, 61 )
sexofmorris <- c( 1, 0, 0, 1, 1, 1, 0, 1, 1, 1, 0, 0, 1, 0, 0, 0, 1, 0, 0, 0, 1, 0, 1, 0, 1, 1, 0, 1, 0,
1, 0, 0, 0, 1, 1, 0, 1, 1, 1, 0, 0, 0, 0, 0, 0, 1, 1, 0, 1, 0, 1, 0, 1, 1, 1, 1, 1) 
sexofmorris<-factor(sexofmorris)
morrisdesignwithageandsex<-model.matrix(~1+ageofmorris+sexofmorris)
# To map probes on the array to Genbank Accessions and gene symbols
oligoIDmappingtoacccoandsymbol<-read.table('platefile_set_human_40k_a updated 24-7(mapbetweenoligoIDandacccoandsymbol).txt', header=TRUE, comment.char="",
fill=TRUE, as.is=TRUE)
position<-NA
morrisnormaldata$genes[,6]<-" morrisnormaldata$genes[,7]<-"
for (count in 1:18725) {
    position<-match(morrisnormaldata$genes[count,]$Name, oligoIDmappingtoaccnoandsymbol[,1]);
    if (is.na(position)==FALSE) {
        morrisnormaldata$genes[count,][6]<-
        oligoIDmappingtoaccnoandsymbol[position,2];
        if (oligoIDmappingtoaccnoandsymbol[position,3]!='') {
            morrisnormaldata$genes[count,][7]<-
            oligoIDmappingtoaccnoandsymbol[position,3];
        }
    }
}

fit<-lmFit(morrisnormaldata, morrisdesignwithageandsex, method='robust',
weights=morrisnormaldata$weights)
fit_morris_robust_ageandsex<-eBayes(fit)

# Genes DE by age in morрис dataset

> topTable(fit_morris_robust_ageandsex, coef=2, adjust.method='fdr', sort.by='p')

<table>
<thead>
<tr>
<th>Block</th>
<th>Column</th>
<th>Row</th>
<th>Name</th>
<th>ID</th>
<th>X</th>
<th>Y</th>
</tr>
</thead>
<tbody>
<tr>
<td>12138</td>
<td>16</td>
<td>34</td>
<td>obshum40K:A#03640 R087X31N20</td>
<td>AC099849</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3497</td>
<td>5</td>
<td>10</td>
<td>obshum40K:A#38527 R087X24L15</td>
<td>D26309</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5368</td>
<td>7</td>
<td>22</td>
<td>obshum40K:A#16129 R087X44H04</td>
<td>NM_152729</td>
<td>NT5C2L1</td>
<td></td>
</tr>
<tr>
<td>12005</td>
<td>16</td>
<td>23</td>
<td>obshum40K:A#02956 R087X23J08</td>
<td>NM_004861</td>
<td>GAL3ST1</td>
<td></td>
</tr>
<tr>
<td>13570</td>
<td>18</td>
<td>36</td>
<td>obshum40K:A#15573 R087X24M09</td>
<td>NM_005990</td>
<td>STK10</td>
<td></td>
</tr>
<tr>
<td>18584</td>
<td>24</td>
<td>18</td>
<td>obshum40K:A#15414 R087X44E12</td>
<td>AL591049</td>
<td>SMAP1</td>
<td></td>
</tr>
<tr>
<td>16043</td>
<td>21</td>
<td>12</td>
<td>obshum40K:A#05139 R087X33O11</td>
<td>AJ697971</td>
<td>DMN</td>
<td></td>
</tr>
<tr>
<td>11917</td>
<td>16</td>
<td>11</td>
<td>obshum40K:A#03540 R087X17J08</td>
<td>NM_006068</td>
<td>TLR6</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Autoflag</th>
<th>logFC</th>
<th>Ave Expr</th>
<th>t</th>
<th>P.Value</th>
<th>adj.P.Val</th>
</tr>
</thead>
<tbody>
<tr>
<td>12138</td>
<td>-0.007973793</td>
<td>9.103289</td>
<td>-6.266751</td>
<td>4.487831e-08</td>
<td>0.0008403464</td>
</tr>
<tr>
<td>3497</td>
<td>0.008644958</td>
<td>7.904247</td>
<td>5.447169</td>
<td>1.022341e-06</td>
<td>0.0054040497</td>
</tr>
<tr>
<td>5368</td>
<td>0.006477321</td>
<td>8.358934</td>
<td>5.498570</td>
<td>1.041983e-06</td>
<td>0.0054040497</td>
</tr>
<tr>
<td>15707</td>
<td>0.007603734</td>
<td>10.307823</td>
<td>5.435749</td>
<td>1.154403e-06</td>
<td>0.0054040497</td>
</tr>
<tr>
<td>12005</td>
<td>0.005390781</td>
<td>9.204738</td>
<td>5.102795</td>
<td>3.672385e-06</td>
<td>0.0011294149</td>
</tr>
<tr>
<td>13570</td>
<td>0.007734021</td>
<td>10.687705</td>
<td>5.073405</td>
<td>4.222110e-06</td>
<td>0.0011294149</td>
</tr>
<tr>
<td>18584</td>
<td>0.005920532</td>
<td>9.826926</td>
<td>4.938095</td>
<td>7.112319e-06</td>
<td>0.00133314569</td>
</tr>
<tr>
<td>16043</td>
<td>0.012258514</td>
<td>7.939211</td>
<td>4.902338</td>
<td>7.854412e-06</td>
<td>0.00133314569</td>
</tr>
<tr>
<td>11917</td>
<td>-0.008534027</td>
<td>7.175146</td>
<td>4.904815</td>
<td>8.261021e-06</td>
<td>0.00133314569</td>
</tr>
</tbody>
</table>

B
12138 6.773747
3497 3.684540
15707 3.722891
18150 3.559436
# Genes DE by gender in morris dataset

topTable(fit_morris_robust_ageandsex, coef=3, adjust.method='fdr', sort.by='p')
>
<table>
<thead>
<tr>
<th>Block</th>
<th>Column</th>
<th>Row</th>
<th>Name</th>
<th>ID</th>
<th>GenBankAccession</th>
<th>Symbol</th>
<th>EntrezID</th>
</tr>
</thead>
<tbody>
<tr>
<td>10545</td>
<td>14</td>
<td>7</td>
<td>obshum40K:A#00370</td>
<td>R087X27I19</td>
<td>NM_004653</td>
<td>KDM5D</td>
<td></td>
</tr>
<tr>
<td>3917</td>
<td>5</td>
<td>15</td>
<td>obshum40K:A#31044</td>
<td>R087X51D15</td>
<td>NM_032490</td>
<td>C14orf142</td>
<td></td>
</tr>
<tr>
<td>4756</td>
<td>2</td>
<td>1</td>
<td>obshum40K:A#33211</td>
<td>R087X03K16</td>
<td>NM_004555</td>
<td>NFATC3</td>
<td></td>
</tr>
<tr>
<td>1163</td>
<td>31</td>
<td>10</td>
<td>obshum40K:A#42285</td>
<td>R087X26A13</td>
<td>NM_001007</td>
<td>RPS4X</td>
<td></td>
</tr>
<tr>
<td>4100</td>
<td>6</td>
<td>16</td>
<td>obshum40K:A#32152</td>
<td>R087X12N03</td>
<td>NM_016509</td>
<td>CLEC1B</td>
<td></td>
</tr>
<tr>
<td>12878</td>
<td>17</td>
<td>40</td>
<td>obshum40K:A#01936</td>
<td>R087X29O21</td>
<td>AC073909</td>
<td>&lt;NA&gt;</td>
<td></td>
</tr>
<tr>
<td>14033</td>
<td>19</td>
<td>18</td>
<td>obshum40K:A#04213</td>
<td>R087X02G10</td>
<td>NM_002705</td>
<td>PPL</td>
<td></td>
</tr>
<tr>
<td>11108</td>
<td>15</td>
<td>41</td>
<td>obshum40K:A#07008</td>
<td>R087X14C08</td>
<td>NM_001935</td>
<td>DPP4</td>
<td></td>
</tr>
<tr>
<td>4088</td>
<td>6</td>
<td>5</td>
<td>obshum40K:A#41890</td>
<td>R087X11M15</td>
<td>AC114728</td>
<td>&lt;NA&gt;</td>
<td></td>
</tr>
<tr>
<td>5905</td>
<td>8</td>
<td>30</td>
<td>obshum40K:A#02247</td>
<td>R087X29E04</td>
<td>NM_005210</td>
<td>CRYGB</td>
<td></td>
</tr>
</tbody>
</table>

Autoflag logFC AveExpr t P.Value adj.P.Val B
| 10545 | 0.4963692 | 7.280682 | 9.868835 | 4.525816e-14 | 18.259027 |
| 3917  | 0.5545661 | 6.474389 | 6.367254 | 7.317344e-07 | 3.423946 |
| 4756  | 0.2773713 | 7.334251 | 5.297860 | 1.114353e-01 | 4.492733 |
| 1163  | -0.4079773 | 10.538454 | -5.137514 | 3.232005e-06 | 3.062178 |
| 4100  | -0.8186268 | 6.296784 | -6.081798 | 1.051316e-05 | 4.888737 |
| 12878 | -0.2539324 | 8.244406 | -4.566494 | 2.597469e-05 | 2.205352 |
| 14033 | 0.3979821 | 6.884447 | 4.546872 | 5.020707e-05 | 1.6798917 |
| 11108 | 0.3532902 | 7.206123 | 4.317571 | 6.08523e-05 | 1.511920 |
| 4088  | -0.3382081 | 11.419787 | -4.216866 | 8.542671e-05 | 1.245401 |
| 5905  | 0.3452389 | 6.697315 | 4.378050 | 8.843816e-05 | 0.9588518 |

# To prepare results in R in a format suitable for GSEA

```r
morris_unparsedgsealist<-topTable(fit_morris_robust_ageandsex, coef=2, number=18725, adjust.method='fdr', sort.by='p')[,c(6,59)]
morris_parsedgsealist_accessions<-vector()
morris_parsedgsealist_tstat<-vector()
count2<-1
for (count in 1:18725){
  if(morris_unparsedgsealist[count,1]!="") {
    morris_parsedgsealist_accessions[count2]<-
    morris_unparsedgsealist[count,1];
    morris_parsedgsealist_tstat[count2]<-morris_unparsedgsealist[count,2];
    count2<-count2+1;
  }
```
write.table(cbind(morris_parsedgsealist_accessions, morris_parsedgsealist_tstat),
file='morris_outputforGSEA_12-12-12.txt', quote=FALSE)
morris_sex_unparsedgsealist<-topTable(fit_morris_robust_ageandsex, coef=3,
number=18725, adjust.method='fdr', sort.by='p')[,c(6,59)]
morris_sex_parsedgsealist_accessions<-vector()
morris_sex_parsedgsealist_tstat<-vector()
count2<-1
for (count in 1:18725){
  if(morris_sex_unparsedgsealist[count,1]!='') {
    morris_sex_parsedgsealist_accessions[count2]<-
    morris_sex_unparsedgsealist[count,1];
    morris_sex_parsedgsealist_tstat[count2]<-
    morris_sex_unparsedgsealist[count,2];
    count2<-count2+1;
  }
}
write.table(cbind(morris_sex_parsedgsealist_accessions, morris_sex_parsedgsealist_tstat),
file='morris_sex_outputforGSEA_12-12-12.txt', quote=FALSE)
Appendix C: R code to Generate Results of Chapter 4

# Code to analyse the raw data of the datasets downloaded from GEO, note 'kittle'
# refers to GSE1869, and that CEL file data was not available, however, RMA
# normalised data was available and was used
kittledata<-read.table("C:\Users\PC\PhD Work\FailingDatasets\Kittleson\GSE1869_series_matrix.txt", header=TRUE, row.names=1,skip=54, comment.char='',nrow=22283)
kittledesign_donorvsidcm<-cbind(1, c(rep(1,times=19), rep(0,times=6), rep(1, times=2)))
kittledesign_donorvsiscm<-cbind(1, c(rep(1,times=3), rep(0,times=6), rep(1, times=7)))
fit_kittle_robust_donorvsidcm<-lmFit(kittledata[, c(1:6, 8, 10, 12:28, 30, 37)],
fit_kittle_robust_donorvsidcm<-eBayes(fit_kittle_robust_donorvsidcm)
fit_kittle_robust_donorvsiscm<-lmFit(kittledata[, c(7, 9, 11, 23:29, 31:36)],
kittledesign_donorvsidcm, method='robust')
fit_kittle_robust_donorvsiscm<-eBayes(fit_kittle_robust_donorvsidcm)
fit_kittle_robust_donorvsiscm<-eBayes(fit_kittle_robust_donorvsiscm)

# Genes DE in end-stage ISCM versus NF LV
> topTable(fit_kittle_robust_donorvsidcm,coef=2,adjust="fdr")

<table>
<thead>
<tr>
<th>ID</th>
<th>logFC</th>
<th>AveExpr</th>
<th>t</th>
<th>P.Value</th>
<th>adj.P.Val</th>
</tr>
</thead>
<tbody>
<tr>
<td>B</td>
<td>17.557800</td>
<td>0.8840175</td>
<td>8.986377</td>
<td>12.399066</td>
<td>8.616383e-13</td>
</tr>
<tr>
<td>1259</td>
<td>8.822386</td>
<td>7.952384</td>
<td>1097</td>
<td>0.068261</td>
<td>6.989015</td>
</tr>
<tr>
<td>1904</td>
<td>7.917624</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
# Genes DE in end-stage ISCM versus NF LV

```r
> topTable(fit_kittle_robust_donorvsiscm, coef=2, adjust="fdr")

ID logFC AveExpr t P.Value adj.P.Val
19153 219789_at 1.7538746 7.242307 14.16983 6.884091e-11 1.105273e-06
1445 201917_s_at 1.0245718 7.308838 13.84627 9.920328e-11 1.105273e-06
17677 218312_s_at 0.8518534 8.883968 12.05124 8.661532e-10 6.433497e-06
18815 219451_at 0.6267327 6.844433 11.48202 1.818765e-09 1.013188e-05
3083 203556_at -1.6313376 7.400374 -10.85552 4.257585e-09 1.514361e-05
12079 212694_s_at 0.8589279 8.882740 10.77576 4.757226e-09 1.514361e-05
8500 209006_s_at 0.9014571 5.874810 10.30646 9.257348e-09 2.355099e-05
2030 202502_at 0.8670620 9.408597 10.28765 9.512136e-09 2.355099e-05
12116 212731_at 1.1135040 6.843720 10.08618 1.275194e-08 2.584573e-05
```

