GENETICS AND GENOMICS OF AORTIC FORM AND FUNCTION

Submitted to Imperial College London

for the degree of Doctor of Philosophy

by

Dr Catherine Francis

MA (Cantab) BM BCh (Oxon) MRCP (UK)

May 2019

Cardiovascular Genetics and Genomics
National Heart and Lung Institute

1. ABSTRACT

The thoracic aorta is a dynamic organ which adapts and remodels throughout life. Thoracic aortic size, shape and function are important contributors to both cardiovascular health and disease and risk of aortic disease. A complex interaction of environmental, genetic and haemodynamic factors is mediated by cells of the aortic wall.

This thesis presents aortic phenotyping, genotyping and genome-wide associations of aortic traits in a large healthy cohort of 1218 volunteers. This is the largest study to report normal parameters for healthy thoracic aortic size, shape and function derived from cardiovascular magnetic resonance imaging. Anthropometric and cardiovascular risk factors such as age, gender, body fat mass and lipid profile are identified as significant determinants of aortic phenotype. The work suggests that cardiovascular risk factors could impair normal adaptive aortic remodelling with age.

Genome-wide association studies of aortic dimensions and function identify new common variants, genes and pathways which could be important in aortic biology and cardiovascular risk. These include genes involved in cardiovascular development (eg *PCDH7* and *SON* associated with aortic root diameter), autonomic cardiovascular responses (eg GABA receptor genes associated with aortic root diameter), fibrosis (eg *ACTC1, AGTR1* associated with ascending aortic distensibility, *BAMBI* and *MYOD* associated with descending aortic distensibility) and obesity (eg *ARID5B* and *IRX3* associated with aortic pulse wave velocity and ascending aortic area respectively). Multiple regulatory pathways including TGF-ß and IGF signalling (*IGF1R, IGF2R*), are identified which are associated with aortic dimensions and function. Joint trait analysis of aortic root dimensions identifies a new genome-wide significant association with *TENM4*, a key driver of early mesodermal development, and suggestive association with *PTN*, which is functionally related and plays a key role in angiogenesis.

The primary analyses are complemented by exploratory assessment of rare genetic variation in bicuspid aortic valve (BAV) using panel sequencing in 177 patients. Rare variants might cause, or modify phenotype in BAV, but the clinical utility of panel sequencing remains poor.

A further complementary study investigates the interaction of haemodynamics with aortic cellular phenotype, using microarray assessment of aortic endothelial cell transcriptomic response to shear stress pattern. Several genes of interest in atherosclerosis and aortic disease are differentially expressed with shear stress pattern, such as *FABP4, ANGPT2, FILIP1, KIT, DCHS1*, *TGFBR3* and *LOX*.

This work yields new insights into aortic phenotype, identifies key loci which might determine aortic traits and explores the complex interdependence of genetics, haemodynamics and environmental variables in aortic biology.

2. CONTENTS

3. DECLARATION OF ORIGINALITY

This thesis and the work presented within it are my own. Many others have of course contributed to the research projects described, and I have duly acknowledged and referenced their input and invaluable contribution in the relevant chapters.

4. COPYRIGHT STATEMENT

5. ACKNOWLEDGEMENTS

I am forever indebted to my supervisors for their support and guidance: Jane Mitchell, whose wisdom about science is equalled by her wisdom about life in general; John Pepper whose dedication to improving lives for patients and to science is an inspiration; and Stuart Cook whose scientific vision, clarity of thinking and ambition is truly awesome.

To Declan O'Regan and his team - in particular, Tamara Diamond, Marina Quinlan and Ben Statton for their help with phenotyping. To Tim Dawes and Antonio de Marvao whose scientific nous, statistical knowledge, clarity of thought and general great humour benefitted me greatly.

To James Ware for his advice, encouragement and brilliant scientific mind. To Paul Barton, the unsung hero of the Cardiovascular Genetics and Genomics team, for his immense support, care and insight into life and science. To the rest of the team – especially Rachel Buchan, Nicky Whiffin, Leanne Felkin, Angharad Roberts and Roddy Walsh for help with gene sequencing, genotyping, advice and for all the fun. To Mian Asbat Ahmad for his patience and help with navigating cluster computing.

To Hannah Meyer at the EBI for her generous and patient help with the GWAS coding.

To Matina Prapa and Nada Abdulkareem for their friendly help with the BAV project.

To Hime Gashaw, Daniel Reed, Mark Paul-Clark and Maxim Freydin for feeding my cells and helping with transcriptomics.

This work was supported by a British Heart Foundation Clinical Research Fellowship. Additional support came from the NIHR Cardiovascular Biomedical Research Unit at Royal Brompton & Harefield NHS Foundation Trust.

To the volunteers and patients who generously gave their time and DNA to be part of these studies.

To Joseph Francis for ensuring life is never dull.

Most of all to my parents for their unconditional love, support, encouragement and faith in me, and for all the mangoes. I owe all this to Wendy and John Fletcher, and am forever grateful.

And finally to my greatest inspirations: my children Hannah and Benjamin – the best creations of my PhD fellowship – who teach me every day about the innate desire of human beings to learn, to enquire and explore, about joy, laughter and mischief, and about the infinite capacity of the human heart for love.

6. ABBREVIATIONS AND ACRONYMS

Abbreviations used in this thesis are listed below. An exception is gene symbols, which are too numerous to list here. Standard HGNC (Human Genome Organisation Gene Nomenclature Committee) gene symbols have been used where appropriate.

3D	3-Dimensional
4D	4-Dimensional
AA	Ascending Aorta
AAo	Ascending Aorta
ACMG	The American College of Medical Genetics and Genomics
AHA	American Heart Association
AR	Aortic Regurgitation
AS	Aortic Stenosis
AV	Aortic Valve
b-SSFP	balanced Steady-State Free Precession
BAV	Bicuspid Aortic Valve
BMI	Body Mass Index
BMP	Bone Morphogenetic Proteins
bp	base pairs – the length of a (double-stranded) DNA sequence is reported in base pairs. Standard prefixes are used to indicate larger units, such as kbp and Mbp
BP	Blood Pressure
BSA	Body Surface Area
CADD Score	Combined Annotation Dependent Depletion Score
cfPWV	carotid-femoral Pulse Wave Velocity
CHiP-Seq	Chromatin Immunoprecipitation with DNA sequencing
Chr	Chromosome
CMR	Cardiovascular Magnetic Resonance Imaging
CO(i)	Cardiac Output (indexed to body surface area)
CRISPR	Clustered Regularly Interspaced Short Palindromic Repeats

CT	Computed Tomography
DA	Descending Aorta
DAo	Descending Aorta
DBP	Diastolic Blood Pressure
dbSNP	The Single Nucleotide Polymorphism Database
DHP	Digital Heart Project
Distens	Distensibility
DNA	Deoxyribonucleic Acid
EC	Endothelial cell
ECG	Electrocardiogram
EDS	Ehlers-Danlos Syndrome
EGF	Epidermal Growth Factor
ENCODE	Encyclopedia of DNA Elements
eNOS	Endothelial Nitric Oxide Synthase
eQTL	expression Quantitative Trait Locus
ESC	European Society of Cardiology
ExAC	Exome Aggregation Consortium
Fatmass	Percentage body fat
FBN1	Fibrillin-1
FDR	False Detection Rate: a statistical correction applied to account for multiple testing
FGF	Fibroblast Growth Factor
FTAAD	Familial Thoracic Aortic Aneurysm and Dissection
GABA	Gamma-Aminobutyric Acid
GCTA	Genome-Wide Complex Trait Analysis
Genetics	The study of heredity, broken down into single units or genes
Genome-Wide Association	The process of generating statistical associations with phenotype for genetic variants across the genome
Genomics	The study of information from across the genome, often at population-scale
Genotyping	The process of identifying genetic variation in an individual genome
gnomAD	The Genome Aggregation Database

GO	Gene Ontology
GREML	Genetic Restricted Maximum Likelihood
GTEx	Genotype-Tissue Expression
GWAS	Genome-Wide Association Study
HAEC	Human Aortic Endothelial Cell
HDL	High-Density Lipoprotein
Hi-C	High-throughput sequencing of ligated fragments from chromatin conformation capture
HLHS	Hypoplastic left Heart Syndrome
HR	Heart Rate
HUVEC	Human Umbilical Vein Endothelial Cell
HW ratio	Height:Width ratio of the aortic arch
HWE	Hardy-Weinberg Equilibrium. The Hardy-Weinberg principle states that, in the absence of selection pressures, the proportion of homozygotes and heterozygotes will remain constant according to the frequency of the SNP in the general population.
IBD	Identity By Descent
ICC	Intra-Class Correlation Coefficient
Indel	Short insertions or deletions
KEGG	Kyoto Encyclopaedia of Genes and Genomes
LD	Linkage Disequilibrium
LDL	Low-Density Lipoprotein
LINC	Long Intervening Non-Coding RNA
LTBP	Latent TGF-ß Binding Protein
LV	Left Ventricle
LVCI	Left Ventricular Concentricity Index
LVEDV(i)	Left Ventricular End Diastolic Volume (indexed to body surface area)
LVEF	Left Ventricular Ejection Fraction
LVESV(i)	Left Ventricular End Systolic Volume (indexed to body surface area)
LVM(i)	Left Ventricular Mass (indexed to body surface area)
LVOT	Left Ventricular Outflow Tract
LVOT	Left Ventricular Outflow Tract

LVSV(i)	Left Ventricular Stroke Volume (indexed to body surface area)
MAF	Minor Allele Frequency
MAP	Mean Arterial Pressure
MAPK	Mitogen Activated Protein Kinase
Mendelian	a Mendelian trait is one that is controlled by a single locus, inherited in a defined pattern
MESA	Multi-Ethnic Study of Atherosclerosis study
MFS	Marfan Syndrome
MGI	Mouse Genome Informatics
MMP	Matrix Metalloproteinase
mPWV	CMR-derived Pulse Wave Velocity
MRI	Magnetic Resonance Imaging
MV	Mitral Valve
MVP	Mitral Valve Prolapse
NFATC1	Nuclear factor of activated T-cells, cytoplasmic 1
NGS	Next Generation Sequencing
NS	Non-significant
PCA	Principal Component Analysis
PCR	Polymerase Chain Reaction
Penetrance	Proportion of individuals carrying a particular variant who also express an associated phenotype
Phenotype	Observable characteristics, such as morphologic, biochemical or physiological traits or disease. A phenotype results from the gene expression as well as the influence of environmental factors and the interactions between the two
PP	Pulse Pressure
PRS	Polygenic Risk Score
PWV	Pulse Wave Velocity
QC	Quality Control
QQ plots	Quantile:Quantile plots
RA	Retinoic Acid
Ref allele	Reference allele
RN	Rank-normalised: observations are ranked and then a normal transformation is applied to the ranked data.