# To prepare results in R in a format suitable for GSEA

```r
kittle_robust_donorvsidcm_gsealist<-topTable(fit_kittle_robust_donorvsidcm, coef=2, number=22283, adjust.method='fdr', sort.by='p')[,c(1,4)]
write.table(kittle_robust_donorvsidcm_gsealist, file='kittle_robust_donorvsidcm_forGSEA_4-6-13.txt', quote=FALSE)
kittle_robust_donorvsiscm_gsealist<-topTable(fit_kittle_robust_donorvsiscm, coef=2, number=22283, adjust.method='fdr', sort.by='p')[,c(1,4)]
write.table(kittle_robust_donorvsiscm_gsealist, file='kittle_robust_donorvsiscm_forGSEA_4-6-13.txt', quote=FALSE)
```

# Code used to analyse the GSE5406 dataset, referred to as 'hannen' in the R code.
# RMA normalised data was available on GEO, but not the original CEL files.
```r
orderofhannen<-read.table('testfilewithorderofhannen.csv', header=TRUE, sep='/t')
idcmcolumn<-vector(length=210, mode='integer')
ismcolumn<-vector(length=210, mode='integer')
```
normalcolumn<-vector(length=210, mode='integer')
for (count in 1:210) {
    if (orderofhannen[count]=='systolic heart failure due to idiopathic dilated cardiomyopathy') { idcmcolumn[count]<-1 }
    if (orderofhannen[count]=='systolic heart failure due to ischemic cardiomyopathy') { iscmcolumn[count]<-1 }
    if (orderofhannen[count]=='normally functioning myocardium from unused donor heart') { normalcolumn[count]<-1 }
}
hannendesign <- cbind(idcm=idcmcolumn, iscm=iscmcolumn, donor=normalcolumn)
hannendata<-read.table('GSE5406_series_matrix_onlydata.csv', header=TRUE, sep='t', row.names=1)
fit_hannen_robust<-lmFit(hannendata, hannendesign, method='robust')
hannen_cont.matrix_donorvsidcm<-makeContrasts(donorvsidcm=idcm-donor,levels=hannendesign)
fit_hannen_robust_donorvsidcm<-contrasts.fit(fit_hannen_robust,hannen_cont.matrix_donorvsidcm)
fit_hannen_robust_donorvsidcm <-eBayes(fit_hannen_robust_donorvsidcm)
topTable(fit_hannen_robust_donorvsidcm,adjust='fdr')

# Genes DE in end-stage IDCM versus NF LV
> topTable(fit_hannen_robust_donorvsidcm,adjust='fdr')

    ID    logFC AveExpr t P.Value adj.P.Val
1  7047  207526_s_at -1.8091431  4.573276 -19.36018 1.187360e-48 2.645795e-44 95.49835
2  21088 221728_x_at -1.5885808  5.653916 -17.79978 7.232666e-44 8.058275e-40 86.69573
3  1437  201909_at  2.5880562  9.010048  17.47978 7.114530e-43 5.284436e-39 84.58230
4  19092   201744_s_at  2.2434212  8.319699  15.28665 5.434264e-36 2.421834e-32 70.19260
5  8164  208668_x_at  0.6765385 10.201308  15.16812 1.287892e-35 4.783018e-32 69.40594
6  1904  202376_at -1.8091431  4.573276 -19.36018 1.187360e-48 2.645795e-44 95.49835
7 1437   201744_s_at  2.2434212  8.319699  15.28665 5.434264e-36 2.421834e-32 70.19260
8 1272  201909_at -1.5885808  5.653916 -17.79978 7.232666e-44 8.058275e-40 86.69573
9 21312 221952_x_at  0.6948184  9.372044  14.08513 3.452314e-32 9.615988e-29 61.92468
10 18451  219087_at -0.8187126  7.713927 -12.55883 2.264182e-27 5.045353e-24 51.55946
11 17485 218120_s_at -0.8187126  7.713927 -12.55883 2.264182e-27 5.045353e-24 51.20195

hannen_cont.matrix_donorvsiscm<-makeContrasts(donorvsiscm=iscm-donor,levels=hannendesign)
fit_hannen_robust_donorvsiscm<-
contrasts.fit(fit_hannen_robust_donorvsiscm)
fit_hannen_robust_donorvsiscm<-eBayes(fit_hannen_robust_donorvsiscm)
topTable(fit_hannen_robust_donorvsiscm,adjust="fdr")

# Genes DE in end-stage ISCM versus NF LV
> topTable(fit_hannen_robust_donorvsiscm,adjust="fdr")

<table>
<thead>
<tr>
<th>ID</th>
<th>logFC</th>
<th>AveExpr</th>
<th>t</th>
<th>P.Value</th>
<th>adj.P.Val</th>
</tr>
</thead>
<tbody>
<tr>
<td>7047</td>
<td>-1.8482103</td>
<td>4.573276</td>
<td>-19.90828</td>
<td>2.622899e-50</td>
<td>5.844607e-46</td>
</tr>
<tr>
<td>21088</td>
<td>-1.6889882</td>
<td>5.653916</td>
<td>-19.33197</td>
<td>1.445975e-48</td>
<td>1.611033e-44</td>
</tr>
<tr>
<td>1437</td>
<td>2.7109691</td>
<td>9.010048</td>
<td>18.68814</td>
<td>1.327551e-46</td>
<td>9.860607e-43</td>
</tr>
<tr>
<td>1272</td>
<td>2.2228390</td>
<td>8.319699</td>
<td>15.34572</td>
<td>1.575582e-32</td>
<td>70.51914</td>
</tr>
<tr>
<td>8164</td>
<td>0.6489695</td>
<td>10.201308</td>
<td>14.75237</td>
<td>2.663076e-31</td>
<td>9.890222e-31</td>
</tr>
<tr>
<td>5393</td>
<td>-1.3053055</td>
<td>6.948432</td>
<td>-12.56917</td>
<td>2.101093e-27</td>
<td>54.854837</td>
</tr>
</tbody>
</table>

# Code to prepare data in R in a format suitable for GSEA
hannen_robust_donorvsidcm_gsealist<-topTable(fit_hannen_robust_donorvsidcm,
number=22283, adjust.method='fdr', sort.by='p')[,c(1,4)]
hannen_robust_donorvsidcm_gsealist[1:10,]
hannen_robust_donorvsiscm_gsealist<-topTable(fit_hannen_robust_donorvsiscm,
number=22283, adjust.method='fdr', sort.by='p')[,c(1,4)]
hannen_robust_donorvsiscm_gsealist[1:10,]
write.table(hannen_robust_donorvsidcm_gsealist,
file='hannen_robust_donorvsidcm_forGSEA_4-6-13.txt', quote=FALSE)
write.table(hannen_robust_donorvsiscm_gsealist,
file='hannen_robust_donorvsiscm_forGSEA_4-6-13.txt', quote=FALSE)

# Code to analyse GSE1145, referred to as 'tsubaki' in the code, CEL file data used
setwd('C:/Users/PC/PhD Work/FailingDatasets/tsubaki/tsubakirawdata')
tsubakidata_prerma<--ReadAffy()
tsubakidata_postrma<-rma(tsubakidata_prerma)
> sampleNames(tsubakidata_postrma)
[1] "GSM18422_PA-D_102.cel.gz" "GSM18423_PA-D_132.cel.gz" "GSM18424_PA-
D_206.cel.gz"
[4] "GSM18425_PA-D_216.cel.gz" "GSM18426_PA-D_225.cel.gz" "GSM18427_PA-
D_229.cel.gz"
[7] "GSM18428_PA-D_317.cel.gz" "GSM18429_PA-D_319.cel.gz" "GSM18430_PA-
D_348.cel.gz"
[10] "GSM18431_PA-D_374.cel.gz" "GSM18432_PA-D_64.cel.gz" "GSM18433_PA-
D_82.cel.gz"
[13] "GSM18434_PA-D_85.cel.gz" "GSM18435_PA-D_93_2.cel.gz" "GSM18436_PA-
D_98.cel.gz"
[16] "GSM18442_PA-N_112-1.cel.gz" "GSM18443_PA-N_118.cel.gz" "GSM18444_PA-
N_148.cel.gz"
[19] "GSM18445_PA-N_200.cel.gz" "GSM18446_PA-N_249.cel.gz" "GSM18447_PA-
N_291.cel.gz"
[22] "GSM18448_PA-N_294.cel.gz" "GSM18449_PA-N_300.cel.gz" "GSM18450_PA-
N_322.cel.gz"
[25] "GSM18451_PA-N_325.cel.gz" "GSM18452_PA-N_326.cel.gz" "GSM18457_PA-
S_115.cel.gz"
[28] "GSM18478_PA-S_224.cel.gz" "GSM18479_PA-S_233.cel.gz" "GSM18480_PA-
S_242_2.cel.gz"
[31] "GSM18481_PA-S_267.cel.gz" "GSM18482_PA-S_281.cel.gz" "GSM18483_PA-
S_297.cel.gz"
[34] "GSM18484_PA-S_314.cel.gz" "GSM18485_PA-S_342.cel.gz" "GSM18486_PA-
S_386.cel.gz"
[37] "GSM18487_PA-S_464.cel.gz"

tsubakidesign_donorvsidcm<-cbind(1, c(rep(1, times=15), rep(0,times=11)))
fit_tsubaki_robust_donorvsidcm<-lmFit(tsubakidata_postrma[,c(1:26)],
tsubakidesign_donorvsidcm, method='robust')
fit_tsubaki_robust_donorvsidcm<-eBayes(fit_tsubaki_robust_donorvsidcm)

# Genes DE in end-stage IDCM versus NF LV
> topTable(fit_tsubaki_robust_donorvsidcm, coef=2)

<table>
<thead>
<tr>
<th>ID</th>
<th>logFC</th>
<th>AveExpr</th>
<th>t</th>
<th>P.Value</th>
<th>adj.P.Val</th>
<th>B</th>
</tr>
</thead>
<tbody>
<tr>
<td>45773</td>
<td>236523</td>
<td>-4.6227193</td>
<td>7.567492</td>
<td>-27.82092</td>
<td>1.288230e-21</td>
<td>7.043396e-17</td>
</tr>
<tr>
<td>39113</td>
<td>229858</td>
<td>-1.4938703</td>
<td>8.011914</td>
<td>-25.09366</td>
<td>1.959383e-20</td>
<td>5.35643e-16</td>
</tr>
<tr>
<td>23030</td>
<td>213729</td>
<td>-0.9763122</td>
<td>9.862392</td>
<td>-22.13793</td>
<td>5.183982e-19</td>
<td>7.085955e-15</td>
</tr>
<tr>
<td>17099</td>
<td>207657</td>
<td>-1.1435069</td>
<td>9.750096</td>
<td>-22.31793</td>
<td>1.288230e-21</td>
<td>1.337970e-14</td>
</tr>
<tr>
<td>48769</td>
<td>239519</td>
<td>1.2599927</td>
<td>6.694475</td>
<td>21.41696</td>
<td>1.223566e-18</td>
<td>1.337970e-14</td>
</tr>
<tr>
<td>36.93707</td>
<td>35.76361</td>
<td>32.92836</td>
<td>32.02839</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
```r
tsubakidesign_donorvsiscm <- cbind(1, c(rep(0, times=11), rep(1,times=11)))
fitted_tsubaki_robust_donorvsiscm <- lmFit(tsubakidata_postrma[,c(16:37)],
                                          tsubakidesign_donorvsiscm, method='robust')
fitted_tsubaki_robust_donorvsiscm <- eBayes(fitted_tsubaki_robust_donorvsiscm)

# Genes DE in end-stage ISCM versus NF LV
> topTable(fitted_tsubaki_robust_donorvsiscm, coef=2)
   ID      logFC   AveExpr         t      P.Value     adj.P.Val     B
21410   212103_at  0.7839567  8.949356  19.92283 5.482843e-16 2.997744e-11
25.65043
17626   208200_at -0.5864127  6.097769 -18.78325 1.966450e-15 5.125231e-11
24.47230
49061   239811_at  0.8839336  7.458833  18.47460 2.812198e-15 5.125231e-11
24.44223
41074   231819_at  1.5305821  7.467291  18.14914 4.124386e-15 5.637520e-11
24.01004
23030   213729_at -0.6171675 10.111086 -17.95400 5.203967e-15 5.690538e-11
23.63777
19667 210266_s_at  0.9479386  9.436008  16.73501 2.341908e-14 1.829197e-10
22.43317
28687 219402_s_at -0.8458126  9.426742 -16.81418 2.117908e-14 1.829197e-10
22.40560
13972   204524_at  0.8462012  8.018265  16.39469 3.624377e-14 2.301891e-10
22.16183
35669   226412_at  0.7967187  8.389042  16.36038 3.789122e-14 2.301891e-10
21.98914
44972   235722_at  2.0693767  8.012745  16.12975 5.119889e-14 2.544818e-10
21.78297

# Code to prepare data in R in a format suitable for GSEA
fitted_tsubaki_robust_donorvsidcm_gsealist <- topTable(fitted_tsubaki_robust_donorvsidcm,
                                                        coef=2, number=54675, sort.by='p')[,c(1,4)]
fitted_tsubaki_robust_donorvsiscm_gsealist <- topTable(fitted_tsubaki_robust_donorvsiscm,
                                                        coef=2, number=54675, sort.by='p')[,c(1,4)]
```

174
```
write.table(tsubaki_robust_donorvsidcm_gsealist,
file='tsubaki_robust_donorvsidcm_forGSEA_4-6-13.txt', quote=FALSE)
write.table(tsubaki_robust_donorvsismc_gsealist,
file='tsubaki_robust_donorvsismc_forGSEA_4-6-13.txt', quote=FALSE)

# Code to analyse GSE430, referred to as 'chen' in the code, only RMA normalised
# data available
chendata<-read.table('GSE430_series_matrix_parsed.txt', header=TRUE, row.names=1,
skip=0, comment.char='",nrow=22283)
targets<-data.frame(FileName=c('GSM6540_HF1', 'GSM6541_LVAD1',
'GSM6542_HF2', 'GSM6543_LVAD2', 'GSM6545_HF4', 'GSM6546_LVAD4',
'GSM6547_HF5', 'GSM6548_LVAD5', 'GSM6549_HF6', 'GSM6550_LVAD6',
'GSM6551_HF3', 'GSM6552_LVAD3', 'GSM6553_HF7', 'GSM6554_LVAD7'),
Subjects=c(1,1,2,2,3,3,4,4,5,5,6,6,7,7), Time=c(0,1,0,1,0,1,0,1,0,1,0,1,0,1))
chensubs<-factor(chentargets$Subjects)
chentiming<-factor(chentargets$Time, levels=c(0,1))
chendesign<-model.matrix(~chensubs+chentiming)
fit_chen_robust_beforevsafterLVAD <- lmFit(chendata, chendesign, method='robust')
fit_chen_robust_beforevsafterLVAD <- eBayes(fit_chen_robust_beforevsafterLVAD)
topTable(fit_chen_robust_beforevsafterLVAD, coef='chentiming1')

# Genes DE by LVAD implantation in IDCM
> topTable(fit_chen_robust_beforevsafterLVAD, coef='chentiming1')

   ID     logFC AveExpr    t    P.Value  adj.P.Val
3070  203543_s_at  307.29776  3.78229 1.330203e-09 0.0056422
11835  212450_at  308.98845  4.17570 1.203307e-09 0.0056422
2415  202887_s_at  248.75859  3.18614 1.863755e-03 0.0056422
8286  205289_at   349.21663  4.14940 1.347075e-09 0.0056422
4816   210121_at     5.3198  5.25975 8.233090e-06 0.0056422
9605  210121_at  -100.47395  94.77143 -13.20673 4.29693e-06 0.0056422
14408 215034_s_at  -140.17641 194.27143  -9.94056 2.71750e-05 0.0248616
7034  207513_s_at  255.79866 272.20000  11.43013 1.10403e-05 0.0153757
3235  203708_at  -198.41954 154.40714 -10.52895 1.87813e-05 0.0190229

B
3070  -0.4475709
11835  -0.5448950
2415  -0.5691934
8286  -0.6391338
4816  -0.6394862
9605  -0.6701487
14408  -0.7038212
11338  -0.8048262
7034  -0.8861980
```
# Code to prepare R results in a format suitable for GSEA
chenaffy_robust_beforevsafterLVAD_gsealist<-topTable(fit_chen_robust_beforevsafterLVAD_AD, coef='chentiming1', number=22283, adjust.method='fdr', sort.by='p')[,c(1,7)]
write.table(chenaffy_robust_beforevsafterLVAD_gsealist, file='chenaffy_robust_beforevsafterLVAD_forGSEA_12-12-12.txt', quote=FALSE)

# Code to analyse GSE974, referred to as 'hall' in code, note CEL file data was available
library(affy)
halldata <- ReadAffy()

halltargets_ISCM<-data.frame(FileName=c("GSM14844.CEL", "GSM14936.CEL", "GSM14937.CEL", "GSM14938.CEL", "GSM14939.CEL", "GSM14940.CEL", "GSM14941.CEL", "GSM14942.CEL", "GSM14943.CEL", "GSM14944.CEL"), Subjects=c(1,1,2,3,3,4,4,4,5,5), Time=c(0,1,1,0,1,0,1,0,1,0))
halltargets_MI<-data.frame(FileName=c("GSM14945.CEL", "GSM14946.CEL", "GSM14947.CEL", "GSM14948.CEL", "GSM14949.CEL", "GSM14950.CEL", "GSM14951.CEL", "GSM14952.CEL", "GSM14953.CEL", "GSM14954.CEL", "GSM14955.CEL", "GSM14956.CEL"), Subjects=c(1,1,2,2,3,3,4,4,5,5,6,6), Time=c(1,0,1,0,1,0,1,0,1,0,1,0))
halltargets_nonISCM<-data.frame(FileName=c("GSM14957.CEL", "GSM14958.CEL", "GSM14959.CEL", "GSM14960.CEL", "GSM14961.CEL", "GSM14962.CEL", "GSM14963.CEL", "GSM14965.CEL", "GSM14966.CEL", "GSM14967.CEL", "GSM14968.CEL", "GSM14969.CEL", "GSM14970.CEL", "GSM14971.CEL", "GSM14972.CEL", "GSM14973.CEL"), Subjects=c(1,1,2,2,3,3,4,4,5,5,6,6,7,7,8,8), Time=c(1,0,1,0,1,0,1,0,1,0,1,0,1,0,1,0))
hallsubs_ISCM<-factor(halltargets_ISCM$Subjects)
hallsubs_MI<-factor(halltargets_MI$Subjects)
hallsubs_nonISCM<-factor(halltargets_nonISCM$Subjects)
halltiming_ISCM<-factor(halltargets_ISCM$Time, levels=c(0,1))
halltiming_MI<-factor(halltargets_MI$Time, levels=c(0,1))
halltiming_nonISCM<-factor(halltargets_nonISCM$Time, levels=c(0,1))
halldesign_ISCM<-model.matrix(~hallsubs_ISCM+halltiming_ISCM)
halldesign_MI<-model.matrix(~hallsubs_MI+halltiming_MI)
halldesign_nonISCM<-model.matrix(~hallsubs_nonISCM+halltiming_nonISCM)
halldata<-rma(halldata)
fit_hall_robust_ISCM <- lmFit(halldata[,1:10], halldesign_ISCM, method='robust')
fit_hall_robust_ISCM <- eBayes(fit_hall_robust_ISCM )
topTable(fit_hall_robust_ISCM , coef='halltiming_ISCM1')

# Genes DE by LVAD implantation in end-stage ISCM
> topTable(fit_hall_robust_ISCM, coef="halltiming_ISCM1")
  ID  KEGG EntrezID Symbol      logFC   AveExpr         t
12050  212665_at <NA>    TIPARP  1.4883617  8.667899  22.65422
9443   209957_s_at <NA>   NPPA -1.4370534 10.298998 -23.89719
3069   203542_s_at <NA>   KLF9  1.3694831  8.204960  21.31384
16569  217202_s_at <NA>   GLUL  0.8320695  7.982406  23.10767
12109  212724_at <NA>   RND3  0.7781892  9.192215  22.91680
9443   209957_s_at <NA>   NPPA -1.4370534 10.298998 -23.89719
3069   203542_s_at <NA>   KLF9  1.3694831  8.204960  21.31384
16569  217202_s_at <NA>   GLUL  0.8320695  7.982406  23.10767
12109  212724_at <NA>   RND3  0.7781892  9.192215  22.91680
3069   203542_s_at <NA>   KLF9  1.3694831  8.204960  21.31384
16569  217202_s_at <NA>   GLUL  0.8320695  7.982406  23.10767
12109  212724_at <NA>   RND3  0.7781892  9.192215  22.91680

P.Value   adj.P.Val        B
12050  7.384297e-07 0.004113607 6.164353
9443   5.441535e-07 0.004113607 6.106631
3069   1.046001e-06 0.004661610 5.931717
16569  6.593726e-07 0.004113607 5.852854
12109  6.913764e-07 0.004113607 5.849611
3659   1.675202e-06 0.006221422 5.433321
17106  2.039419e-06 0.006322548 5.195287
8848   2.955846e-06 0.006322548 5.182567
4817   2.697809e-06 0.006322548 5.162461
10768  3.121125e-06 0.006322548 5.162461

> fit_hall_robust_nonISCM <- lmFit(halldata[,23:38], halldesign_nonISCM, method="robust")
> fit_hall_robust_nonISCM <- eBayes(fit_hall_robust_nonISCM)

# Genes DE by LVAD implantation in end-stage IDC
> topTable(fit_hall_robust_nonISCM, coef="halltiming_nonISCM1")
  ID  KEGG EntrezID Symbol      logFC   AveExpr         t
18418  219054_at <NA>     NPR3  -1.0279711  8.104466 -20.38712
4595   205068_s_at <NA>    ARHGAP26  0.6043282  7.196266  15.84799
3233   203706_s_at <NA>     FZD7  -0.7180439  7.664089 -13.89511
7961    208463_at <NA>    GABRA4  -0.7448972  7.062312 -12.34901
20673  221310_at <NA>     FGF14  -0.4141530  6.113029 -11.96232
18577  219213_at <NA>     JAM2  -0.5285529  8.098972 -11.78985
9755   210275_s_at <NA>    ZFAND5  0.7831544 10.581024  11.45944
6804   207279_s_at <NA>    NEBL  -0.3836398  8.755252 -11.44953
8286  208791_at <NA>     CLU  0.7965801  8.636077  11.15729
3658   204131_s_at <NA>    FOXO3  0.9941460  9.585396  10.42132

P.Value   adj.P.Val        B
18418  4.065856e-09 9.059947e-05 9.507821
4595  4.107842e-08 4.576752e-04 8.471948
3233  1.354918e-07 1.006388e-03 7.417613
7961   3.908068e-07 2.132351e-03 6.829055
fit_hall_robust_MI <- lmFit(halldata[,11:22], halldesign_MI, method='robust')
fithall_robust_MI <- eBayes(fit_hall_robust_MI)
topTable(fit_hall_robust_MI, coef='halltiming_MI1')

# Genes DE by LVAD implantation in within 10 days of MI causing HF
> topTable(fit_hall_robust_MI, coef='halltiming_MI1')

ID    KEGG EntrezID Symbol   logFC AveExpr          t    P.Value  adj.P.Val
B
4918  205391_x_at <NA>      286    ANK1  1.0915109 8.064126  13.156658 3.256923e-06 0.04159778 4.405422
20934 221571_at <NA>       7187   TRAF3 -0.7160073 6.785864 -11.727364 7.092089e-06 0.04159778 3.968361
5217  205690_s_at <NA>     8896   BUD31 -0.7103881 7.052923 -12.219485 5.374276e-06 0.04159778 3.836787
5487  205960_at <NA>      5166    PDK4  1.7644725 8.764631  11.581702 7.714623e-06 0.04159778 3.742569
12905 213524_s_at <NA>   50486    G0S2 -1.9676300 8.534494 -10.748997 1.272289e-05 0.04159778 3.592579
12354 212971_at <NA>       833    CARS -0.7570041 8.958328 -10.947517 1.125747e-05 0.04159778 3.564874
17616 218251_at <NA>    58526 MID1IP1 -0.8374364 9.071080 -10.617586 1.381187e-05 0.04159778 3.458033
14868 215495_s_at <NA>   23034   SAMD4A  0.6174235 7.287960  10.493898 1.493435e-05 0.04159778 3.375100
11686 212300_at <NA>    200081   TXLNA -0.6311937 6.154018  -9.920328 2.169489e-05 0.05081869 2.976260
533   201005_at <NA>      928    CD9 -0.5667625 9.065714  -9.845751 2.280604e-05 0.05081869 2.875080

# Code to prepare R results in a format suitable for GSEA
hall_robust_ISCM_gsealist<-topTable(fit_hall_robust_ISCM, coef='halltiming_ISCM1', number=22283, adjust.method='fdr', sort.by='p')[,c(1,7)]
write.table(hall_robust_ISCM_gsealist, file='hall_robust_ISCM_forGSEA_13-12-12.txt', quote=FALSE)
hall_robust_nonISCM_gsealist<-topTable(fit_hall_robust_nonISCM, coef='halltiming_nonISCM1', number=22283, adjust.method='fdr', sort.by='p')[,c(1,7)]
write.table(hall_robust_nonISCM_gsealist, file='hall_robust_nonISCM_forGSEA_13-12-12.txt', quote=FALSE)
hall_robust_MI_gsealist<-topTable(fit_hall_robust_MI, coef='halltiming_MI1',
number=22283, adjust.method='fdr', sort.by='p')[,c(1,7)]
write.table(hall_robust_MI_gsealist, file='hall_robust_MI_forGSEA_13-12-12.txt',
quote=FALSE)

# Code to analyse GSE3585, referred to as 'barth', raw CEL data was available
setwd('C:/Users/Owner/PhD Work/FailingDatasets/barthaffy')
barthdata <- ReadAffy()
barthdata<-rma(barthdata)
barthdesign <- cbind(rep(1,times=12), c(rep(0,times=5), rep(1,times=7)))
fit_barth_robust <- lmFit(barthdata, barthdesign, method='robust')
fit_barth_robust <- eBayes(fit_barth_robust)
topTable(fit_barth_robust, coef=2)

# Genes DE in end-stage IDCm versus NF LV
> topTable(fit_barth_robust, coef=2)
  ID      logFC   AveExpr          t      P.Value    adj.P.Val
6327    206801_at  4.9991362 12.039023  20.241215 6.150522e-11 1.370521e-06
3993  204466_s_at  1.2267914  9.058065  15.387211 1.677911e-09 1.869444e-05
9530    210045_at -0.8351628  9.788719 -11.675736 4.294232e-08 3.189612e-04
17362   217997_at  1.8529902  8.266368  11.082782 7.813620e-08 4.352772e-04
1033  201505_at  0.7461697  8.298119  10.849146 9.964839e-08 4.440930e-04
11843   212458_at  0.5752567  7.917872  10.411846 5.807346e-07 1.589548e-04
12291  213509_x_at -0.5644735  9.783237  -9.612384 3.894416e-07 8.677928e-04

# Code to prepare R results in a format suitable for GSEA
barth_robust_gsealist<-topTable(fit_barth_robust, coef=2, number=22283,
sort.by='p')[,c(1,7)]
write.table(barth_robust_gsealist, file='barth_robust_forGSEA_4-6-13.txt',
quote=FALSE)
# Code to analyse GSE3586, referred to as 'barth2colour', raw data available in the # form of GPR files
setwd('C:/Users/PC/PhD Work/FailingDatasets/barth2colour')
library(limma)
barth2colourtargs <- readTargets()
RG <- read.maimages(barth2colourtargs,source="genepix.median",
columns=list(G='F635 Median', R='F532 Median', Gb='B635 Median', Rb='B532
Median'), wt.fun=wtflags(weight=0,cutoff=-40)
)
RG$genes <- readGAL('GSM82393_1.gpr')
RG$printer <- getLayout(RG$genes)
RG <- backgroundCorrect(RG, method="normexp", offset=50)
MA <- normalizeWithinArrays(RG)
remove(RG)
barth2colourdata <- normalizeBetweenArrays(MA, method='scale')
remove(MA)

# To annotate probes with Genbank accession numbers
barth2colourannots<-read.table(file='arrayplatefile_modifiedforR.gal.csv',
header=TRUE, sep='"t"
for (count in 1:38976) {
    position<-match(barth2colourdata$genes[count,4], barth2colourannots[,1]);
    print(count)
    if (is.na(position)==FALSE) {
        barth2colourdata$genes[count,5]<-
        as.vector(barth2colourannots[position,2]);
        barth2colourdata$genes[count,6]<-
        as.vector(barth2colourannots[position,3]);
        barth2colourdata$genes[count,7]<-
        as.vector(barth2colourannots[position,4]);
        barth2colourdata$genes[count,8]<-
        as.vector(barth2colourannots[position,5]);
        barth2colourdata$genes[count,9]<-
        as.vector(barth2colourannots[position,6]);
    }
}