RNA	Ribonucleic acid
SBP	Systolic Blood Pressure
SCMR	Society of Cardiovascular Magnetic Resonance Imaging
SD	Standard Deviation
SE	Standard Error
SHF	Secondary Heart Field
SNP	Single-Nucleotide Polymorphism
SoV	Sinuses of Valsalva
STEMI	ST-Elevation Myocardial Infarction
STJ	Sino-Tubular Junction
TAA	Thoracic Aortic Aneurysm
TAAD	Thoracic Aortic Aneurysm and Dissection
Tchol	Total Cholesterol
TGF-ß	Transforming Growth Factor Beta
TGs	Triglycerides
VEGF	Vascular Endothelial Growth Factor
VENC	Velocity Encoded gradient echo imaging
VG/Vp	Trait variance explained by genotype as a proportion of the overall trait variance
VSD	Ventricular Septal Defect
VSMC	Vascular Smooth Muscle Cell
VUS	Variant of Uncertain Significance
WES	Whole Exome Sequencing
WGS	Whole Genome Sequencing

7. LIST OF FIGURES

8. LIST OF TABLES

9. AIMS OF THESIS

- To define healthy thoracic aortic form and function
- To identify common genetic variants, genes and pathways which influence aortic form and function
- To define the role of panel sequencing in bicuspid aortic valve and identify candidate rare, pathogenic variants
- To define the impact of shear stress pattern on gene expression and identify key genes which might drive the phenotypic response to haemodynamics and genetic variation

10: INTRODUCTION

"I am firmly convinced that the best book in medicine is the book of Nature, as writ large in the bodies of men" - Sir William Osler

10.1 OVERVIEW

The aorta functions as both conduit and cushion[1]. It conveys blood from the left ventricle to the rest of the body via its branches. Its distensibility allows the aorta to absorb the strain of each cardiac ejection, and its elastic recoil maintains perfusion pressures during diastole, particularly to the coronary circulation.

This elastic function of the aorta has long been recognised as vitally important. Elastic function deteriorates with vascular ageing, and this decline is an early cardiovascular risk marker which is independent of classical risk factors such as hypertension, hyperlipidaemia and diabetes[2-7]. Aortic stiffness may also contribute to microvascular dysfunction in distal circulations such as the renal and cerebral vasculature, and is associated with cognitive decline and renal dysfunction[8-11]. There is also increasing evidence that impaired aortic elastic function may be of prognostic significance in aortopathies[12, 13]. Inextricably linked to the elastic function of the aorta are other metrics of size, shape and haemodynamics – most of which are prognostic in the general population and in aortic disease[1, 14-20].

Over the past decades, a paradigm shift has occurred in our understanding of how this vital elastic function might be regulated. No longer do we view the aorta as a passive elastic tube. Instead, we have come to regard the aorta as a neglected organ in its own right; a dynamic and complex structure with a host of regulatory and adaptive roles. Cells within the aortic wall – both endothelial and vascular smooth muscle cells - are exquisitely sensitive to environmental stimuli; they sense changing haemodynamic conditions and the paracrine and endocrine milieu and adjust their phenotype accordingly[21-24]. The aorta is a perfect example of the key interface between mechanical forces and cell function, and we are just starting to learn how this mechanotransduction is vital to homeostasis within the aortic wall and maintenance of aortic function.

At the same time as this advance in understanding of aortic function, has come huge progress in the science of genetics and genomics and increasingly precise tools to measure phenotype at molecular, cellular and whole-organism levels. We therefore have unprecedented opportunities to learn about the molecular and genetic basis of aortic biology, but also face the myriad challenges of collecting, quality-assuring, interpreting and integrating "big data".

This thesis aims to bring together work on environmental and genetic variables which might contribute to differences in aortic structure and function between individuals. I will explore how common and rare genetic variation affect aortic phenotype. I will use SNP array genotyping with imputation for genome-wide association studies; panel-based sequencing for rare variant analysis and micro-array transcriptomics to assess how the haemodynamic environment might regulate gene expression. I will also use publicly available large-scale datasets such as dbSNP[25], gnomAD[26, 27], GTEx[28, 29], CADD[30], ENCODE[31], RegulomeDB[32], Haploreg[33, 34], GRASP[35], GWAS catalog[36, 37], and MGI database[38, 39] for variant annotation and functional characterisation.

I will explore what we can learn from each of these techniques and resources about aortic biology and aortic health and how this knowledge might be translated to clinical use.

10.2 THE HISTORY OF THE AORTA: FROM GROSS STRUCTURE TO MOLECULAR MECHANISMS

Modern understanding of the aorta and the importance of its elastic function has built upon many centuries of knowledge. It is possible to date the recognition of aortic aneurysm to 1500BC, when the convergence of vessels on the heart was documented in the Ebers papyrus, along with the description of aortic aneurysm as a pulsatile "*tumour*"[40]. The Greek physician, Praxagoras, born around 340BC, was amongst the first to describe the aortic structure itself: "*The largest artery has its origin in the left hollow of the heart like a trunk of all the arteries in the animal."*[41] It took almost 5 centuries, to the time of Galen, to appreciate that the aorta conveyed blood. However, his view of the circulatory system as two separate circulations through arteries and veins persisted until William Harvey's work in the early 17th century to understand aortic function within the context of the circulatory system as a whole. It is not until the last 20-30 years that a paradigm shift has occurred, from examining the macro-

structure of the aorta, to unravelling the molecular-level signalling pathways and interactions which contribute to aortic function.

10.3 AORTIC STRUCTURE AND FUNCTION

10.3.1 Basic functions of the aorta

The most obvious function of the aorta is to deliver blood from the heart to the rest of the body. The thoracic aorta also acts as a "Windkessel" – a term originally used to describe an air chamber in 18th century German fire engines which converted pulsatile into continuous flow. The term "Windkessel" was used by Otto Frank in the late 19th century to describe the capacitance function of the aorta. Whilst it is now recognised as a simplistic model of aortic haemodynamics[42], it is a useful illustration of the principles at play. As the left ventricle ejects a bolus of blood during systole, the proximal aorta distends to accommodate the increased volume. The elastic recoil of the aortic wall then comes into effect during ventricular diastole. This buffering capacity firstly protects the peripheral circulation from high pulsatile pressures, and also maintains peripheral and coronary blood flow during diastole, smoothing the flow profile to distal circulations. In addition, the elasticity of the aorta allows the propagating pressure wave to travel at steady speeds. This ensures that wave reflections from aortic branch points and diffuse reflections from stiffer distal regions augment pressure during diastole (rather than earlier during systole in a stiffened aorta), again contributing to the maintenance of diastolic coronary perfusion. The ventriculo-aortic coupling additionally regulates the afterload of the left ventricle, ensuring that it can function efficiently to maintain cardiac output. The structure of the aorta at both a macroscopic and microscopic level, is superbly adapted to fulfil both the joint conductance and capacitance functions of the aorta[43, 44].

10.3.2 Anatomy of the thoracic aorta and aortic valve

As the aorta leaves the heart at the aortic valve, it passes cranially, twisting anteriorly and to the right around the pulmonary trunk as it does so, and loops from anterior to posterior over the left pulmonary artery to form the aortic arch posteriorly and caudally to form the descending thoracic aorta[45] (see Figure 10.1). The aortic valve is usually formed of 3 leaflets, or cusps, giving the valve the classic "Mercedes-Benz" appearance on imaging. The valve

opens and closes passively with the cardiac cycle, opening as the left ventricular pressure rises during ventricular systole to allow ejection of blood from the heart, and closing during ventricular relaxation to prevent back-flow of blood during diastole. The aortic root sits just above the valve, and is formed of the portion of aorta between the aortic valve and the tubular portion of the ascending aorta. This root bulges outwards to form three sinuses of Valsalva. Two of these give rise to the main coronary arteries – termed the left and right coronary sinuses; the third is the non-coronary sinus[46].

FIGURE 10.1: Anatomy of the thoracic aorta

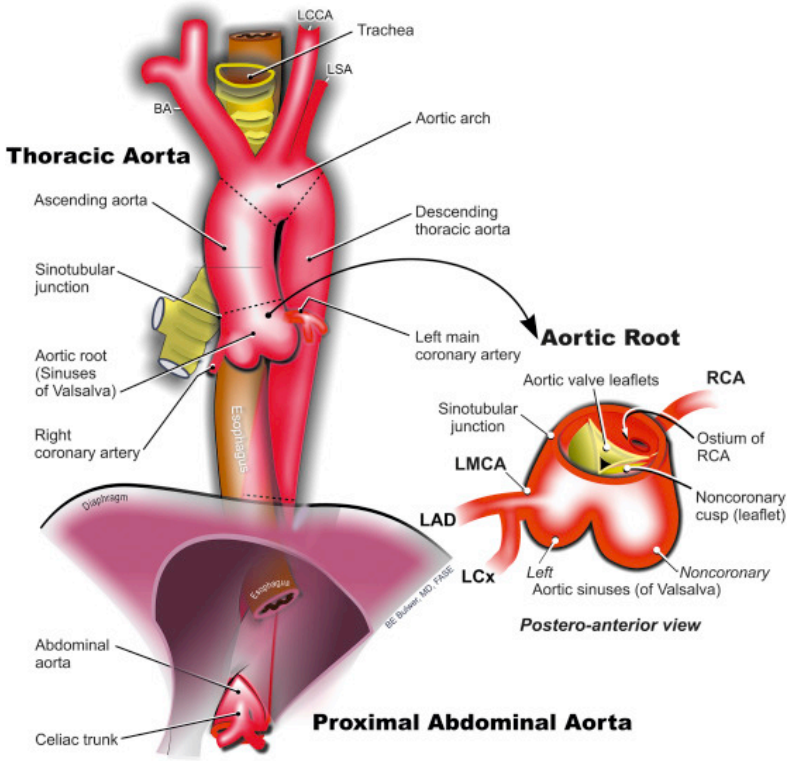


Diagram from ScienceDirect.com, showing thoracic aortic anatomy. BA: Brachiocephalic Artery; LCCA: Left Common Carotid Artery; LSA: Left Subclavian Artery; RCA: Right Coronary Artery; LMCA: Left Main Coronary Artery; LCx: Circumflex Artery; LAD: Left Anterior Descending Artery

The anatomic demarcation between the left ventricle (LV) and the aorta is not quite as clear-cut as it might appear from classical diagrams; there is fibrous continuity between the fibrous core of the aortic valve and the mitral valve for example, and the nadir of the insertion points of the aortic cusps leads to the inclusion of some histological myocardium into anatomically aortic root[46].

This anatomical coupling between aorta and left ventricle also reflects the importance of the functional relationship between the two structures. For maximum efficiency of left ventricular work, the LV elastance and aortic elastance (reflecting volume changes for given pressures) must be closely matched. In disease states, ventricular and aortic elastance or compliance can become mismatched, resulting in increased amounts of "wasted" work by the left ventricle. Thus, increasing aortic stiffness (a shallower pressure:volume relationship) requires generation of higher LV end systolic pressures to eject a similar volume of blood. This in turn can cause LV concentric remodelling and, over time, impaired LV function[18, 47, 48].

10.3.3 Cellular and molecular structure of the aorta

Histologically, the aorta is made up of 3 distinct layers: the tunica intima, tunica media and adventitia. The tunica intima lines the lumen of the aorta and is comprised primarily of endothelial cells, basement membrane and a thin layer of connective tissue, bounded by the internal elastic lamina. The tunica media is the layer which contributes most to the elastic function of the aorta, and is the site where many pathologies disrupt the normal aortic microstructure[49]. It contains elastic fibres, bound together in sheets or elastic laminae, arranged circumferentially, and connected diagonally between layers with a herringbone pattern of thinner elastic fibres. In between the elastic laminae are not only thinner elastic fibres, but also collagen fibres, layers of proteoglycans, glycosaminoglycans and vascular smooth muscle cells. The outermost layer of the aortic wall, the tunica adventitia, is bounded by the external elastic lamina. This outer layer consists of sparser collagenous extracellular matrix, fibroblasts and fibrocytes and the vasa vasorum - small blood vessels which supply the outer layers of the aorta with nutrients and oxygen[50].

The primary non-cellular components of the aortic tunica media are elastic fibres, collagen, proteoglycans and glycosaminoglycans[49]. Elastic fibres are formed into sheets, or lamellae, oriented circumferentially, and these are inter-connected by thinner elastic fibres and vascular smooth muscle cells. The main component of these elastic fibres is elastin. This is a highly elastic polymer, formed from extensive post-translational cross-linking of tropoelastin monomers. Elastin is crucial for the normal functioning of the aorta, providing the mechanical properties necessary for the Windkessel effect[44]. Many aortic pathologies – both genetic and acquired – share a final common pathway of degradation of elastic fibres and elastin in the

aortic media[51-53]. In aortopathies such as that seen in Marfan or Loeys Dietz syndromes, there are classical hallmark changes in the aortic media which include elastin fibre degradation[53, 54]. The products of elastolysis as well as the soluble elastin precursor, tropoelastin, also have important signalling roles in maintaining aortic wall homeostasis and determining vascular smooth muscle cell phenotype[55]. There has been much interest in the roles of elastin breakdown products as a biomarker for cardiovascular diseases and, more controversially, as a potential modifier or driver of worsening vascular function[55, 56]. Loss-of-function mutations of the elastin gene (*ELN*) can cause significant vascular pathology; most notably from an aortic standpoint, supravalvular aortic stenosis[44, 57-59]. Elastin is also crucial for cardiovascular development[60, 61].

Other key molecules involved in elastic fibre synthesis or structure include the lysyl oxidase family (notably *LOXL2*[62], *LOX*[63]) and components of the microfibrillar scaffold which forms the other major part of elastic fibres. This consists of nearly 30 different proteins; major contributors including fibrillins 1 and 2 and fibulins 4 and 5. Many of these molecules have also been implicated in genetic aortic disease[64] (for example, *FBN1* mutations cause Marfan syndrome[65, 66], *FBLN4* (aka *EFEMP2*) variants cause a diffuse arteriopathy[67, 68] and *LOX1* mutations cause isolated familial thoracic aortic aneurysm[63]).

Whilst it is mostly the elastic fibres which determine the circumferential distensibility of the aorta at physiological loads, the collagen fibres between the elastic laminae confer tensile strength at higher loads[44]. In the aortic wall, collagens III and I predominate. These collagen fibres protect the integrity of the aortic wall against over-deformation.

The passive mechanical properties of the aorta are determined mainly by the elastic fibre and collagen content of the aortic media. Vascular smooth muscle cells (VSMCs) are vital for the synthesis of these molecules and also for the autonomic determination of aortic tone. They are physically linked to the elastic fibres via microfibrils, which attach directly to dense plaques within the VSMC wall[64]. VSMCs dynamically regulate the balance of proteolysis and synthesis within the aortic wall. Vascular smooth muscle cell survival depends on adhesion and tension within the cell; loss of contact with the extracellular matrix leads to anoikis, or apoptosis of the vascular smooth muscle cell. VSMCs also play a key role in mechanotransduction, sensing stretch and strain within the aortic wall. Components of the

contractile apparatus within VSMCs are vital to maintain aortic integrity; mutations in genes encoding these proteins (e.g. *ACTA2*[69], *MYH11*[70, 71,72], *MYLK*[73], *PRKG1*[74]) cause heritable aortopathies. Phenotypic switching to a secretory, more pro-inflammatory phenotype from a more quiescent contractile phenotype, has also been implicated in many vascular disorders including atherosclerosis and several aortopathies[75]. This phenotypic switching may occur in response to mechanical or local biochemical stimulus.

Endothelial cells which line the aortic lumen also have a crucial role in regulation of aortic tone and VSMC phenotype. Local mechanical interactions between the cell types, and EC-driven biochemical signalling through signalling pathways such as TGF-ß, nitric oxide and angiotensin can drive phenotypic switching, proliferation and migration of VSMCs[76]. ECs themselves dynamically regulate the permeability of the aortic wall to blood-borne proteins, inflammatory cells and signalling molecules[77, 78]. ECs are uniquely placed to sense and respond to changing haemodynamic conditions, and alter their phenotype and that of the cells around them[77, 79]. This heterotypic signalling directly influences the contractile status of VSMCs. Aberrant interaction between ECs and VSMCs contributes to pathogenesis of atherosclerosis, where endothelial dysfunction is very well-described to initiate a cycle of changes in vascular reactivity and composition of the vessel wall[76, 79]. In addition to the classical paradigms of atherosclerosis, there is increasing interest in the role of ECs in conditions such as bicuspid aortic valve-related aortopathy (see section 10.11.3). Patients with BAV exhibit widespread endothelial dysfunction, the cause of which is as yet undefined[80]. The ability of aortic ECs to sense shear stress places them in a key role in mechanotransduction in the aorta, and their ability to dictate VSMC phenotype makes them a key cell of interest both for understanding pathogenesis and potentially for directing treatment. The role of endothelial cells in the aorta is discussed further in Chapter 16.

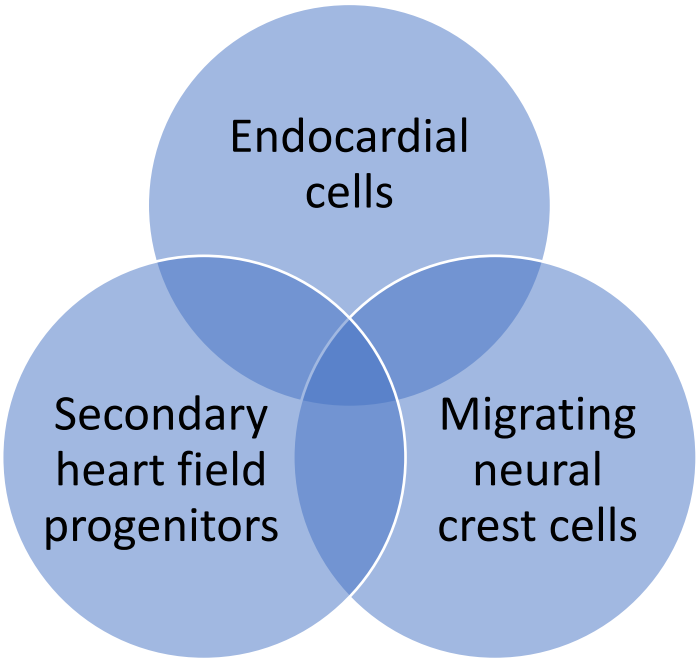
10.4 AORTIC VALVE AND THORACIC AORTIC DEVELOPMENT

In order to understand the different genetic influences on aortic form and function, it is important to understand how the aortic structure develops. It is particularly important for clinical translation to understand which genetic influences occur solely during development, and which represent ongoing homeostatic processes. The former are not likely to be amenable to drug intervention; whereas the latter might be modifiable targets.

10.4.1 Aortic valve development

The primitive linear heart tube forms by around 3 weeks in the human embryo, expressing cardiac-specific genes such as *NKX2.5* and *GATA4*. This is followed at weeks 4-5 by looping of the cardiac tube[81]. At around this time, the primary myocardium secretes a gelatinous matrix, called cardiac jelly, forming swellings at the atrioventricular junction and outflow tract[82]. These swellings form the endocardial cushions. The endocardial cells lining the inside surface of the heart tube at these points undergo endothelial-mesenchymal transformation[81]; these cells proliferate and grow inwards, and are joined by both secondary heart field progenitors (positioned in the pharyngeal mesoderm) and migrating cardiac neural crest cells. Contact between the two endocardial cushions results in fusion, followed by formation of the thin valve leaflets. The neural crest cells create a crescent-shaped septal wedge which divides the outflow tract into aortic and pulmonary components[83]. The interaction with neural crest cells and with secondary heart field cells is crucial for normal outflow tract development, and it is hypothesised that defects at different stages of valve and outflow tract development may be in part responsible for the different phenotypes (differing valve morphologies and aortic dilatation at different levels) observed in bicuspid aortic valve (BAV; see section 10.11.3)[84]. New evidence this year also implicates lineage-specific responses to aberrant TGF-ß signalling in pathogenesis of aortic complications in Loeys-Dietz syndrome[85] (see Section 10.11.1.1 below).

FIGURE 10.2: The 3 major cell lineages which contribute to valvulogenesis & outflow tract development



The 3 cell lineages thought to contribute to valvulogenesis & outflow tract development; defects in the function of each may contribute to different phenotypes seen in BAV and other proximal aortic diseases

Multiple different signalling pathways are involved in the complex process of valve formation; over a hundred genes have been identified which contribute to the process of endothelial-mesenchymal transformation alone, with most research focussing on chick and mouse embryo development[81, 86-90]. These signalling pathways include those controlled by: Vascular Endothelial Growth Factor (VEGF), Transforming Growth Factor-β (TGF- β), Epidermal Growth Factor (EGF), Notch1, Wnt/ß-catenin, Nuclear factor of activated T-cells, cytoplasmic 1 (NFATC1), Mitogen Activated Protein Kinase (MAPK),Fibroblast Growth Factors (FGFs), Bone Morphogenetic Proteins (BMPs), along with transcription factors, including Twist1, Tbx20, Msx1/2, and Sox9[81, 82, 84, 91, 92]. These are active during early valvulogenesis and maturation of the valve leaflets. The exquisite temporal distribution of these factors, along with cross-talk and interaction between the pathways allows normal valve development. It is no surprise, therefore, that small perturbations in levels of key factors can lead to significant phenotypic aberrations, and these pathways have become the subjects of investigation in BAV.

10.4.2 Aortic arch development

In parallel, the structures which form the aortic arch and its branches develop. The aortic arch itself is formed largely from the left 4th pharyngeal arch. Between the second and seventh weeks of gestation, paired branchial (or pharyngeal) arches develop symmetrically from the mesoderm sequentially from the most cranial to the most caudal. They undergo a complex process of regression and remodelling in turn, leaving the asymmetric arrangement present in the mature cardiovascular system. Each of these arches gives rise to a vessel linking the paired dorsal and ventral aortae on either side of the midline (see Figure 10.3). There are 6 paired arches in humans; the fifth develops only transiently. The fourth arch persists on the left to form the definitive aortic arch; the right fourth arch regresses to form the brachiocephalic trunk and proximal right subclavian artery. The sixth arch forms the distal pulmonary artery and ductus arteriosus. The paired ventral aortae fuse to form the ascending aorta down to the developing outflow tract, and the left dorsal aorta forms both the descending aorta and the internal carotid artery, regressing only between the 3rd and 4th pharyngeal arches. Initial pharyngeal arch formation is driven by molecules such as the transcription factor Tbx1, retinoic acid (RA), Fibroblast Growth Factor (FGF) and Wnt signalling pathways and laterality and specific differentiation is driven by other overlapping pathways such as those driven by endothelin1 / Hox transcription factors / Hand 1 & 2 (Heart And Neural crest Derivatives expressed proteins)[93]. The crucially-timed sequence of growth and differentiation is highly sensitive to epigenetic or genetic insults, with aberrant aortic arch morphogenesis apparent in around 20% of all patients with congenital heart disease[94].