# To take the mean of the two replicate arrays hybridised with labelled mRNA from the
# same LV
# sample
barth2colourmeandata_M<-barth2colourdata$M
original<-barth2colourdata$M
barth2colourmeandata_M[,c(1:28)]<-NA
for (count in 1:38976) {
    barth2colourmeandata_M[count,1]<-mean(original[count,1:2]);
    barth2colourmeandata_M[count,2]<-mean(original[count,3:4]);
    barth2colourmeandata_M[count,3]<-mean(original[count,5:6]);
barth2colourmeandata_M[count,4] <- mean(original[count,7:8]);
barth2colourmeandata_M[count,5] <- mean(original[count,9:10]);
barth2colourmeandata_M[count,6] <- mean(original[count,11:12]);
barth2colourmeandata_M[count,7] <- mean(original[count,13:14]);
barth2colourmeandata_M[count,8] <- mean(original[count,15:16]);
barth2colourmeandata_M[count,9] <- mean(original[count,17:18]);
barth2colourmeandata_M[count,10] <- mean(original[count,19:20]);
barth2colourmeandata_M[count,11] <- mean(original[count,21:22]);
barth2colourmeandata_M[count,12] <- mean(original[count,23:24]);
barth2colourmeandata_M[count,13] <- mean(original[count,25:26]);
barth2colourmeandata_M[count,14] <- mean(original[count,27:28]);
barth2colourmeandata_M[count,15] <- mean(original[count,29:30]);
barth2colourmeandata_M[count,16] <- mean(original[count,31:32]);
barth2colourmeandata_M[count,17] <- mean(original[count,33:34]);
barth2colourmeandata_M[count,18] <- mean(original[count,35:36]);
barth2colourmeandata_M[count,19] <- mean(original[count,37:38]);
barth2colourmeandata_M[count,20] <- mean(original[count,39:40]);
barth2colourmeandata_M[count,21] <- mean(original[count,41:42]);
barth2colourmeandata_M[count,22] <- mean(original[count,43:44]);
barth2colourmeandata_M[count,23] <- mean(original[count,45:46]);
barth2colourmeandata_M[count,24] <- mean(original[count,47:48]);
barth2colourmeandata_M[count,25] <- mean(original[count,49:50]);
barth2colourmeandata_M[count,26] <- mean(original[count,51:52]);
barth2colourmeandata_M[count,27] <- mean(original[count,53:54]);
barth2colourmeandata_M[count,28] <- mean(original[count,55:56]);
print(count);
}

barth2colourmeandata_M <- barth2colourmeandata_M[,1:28]
barth2colourdata$M <- barth2colourmeandata_M
barth2colourmeandata_A <- barth2colourdata$A
original <- barth2colourdata$A
barth2colourmeandata_A[,c(1:28)] <- NA
for (count in 1:38976) {
    barth2colourmeandata_A[count,1] <- mean(original[count,1:2]);
    barth2colourmeandata_A[count,2] <- mean(original[count,3:4]);
    barth2colourmeandata_A[count,3] <- mean(original[count,5:6]);
    barth2colourmeandata_A[count,4] <- mean(original[count,7:8]);
    barth2colourmeandata_A[count,5] <- mean(original[count,9:10]);
    barth2colourmeandata_A[count,6] <- mean(original[count,11:12]);
    barth2colourmeandata_A[count,7] <- mean(original[count,13:14]);
    barth2colourmeandata_A[count,8] <- mean(original[count,15:16]);
    barth2colourmeandata_A[count,9] <- mean(original[count,17:18]);
    barth2colourmeandata_A[count,10] <- mean(original[count,19:20]);
    barth2colourmeandata_A[count,11] <- mean(original[count,21:22]);
    barth2colourmeandata_A[count,12] <- mean(original[count,23:24]);
    barth2colourmeandata_A[count,13] <- mean(original[count,25:26]);
barth2colourmeandata_A[count,14]<-mean(original[count,27:28]);
barth2colourmeandata_A[count,15]<-mean(original[count,29:30]);
barth2colourmeandata_A[count,16]<-mean(original[count,31:32]);
barth2colourmeandata_A[count,17]<-mean(original[count,33:34]);
barth2colourmeandata_A[count,18]<-mean(original[count,35:36]);
barth2colourmeandata_A[count,19]<-mean(original[count,37:38]);
barth2colourmeandata_A[count,20]<-mean(original[count,39:40]);
barth2colourmeandata_A[count,21]<-mean(original[count,41:42]);
barth2colourmeandata_A[count,22]<-mean(original[count,43:44]);
barth2colourmeandata_A[count,23]<-mean(original[count,45:46]);
barth2colourmeandata_A[count,24]<-mean(original[count,47:48]);
barth2colourmeandata_A[count,25]<-mean(original[count,49:50]);
barth2colourmeandata_A[count,26]<-mean(original[count,51:52]);
barth2colourmeandata_A[count,27]<-mean(original[count,53:54]);
barth2colourmeandata_A[count,28]<-mean(original[count,55:56]);
print(count);
}
barth2colourmeandata_A<-barth2colourmeandata_A[,1:28]
barth2colourdataA<-barth2colourmeandata_A
barth2colordesign<-cbind(1, c(rep(0,times=15), rep(1,times=13)))
fit_barth2colour_robust_idcmvsdonor<-lmFit(barth2colourdata,
design=barth2colordesign, weights=1, method='robust')
fit_barth2colour_robust_idcmvsdonor<-eBayes(fit_barth2colour_robust_idcmvsdonor)

# Genes DE in end-stage IDCM vs NF LV
> topTable(fit_barth2colour_robust_idcmvsdonor, coef=2)

<table>
<thead>
<tr>
<th>Block</th>
<th>Column</th>
<th>Row</th>
<th>ID</th>
<th>Name</th>
</tr>
</thead>
<tbody>
<tr>
<td>30196</td>
<td>38</td>
<td>7</td>
<td>6</td>
<td>IMAGp998L10643, N78562, N73714, BX119658 PDHX</td>
</tr>
<tr>
<td>9972</td>
<td>13</td>
<td>25</td>
<td>8</td>
<td>IMAGp998P10840, AA024968, AA024662, BX094547 URB, ANKH</td>
</tr>
<tr>
<td>3168</td>
<td>7</td>
<td>26</td>
<td>RZPDp202C071D BC009242, BX283465, BE299007 DMD</td>
<td></td>
</tr>
<tr>
<td>8155</td>
<td>11</td>
<td>6</td>
<td>2</td>
<td>IMAGp998H04207, R67351, R66519, BX090461 CPT1B</td>
</tr>
<tr>
<td>2650</td>
<td>4</td>
<td>11</td>
<td>8</td>
<td>IMAGp998I23929, BX096554, AA732982</td>
</tr>
<tr>
<td>26683</td>
<td>33</td>
<td>3</td>
<td>25</td>
<td>RZPDp201D0527D, BX282272, BI861738 CDC42EP3</td>
</tr>
<tr>
<td>4884</td>
<td>7</td>
<td>12</td>
<td>1</td>
<td>IMAGp998A08121, R02205, R02093, BX093967 DLAT</td>
</tr>
<tr>
<td>10372</td>
<td>13</td>
<td>19</td>
<td>22</td>
<td>IMAGp998J225591, BX089301, AI620570</td>
</tr>
<tr>
<td>26697</td>
<td>33</td>
<td>17</td>
<td>25</td>
<td>RZPDp201F0519D, BG674811, BX284055 SCP2</td>
</tr>
<tr>
<td>34648</td>
<td>43</td>
<td>22</td>
<td>19</td>
<td>IMAGp998B2173, R39190, T75308, BX111273 ZMYND11</td>
</tr>
</tbody>
</table>

# Code to prepare R results for GSEA, note that given that multiple probes were present on the microarray that could hybridise to the same gene sequence, the mean of all the t-statistics from said probes generated from the linear model applied, was used for GSEA for each particular gene
temp<-fit_barth2colour_robust_idcmvsdonor
tempdataframe<-data.frame()
tempdataframe[1:38976,1]<-NA
tempdataframe[1:38976,2]<-NA
count2<-1
for (count in 1:38976) {
currentprobe<-temp$genes[count,9];
splitcurrentprobe<-strsplit(currentprobe, ',')[[1]][1];
if (currentprobe!="" & is.na(match(splitcurrentprobe, tempdataframe[,1]))==TRUE & is.na(as.numeric(currentprobe))==TRUE) {
currenttstats<-
temp$t[temp$genes[,9]==currentprobe&is.na(temp$genes[,9])==FALSE,2];
tempdataframe[count2,1]<-splitcurrentprobe;
tempdataframe[count2,2]<-mean(currenttstats);
count2<-count2+1;
print(count);
}
collapsedfit_barth2colour_robust_idcmvsdonor_genbank<-tempdataframe[1:15478,]
colnames(collapsedfit_barth2colour_robust_idcmvsdonor_genbank)<-c('Name', 't')
write.table(collapsedfit_barth2colour_robust_idcmvsdonor_genbank, file='barth2colour_idcmvsdonor_genbank_gsealist_15-8-13.txt', quote=FALSE)

# Code used to analyse GSE21610, referred to as 'schweintek', raw CEL file data
# was available. The code below was used to analyse DE in end-stage IDCM and ISCM versus NF LV, as well as the effect of LVAD implantation on LV from patients with end-stage IDCM and ISCM
library(affy)
setwd('C:/Users/Owner/PhD Work/FailingDatasets/schweintek')
schweintekdata<-justRMA()
designforschweintek_donorsvsidcm<-cbind(1, c(0,0,0,0,0,0,0,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1))
fit_schweintek_donorsvidcm<-lmFit(schweintekdata[,c(1:8, 9:13, 15, 17, 18, 21, 22, 25:30, 32:34, 36, 37)], designforschweintek_donorsvsidcm)
fit_schweintek_donorsvidcm<-eBayes(fit_schweintek_donorsvidcm)

# Genes DE in end-stage IDCM versus NF LV
> topTable(fit_schweintek_donorsvidcm, coef=2, sort.by='p')
   ID logFC AveExpr   t P.Value adj.P.Val B
42823 233571_x_at 1.2516305 9.576656 9.249824 2.469947e-10 1.350443e-05 13.138533
4238 1558487_a_at 1.0637660 8.281482 8.883942 6.097263e-10 1.447984e-05 12.336527
35275 226018_at 1.6070617 6.531544 8.674059 6.097263e-10 1.447984e-05 11.866949
<table>
<thead>
<tr>
<th>ID</th>
<th>logFC</th>
<th>AveExpr</th>
<th>t</th>
<th>P.Value</th>
<th>adj.P.Val</th>
<th>B</th>
</tr>
</thead>
<tbody>
<tr>
<td>155</td>
<td>2.1234653</td>
<td>5.209639</td>
<td>10.306698</td>
<td>4.112816e-09</td>
<td>0.0002248682</td>
<td>10.466263</td>
</tr>
<tr>
<td>33848</td>
<td>-6.1298460</td>
<td>7.448365</td>
<td>-9.085146</td>
<td>2.954102e-08</td>
<td>0.0005310377</td>
<td>8.809037</td>
</tr>
<tr>
<td>10348</td>
<td>0.8839277</td>
<td>8.070031</td>
<td>8.872926</td>
<td>4.236028e-08</td>
<td>0.0005310377</td>
<td>8.499415</td>
</tr>
<tr>
<td>16071</td>
<td>2.1795620</td>
<td>4.691875</td>
<td>8.832453</td>
<td>4.540298e-08</td>
<td>0.0005310377</td>
<td>8.439606</td>
</tr>
<tr>
<td>12377</td>
<td>0.5654370</td>
<td>7.589190</td>
<td>8.793307</td>
<td>4.856312e-08</td>
<td>0.0005310377</td>
<td>8.381524</td>
</tr>
<tr>
<td>37601</td>
<td>2.28346</td>
<td>6.285383</td>
<td>8.394093</td>
<td>9.751779e-08</td>
<td>0.0008886308</td>
<td>7.775938</td>
</tr>
<tr>
<td>24462</td>
<td>2.15167</td>
<td>6.281897</td>
<td>8.206945</td>
<td>1.361639e-07</td>
<td>0.0010635376</td>
<td>7.483610</td>
</tr>
<tr>
<td>13671</td>
<td>1.9928544</td>
<td>7.214844</td>
<td>7.916570</td>
<td>2.306345e-07</td>
<td>0.0010752420</td>
<td>7.019198</td>
</tr>
<tr>
<td>37089</td>
<td>2.27834</td>
<td>6.252234</td>
<td>7.842163</td>
<td>2.644503e-07</td>
<td>0.0010752420</td>
<td>6.898054</td>
</tr>
<tr>
<td>37024</td>
<td>2.27769</td>
<td>5.778544</td>
<td>7.839281</td>
<td>2.658595e-07</td>
<td>0.0010752420</td>
<td>6.893343</td>
</tr>
</tbody>
</table>

# Genes DE in end-stage ISCM versus ND LV

> topTable(fit_schweintek_donorsviscm, coef=2, sort.by='p')

# Code to prepare R results of comparisons of end-stage IDCM and ISCM versus NF LV for GSEA
schweintek_donorsvidcm_gsealist <- topTable(fit_schweintek_donorsvidcm, coef=2, number=54675, sort.by='p')[,c(1,4)]
write.table(schweintek_donorsvidcm_gsealist, file='schweintek_donorsvidcm_forGSEA_4-6-13.txt', quote=FALSE)
schweintek_donorsviscm_gsealist <- topTable(fit_schweintek_donorsviscm, coef=2, number=54675, sort.by='p')[,c(1,4)]
write.table(schweintek_donorsviscm_gsealist, file='schweintek_donorsviscm_forGSEA_4-6-13.txt', quote=FALSE)

# Code to analyse the effects of LV AD implantation in end-stage IDCM and ISCM
SchweintekFileNames <- c("GSM545665.cel", "GSM545666.cel", "GSM545667.cel", "GSM545668.cel", "GSM545669.cel", "GSM545670.cel", "GSM545671.cel", "GSM545672.cel", "GSM545673.cel", "GSM545674.cel", "GSM545675.cel", "GSM545676.cel", "GSM545677.cel", "GSM545678.cel", "GSM545679.cel", "GSM545680.cel", "GSM545681.cel", "GSM545682.cel", "GSM545683.cel", "GSM545684.cel", "GSM545685.cel", "GSM545686.cel", "GSM545687.cel", "GSM545688.cel", "GSM545689.cel", "GSM545690.cel", "GSM545691.cel", "GSM545692.cel", "GSM545693.cel", "GSM545694.cel", "GSM545695.cel", "GSM545696.cel", "GSM545697.cel", "GSM545698.cel", "GSM545699.cel", "GSM545700.cel", "GSM545701.cel", "GSM545702.cel", "GSM545703.cel", "GSM545704.cel", "GSM545705.cel", "GSM545706.cel", "GSM545707.cel", "GSM545708.cel", "GSM545709.cel", "GSM545710.cel", "GSM545711.cel", "GSM545712.cel", "GSM545713.cel", "GSM545714.cel", "GSM545715.cel", "GSM545716.cel", "GSM545717.cel", "GSM545718.cel", "GSM545719.cel", "GSM545720.cel", "GSM545721.cel", "GSM545722.cel", "GSM545723.cel", "GSM545724.cel")
schweintektargets_iscm <- data.frame(Filenames=SchweintekFileNames[vectorofphenotypesforschweintekfailingph
enotype=='i'], Subjects=c(1:9,1:9), Time=c(rep(0,times=9), rep(1, times=9)))
schweintektargets_idcm <- data.frame(Filenames=SchweintekFileNames[vectorofphenotypesforschweintekfailingph
enotype=='d'], Subjects=c(1:21,1:21), Time=c(rep(0,times=21), rep(1, times=21)))
schweintekdesign_idcm <- model.matrix(~schweinteksub_idcm+schweintektiming_idcm)
schweinteksubs_iscm<-factor(schweintektargets_iscm$Subjects)
schweintektiming_iscm<-factor(schweintektargets_iscm$Time, levels=c(0,1))
schweintekdesign_iscm<-model.matrix(~schweinteksubs_iscm+schweintektiming_iscm)
schweintek_failing<-schweintekdata[,9:68]
fit_schweintek_robust_LVADexvsim_idcm <-
  lmFit(schweintek_failing[,vectorofphenotypesforschweintekfailingphenotype=='d'],
         schweintekdesign_idcm, method='robust')
fit_schweintek_robust_LVADexvsim_idcm <-
eBayes(fit_schweintek_robust_LVADexvsim_idcm)

# Genes DE by LVAD implantation in end-stage IDCM
> topTable(fit_schweintek_robust_LVADexvsim_idcm, coef='schweintektiming_idcm1')
  ID      logFC AveExpr         t     P.Value adj.P.Val  B
16955 207513_s_at 1.3425099 8.499935 30.93333 1.156314e-20 6.322147e-16 33.75298
28008 218723_s_at 1.2706389 9.564207 26.02777 1.687587e-14 30.93276
15331 205883_at 1.3421178 8.420752 21.20945 6.499152e-17 1.184471e-12 27.20863
17820 208398_s_at 0.8352029 7.625273 20.23330 1.875956e-16 2.564197e-12 26.47265
34148 224889_at 0.9825554 8.681649 19.75144 3.220050e-16 3.521125e-12 26.02252
35312 226055_at 0.8140701 7.456819 19.56002 4.003952e-16 3.648601e-12 25.94579
34115 224856_at 1.9608130 9.09509 19.30856 5.346344e-16 1.75876e-12 25.69509
32636 223358_s_at 0.9886052 8.639506 18.52501 1.345267e-15 9.194057e-12 24.86555
12337 202887_s_at 1.8455939 8.269095 17.81154 3.212563e-15 1.816861e-11 24.16195
18179 208763_s_at 1.4193679 8.097330 17.78433 3.323019e-15 1.816861e-11 24.07355

fit_schweintek_robust_LVADexvsim_iscm <-
  lmFit(schweintek_failing[,vectorofphenotypesforschweintekfailingphenotype=='i'],
         schweintekdesign_iscm, method='robust')
fit_schweintek_robust_LVADexvsim_iscm<-
eBayes(fit_schweintek_robust_LVADexvsim_iscm)

# Genes DE by LVAD implantation in end-stage ISCM
> topTable(fit_schweintek_robust_LVADexvsim_iscm, coef='schweintektiming_iscm1')
  ID      logFC AveExpr         t     P.Value adj.P.Val  B
34099 224840_at  1.9868161 10.116763 22.12817 2.652294e-10 1.450142e-05 12.53339
18179 208763_s_at  1.7879048 8.597411 19.08687 1.250905e-09 1.709831e-05 11.64550
29299 220014_at  1.0520837 5.218632 19.19835 1.177018e-09 1.709831e-05 11.38350
53948 244697_at  1.1226816 6.984085 19.58219 9.568398e-10 1.709831e-05 11.34034
34115 224856_at  1.9016837 9.065395 18.30444 1.936577e-09 1.733020e-05 11.27896
35717 226460_at  0.9284098 8.963066 18.61211 1.627413e-09 1.733020e-05 11.00028
# Code to prepare results in R regarding LVAD implantation for GSEA

schweintek_LVADexvsim_ismc_gsealist <-
topTable(fit_schweintek_robust_LVADexvsim_ismc, coef='schweintektiming_ismc1', number=54675, sort.by='p')[,c(1,4)]
write.table(schweintek_LVADexvsim_ismc_gsealist, file='schweintek_LVADexvsim_ismc_forGSEA_18-12-12.txt', quote=FALSE)

schweintek_LVADexvsim_idcm_gsealist <-
topTable(fit_schweintek_robust_LVADexvsim_idcm, coef='schweintektiming_idcm1', number=54675, sort.by='p')[,c(1,4)]
write.table(schweintek_LVADexvsim_idcm_gsealist, file='schweintek_LVADexvsim_idcm_forGSEA_18-12-12.txt', quote=FALSE)

# R code to generate intersection plots in Figures 4.1 and 4.2 for the IDCM
# Affymetrix datasets - GSE21610 referred to as 'schweintek', GSE5406 referred to as 'hannen', GSE1869 referred to as 'kittle', GSE1145 referred to as 'tsubaki', GSE3585 referred to as 'barth'

ranked_kittle_idcm <- topTable(fit_kittle_robust_donorvsidcm, coef=2, number=22238, adjust.method='fdr', sort.by='p')$ID
ranked_hannen_idcm <- topTable(fit_hannen_robust_donorvsidcm, number=22238, adjust.method='fdr', sort.by='p')$ID
ranked_barth_idcm <- topTable(fit_barth_robust, coef=2, number=22238, adjust.method='fdr', sort.by='p')$ID
ranked_schweintek_idcm_all <- topTable(fit_schweintek_donorsvidcm, coef=2, number=54675, adjust.method='fdr', sort.by='p')$ID
ranked_schweintek_idcm_subset <- vector(length=22238)
count2 <- 1
for (count in 1:54675) {
  if (is.na(match(ranked_schweintek_idcm_all[count], ranked_kittle_idcm)) == FALSE) {
    ranked_schweintek_idcm_subset[count2] <- ranked_schweintek_idcm_all[count]
    count2 <- count2 + 1
  }
}

10.64936

# R code to generate intersection plots in Figures 4.1 and 4.2 for the IDCM
# Affymetrix datasets - GSE21610 referred to as 'schweintek', GSE5406 referred to as 'hannen', GSE1869 referred to as 'kittle', GSE1145 referred to as 'tsubaki', GSE3585 referred to as 'barth'
ranked_kittle_idcm <- topTable(fit_kittle_robust_donorvsidcm, coef=2, number=22238, adjust.method='fdr', sort.by='p')$ID
ranked_hannen_idcm <- topTable(fit_hannen_robust_donorvsidcm, number=22238, adjust.method='fdr', sort.by='p')$ID
ranked_barth_idcm <- topTable(fit_barth_robust, coef=2, number=22238, adjust.method='fdr', sort.by='p')$ID
ranked_schweintek_idcm_all <- topTable(fit_schweintek_donorsvidcm, coef=2, number=54675, adjust.method='fdr', sort.by='p')$ID
ranked_schweintek_idcm_subset <- vector(length=22238)
count2 <- 1
for (count in 1:54675) {
  if (is.na(match(ranked_schweintek_idcm_all[count], ranked_kittle_idcm)) == FALSE) {
    ranked_schweintek_idcm_subset[count2] <- ranked_schweintek_idcm_all[count]
    count2 <- count2 + 1
  }
}
ranked_tsubaki_idcm_all<-topTable(fit_tsubaki_robust_donorvsidcm, coef=2, number=54675, adjust.method='fdr', sort.by='p')$ID
ranked_tsubaki_idcm_subset<-vector(length=22238)
count2<-1
for (count in 1:54675) {
    if (is.na(match(ranked_tsubaki_idcm_all[count], ranked_kittle_idcm))==FALSE) {
        ranked_tsubaki_idcm_subset[count2]<-ranked_tsubaki_idcm_all[count];
        count2<-count2+1;
    }
}
kittlevshannen_idcm_intersect<-vector()
kittlevsbarth_idcm_intersect<-vector()
kittlevstsubaki_idcm_intersect<-vector()
hannenvsbarth_idcm_intersect<-vector()
hannenvsstsubaki_idcm_intersect<-vector()
hannenvsschweintek_idcm_intersect<-vector()
barthvstsubaki_idcm_intersect<-vector()
barthvsschweintek_idcm_intersect<-vector()
tsubakivsschweintek_idcm_intersect<-vector()
for (count in 1:22283) {
    kittlevshannen_idcm_intersect[count]<-
    length(intersect(ranked_kittle_idcm[1:count], ranked_hannen_idcm[1:count]))/count;
    kittlevsbarth_idcm_intersect[count]<-
    length(intersect(ranked_kittle_idcm[1:count], ranked_barth_idcm[1:count]))/count;
    kittlevstsubaki_idcm_intersect[count]<-
    length(intersect(ranked_kittle_idcm[1:count], ranked_tsubaki_idcm_subset[1:count]))/count;
    kittlevsschweintek_idcm_intersect[count]<-
    length(intersect(ranked_kittle_idcm[1:count], ranked_schweintek_idcm_subset[1:count]))/count;
    hannenvsbarth_idcm_intersect[count]<-
    length(intersect(ranked_hannen_idcm[1:count], ranked_barth_idcm[1:count]))/count;
    hannenvsstsubaki_idcm_intersect[count]<-
    length(intersect(ranked_hannen_idcm[1:count], ranked_tsubaki_idcm_subset[1:count]))/count;
    hannenvsschweintek_idcm_intersect[count]<-
    length(intersect(ranked_hannen_idcm[1:count], ranked_schweintek_idcm_subset[1:count]))/count;
    barthvstsubaki_idcm_intersect[count]<-
    length(intersect(ranked_barth_idcm[1:count], ranked_tsubaki_idcm_subset[1:count]))/count;
    barthvsschweintek_idcm_intersect[count]<-
    length(intersect(ranked_barth_idcm[1:count], ranked_schweintek_idcm_subset[1:count]))/count;
```r
tsubakivsschweintek_idcm_intersect[count] <- length(intersect(ranked_tsubaki_idcm_subset[1:count], ranked_schweintek_idcm_subset[1:count]))/count;
}

# To plot the intersection plots for IDCM Affymetrix datasets (all 22283 probes)
plot(1:22283, kittlevshannen_idcm_intersect, type='l', col='blue', xlim=c(0,22283), ylim=c(0, 1), ylab='Proportion of Overlapping Genes', xlab='Gene Rank')
par(new=TRUE)
plot(1:22283, kittlevsbarth_idcm_intersect, type='l', col='yellow', xlim=c(0,22283), ylim=c(0, 1), ylab='', xlab='')
par(new=TRUE)
plot(1:22283, kittlevstsubaki_idcm_intersect, type='l', col='red', xlim=c(0,22283), ylim=c(0, 1), ylab='', xlab='')
par(new=TRUE)
plot(1:22283, kittlevsschweintek_idcm_intersect, type='l', col='purple', xlim=c(0,22283), ylim=c(0, 1), ylab='', xlab='')
par(new=TRUE)
plot(1:22283, hannenvsbarth_idcm_intersect, type='l', col='orange', xlim=c(0,22283), ylim=c(0, 1), ylab='', xlab='')
par(new=TRUE)
plot(1:22283, hannenvstsubaki_idcm_intersect, type='l', col='green', xlim=c(0,22283), ylim=c(0, 1), ylab='', xlab='')
par(new=TRUE)
plot(1:22283, hannenvschweintek_idcm_intersect, type='l', col='brown', xlim=c(0,22283), ylim=c(0, 1), ylab='', xlab='')
par(new=TRUE)
plot(1:22283, barthvstsubaki_idcm_intersect, type='l', col='grey', xlim=c(0,22283), ylim=c(0, 1), ylab='', xlab='')
par(new=TRUE)
plot(1:22283, barthvsschweintek_idcm_intersect, type='l', col='black', xlim=c(0,22283), ylim=c(0, 1), ylab='', xlab='')
par(new=FALSE)

# To plot the intersection plots for IDCM Affymetrix datasets (200 most highly ranked probes in order of strength of DE)
plot(1:2000, kittlevshannen_idcm_intersect[1:2000], type='l', col='blue', xlim=c(0,2000), ylim=c(0, 1), ylab='Proportion of Overlapping Genes', xlab='Gene Rank')
par(new=TRUE)
plot(1:2000, kittlevsbarth_idcm_intersect[1:2000], type='l', col='yellow', xlim=c(0,2000), ylim=c(0, 1), ylab='', xlab='')
par(new=TRUE)
```

189
```r
plot(1:2000, kittlevstsubaki_idcm_intersect[1:2000], type='l', col='red', xlim=c(0,2000), ylim=c(0, 1), ylab='', xlab='')
par(new=TRUE)
plot(1:2000, kittlevsschweintek_idcm_intersect[1:2000], type='l', col='purple', xlim=c(0,2000), ylim=c(0, 1), ylab='', xlab='')
par(new=TRUE)
plot(1:2000, hannenvsbarth_idcm_intersect[1:2000], type='l', col='orange', xlim=c(0,2000), ylim=c(0, 1), ylab='', xlab='')
par(new=TRUE)
plot(1:2000, hannenvstsubaki_idcm_intersect[1:2000], type='l', col='green', xlim=c(0,2000), ylim=c(0, 1), ylab='', xlab='')
par(new=TRUE)
plot(1:2000, hannenvschweintek_idcm_intersect[1:2000], type='l', col='brown', xlim=c(0,2000), ylim=c(0, 1), ylab='', xlab='')
par(new=TRUE)
plot(1:2000, barthvstsubaki_idcm_intersect[1:2000], type='l', col='grey', xlim=c(0,2000), ylim=c(0, 1), ylab='', xlab='')
par(new=TRUE)
plot(1:2000, barthvsschweintek_idcm_intersect[1:2000], type='l', col='black', xlim=c(0,2000), ylim=c(0, 1), ylab='', xlab='')
par(new=TRUE)
plot(1:2000, tsubakivsschweintek_idcm_intersect[1:2000], type='l', col='cyan', xlim=c(0,2000), ylim=c(0, 1), ylab='', xlab='')
par(new=FALSE)

# R code to generate intersection plots in Figures 4.3 and 4.4 for the ISCM
# Affymetrix datasets - GSE21610 referred to as 'schweintek', GSE5406 referred to as 'hannen', GSE1869 referred to as 'kittle', GSE1145 referred to as 'tsubaki'
ranked_kittle_iscm <- topTable(fit_kittle_robust_donorvsiscm, coef=2, number=22238, adjust.method='fdr', sort.by='p')$ID
ranked_hannen_iscm <- topTable(fit_hannen_robust_donorvsiscm, coef=2, number=22238, adjust.method='fdr', sort.by='p')$ID
ranked_schweintek_iscm_all <- topTable(fit_schweintek_donorsviscm, coef=2, number=54675, adjust.method='fdr', sort.by='p')$ID
ranked_schweintek_iscm_subset <- vector(length=22238)
count2 <- 1
for (count in 1:54675) {
  if (is.na(match(ranked_schweintek_iscm_all[count], ranked_kittle_iscm)) == FALSE) {
    ranked_schweintek_iscm_subset[count2] <- ranked_schweintek_iscm_all[count];
    count2 <- count2 + 1;
  }
}
```