FIGURE 10.3: Development and remodelling of the pharyngeal arches to form the aortic arch and related structures

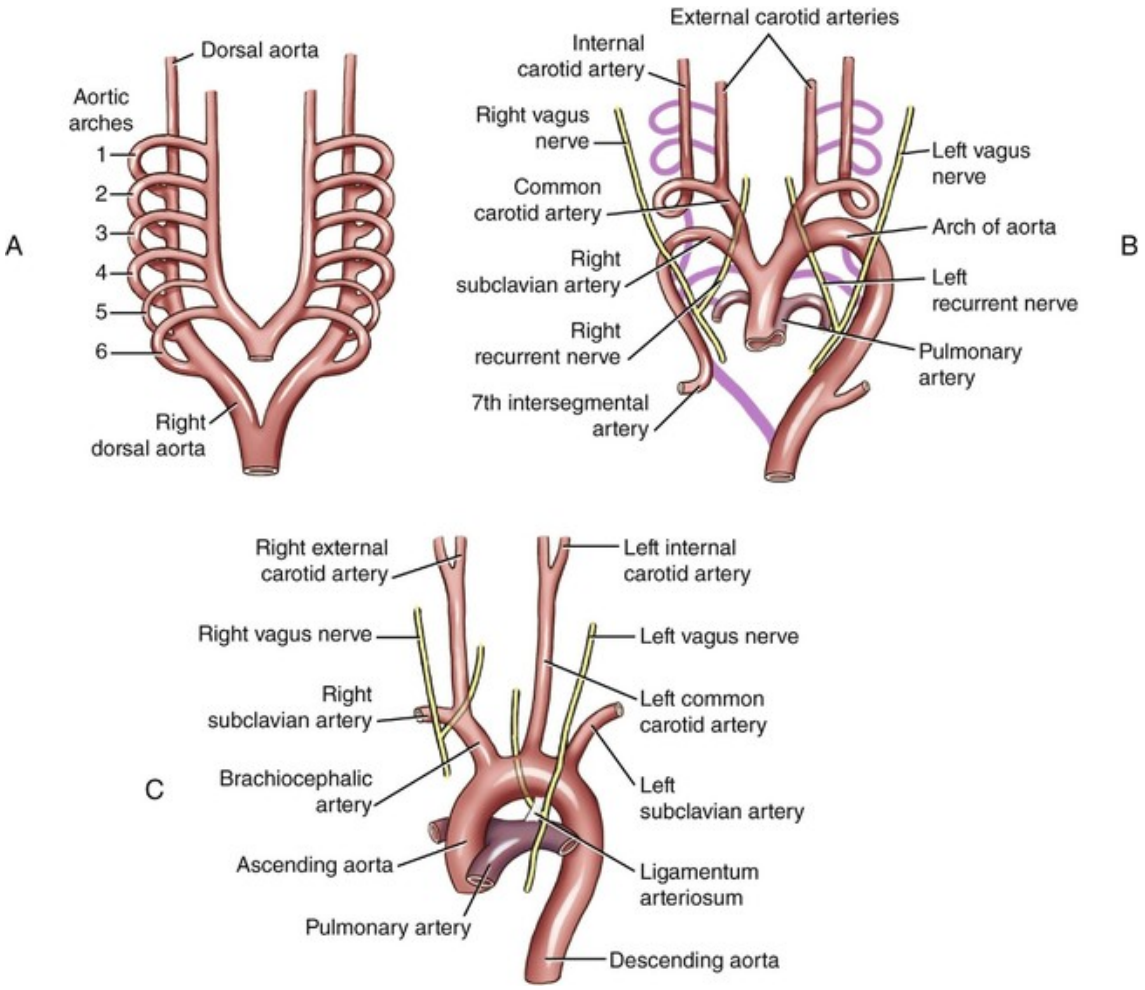


Figure (from RadiologyKey.com95) showing the pattern of formation and sequential regression /remodelling of the pharyngeal arches to contribute to aortic structure. A: The symmetrical aortic arches which form sequentially. B: The structure of the arches after transformation. Purple lines represent regressed structures. C: The adult aortic arch structure.

So different portions of the thoracic aorta are derived from the secondary heart field, from the fused paired ventral aortae, from the 4th pharyngeal arch and from the left dorsal aorta. Disruptions to this complex developmental pattern explain many of the anatomical variants and anomalies seen in the aortic arch and its branches. For example, right-sided aorta, seen in approximately 0.1% of the population, arises due to abnormal persistence of the right dorsal aorta rather than the left dorsal aorta.

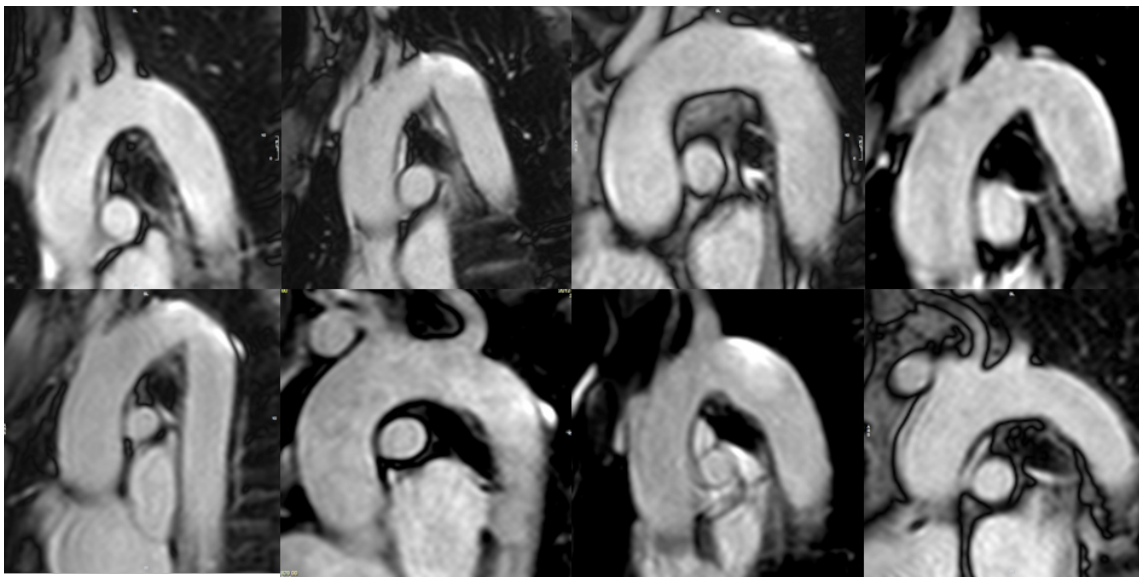
Similarly important in determining clinical presentation are the origins of the cells which form the aortic wall. Fate-mapping studies have demonstrated that the vascular smooth muscle

cells which form the aortic root, aortic arch and descending aorta, derive from different embryological tissues: namely the secondary heart field, neural crest and presomitic mesoderm, respectively[96]. This is not merely of academic interest: the boundaries between these different cell populations are very strictly demarcated, and these boundaries have higher risk of dissecting[97, 98].

10.5 VARIABILITY IN AORTIC SHAPE, SIZE AND FUNCTION

Whilst the developmental processes described above give rise to the same basic blueprint for aortic structure, there is significant inter-individual variability in shape and function. With advanced cardiovascular imaging has come increased understanding of the differences between individuals' aortas and the implications that this might have for blood flow and for the risk of cardiovascular and aortic disease. Figure 10.4, below, serves to illustrate the wide variation in gross aortic arch morphology, even amongst healthy volunteers. The dimensions, taper, angulation, branching pattern and length of the aortic arch are all well-recognised to have significant implications for blood flow through the aorta and to expose the aortic wall to different patterns of strain and shear stress[1, 17, 18]. Both during embryological development and throughout adult life, these haemodynamic forces define and refine the structural and elastic properties of the aortic wall itself[22, 99, 100]. This mechanotransduction, and the adaptation or maladaptation of the aortic wall to the prevailing haemodynamic conditions, underlies the normal cardiovascular ageing process, with the aorta becoming progressively less elastic throughout life[1]. In keeping with the concept of the aorta as a dynamic organ, changes in aortic morphology are also seen with ageing – namely an elongation of the thoracic aorta with concomitant reduction in aortic arch angulation and increased arch taper[18].

FIGURE 10.4: Individual aortic arch shapes in healthy volunteers



Cardiovascular magnetic resonance images showing the aortic arch in the classic "hockey-stick" view in 8 illustrative healthy volunteers. Note the marked difference in shape, angulation and branching of the aortic arch.

10.6 THE SIGNIFICANCE OF AORTIC TRAITS

10.6.1 The aorta in ageing

Across the whole cardiovascular system, function and morphology alter with ageing[1, 18, 101, 102]. This occurs at both a macroscopic and a cellular level. In the aorta, the changes with age are both reflective of, and contributory to, the ageing of the cardiovascular system as a whole, mirroring the endothelial dysfunction and the inflammatory activation seen in more athero-prone regions of the vascular tree[1, 103-105]. The aorta becomes stiffer, meaning that for a given stroke volume, the systolic pressure generated by the left ventricle must be higher. This impairs the normal efficiency of ventriculo-aortic coupling, increasing the work of the left ventricle during ejection[106]. Increased stiffness also increases the speed of travel of the pulse wave, meaning that reflected waves return early. This further augments the systolic pressure, reduces stroke volume, and reduces diastolic perfusion of the coronary circulation. The increased stiffness and therefore reduction in buffering capacity, exposes the peripheral circulation to much higher pulse pressure and shear stress, resulting in endothelial dysfunction and higher risk of atherosclerosis as well as end-organ damage[8, 107-109]. The

increasing stiffness with age arises from a relative increase in the collagen content of the aorta (the absolute amount decreases), coupled with a degradation of the elastic fibres and decreased cellularity[43, 44, 51]. In addition, the matrix proteins themselves undergo conformational and biochemical changes which alter their passive mechanical properties. For example, elastic lamellae become sparser, and fragment, and elastic fibres can become calcified, whilst collagen fibres acquire progressively more cross-links, increasing their stiffness. The causes of this remodelling have been the subject of much research.