# R code to generate intersection plots in Figures 4.3 and 4.4 for the ISCM
# Affymetrix datasets - GSE21610 referred to as 'schweintek', GSE5406 referred to as 'hannen', GSE1869 referred to as 'kittle', GSE1145 referred to as 'tsubaki'
ranked_kittle_iscm <- topTable(fit_kittle_robust_donorvsiscm, coef=2, number=22238, adjust.method='fdr', sort.by='p')$ID
ranked_hannen_iscm <- topTable(fit_hannen_robust_donorvsiscm, coef=2, number=22238, adjust.method='fdr', sort.by='p')$ID
ranked_schweintek_iscm_all <- topTable(fit_schweintek_donorsviscm, coef=2, number=54675, adjust.method='fdr', sort.by='p')$ID
ranked_schweintek_iscm_subset <- vector(length=22238)
count2 <- 1
for (count in 1:54675) {
  if (is.na(match(ranked_schweintek_iscm_all[count], ranked_kittle_iscm)) == FALSE) {
    ranked_schweintek_iscm_subset[count2] <- ranked_schweintek_iscm_all[count];
    count2 <- count2 + 1;
  }
}
ranked_tsubaki_iscm_all <- topTable(fit_tsubaki_robust_donorvsiscm, coef=2, number=54675, adjust.method='fdr', sort.by='p')$ID
ranked_tsubaki_iscm_subset <- vector(length=22238)
count2 <- 1
for (count in 1:54675) {
  if (is.na(match(ranked_tsubaki_iscm_all[count], ranked_kittle_iscm)) == FALSE) {
    ranked_tsubaki_iscm_subset[count2] <- ranked_tsubaki_iscm_all[count];
    count2 <- count2 + 1;
  }
}
kittlevshannen_iscm_intersect <- vector()
kittlevtsubaki_iscm_intersect <- vector()
kittlevschweintek_iscm_intersect <- vector()
hannenvtsubaki_iscm_intersect <- vector()
hannenvschweintek_iscm_intersect <- vector()
tsubakivsschweintek_iscm_intersect <- vector()
for (count in 1:22283) {
  kittlevshannen_iscm_intersect[count] <-
  length(intersect(ranked_kittle_iscm[1:count], ranked_hannen_iscm[1:count]))/count;
  kittlevtsubaki_iscm_intersect[count] <-
  length(intersect(ranked_kittle_iscm[1:count], ranked_tsubaki_iscm_subset[1:count]))/count;
  kittlevschweintek_iscm_intersect[count] <-
  length(intersect(ranked_kittle_iscm[1:count], ranked_schweintek_iscm_subset[1:count]))/count;
  hannenvtsubaki_iscm_intersect[count] <-
  length(intersect(ranked_hannen_iscm[1:count], ranked_tsubaki_iscm_subset[1:count]))/count;
  hannenvschweintek_iscm_intersect[count] <-
  length(intersect(ranked_hannen_iscm[1:count], ranked_schweintek_iscm_subset[1:count]))/count;
  tsubakivsschweintek_iscm_intersect[count] <-
  length(intersect(ranked_tsubaki_iscm_subset[1:count], ranked_schweintek_iscm_subset[1:count]))/count;
}

# To plot the intersection plots for ISCM Affymetrix datasets (all 22238 probes)
plot(1:22283, kittlevshannen_iscm_intersect, type='l', col='blue', xlim=c(0,22283), ylim=c(0,1), ylab='Proportion of Overlapping Genes', xlab='Gene Rank')
par(new=TRUE)
plot(1:22283, kittlevtsubaki_iscm_intersect, type='l', col='red', xlim=c(0,22283), ylim=c(0,1), ylab='', xlab='')
par(new=TRUE)
```r
# To plot the intersection plots for ISCM Affymetrix datasets (200 most highly ranked probes in order of strength of DE)
plot(1:2000, kittlevshannen_iscm_intersect[1:2000], type='l', col='blue', xlim=c(0,2000), ylim=c(0, 1), ylab='Proportion of Overlapping Genes', xlab='Gene Rank')
par(new=TRUE)
plot(1:2000, kittlevstsubaki_iscm_intersect[1:2000], type='l', col='red', xlim=c(0,2000), ylim=c(0, 1), ylab='')
par(new=TRUE)
plot(1:2000, kittlevsschweintek_iscm_intersect[1:2000], type='l', col='purple', xlim=c(0,2000), ylim=c(0, 1), ylab='')
par(new=TRUE)
plot(1:2000, hannenvstsubaki_iscm_intersect[1:2000], type='l', col='green', xlim=c(0,2000), ylim=c(0, 1), ylab='')
par(new=TRUE)
plot(1:2000, hannenvschweintek_iscm_intersect[1:2000], type='l', col='brown', xlim=c(0,2000), ylim=c(0, 1), ylab='')
par(new=TRUE)
plot(1:2000, tsubakivsschweintek_iscm_intersect[1:2000], type='l', col='cyan', xlim=c(0,2000), ylim=c(0, 1), ylab='')
par(new=FALSE)
```

```r
# R code to generate intersection plots in Figures 4.5 and 4.6 for the IDCM
# Affymetrix datasets - GSE21610 referred to as 'schweintek', GSE5406 referred to as 'hannen', GSE1869 referred to as 'kittle', GSE1145 referred to as 'tsubaki', GSE3585 referred to as 'barth'
ranked_kittle_idcm_upregulated<-topTable(fit_kittle_robust_donorvsidcm, coef=2, number=22238, adjust.method='fdr', sort.by='logFC')
```
ranked_kittle_idcm_downregulated <- topTable(fit_kittle_robust_donorvsidcm, coef=2, number=22238, adjust.method='fdr', sort.by='logFC')[
  topTable(fit_kittle_robust_donorvsidcm, coef=2, number=22238, adjust.method='fdr', sort.by='logFC')[,2]<0[,]$ID
]

ranked_hannen_idcm_upregulated <- topTable(fit_hannen_robust_donorvsidcm, number=22238, adjust.method='fdr', sort.by='logFC')[
  topTable(fit_hannen_robust_donorvsidcm, number=22238, adjust.method='fdr', sort.by='logFC')[,5]>0[,]$ID
]

ranked_hannen_idcm_downregulated <- topTable(fit_hannen_robust_donorvsidcm, number=22238, adjust.method='fdr', sort.by='logFC')[
  topTable(fit_hannen_robust_donorvsidcm, number=22238, adjust.method='fdr', sort.by='logFC')[,5]<0[,]$ID
]

ranked_barth_idcm_upregulated <- topTable(fit_barth_robust, coef=2, number=22238, adjust.method='fdr', sort.by='logFC')[
  topTable(fit_barth_robust, coef=2, number=22238, adjust.method='fdr', sort.by='logFC')[,5]>0[,]$ID
]

ranked_barth_idcm_downregulated <- topTable(fit_barth_robust, coef=2, number=22238, adjust.method='fdr', sort.by='logFC')[
  topTable(fit_barth_robust, coef=2, number=22238, adjust.method='fdr', sort.by='logFC')[,5]<0[,]$ID
]

ranked_schweintek_idcm_upregulated <- topTable(fit_schweintek_donorsvidcm, coef=2, number=54675, adjust.method='fdr', sort.by='logFC')[
  topTable(fit_schweintek_donorsvidcm, coef=2, number=54675, adjust.method='fdr', sort.by='logFC')[,2]>0[,]$ID
]

ranked_schweintek_idcm_downregulated <- topTable(fit_schweintek_donorsvidcm, coef=2, number=54675, adjust.method='fdr', sort.by='logFC')[
  topTable(fit_schweintek_donorsvidcm, coef=2, number=54675, adjust.method='fdr', sort.by='logFC')[,2]<0[,]$ID
]

ranked_schweintek_idcm_upregulated_subset <- vector(length=22238)
count2<-1
for (count in 1:length(ranked_schweintek_idcm_upregulated)) {
    if (is.na(match(ranked_schweintek_idcm_upregulated[count], ranked_kittle_idcm))===FALSE) {
        ranked_schweintek_idcm_upregulated_subset[count2]<-
        ranked_schweintek_idcm_upregulated[count];
        count2<-count2+1;
    }
}

ranked_schweintek_idcm_downregulated_subset <- vector(length=22238)
count2<-1
for (count in 1:length(ranked_schweintek_idcm_downregulated)) {
    if (is.na(match(ranked_schweintek_idcm_downregulated[count], ranked_kittle_idcm))===FALSE) {
        ranked_schweintek_idcm_downregulated_subset[count2]<-
        ranked_schweintek_idcm_downregulated[count];
        count2<-count2+1;
    }
}
ranked_tsubaki_idcm_upregulated<-topTable(fit_tsubaki_robust_donorvsidcm, coef=2, number=54675, adjust.method='fdr', sort.by='logFC')[topTable(fit_tsubaki_robust_donorvsidcm, coef=2, number=54675, adjust.method='fdr', sort.by='logFC')[,2]>0,]$ID
ranked_tsubaki_idcm_downregulated<-topTable(fit_tsubaki_robust_donorvsidcm, coef=2, number=54675, adjust.method='fdr', sort.by='logFC')[topTable(fit_tsubaki_robust_donorvsidcm, coef=2, number=54675, adjust.method='fdr', sort.by='logFC')[,2]<0,]$ID
ranked_tsubaki_idcm_upregulated_subset<-vector(length=22238)
count2<-1
for (count in 1:length(ranked_tsubaki_idcm_upregulated)) {
  if (is.na(match(ranked_tsubaki_idcm_upregulated[count], ranked_kittle_idcm))){FALSE) {
    ranked_tsubaki_idcm_upregulated_subset[count2]<-
    ranked_tsubaki_idcm_upregulated[count];
    count2<-count2+1;
  }
}
ranked_tsubaki_idcm_downregulated_subset<-vector(length=22238)
count2<-1
for (count in 1:length(ranked_tsubaki_idcm_downregulated)) {
  if (is.na(match(ranked_tsubaki_idcm_downregulated[count], ranked_kittle_idcm))){FALSE) {
    ranked_tsubaki_idcm_downregulated_subset[count2]<-
    ranked_tsubaki_idcm_downregulated[count];
    count2<-count2+1;
  }
}
kittlevshannen_idcm_intersect_down<-vector()
kittlevsbarth_idcm_intersect_down<-vector()
kittlevstsubaki_idcm_intersect_down<-vector()
kittlevsschweintek_idcm_intersect_down<-vector()
hannenvsbarth_idcm_intersect_down<-vector()
hannenvstsubaki_idcm_intersect_down<-vector()
hannenvschweintek_idcm_intersect_down<-vector()
barthvstsubaki_idcm_intersect_down<-vector()
barthvsschweintek_idcm_intersect_down<-vector()
tsubakivsschweintek_idcm_intersect_down<-vector()
kittlevshannen_idcm_intersect_up<-vector()
kittlevsbarth_idcm_intersect_up<-vector()
kittlevstsubaki_idcm_intersect_up<-vector()
kittlevsschweintek_idcm_intersect_up<-vector()
hannenvsbarth_idcm_intersect_up<-vector()
hannenvstsubaki_idcm_intersect_up<-vector()
hannenvschweintek_idcm_intersect_up<-vector()
barthvstsubaki_idcm_intersect_up<-vector()
barthvsschweintek_idcm_intersect_up<-vector()
tsubakivsschweintek_idcm_intersect_up<-vector()
for (count in 1:2000) {
    kittlevshannen_idcm_intersect_up[count]<-
    length(intersect(ranked_kittle_idcm_upregulated[1:count],
    ranked_hannen_idcm_upregulated[1:count]))/count;
    kittlevsbarth_idcm_intersect_up[count]<-
    length(intersect(ranked_kittle_idcm_upregulated[1:count],
    ranked_barth_idcm_upregulated[1:count]))/count;
    kittlevstsubaki_idcm_intersect_up[count]<-
    length(intersect(ranked_kittle_idcm_upregulated[1:count],
    ranked_tsubaki_idcm_upregulated_subset[1:count]))/count;
    kittlevsschweintek_idcm_intersect_up[count]<-
    length(intersect(ranked_kittle_idcm_upregulated[1:count],
    ranked_schweintek_idcm_upregulated_subset[1:count]))/count;
    hannenvsbarth_idcm_intersect_up[count]<-
    length(intersect(ranked_hannen_idcm_upregulated[1:count],
    ranked_barth_idcm_upregulated[1:count]))/count;
    hannenvstsubaki_idcm_intersect_up[count]<-
    length(intersect(ranked_hannen_idcm_upregulated[1:count],
    ranked_tsubaki_idcm_upregulated_subset[1:count]))/count;
    hannenvschweintek_idcm_intersect_up[count]<-
    length(intersect(ranked_hannen_idcm_upregulated[1:count],
    ranked_schweintek_idcm_upregulated_subset[1:count]))/count;
    barthvstsubaki_idcm_intersect_up[count]<-
    length(intersect(ranked_barth_idcm_upregulated[1:count],
    ranked_tsubaki_idcm_upregulated_subset[1:count]))/count;
    barthvsschweintek_idcm_intersect_up[count]<-
    length(intersect(ranked_barth_idcm_upregulated[1:count],
    ranked_schweintek_idcm_upregulated_subset[1:count]))/count;
    tsubakivsschweintek_idcm_intersect_up[count]<-
    length(intersect(ranked_tsubaki_idcm_upregulated_subset[1:count],
    ranked_schweintek_idcm_upregulated_subset[1:count]))/count;
}
kittlevshannen_idcm_intersect_down[count]<-
length(intersect(ranked_kittle_idcm_downregulated[1:count],
ranked_hannen_idcm_downregulated[1:count]))/count;
kittlevsbarth_idcm_intersect_down[count]<-
length(intersect(ranked_kittle_idcm_downregulated[1:count],
ranked_barth_idcm_downregulated[1:count]))/count;
kittlevstsubaki_idcm_intersect_down[count]<-
length(intersect(ranked_kittle_idcm_downregulated[1:count],
ranked_tsubaki_idcm_downregulated_subset[1:count]))/count;
kittlevsschweintek_idcm_intersect_down[count]<-
length(intersect(ranked_kittle_idcm_downregulated[1:count],
ranked_schweintek_idcm_downregulated_subset[1:count]))/count;
hannenvsbarth_idcm_intersect_down[count]<-
length(intersect(ranked_hannen_idcm_downregulated[1:count],
ranked_barth_idcm_downregulated[1:count]))/count;
hannenvstsubaki_idcm_intersect_down[count]<-
length(intersect(ranked_hannen_idcm_downregulated[1:count],
ranked_tsubaki_idcm_downregulated_subset[1:count]))/count;
hannenvschweintek_idcm_intersect_down[count]<-
length(intersect(ranked_hannen_idcm_downregulated[1:count],
ranked_schweintek_idcm_downregulated_subset[1:count]))/count;
barthvstsubaki_idcm_intersect_down[count]<-
length(intersect(ranked_barth_idcm_downregulated[1:count],
ranked_tsubaki_idcm_downregulated_subset[1:count]))/count;
barthvsschweintek_idcm_intersect_down[count]<-
length(intersect(ranked_barth_idcm_downregulated[1:count],
ranked_schweintek_idcm_downregulated_subset[1:count]))/count;
tsubakivsschweintek_idcm_intersect_down[count]<-
length(intersect(ranked_tsubaki_idcm_downregulated_subset[1:count],
ranked_schweintek_idcm_downregulated_subset[1:count]))/count;
kittelvsschweintek_idcm_intersect_down[count]<-
length(intersect(ranked_kittle_idcm_downregulated[1:count],
ranked_schweintek_idcm_downregulated_subset[1:count]))/count;