Elastin production occurs at very low levels in adults (it is almost undetectable in mice >3weeks old)[110]. Elastin has a half-life of around 40 years, and, like all elastic materials, it is fatigable. Therefore the mechanical strain from repeated cycles of stretching may cause material fatigue. In vitro experiments have demonstrated elastin fragmentation and rupture after repetitive cyclical stretching[43], and shown that this "fatigue failure" occurs earlier with higher degrees of stretch. This concept provides neat explanations for the widely demonstrated relationship of heart rate with cardiovascular outcomes – the thesis being that the more times the elastin is stretched, the earlier it will fatigue and therefore the earlier cardiovascular "ageing" will be manifest. It additionally explains the acceleration of cardiovascular ageing with systemic hypertension – the additional stretch causing earlier fatigue. Fragmentation and failure of the elastic fibres eventually results in more and more of the haemodynamic load falling on the collagen fibres – structures which are 100-fold stiffer than elastin. A cycle of increased stiffness leading to higher systolic pressures to maintain stroke volume, causes a rapidly worsening spiral of aortic stiffness and hypertension. This places increased workload on the left ventricle and can lead to cardiac failure[43].

The lack of elastin synthesis in the adult also hints at the vital importance of vascular development in determining the functional characteristics of arteries later in life. It has long been recognised that the developing vascular system is exquisitely sensitive to flow and haemodynamics. There is now good evidence that perturbations in flow during foetal development can result in lower-than-normal elastin content of the aortic wall – an abnormality which will significantly influence the aortic stiffness later in life[111]. The hypothesis that adult-onset diseases such as cardiovascular disease and hypertension might have their origins in early development and foetal life is borne out by a number of epidemiological and

animal studies[112]. Particularly in the cardiovascular system, it seems intuitive that there should be an interaction between very early flow and vascular development; the effects of which might persist into adult life and ageing.

The process of elastin fragmentation can be accelerated by multiple factors including oxidative stress and inflammation, as well as exposure to advanced glycation end-products. Indeed, a cycle of matrix degradation, inflammatory pathway upregulation from the degradation products, and induction of further protease activity is hypothesised to underlie an accelerated process of vascular ageing – called "inflammageing"[113]. The normal tight regulation of protease/anti-protease activity in the aortic media can become imbalanced – in ageing, there is a gradual increase of certain matrix metalloproteinases (MMPs) – for example, MMP2. However, in pathology, the balance is shifted so that there is a reduction in MMP activity, and there is overproduction of matrix proteins, with intimal and medial hypertrophy. Conversely, in thoracic aortic aneurysms, there is an increase in MMP activity (MMPs 2,9 and 12 being particularly implicated), and subsequent medial degeneration[113, 114].

The gradual stiffening with age exposes endothelial cells to altered shear stress patterns. This results in phenotypic modifications and endothelial dysfunction – pathways discussed in more detail in Chapter 16. In the more distal resistance vessels, this endothelial cell response causes a reduction in flow-mediated vasodilatation, and local signalling leads to an increase in VSMC tone. This peripheral vasoconstriction augments the increase in arterial pressure, further exacerbating the cycle of arterial stiffness and hypertension. In the aorta, this increase in VSMC tone may not play a large direct role in determining aortic stiffness. However, it does cause phenotypic switching of VSMCs, from a quiescent state to a synthetic and proliferative phenotype, resulting in medial remodelling. ECs and VSMCs thus play tightly interlinked roles in moderating and mediating aortic stiffness, and dysfunction or altered signalling between the two cell types is an integral part of aortic and vascular ageing.

The rate at which these age-related changes occur differs between individuals. These inter-individual differences are influenced by classical risk factors such as diabetes or hypertension, as well as by developmental and genetic background[43]. At present, there are few treatments which directly target aortic stiffness – largely because the mechanisms are not well-described.

In addition to the functional aortic changes which occur with advancing age, gross morphological changes appear – perhaps largely as an adaptive response to the increased stiffness. The thoracic aorta lengthens, unfolds and dilates, allowing it to accommodate more volume and therefore offset some of the loss of buffering capacity resulting from the loss of elasticity with age. The lengthening and dilatation occurs predominantly in the ascending aorta[115], and is associated with changes in LV mass and concentric LV remodelling[18]. This "unfolding" of the aorta causes a widening of the aortic arch, and increases the storage capacity of the proximal ascending aorta to compensate for the loss of elastic function. These morphological changes with age therefore diminish the efficiency of cardiac output. Indeed, aortic root dilatation has been linked with risk of incident heart failure in older adults in the Framingham Heart Study[16, 116], suggesting that the loss of efficient ventriculo-aortic coupling, along with the structural adaptations to this loss, might place additional stress on an ageing left ventricle.

10.6.2 Aortic elastic function as a marker of cardiovascular risk

There is strong evidence that aortic traits are important predictors of adverse cardiovascular outcomes, both in the general population[2, 5, 117, 118] and in specific aortic diseases[36,74,136 13, 119, 120]. Aortic elastic function, as measured by pulse wave velocity (PWV; see section 10.7 below), or by local aortic distensibility, is a marker of cardiovascular risk and vascular ageing. Importantly, aortic stiffness exerts its effects on cardiovascular risk independently of known atherosclerotic risk factors[2, 118, 121]. This implies that there are novel mechanisms of vascular risk that are, thus far, poorly understood. It also suggests that aortic elasticity might itself be a valid therapeutic target.

Most population studies have used carotid-femoral PWV (cfPWV) as a marker for aortic elasticity – due to its ease of measurement in larger cohorts (see section 10.7). The Framingham Heart Study, showed a 48% increased risk of incident cardiovascular disease per standard deviation increase in cfPWV[122]. Many others have shown increased risk of mortality, renal impairment, or cardiovascular risk with increasing cfPWV in a range of ages and populations[4, 123-126]. Studies examining PWV derived from cardiovascular magnetic resonance imaging (mPWV; see section 10.7 below) in relation to cardiovascular risk have tended to focus on higher-risk populations in whom imaging is mandated for follow-up.

This relationship between aortic stiffness and cardiovascular risk is particularly apparent in these higher-risk populations, such as patients with end-stage renal failure (ESRF). In this population, a 1m/s increase in cfPWV was associated with a 39% increase in cardiovascular mortality[127]. Similar associations between ascending aortic distensibility and mortality have also been found in patients with ESRF[128]. A similarly high-risk population can be found in patients after ST-elevation myocardial infarction (STEMI). After STEMI, an increase in aortic stiffness, as measured by mPWV, was a significant independent predictor of major adverse cardiac and cerebrovascular events[129], and its addition to standard risk assessment significantly improved early post-STEMI risk stratification.

The high-risk ESRF population has also provided some evidence to suggest that modification of aortic stiffness could improve prognosis – patients whose cfPWV failed to improve following treatment of hypertension had a hazard ratio for all-cause mortality of 2.59 [95% confidence interval (CI) 1.51–4.43] compared to those whose cfPWV did improve[127, 130]. Of course, this may reflect differences in underlying aetiology or additional factors affecting both cfPWV and mortality. However, it is tempting to suppose that attempting to identify treatments which could reduce aortic stiffness might be of benefit, particularly in high-risk populations.

However, two large studies have examined the impact of mPWV on cardiovascular risk in a general population. The first of these, the Dallas Heart Study[131], found no significant predictive effect of any MRI-derived measure of aortic stiffness on total incident cardiovascular events. It should be noted that this was a relatively healthy population, with a mean age of 44, and excluded those with previous cardiovascular disease so the event rate was very low. Despite this, mPWV was significantly associated with non-fatal extra-cardiac vascular events, and aortic distensibility with non-fatal cardiac events. mPWV did modestly improve risk scoring when incorporated into a model with the Framingham Risk Score (C statistic 0.771 vs. 0.755, p = 0.05). The second of these, the Multi-Ethnic Study of Atherosclerosis (MESA)[6], which recruited older participants (>45 years), examined the impact of mPWV on risk of incident cardiovascular events. Here, the study found a statistically significant impact of mPWV on CV risk in patients between 45 and 54 years of age, but not in the older age group. The lack of such a strong association between mPWV and cardiovascular

risk in the elderly population may be explained by a number of different considerations. Firstly, different mechanisms may account for cardiovascular risk in the younger versus the older age group. Secondly, the study may be underpowered, and thirdly, perhaps the elastic properties of the aortic arch alone are simply not as strongly associated with cardiovascular risk in the elderly as either the proximal ascending aorta (assessed with distensibility measurements) or the carotid-femoral arterial tree (assessed with cfPWV). These different measures may also contribute differentially to various components of cardiovascular risk; the Dallas Heart Study suggests that aortic distensibility may be a better predictor of non-fatal cardiovascular events, whereas PWV may be a better predictor of stroke. Anatomically speaking, this conclusion makes sense, although this distinction has not been replicated in many studies, and as yet, mechanisms underlying this observation seem elusive.

Variation in aortic morphology is also important in a range of diseases. Chuang et al found that increased arch width was an independent predictor of cardiovascular events in the Framingham Heart Study Offspring cohort[117], with increased aortic width acting as an independent predictor of cardiovascular events. Increased angle of arch curvature and increased aortic taper is seen in patients with so-called “spontaneous” aortic dissection compared with those patients with traumatic dissections[132]. Numerous studies demonstrate that variations in morphology can dramatically influence the forces to which the aortic wall is exposed, and therefore may contribute to the site and risk of dissection. The branching pattern of the aortic arch may also be linked with increased cardiovascular risk and risk of aortic dissection[133].

10.6.3 Aortic traits and local aortic risk – thoracic aortic aneurysm and dissection

In Marfan syndrome, local distensibility has been identified as an independent predictor of progressive descending aortic dilatation[134] and there is good evidence for aortic stiffness being an early marker of aortic disease in patients with MYH11 variants[119] or with Marfan syndrome[13]. Data from patients with non-syndromic inherited aortopathies also strongly associates higher mPWV with larger aortic diameters[135], suggesting that aortic stiffness may be of prognostic relevance.

Whether a higher PWV or lower distensibility in the context of a genetic aortopathy confers a higher risk of dissection remains unproven, although haemodynamic modelling would suggest this could be the case.

10.7 MEASURING AORTIC TRAITS

Arguably, our ability to derive accurate phenotypes from populations of interest has lagged behind advances in genomics and cellular biology. Echocardiography was, until recently, the most widely used cardiovascular imaging technique. However, its application to aortic arch imaging in particular is greatly limited by the available acoustic windows, and by geometric assumptions to calculate volumes and masses.

For aortic imaging, cross-sectional imaging such as CT (Computed Tomography) or MRI has provided much more detailed pictures. Not only are they unlimited by narrow acoustic windows, but the already well-defined boundaries between blood pool and tissues can be accentuated by contrast agents. The ability to reconstruct individual aortas in 3 dimensions has given new insights into the morphology of the aortic arch. However, CT involves significant radiation exposure, which limits its use both in serial imaging and in healthy population studies. Cardiovascular MRI, on the other hand, is emerging as a precise, reproducible and powerful method to interrogate structures and assess flow within the aorta and other areas of the vasculature.

10.7.1 Aortic MRI

Magnetic resonance imaging allows integration of structural and functional measurement. Aortic form and function can be assessed with different sequences. Structural characteristics such as cross-sectional area can be assessed using standard 2D or 3D imaging, using balanced steady-state free precession (b-SSFP) sequences, aligned carefully to the anatomical landmarks obtained in scout images or previous sequences. These enable highly replicable measurements to be undertaken at pre-specified anatomical locations and orientations. MRI has the benefit of also acquiring functional information. Cine imaging can be used to track the motion of structures within the plane of the imaging, and velocity encoded gradient echo

imaging (VENC; also called phase contrast (PC) imaging) allows through-plane velocity of each point in the image to be determined. This is achieved by acquiring two phase contrast images at different flow sensitivities, and then subtracting the two resulting images for quantitative assessment. This process is repeated at multiple phases of the cardiac cycle, to create a velocity cine. Care must be taken to align these planes of imaging exactly perpendicular to blood flow; then detailed information about the velocity of blood flow at particular anatomical locations can be derived. Thus, transit time of the pulse wave can be assessed from the temporal shift between the two velocity waveforms acquired at the ascending and descending aorta, with the path length usually obtained from a direct measurement of the aortic midline in the "hockey stick" sagittal oblique view of the aortic arch. Path length is divided by transit time for calculation of the aortic arch PWV.

One of the major limitations of this technique is the temporal resolution, which is typically around 20ms but which can be improved to around 10ms by interpolation. However, this is sufficient to discriminate meaningful differences in pulse wave velocity, and this technique provides PWV measurements in concordance with the gold standard of invasively measured PWV[136].

10.7.2 Measurements of aortic elastic function

With the understanding of the importance of aortic elastic function has come a plethora of ways to measure this. It is important to note that the different techniques for measuring aortic "elasticity" in fact measure very different components of aortic elastic function, as summarised in Table 10.1 below.

TABLE 10.1: different aortic elastic function phenotypes

PHENOTYPE	ANATOMICAL SITES ASSESSED	COMPONENT OF ELASTIC FUNCTION ASSESSED
Distensibility	Single plane	Relative change in area per unit of pressure (inverse of elastic modulus)
MRI-derived Pulse wave velocity (mPWV)	Whole arch	Integration of stiffness of whole aortic arch
Carotid-femoral pulse wave velocity (cfPWV)	Many components of arterial tree from carotid to femoral arteries	Integration of stiffness of carotid artery, aortic arch, descending aorta, abdominal aorta, iliac artery and femoral artery
Compliance	Any pre-defined length of aorta	Change in aortic volume for a given change in pressure: highly dependent on starting diameter
Augmentation Index	Whole arch and vessels to sites of wave reflection	Reflection of forward pulse wave from branch points etc. Related to pulse wave velocity

All of these varied measures can be associated with cardiovascular outcomes. In the Dallas Heart study[131], MRI-derived aortic arch pulse wave velocity, ascending aortic distensibility and total arterial compliance (left ventricular stroke volume divided by average pulse pressure) were all associated with incident cardiovascular events in univariate analysis, and with subgroups of cardiovascular events by multivariate analysis. As discussed in chapter 12, the use of multivariate analysis including blood pressure metrics must be interpreted with caution, as there seems to be a bidirectional causal relationship between blood pressure (particularly systolic and pulse pressure) and measures of aortic elastic function – changes in aortic distensibility will lead to compensatory increase in systolic and pulse pressure, and over time, increased blood pressure can cause the aortic remodelling which increases stiffness.

Distensibility is widely used to measure local aortic elastic function. This simply refers to the relative change in lumen area divided by the pulse pressure. It is, in theory, relatively straightforward to measure using any cross-sectional imaging modality. However, obtaining an accurate central pulse pressure at the exact timing of the imaging study is much more challenging. Therefore, studies have tended to use peripherally-measured blood pressure at a time point as close to the imaging acquisition as feasible to calculate distensibility. This limits the accuracy of the technique when applied at large scale. Nevertheless, studies of

distensibility have clearly demonstrated the predictive power of this measurement for assessment of cardiovascular risk. Over 7.2 years of follow-up, the Multi-Ethnic Study of Atherosclerosis[5] found a hazard ratio for all-cause mortality of 2.7 (p=0.008) for the lowest versus the highest quintile of ascending aortic distensibility.

Carotid-femoral pulse wave velocity has also traditionally been used to measure aortic and large arterial stiffness. This is most commonly measured using applanation tonometry or doppler ultrasound to record pressure waveforms. Propagation time is measured from a point (either the foot or the peak) of the carotid waveform to that of the femoral waveform, using sequential readings referenced to the ECG. The path length is measured simply by a tape measure from the site of the carotid to the site of the femoral recordings. This method of measurement is hampered by inaccuracies in path length measurement – particularly for overweight individuals, or those with more tortuous arterial systems[136]. This can lead to errors of up to 30%. The measurement also integrates the properties of the aorta and those of the carotid, iliac and femoral arteries. Whilst this measurement has been robustly associated with cardiovascular outcomes, it is certainly much less accurate a reflection of the aortic function itself than other measurement techniques. Nevertheless, because of the simplicity of the equipment required, it can be applied at larger scale than, for example, mPWV. Wearable devices able to measure pulse wave velocity over a 24-hour period have been developed, and there has even been interest in developing a bathroom scale which is capable of deriving pulse wave velocity measurements[137]! Clearly, these different techniques assess different parts of the arterial tree, and therefore the relative contributions of elastic and muscular arteries differ greatly.

When one is trying to isolate biological processes and pathways underlying cardiovascular risk, it becomes much more important to identify measurements which can pin down particular contributions to that risk. There has, as a result, been more interest in MRI-derived measurement of pulse wave velocity as a more accurate assessment of “pure” aortic function[136]. This method is described in detail in Chapter 12. It has the benefit of greater anatomical specificity, but lower temporal resolution. It is important to note that the reading is not directly comparable to the cfPWV method, for the reasons described above; indeed, mean values are generally lower than for cfPWV[121, 136]. Studies have confirmed that MRI-

derived PWV is also associated with cardiovascular risk; albeit perhaps slightly less strongly[6] (see section 12.1 below). This is perhaps not surprising – cfPWV incorporates assessment of the properties of many different components of the arterial tree, and therefore perhaps gathers more global information about vascular health. Studies which use mPWV as a predictor, on the other hand, focus more exclusively on the contribution of the aorta to cardiovascular health. In both cases, this relationship is likely to be bidirectional: aortic stiffness will tend to cause an increase in vascular risk by exposing the peripheral circulation to higher pressures and pulsatile flow, whilst atherosclerotic or arteriosclerotic processes and hypertension will cause adverse aortic remodelling in turn. There is thus a spiral of worsening aortic and general vascular health.

10.7.3 Measurements of aortic morphology

Apart from the standard aortic dimension measurements, there are no accepted best methods for quantifying or assessing aortic arch geometry. Even basic parameters such as arch branching pattern are usually ignored in standard clinical reporting. Major population studies have often not acquired detailed enough aortic views to undertake precise aortic morphometric assessment, and therefore there is limited data on their prognostic significance in the community. A number of morphometrics have emerged as being of interest in particular clinical settings. Simple two-dimensional measurements of aortic height, width and angle have been associated with cardiovascular risk and with aortic and left ventricular functional parameters[18, 117], whilst arch curvature has been shown in mathematical modelling to be relatively more important in determining the forces acting on the aortic wall than other considerations such as aortic diameter, blood pressure or body size[138]. Aortic tortuosity is also increased in patients with Marfan syndrome, and may indicate worse aortic phenotype. However, measurement of arch curvature for example is not mathematically straightforward, and differences in methodology (spline interpolation vs simpler angle-based measures for example) can yield differing results. Qualitative classifications of arch shape have been described in different conditions. For example, relatives of patients with bicuspid aortic valve were found to have a higher-than-expected prevalence of “cubic” or “gothic” aortic arch[139], and the patterns of aortic flow were found to be different in these aortic configurations. This

observation could be related to the increased risk of aortic dilatation and dissection in first-degree relatives of BAV patients[140-142].

Perhaps the most promising metrics of aortic morphology are those obtained from 3D aortic imaging, particularly those which can incorporate functional data such as aortic flow from 4D flow MRI. Intricate models can be built which minimise data loss from the 3D structures and aortic flow patterns, and which can be used with machine learning techniques to extract features which contribute to risk prediction. The drawback of this method is that the sequences are time-consuming to acquire, and require specialised software for processing and interpretation – factors which limit its use in large population studies, or indeed in resource-poor clinical settings.

10.8 THORACIC AORTIC PATHOLOGY

In the words of William Osler, "There is no disease more conducive to clinical humility than aneurysm of the aorta". Aortic disease remains difficult to diagnose, difficult to treat and difficult to prevent. Aortic aneurysm is defined as abnormal dilatation of the aorta. In the worst scenarios, this dilatation can progress and lead to aortic dissection; a tearing of one or more layers of the aortic wall which carries a very high mortality rate. Around 20% of patients die in the pre-hospital setting, and the in-hospital mortality rate for a Type A dissection is around 21% in the modern era[143].

Aortic aneurysm is usually asymptomatic and is often diagnosed incidentally, or in combination with syndromic features (see below). At present, management of thoracic aortic aneurysm largely rests on the decision of when to undertake prophylactic aortic surgery to replace the aortic root, ascending aorta or undertake major aortic arch surgery.

A number of different conditions can cause thoracic aortic aneurysm and dissection. These range from the largely "environmental" influences contributing to aortic atherosclerosis, to pure Mendelian forms of inherited connective tissue disorders such as Marfan syndrome. Somewhere along this spectrum lies Bicuspid Aortic Valve; a common and partly heritable condition in which genetic and haemodynamic influences combine to affect both valvar and aortic phenotype. These conditions affecting aortic integrity are very varied in terms of the

initial insult to aortic function, yet they often share a final common pathway of disrupted aortic elastic function, aortic medial degeneration with elastic fibre fragmentation and accumulation of glycosaminoglycans, aneurysm formation and, in the worst cases, aortic dissection or rupture. An understanding of the commonalities between these conditions as well as the specifics of pathogenesis in each case is vital for the discovery of biomarkers and novel therapeutic strategies. In addition, the pathways and genes implicated in specific aortic diseases provide insights into the determinants of normal aortic function, and of aortic and more general cardiovascular ageing; mechanisms which could be of much wider therapeutic interest.

10.8.1 Mendelian aortopathies

Thoracic aortic aneurysm and dissection often occurs in the older population due to the cumulative effect of many classical cardiovascular risk factors. However, thoracic aortic pathology in a younger patient should always prompt careful phenotypic and genetic screening for syndromes or inherited forms of the disease. There are many syndromic and non-syndromic genetic causes of thoracic aortic pathology. Targeted panel gene sequencing is usually performed, and can greatly influence management. It can help to determine the appropriate thresholds for surgical intervention, and inform family screening, reproductive counselling and even the consideration of preimplantation genetic diagnosis and selective embryo transfer.