hannenvsbarth_idcm_intersect_down[count]<-
length(intersect(ranked_hannen_idcm_downregulated[1:count],
ranked_barth_idcm_downregulated[1:count]))/count;

hannenvstsubaki_idcm_intersect_down[count]<-
length(intersect(ranked_hannen_idcm_downregulated[1:count],
ranked_tsubaki_idcm_downregulated_subset[1:count]))/count;

hannenvschweintek_idcm_intersect_down[count]<-
length(intersect(ranked_hannen_idcm_downregulated[1:count],
ranked_schweintek_idcm_downregulated_subset[1:count]))/count;

barthvstsubaki_idcm_intersect_down[count]<-
length(intersect(ranked_barth_idcm_downregulated[1:count],
ranked_tsubaki_idcm_downregulated_subset[1:count]))/count;

barthvsschweintek_idcm_intersect_down[count]<-
length(intersect(ranked_barth_idcm_downregulated[1:count],
ranked_schweintek_idcm_downregulated_subset[1:count]))/count;

tsubakivsschweintek_idcm_intersect_down[count]<-
length(intersect(ranked_tsubaki_idcm_downregulated_subset[1:count],
ranked_schweintek_idcm_downregulated_subset[1:count]))/count;
}

# To plot the intersection plots for IDCM Affymetrix datasets (200 most highly ranked
# upregulated probes in order of strength of DE)
plot(1:2000, kittlevshannen_idcm_intersect_up[1:2000], type='l', col='blue',
xlim=c(0,2000), ylim=c(0, 1), ylab='Proportion of Overlapping Genes', xlab='Gene
Rank')
par(new=TRUE)
plot(1:2000, kittlevsbarth_idcm_intersect_up[1:2000], type='l', col='yellow',
xlim=c(0,2000), ylim=c(0, 1), ylab='', xlab='')
par(new=TRUE)
plot(1:2000, kittlevstsubaki_idcm_intersect_up[1:2000], type='l', col='red',
xlim=c(0,2000), ylim=c(0, 1), ylab='', xlab='')
par(new=TRUE)
plot(1:2000, kittlevsschweintek_idcm_intersect_up[1:2000], type='l', col='purple',
xlim=c(0,2000), ylim=c(0, 1), ylab='', xlab='')
par(new=TRUE)
plot(1:2000, hannenvsbarth_idcm_intersect_up[1:2000], type='l', col='orange',
xlim=c(0,2000), ylim=c(0, 1), ylab='', xlab='')
par(new=TRUE)
plot(1:2000, hannenvstsubaki_idcm_intersect_up[1:2000], type='l', col='green',
xlim=c(0,2000), ylim=c(0, 1), ylab='', xlab='')
par(new=TRUE)
plot(1:2000, hannenvschweintek_idcm_intersect_up[1:2000], type='l', col='brown',
xlim=c(0,2000), ylim=c(0, 1), ylab='', xlab='')
par(new=TRUE)
plot(1:2000, barthvstsubaki_idcm_intersect_up[1:2000], type='l', col='grey',
xlim=c(0,2000), ylim=c(0, 1), ylab='', xlab="")
par(new=TRUE)
plot(1:2000, barthvsschweintek_idcm_intersect_up[1:2000], type='l', col='black',
xlim=c(0,2000), ylim=c(0, 1), ylab='', xlab="")
par(new=TRUE)
plot(1:2000, tsubakivsschweintek_idcm_intersect_up[1:2000], type='l', col='cyan',
xlim=c(0,2000), ylim=c(0, 1), ylab='', xlab="")
par(new=FALSE)

# To plot the intersection plots for IDCM Affymetrix datasets (200 most highly ranked
# downregulated probes in order of strength of DE)
plot(1:2000, kittlevshannen_idcm_intersect_down[1:2000], type='l', col='blue',
xlim=c(0,2000), ylim=c(0, 1), ylab='Proportion of Overlapping Genes', xlab='Gene
Rank')
par(new=TRUE)
plot(1:2000, kittlevsbarth_idcm_intersect_down[1:2000], type='l', col='yellow',
xlim=c(0,2000), ylim=c(0, 1), ylab='', xlab="")
par(new=TRUE)
plot(1:2000, kittlevstsubaki_idcm_intersect_down[1:2000], type='l', col='red',
xlim=c(0,2000), ylim=c(0, 1), ylab='', xlab="")
par(new=TRUE)
plot(1:2000, kittlevsschweintek_idcm_intersect_down[1:2000], type='l', col='purple',
xlim=c(0,2000), ylim=c(0, 1), ylab='', xlab="")
par(new=TRUE)
plot(1:2000, hannenvsbarth_idcm_intersect_down[1:2000], type='l', col='orange',
xlim=c(0,2000), ylim=c(0, 1), ylab='', xlab="")
par(new=TRUE)
plot(1:2000, hannenvstsubaki_idcm_intersect_down[1:2000], type='l', col='green',
xlim=c(0,2000), ylim=c(0, 1), ylab='', xlab="")
par(new=TRUE)
plot(1:2000, hannenvschweintek_idcm_intersect_down[1:2000], type='l', col='brown',
xlim=c(0,2000), ylim=c(0, 1), ylab='', xlab="")
par(new=TRUE)
plot(1:2000, barthvstsubaki_idcm_intersect_down[1:2000], type='l', col='grey',
xlim=c(0,2000), ylim=c(0, 1), ylab='', xlab="")
par(new=TRUE)
plot(1:2000, barthvsschweintek_idcm_intersect_down[1:2000], type='l', col='black',
xlim=c(0,2000), ylim=c(0, 1), ylab='', xlab="")
par(new=TRUE)
plot(1:2000, tsubakivsschweintek_idcm_intersect_down[1:2000], type='l', col='cyan',
xlim=c(0,2000), ylim=c(0, 1), ylab='', xlab="")
par(new=FALSE)
# R code to generate intersection plots in Figures 4.7 and 4.8 for the ISCM
# Affymetrix datasets - GSE21610 referred to as 'schweintek', GSE5406 referred to
# as 'hannen', GSE1869 referred to as 'kittle', GSE1145 referred to as 'tsubaki'

ranked_kittle_iscm_upregulated <- topTable(fit_kittle_robust_donorvsiscm, coef=2,
number=22238, adjust.method='fdr', sort.by='logFC')[
  topTable(fit_kittle_robust_donorvsiscm, coef=2, number=22238, adjust.method='fdr',
sort.by='logFC')[,2]>0,]$ID

ranked_kittle_iscm_downregulated <- topTable(fit_kittle_robust_donorvsiscm, coef=2,
number=22238, adjust.method='fdr', sort.by='logFC')[
  topTable(fit_kittle_robust_donorvsiscm, coef=2, number=22238, adjust.method='fdr',
sort.by='logFC')[,2]<0,]$ID

ranked_hannen_iscm_upregulated <- topTable(fit_hannen_robust_donorvsiscm,
number=22238, adjust.method='fdr', sort.by='logFC')[
  topTable(fit_hannen_robust_donorvsiscm, number=22238, adjust.method='fdr',
sort.by='logFC')[,5]>0,]$ID

ranked_hannen_iscm_downregulated <- topTable(fit_hannen_robust_donorvsiscm,
number=22238, adjust.method='fdr', sort.by='logFC')[
  topTable(fit_hannen_robust_donorvsiscm, number=22238, adjust.method='fdr',
sort.by='logFC')[,5]<0,]$ID

ranked_schweintek_iscm_upregulated <- topTable(fit_schweintek_donorsviscm, coef=2,
number=54675, adjust.method='fdr', sort.by='logFC')[
  topTable(fit_schweintek_donorsviscm, coef=2, number=54675, adjust.method='fdr',
sort.by='logFC')[,2]>0,]$ID

ranked_schweintek_iscm_downregulated <- topTable(fit_schweintek_donorsviscm, coef=2,
number=54675, adjust.method='fdr', sort.by='logFC')[
  topTable(fit_schweintek_donorsviscm, coef=2, number=54675, adjust.method='fdr',
sort.by='logFC')[,2]<0,]$ID

ranked_schweintek_iscm_upregulated_subset <- vector(length=22238)
count2<-1
for (count in 1:length(ranked_schweintek_iscm_upregulated)) {
  if (is.na(match(ranked_schweintek_iscm_upregulated[count],
  ranked_kittle_idcm))==FALSE) {
    ranked_schweintek_iscm_upregulated_subset[count2]<-
ranked_schweintek_iscm_upregulated[count];
  count2<-count2+1;
  }
}

ranked_schweintek_iscm_downregulated_subset <- vector(length=22238)
count2<-1
for (count in 1:length(ranked_schweintek_iscm_downregulated)) {
  if (is.na(match(ranked_schweintek_iscm_downregulated[count],
  ranked_kittle_idcm))==FALSE) {
    ranked_schweintek_iscm_downregulated_subset[count2]<-
ranked_schweintek_iscm_downregulated[count];
  count2<-count2+1;
  }
}
ranked_tsubaki_iscm_upregulated <- topTable(fit_tsubaki_robust_donorvsiscm, coef=2, number=54675, adjust.method='fdr', sort.by='logFC')[
  topTable(fit_tsubaki_robust_donorvsiscm, coef=2, number=54675, adjust.method='fdr', sort.by='logFC')[,2]>0[,]$ID
ranked_tsubaki_iscm_downregulated <- topTable(fit_tsubaki_robust_donorvsiscm, coef=2, number=54675, adjust.method='fdr', sort.by='logFC')[
  topTable(fit_tsubaki_robust_donorvsiscm, coef=2, number=54675, adjust.method='fdr', sort.by='logFC')[,2]<0[,]$ID
ranked_tsubaki_iscm_upregulated_subset <- vector(length=22238)
count2 <- 1
for (count in 1:length(ranked_tsubaki_iscm_upregulated)) {
  if (is.na(match(ranked_tsubaki_iscm_upregulated[count], ranked_kittle_idcm))==FALSE) {
    ranked_tsubaki_iscm_upregulated_subset[count2] <-
    ranked_tsubaki_iscm_upregulated[count];
    count2 <- count2 + 1;
  }
}
ranked_tsubaki_iscm_downregulated_subset <- vector(length=22238)
count2 <- 1
for (count in 1:length(ranked_tsubaki_iscm_downregulated)) {
  if (is.na(match(ranked_tsubaki_iscm_downregulated[count], ranked_kittle_idcm))==FALSE) {
    ranked_tsubaki_iscm_downregulated_subset[count2] <-
    ranked_tsubaki_iscm_downregulated[count];
    count2 <- count2 + 1;
  }
}
kittlevshannen_iscm_intersect_down <- vector()
kittlevssubaki_iscm_intersect_down <- vector()
kittlevschweintek_iscm_intersect_down <- vector()
hannenvstsubaki_iscm_intersect_down <- vector()
hannenvschweintek_iscm_intersect_down <- vector()
tsubakivsschweintek_iscm_intersect_down <- vector()
kittlevshannen_iscm_intersect_up <- vector()
kittlevssubaki_iscm_intersect_up <- vector()
kittlevschweintek_iscm_intersect_up <- vector()
hannenvstsubaki_iscm_intersect_up <- vector()
hannenvschweintek_iscm_intersect_up <- vector()
tsubakivsschweintek_iscm_intersect_up <- vector()
for (count in 1:2000) {
  kittlevshannen_iscm_intersect_up[count] <-
  length(intersect(ranked_kittle_iscm_upregulated[1:count],
                  ranked_hannen_iscm_upregulated[1:count]))/count;
kittlevstsubaki_iscm_intersect_up[count]<-
length(intersect(ranked_kittle_iscm_upregulated[1:count],
ranked_tsubaki_iscm_upregulated_subset[1:count]))/count;

kittlevsschweintek_iscm_intersect_up[count]<-
length(intersect(ranked_kittle_iscm_upregulated[1:count],
ranked_schweintek_iscm_upregulated_subset[1:count]))/count;

hannenvstsubaki_iscm_intersect_up[count]<-
length(intersect(ranked_hannen_iscm_upregulated[1:count],
ranked_tsubaki_iscm_upregulated_subset[1:count]))/count;

hannenvschweintek_iscm_intersect_up[count]<-
length(intersect(ranked_hannen_iscm_upregulated[1:count],
ranked_schweintek_iscm_upregulated_subset[1:count]))/count;

tsubakivsschweintek_iscm_intersect_up[count]<-
length(intersect(ranked_tsubaki_iscm_upregulated_subset[1:count],
ranked_schweintek_iscm_upregulated_subset[1:count]))/count;

kittlevshannen_iscm_intersect_down[count]<-
length(intersect(ranked_kittle_iscm_downregulated[1:count],
ranked_hannen_iscm_downregulated[1:count]))/count;

kittlevstsubaki_iscm_intersect_down[count]<-
length(intersect(ranked_kittle_iscm_downregulated[1:count],
ranked_tsubaki_iscm_downregulated_subset[1:count]))/count;

kittlevsschweintek_iscm_intersect_down[count]<-
length(intersect(ranked_kittle_iscm_downregulated[1:count],
ranked_schweintek_iscm_downregulated_subset[1:count]))/count;

hannenvstsubaki_iscm_intersect_down[count]<-
length(intersect(ranked_hannen_iscm_downregulated[1:count],
ranked_tsubaki_iscm_downregulated_subset[1:count]))/count;

hannenvschweintek_iscm_intersect_down[count]<-
length(intersect(ranked_hannen_iscm_downregulated[1:count],
ranked_schweintek_iscm_downregulated_subset[1:count]))/count;

tsubakivschweintek_iscm_intersect_down[count]<-
length(intersect(ranked_tsubaki_iscm_downregulated_subset[1:count],
ranked_schweintek_iscm_downregulated_subset[1:count]))/count;

}

# To plot the intersection plots for ISCM Affymetrix datasets (200 most highly ranked
# upregulated probes in order of strength of DE)
plot(1:2000, kittlevshannen_iscm_intersect_up[1:2000], type='l', col='blue',
xlim=c(0,2000), ylim=c(0, 1), ylab='Proportion of Overlapping Genes', xlab='Gene
Rank')
par(new=TRUE)
plot(1:2000, kittlevstsubaki_iscm_intersect_up[1:2000], type='l', col='red',
xlim=c(0,2000), ylim=c(0, 1), ylab='', xlab='')
par(new=TRUE)
plot(1:2000, kittlevsschweintek_iscm_intersect_up[1:2000], type='l', col='purple',
xlim=c(0,2000), ylim=c(0, 1), ylab='', xlab='')
```r
par(new=TRUE)
plot(1:2000, hannenvstsubaki_iscm_intersect_up[1:2000], type='l', col='green',
    xlim=c(0,2000), ylim=c(0, 1), ylab="", xlab="")
par(new=TRUE)
plot(1:2000, hannenvschweintek_iscm_intersect_up[1:2000], type='l', col='brown',
    xlim=c(0,2000), ylim=c(0, 1), ylab="", xlab="")
par(new=TRUE)
plot(1:2000, tsubakivsschweintek_iscm_intersect_up[1:2000], type='l', col='cyan',
    xlim=c(0,2000), ylim=c(0, 1), ylab="", xlab="")
par(new=FALSE)

# To plot the intersection plots for ISCM Affymetrix datasets (200 most highly ranked
# downregulated probes in order of strength of DE)
plot(1:2000, kittlevshannen_iscm_intersect_down[1:2000], type='l', col='blue',
    xlim=c(0,2000), ylim=c(0, 1), ylab='Proportion of Overlapping Genes', xlab='Gene Rank')
par(new=TRUE)
plot(1:2000, kittlevstsubaki_iscm_intersect_down[1:2000], type='l', col='red',
    xlim=c(0,2000), ylim=c(0, 1), ylab="", xlab="")
par(new=TRUE)
plot(1:2000, kittlevsschweintek_iscm_intersect_down[1:2000], type='l', col='purple',
    xlim=c(0,2000), ylim=c(0, 1), ylab="", xlab="")
par(new=TRUE)
plot(1:2000, hannenvstsubaki_iscm_intersect_down[1:2000], type='l', col='green',
    xlim=c(0,2000), ylim=c(0, 1), ylab="", xlab="")
par(new=TRUE)
plot(1:2000, hannenvschweintek_iscm_intersect_down[1:2000], type='l', col='brown',
    xlim=c(0,2000), ylim=c(0, 1), ylab="", xlab="")
par(new=TRUE)
plot(1:2000, tsubakivsschweintek_iscm_intersect_down[1:2000], type='l', col='cyan',
    xlim=c(0,2000), ylim=c(0, 1), ylab="", xlab="")
par(new=FALSE)

# R code to calculate amount and significance of overlap among the Affymetrix
# datasets - GSE21610 referred to as 'schweintek', GSE5406 referred to as 'hannen',
# GSE1869 referred to as 'kittle', GSE1145 referred to as 'tsubaki', GSE3585
# referred to as 'barth'

# Number of overlapping probes in the top 2000 probes ordered by strength of DE among
# the Affymetrix ISCM datasets
> length(intersect(ranked_kittle_iscm[1:2000], ranked_hannen_iscm[1:2000]))
[1] 270
> length(intersect(ranked_kittle_iscm[1:2000], ranked_tsubaki_iscm_subset[1:2000]))
[1] 259
```
> length(intersect(ranked_kittle_iscm[1:2000], ranked_schweintek_iscm_subset[1:2000]))
[1] 344
> length(intersect(ranked_hannen_iscm[1:2000], ranked_tsubaki_iscm_subset[1:2000]))
[1] 593
> length(intersect(ranked_hannen_iscm[1:2000], ranked_schweintek_iscm_subset[1:2000]))
[1] 419
> length(intersect(ranked_tsubaki_iscm_subset[1:2000], ranked_schweintek_iscm_subset[1:2000]))
[1] 279

# Number of overlapping probes in the top 2000 probes ordered by strength of DE among
# the Affymetrix IDCM datasets
> length(intersect(ranked_kittle_idcm[1:2000], ranked_hannen_idcm[1:2000]))
[1] 246
> length(intersect(ranked_kittle_idcm[1:2000], ranked_tsubaki_idcm_subset[1:2000]))
[1] 280
> length(intersect(ranked_kittle_idcm[1:2000], ranked_schweintek_idcm_subset[1:2000]))
[1] 253
> length(intersect(ranked_kittle_idcm[1:2000], ranked_barth_idcm[1:2000]))
[1] 293
> length(intersect(ranked_hannen_idcm[1:2000], ranked_tsubaki_idcm_subset[1:2000]))
[1] 269
> length(intersect(ranked_hannen_idcm[1:2000], ranked_schweintek_idcm_subset[1:2000]))
[1] 478
> length(intersect(ranked_hannen_idcm[1:2000], ranked_barth_idcm[1:2000]))
[1] 386
> length(intersect(ranked_tsubaki_idcm_subset[1:2000], ranked_schweintek_idcm_subset[1:2000]))
[1] 141
> length(intersect(ranked_tsubaki_idcm_subset[1:2000], ranked_barth_idcm[1:2000]))
[1] 177
> length(intersect(ranked_schweintek_idcm_subset[1:2000], ranked_barth_idcm[1:2000]))
[1] 311

# To generate the frequency distribution of amount of overlap expected if 2000 randomly
# selected probes are used to assess overlap between two datasets, i.e. the null
# distribution
vectorofoverlaplength_kittlevshannen_iscm<-vector(length=1000000)
for (count in 1:1000000) {
    currentsamplesromkittle_iscm<-sample(ranked_kittle_iscm, 2000);
currentsamplefromhannen_iscm<-sample(ranked_hannen_iscm, 2000);
vectorofoverlaplength_kittlevshannen_iscm[count]<-
length(intersect(currentsamplefromkittle_iscm, currentsamplefromhannen_iscm));
}

# To calculate the p-value for the observed amount of overlapping probes among the
# Affymetrix datasets, p-values are listed at end of each paragraph of computation which
# corresponds to one comparison between datasets

numbergreaterthan_top2000_kittlevshannen_iscm<-0
for (count in 1:1000000) {
    if (vectorofoverlaplength_kittlevshannen_iscm[count]>=270) {
        numbergreaterthan_top2000_kittlevshannen_iscm<-
        numbergreaterthan_top2000_kittlevshannen_iscm+1;
    }
}
> numbergreaterthan_top2000_kittlevshannen_iscm/1000000
0

numbergreaterthan_top2000_kittlevshannen_idcm<-0
for (count in 1:1000000) {
    if (vectorofoverlaplength_kittlevshannen_iscm[count]>=246) {
        numbergreaterthan_top2000_kittlevshannen_idcm<-
        numbergreaterthan_top2000_kittlevshannen_idcm+1;
    }
}
> numbergreaterthan_top2000_kittlevshannen_idcm/1000000
0

numbergreaterthan_top2000_kittlevstsubaki_iscm<-0
for (count in 1:1000000) {
    if (vectorofoverlaplength_kittlevshannen_iscm[count]>=259) {
        numbergreaterthan_top2000_kittlevstsubaki_iscm<-
        numbergreaterthan_top2000_kittlevstsubaki_iscm+1;
    }
}
> numbergreaterthan_top2000_kittlevstsubaki_iscm/1000000
0

numbergreaterthan_top2000_kittlevstsubaki_idcm<-0
for (count in 1:1000000) {
    if (vectorofoverlaplength_kittlevshannen_iscm[count]>=280) {
        numbergreaterthan_top2000_kittlevstsubaki_idcm<-
        numbergreaterthan_top2000_kittlevstsubaki_idcm+1;
    }
}
> numbergreaterthan_top2000_kittlevstsubaki_idcm/1000000
0

203
0

data:

for (count in 1:1000000) {
    if (vectorofoverlaplength_kittlevshannen_iscm[count] >= 291) {
        numbergreaterthan_top2000_kittlevsbarth_idcm <-
        numbergreaterthan_top2000_kittlevsbarth_idcm + 1;
    }
}

> numbergreaterthan_top2000_kittlevsbarth_idcm/1000000
0

for (count in 1:1000000) {
    if (vectorofoverlaplength_kittlevshannen_iscm[count] >= 344) {
        numbergreaterthan_top2000_kittlevsschweintek_iscm <-
        numbergreaterthan_top2000_kittlevsschweintek_iscm + 1;
    }
}

> numbergreaterthan_top2000_kittlevsschweintek_iscm/1000000
0

for (count in 1:1000000) {
    if (vectorofoverlaplength_kittlevshannen_iscm[count] >= 253) {
        numbergreaterthan_top2000_kittlevsschweintek_idcm <-
        numbergreaterthan_top2000_kittlevsschweintek_idcm + 1;
    }
}

> numbergreaterthan_top2000_kittlevsschweintek_idcm/1000000
0

for (count in 1:1000000) {
    if (vectorofoverlaplength_kittlevshannen_iscm[count] >= 419) {
        numbergreaterthan_top2000_hannenvsschweintek_iscm <-
        numbergreaterthan_top2000_hannenvsschweintek_iscm + 1;
    }
}

> numbergreaterthan_top2000_hannenvsschweintek_iscm/1000000
0

for (count in 1:1000000) {
    if (vectorofoverlaplength_kittlevshannen_iscm[count] >= 478) {
        numbergreaterthan_top2000_hannenvsschweintek_idcm <-
        numbergreaterthan_top2000_hannenvsschweintek_idcm + 1;
    }
}

> numbergreaterthan_top2000_hannenvsschweintek_idcm/1000000
0
numbergreaterthan_top2000_hannenvsschweintek_idcm<-numbergreaterthan_top2000_hannenvsschweintek_idcm+1;
}

> numbergreaterthan_top2000_hannenvsschweintek_idcm/1000000
0

numbergreaterthan_top2000_hannenvstsubaki_iscm<-0
for (count in 1:1000000) {
    if (vectorofoverlaplength_kittlevshannen_iscm[count]>=593) {
        numbergreaterthan_top2000_hannenvstsubaki_iscm<-numbergreaterthan_top2000_hannenvstsubaki_iscm+1;
    }
}

> numbergreaterthan_top2000_hannenvstsubaki_iscm/1000000
0

numbergreaterthan_top2000_hannenvstsubaki_idcm<-0
for (count in 1:1000000) {
    if (vectorofoverlaplength_kittlevshannen_iscm[count]>=269) {
        numbergreaterthan_top2000_hannenvstsubaki_idcm<-numbergreaterthan_top2000_hannenvstsubaki_idcm+1;
    }
}

> numbergreaterthan_top2000_hannenvstsubaki_idcm/1000000
0

numbergreaterthan_top2000_hannenvsbarth_idcm<-0
for (count in 1:1000000) {
    if (vectorofoverlaplength_kittlevshannen_iscm[count]>=386) {
        numbergreaterthan_top2000_hannenvsbarth_idcm<-numbergreaterthan_top2000_hannenvsbarth_idcm+1;
    }
}

> numbergreaterthan_top2000_hannenvsbarth_idcm/1000000
0

numbergreaterthan_top2000_tsubakivsbarth_idcm<-0
for (count in 1:1000000) {
    if (vectorofoverlaplength_kittlevshannen_iscm[count]>=177) {
        numbergreaterthan_top2000_tsubakivsbarth_idcm<-numbergreaterthan_top2000_tsubakivsbarth_idcm+1;
    }
}

pval_top2000_tsubakivsbarth_idcm<-numbergreaterthan_top2000_tsubakivsbarth_idcm/1000000

205
> pval_top2000_tsubakivsbarth_idcm
0.593941

dumbergreaterthan_top2000_schweintekvsbarth_idcm<-0
for (count in 1:1000000) {
    if (vectorofoverlaplength_kittlevshannen_iscm[count]>=311) {
        numbergreaterthan_top2000_schweintekvsbarth_idcm<-numbergreaterthan_top2000_schweintekvsbarth_idcm+1;
    }
}
> numbergreaterthan_top2000_schweintekvsbarth_idcm/1000000
0

dumbergreaterthan_top2000_schweintevstsubaki_iscm<-0
for (count in 1:1000000) {
    if (vectorofoverlaplength_kittlevshannen_iscm[count]>=279) {
        numbergreaterthan_top2000_schweintevstsubaki_iscm<-numbergreaterthan_top2000_schweintevstsubaki_iscm+1;
    }
}
> numbergreaterthan_top2000_schweintevstsubaki_iscm/1000000
0

dumbergreaterthan_top2000_schweintevstsubaki_idcm<-0
for (count in 1:1000000) {
    if (vectorofoverlaplength_kittlevshannen_iscm[count]>=141) {
        numbergreaterthan_top2000_schweintevstsubaki_idcm<-numbergreaterthan_top2000_schweintevstsubaki_idcm+1;
    }
}
> numbergreaterthan_top2000_schweintevstsubaki_idcm/1000000
0.999476
Appendix D: Supplementary Data for

Chapter 5

Figure D.1: Plot illustrating mean nuclear area versus age (in years) in NF hearts (n = 33). LV samples from females in red, males in blue. The least-squares regression lines for relationship between age and mean nuclear area for all NF hearts (yellow), and for subset of NF subjects >= 20 years old (green) are
depicted. Asterisk indicates p-value < 0.05 for linear relationship between age and mean nuclear area (controlled for gender). Nil statistically significant changes of mean nuclear area with age (in subset of NF subjects >= 20 years) and gender were observed (p-values > 0.05 for both analyses).

**Figure D.2:** Plot illustrating mean AgNOR area / nuclear area versus age (in years) in NF hearts (n = 33). LV samples from females in red, males in blue. The least-squares regression lines for relationship between age and mean AgNOR / nuclear area for all NF hearts (yellow), and for subset of NF subjects >= 20 years old (green) are depicted. Nil statistically significant changes of mean AgNOR area / nuclear area with age, age (in subset of NF subjects >= 20 years), and gender were observed (p-values > 0.05 for all analyses).
**Figure D.3:** Plot illustrating mean AgNOR area / nuclear area in the LV from patients with end-stage HF of different aetiologies (ISCM, n = 11; HCM; n = 3, IDCM, n = 19) versus NF subjects (n = 8). Asterisk indicates statistically significant decrease (p-value < 0.05) for comparison with NF subjects.
Figure D.4: Plot illustrating mean nuclear area (in $\mu m^2$) in the LV from patients with end-stage HF (ISCM, $n = 11$; HCM, $n = 3$; IDCM, $n = 19$) versus NF subjects ($n = 8$). Nil statistically significant difference in mean nuclear area comparing end-stage ISCM, HCM, and IDCM with NF subjects (p-value $> 0.05$ for all comparisons).