10.8.1.1 FBN1 and TGF- β: Marfan syndrome and Loeys-Dietz syndrome

The understanding of aortic pathology began in earnest in 1991 with the discovery, by Hal Dietz and colleagues, of *fibrillin-1 (FBN1)* as the gene in which mutations were responsible for Marfan syndrome (MFS)[65]. This is an autosomal dominant condition with characteristic skeletal, ocular and cardiovascular manifestations, including a very high incidence of aortic dissection. This observation fitted partially with the prevailing concept of a purely structural role for fibrillin-1 – defects in this structural component of extracellular matrix would lead to “weakness” in the aortic wall causing aortopathy and “weakness” in the ciliary apparatus causing lens displacement[144, 145]. But this structural role of fibrillin-1 never accounted for the abnormal skeletal growth and morphology apparent in Marfan syndrome, and it became

apparent that the function of fibrillin-1 was more complex than first thought. It has now been shown to be a key regulator of TGF-β bioavailability, and usually functions to sequester LTBP-TGF-β complex[146-148]. In MFS, failure of TGF-β sequestration is thought to increase bioavailability of TGF-β, activating its Smad signalling cascade and increasing expression of matrix metalloproteinases (MMPs), leading to matrix proteolysis[149] and thus further impairment of the elastic properties of the aortic wall. Treatments aimed at reducing TGF-β activity have thus far proven disappointing in the clinical setting, although results have been somewhat variable[150-152].

The TGF-β story continued with the discovery by Loeys, Dietz et al in 2005, of variants in TGF-β receptor genes causing an aggressive aortopathy and connective tissue disorder; Loeys-Dietz Syndrome[153]. This syndrome overlaps phenotypically with Marfan syndrome, but has additional features including characteristic facial appearance with hypertelorism, cranial synostosis, translucent skin, bifid uvula and arterial tortuosity.

These patients, particularly those with more prominent facial features, have extremely aggressive aortopathy, requiring commensurately lower thresholds for surgical intervention - a threshold of 4.2cm is suggested in 2010 AHA guidelines[154]. Subsequently, many different TGF-β pathway components have been implicated in Mendelian aortic disease. Variants in genes encoding TGF-β ligands TGFß2 and TGFß3, and the downstream effectors Smad2 and Smad3, cause different forms of Loeys-Dietz syndrome [155-159]. These different forms all exhibit aortic aneurysm and dissection risk, but subtly different additional clinical features – for example, variants in *SMAD3* often cause an aneurysm-osteoarthritis syndrome. Other related genes cause similar clinical presentations, such as Shprintzen-Goldberg syndrome, caused by variants in *SKI*, a repressor of nuclear translocation of Smad complexes, or less syndromic autosomal dominant familial thoracic aortic aneurysm and dissection (FTAAD), caused by variants in other Smad family members (*SMAD4, SMAD6*).

It is somewhat counterintuitive that Loeys-Dietz syndrome and related "TGF-β-opathies" are generally caused by loss-of-function variants in TGF-β pathway components, when a key driver of pathology seems to be over-activation of the TGF-β pathway. In patients' aortic tissue, an upregulation of downstream effectors of TGF-β is clearly apparent in all the "TGF-β-opathies", with increases in nuclear phosphorylated Smad2 (pSmad2), increased collagen

and connective tissue growth factor expression and increased nuclear phosphorylated ERK (pERK1/2) found in the nuclei of patient VSMCs[153, 155-157, 160]. This so-called "TGF-β paradox" is now thought to result from overactivity of non-canonical TGF-β signalling pathways, including increased signalling through alternative TGF-β receptors, or via alternative ligands, and the activation of additional pathways such as the angiotensin II signalling pathway or Notch1 signalling. There is also recent evidence that cells arising from different developmental lineages might respond to loss of function in pathway components in different ways[85]. The cell-autologous component of TGF-β signalling is reduced in secondary heart field (SHF)-derived VSMCs, as evidenced by a reduction in nuclear pSmad2/3. VSMCs derived from neural crest did not display a reduction in pSmad2/3 and therefore had intact cell-autologous TGF-β signalling pathways. The signalling-impaired SHF-derived VSMCs increase expression of TGF-β ligands, and this increase stimulates overactivity of the TGF-β pathway in signalling-intact neural crest-derived VSMCs. In keeping with this hypothesis, selective *SMAD2* knockout in neural crest rescued the aortic phenotype in a mouse model of Loeys-Dietz syndrome. This paper also confirmed previous observations of increased angiotensin II-mediated ERK signalling in SHF-derived VSMCs[85, 161].

This is perhaps the most elegant explanation yet of the TGF-β paradox: cell lineage-specific responses, combined with an over-activation of non-canonical and alternative signalling pathways, all contribute to Smad overactivity and the pathogenesis of aortic root medial degeneration.

10.8.1.2 Variants in other extracellular matrix components

Apart from Marfan syndrome resulting from defects in fibrillin-1, aortopathy can arise from variants in other extracellular matrix components. For example, collagen synthesis is defective in Ehlers-Danlos-like syndromes. Vascular EDS classically presents with aortic involvement – this is due to variants in the gene encoding collagen 3A1 and results also in the classical hypermobility and skin laxity[162, 163]. It confers a greatly limited life expectancy – around 40 years – largely due to rupture of hollow organs such as the uterus or small intestine. Aortic rupture can also occur, and may happen at normal diameters – so intervention thresholds for prophylactic surgery need to be decided with full awareness of the limited life expectancy and knowledge of likely complications.

Loss-of-function mutations in *LOX*, encoding an enzyme which catalyses the formation of lysine crosslinks in elastin and hydroxylysine cross-links in collagens, can also cause heritable TAAD[63]. Variants in *Biglycan*, a gene encoding a small protein which forms a mechanical link between extracellular matrix components, also cause a Mendelian pattern of autosomal dominant TAAD[164]. As with *FBN1* variants, the functional molecular defect in this disorder is not simply a structural one; biglycan also plays a role in regulating bioavailability of signalling molecules such as TGF-β. Therefore, defects in biglycan increase TGF-β signalling.

10.8.1.3 Smooth muscle cell contractile apparatus

The other big category of familial aortopathies is those caused by variants in genes encoding components of the smooth muscle cell contractile apparatus[165, 166]. These cause non-syndromic forms of TAA, usually with autosomal dominant inheritance, but incomplete penetrance. It is pertinent to clinical care to identify these conditions, as they each have their own particular pattern of associated features and complications. For example, *ACTA2* mutations are associated with increased risk of early stroke and coronary artery disease, and probably confer significantly elevated risk during pregnancy – facts which are important for primary prevention, counselling and follow-up of patients[69, 167]. *MYH11* variants are associated with persistent ductus arteriosus[71] and *MYLK* / *PRKG1* mutations confer risks of hypertension which can be rather tricky to treat[73, 74, 168, 169]. More recently, variants in *FOXE3*, a transcription factor involved in smooth muscle cell differentiation and proliferation has also been implicated in FTAAD.[170]

10.8.1.4 Other pathways

There are a number of additional known genetic causes of TAA, with the list growing longer each year. These include *MAT2A*[171], which encodes an enzyme in the methionine metabolic pathway and which causes familial thoracic aortic aneurysm. Developmental genes such as *HOXA1*[172], *NKX2-5*[173] and Notch signalling pathway genes such as *JAG1*[174] are also implicated in aortic disease, although aortic abnormalities are often part of a more severe developmental syndromic phenotype.

10.8.2 Complexity of aortopathy

As with anything in biological science, things which seem simple and neat are often too good to be true. And whilst the basic classification system outlined above is useful for clinical management of aortopathy conditions (for example, to ensure cerebral vessel imaging for patients with TGF-β pathway genetic variants), the real picture is of a huge amount of overlap between syndromes and interplay between different genetic and molecular mechanisms of disease. Rather than considering components of vascular smooth muscle cell contractile apparatus as totally separate from extracellular matrix proteins, in fact these are often physically as well as functionally linked through cytoskeletal attachments. To separate the TGF-β pathway, for example, from the other signalling pathways such as BMP, Notch or Wnt signalling, and to consider this group as totally distinct from structural protein variants, is to miss the complexity and interdependence of the homeostatic and mechanotransduction pathways which operate in the aortic wall. It would be much more accurate to think of the aortic wall as an ecosystem, with each component dependent on the next – so that blunt tools (like Angiotensin II receptor blockers) are bound to be only partially successful in modifying risk.

Another fly in the ointment of the neat classification system is that there is much overlap between syndromes, and with patients who are non-syndromic. For example, there are many patients reported to have pathogenic variants in *TGFBR2*, associated with Loeys-Dietz, who have isolated aortopathy with none of the other classical findings.[175, 176] And the same is true of many of the classical syndromic genes. There are clearly environmental or additional genetic factors at play which greatly influence expressivity. The concept of "risk alleles" and low-penetrance forms of aortic disease is further discussed in the clinical genetics section below.

10.8.3 Bicuspid Aortic Valve - Haemodynamics and genetics

Sir William Osler's recognition of the clinical importance of the bicuspid aortic valve[177] laid the foundations of our current understanding of BAV as the most common congenital cardiac malformation[178, 179]. The aortic valve in this condition is formed of 2 leaflets, rather than the usual "Mercedes-Benz" arrangement of 3 cusps. These 2 cusps can be orientated in a number of different ways (see Figure 10.5), each of which carries a slightly different risk of

complications. This structural abnormality is frequently picked up as an incidental finding, but may also be identified in patients who present with a complication of the condition.

FIGURE 10.5. Different aortic valve morphologies in bicuspid aortic valve

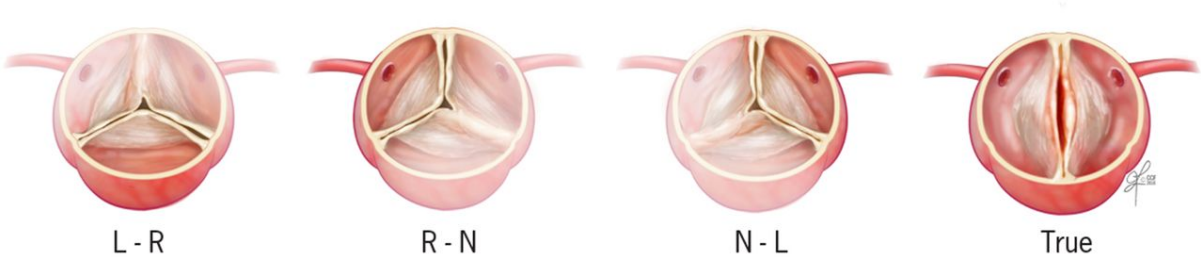


Figure from Masri et al[180]. The left-to-right coronary cusp fusion pattern is the most common conformation, found in approximately 70% of cases, followed by fusion of the right and non-coronary cusps (approx. 20%). Fusion of left and non-coronary cusps (approx. 1-2%) and "true" bicuspid morphology are rare.

BAV is comprised not only of misshapen aortic valves, but also of fundamental abnormalities of haemodynamics and aortic function[181]. It has an estimated prevalence in the general population of 0.5-2%. Major sequelae of BAV include significant valve disease, risk of infective endocarditis, and aortopathy with concomitant risk of aortic dissection; indeed it is responsible for more aortic dissections than any other known genetic disorder [182]. There is a 9-fold increase in the risk of aortic dissection in patients with BAV[183]; an event which carries a very high mortality rate up to around 50% in the modern era[184]. The absolute risk to an individual, however, remains low, and one of the key aims of research into BAV is to identify those patients in whom a higher absolute risk justifies early intervention in the form of aortic root or ascending aortic replacement. Central to this aim is the great debate in the literature over the degree to which aortopathy represents an intrinsic genetic and molecular abnormality of the aorta. The alternative hypothesis is that the aortopathy arises as a result of abnormal blood flow patterns generated by the single initial insult of the abnormally-formed valve.

10.8.3.1 BAV – heritability

BAV is known to be heritable, but it displays incomplete penetrance and phenotypic heterogeneity, making genetic studies problematic [181, 185]. First degree relatives of patients with BAV have around a 10-fold higher risk of BAV than the general population [186, 187]. The

literature reports many multiplex families, in whom BAV follows classical Mendelian inheritance; most commonly autosomal dominant[186]. Indeed over a third of BAV cases (36.7%) appear to be familial and inherited in Mendelian fashion[186] and yet, even in these apparently "straightforward" cases, the genetic basis of the disorder has remained elusive.

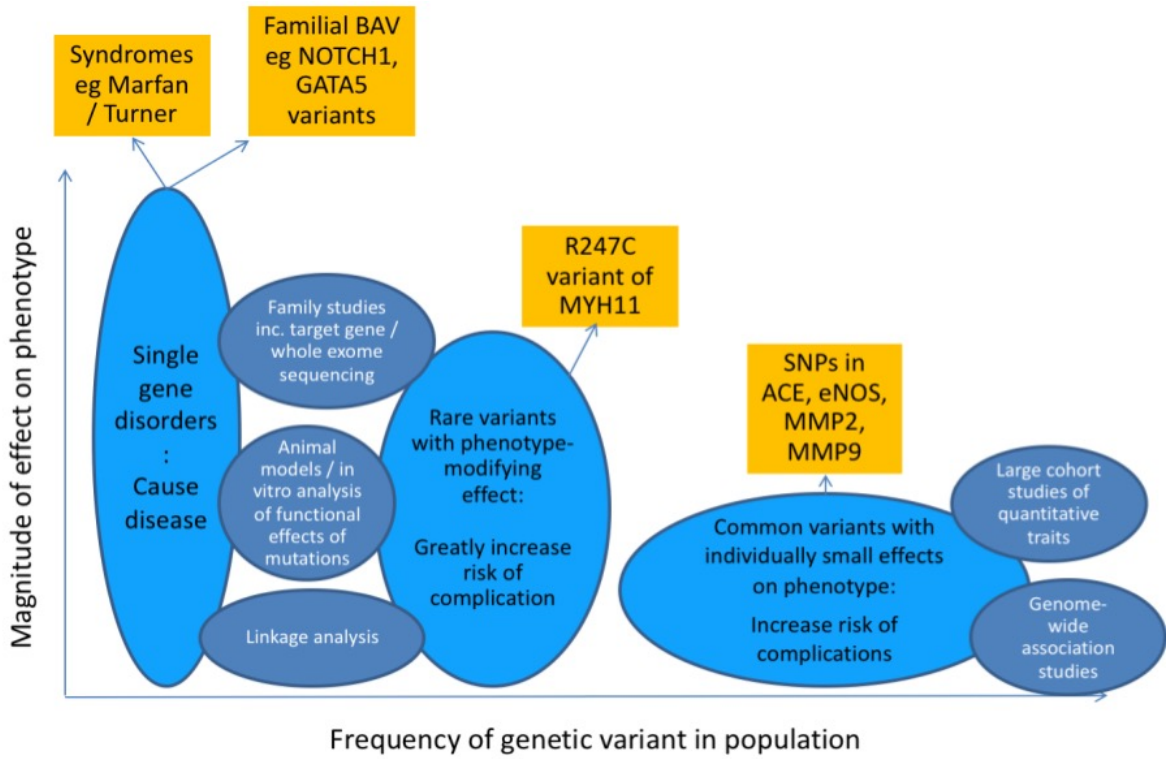
A prospective family screening study by Cripe et al[140] evaluated 50 probands with BAV and their families; a total of 309 subjects. 32% of the families studied had two or more members with BAV or related congenital cardiovascular disease such as persistent ductus arteriosus, coarctation of the aorta, ventricular septal defect, thoracic aortic aneurysm or hypoplastic left heart syndrome. There is marked overlap with other congenital heart defects, with which BAV co-segregates in some cases[140]. Phenotypes such as hypoplastic left heart syndrome(HLHS) are closely related to BAV, along with abnormalities of the aorta including coarctation and thoracic aortic aneurysm (TAA[188])[182, 189, 190]. The latter may be seen in isolation in families of patients with BAV, even in individuals with a normal tricuspid aortic valve[141, 191]. These related phenotypes represent a broad spectrum of developmental abnormalities of the aorta and outflow tract; whether they represent different manifestations of one entity or multiple separate pathophysiological processes is still uncertain. Of interest is one case report of monochorionic twins who displayed distinct phenotypes[192] - one with HLHS, the other with isolated non-stenotic BAV. It was hypothesised that the genetic substrate must be identical in these twins, and therefore the divergent phenotypes arose solely due to exposure to differing haemodynamic profiles in utero. This is certainly a compelling argument and indeed it has long been known that blood flow through the primitive heart during embryonic development plays a crucial role in differentiation of cardiovascular structures. The idea of BAV being a "forme fruste" of HLHS is not a new one[190], and there is clear evidence of increased prevalence of HLHS in BAV kindreds. It is likely that multiple influences; some undoubtedly haemodynamic, some genetic and some environmental, interact to determine expressivity in individuals. In the case of the twins, subtle changes in genetics or haemodynamics may have arisen during development, accounting for the phenotypic difference. Whatever the explanation for this particular anomaly, it serves as a good illustration of the subtlety and complexity of the influences on phenotype in BAV, and therefore the difficulty in elucidating the genetic architecture of the condition.

10.8.3.2 BAV: Molecular genetics

It has therefore become apparent that BAV is not as straightforward as a simple Mendelian genetic disorder. In fact, as with any complex disease, there are multiple influences on phenotype – genetic variants across the frequency spectrum, haemodynamics both in utero and later life, environmental factors and stochastic factors – all of which contribute in a cumulative fashion to the final phenotypic expression in an individual.

In addition to the "causative" single gene mutations underlying syndromic and inherited forms of BAV, it seems that there are also genetic alleles which act as significant disease modifiers. These may be rare, with large phenotypic effects, such as alleles which may cause aortic structural abnormalities; or common polymorphisms, which may, for example, increase susceptibility to aortic "damage" caused by aberrant haemodynamics.

FIGURE 10.6. The varied genetic effects on phenotype in BAV



The varied genetic influences on phenotype in BAV, with examples of each, and experimental strategies to identify these variants. Experimental strategies identified here are examples of those most commonly used in different scenarios - not mutually exclusive or exhaustive, and not necessarily related to the y axis of magnitude of effect. Figure adapted from Prakash et al 2014 [185].

10.8.3.3 BAV - syndromic associations

Syndromic associations with BAV, such as Marfan, Loeys-Dietz and Turner syndromes have provided some clues as to the underlying genetic / molecular basis of the disorder[185, 193]. In particular, 30-40% of patients with Turner syndrome have BAV[194], and these patients have a higher than average incidence of associated aortic abnormalities such as aortic coarctation or aneurysm. The link with Turner syndrome, as well as the 2:1 male preponderance of the condition in the population, provides some evidence for an X-linked pathogenetic mechanism in certain cases. However, the genes responsible are unknown.

The genes which have been discovered in syndromes associated with BAV can also give us insights into the molecular mechanisms underlying the condition, and therefore identify potential treatment targets. The relatively frequent association of a bicuspid aortic valve with aortopathies such as Marfan and Loeys-Dietz also implicates TGF-β signalling in the pathogenesis of BAV[195], and indeed reduced expression of fibrillin-1 has been seen in the aorta of patients with BAV[196]. In addition, the mouse model of Marfan syndrome, with targeted deletion of *FBN1*, displays a bicuspid valve phenotype[145]. Levels of MMP2; a downstream effector of TGF-β signalling, have also been shown to be higher, and levels of TIMP-1 lower in patients with BAV and thoracic aortic aneurysm (TAA) than in patients with a tricuspid valve and TAA[197]. However, the TGF-β signalling pathway certainly does not provide all the answers. To date, no causal variant in *TGFBR1* or *TGFBR2* genes has been detected in patients with isolated BAV[198].

10.8.3.4: BAV: insights from animal models and family segregation analysis

Model mice have also provided insights into BAV pathogenesis. Endothelial Nitric Oxide Synthase (*eNOS*) knockout mice develop a BAV phenotype[199] and, whilst no genetic mutation in *NOS3* has been identified in human BAV, eNOS expression has been found to be lower in the aorta of patients with BAVs [200]. eNOS expression is shear-stress responsive, also creating an interesting theoretical link between the aberrant haemodynamics produced by turbulent flow through a bicuspid valve, and the pathogenetic mechanism itself [82]. GATA5, a transcription factor which induces eNOS expression, has also been implicated in BAV pathogenesis. Evidence for its involvement comes from the *GATA5* knockout mouse, in which there was incomplete penetrance (around 1/4) of a BAV phenotype[201]. GATA5 is expressed in

endothelial and endocardial cells, and it is hypothesised that the tight regulation of eNOS levels by GATA5 during endocardial cushion development is vital for normal tricuspid aortic valve formation[202]. There is compelling biological evidence for the importance of transcription factors GATA5 and NKX2-5 in the pathogenesis of BAV - both from mouse models and also from an understanding of their normal role in cardiac development and valvulogenesis. *In silico* modelling and *in vitro* experiments demonstrate the effects of the identified variants on protein function. Evidence for *GATA5* and *NKX2-5* being key in BAV also comes from human studies, in which rare *GATA5* variants in highly-conserved transcriptional activation domains have been shown to segregate with BAV[203-206].

Evidence from in vitro and mouse experiments also suggested Notch1 signalling as key in valvulogenesis and normal cardiac development[90, 207]. *GATA5* knockout mice display reduced Notch1 signalling [201]. *NOTCH1* knockout mice develop severe cardiac malformations, which result in early death. Endothelial-mesenchymal transformation is dependent upon Notch1[90], which also acts as a mediator of signalling between the secondary heart field and migrating neural crest. Variants in *NOTCH1* were the first genetic culprits identified in families with bicuspid aortic valve, and mutations are associated with increased valve calcification[208] and histopathological abnormalities of the aortic wall[209]. Mutations in *JAG1*, which encodes the Notch1 ligand, also cause BAV in the context of Alagille syndrome[174] Interestingly, there is significant interaction between the Notch1 signalling pathway and both the eNOS and TGFβ pathways, the latter of which is heavily implicated in human aortic disease[87]. This provides biological plausibility for the involvement of genes in all 3 pathways in development of BAV phenotypes.

Specific involvement of cells of the cushion mesenchyme in pathogenesis has also been established by the high prevalence of BAV in a mouse model with tissue-specific deletion of *Alk2* (*ACVR1*)[210].

Other animal models of BAV formation include the *Nkx2-5* haplo-insufficient mouse. *Nkx2-5* knockout induces early embryonic lethality, but heterozygous mice display increased prevalence of BAV, at 11%, along with atrial septation defects[211]. There is significant phenotypic heterogeneity in these mice; a finding mirrored in human studies, in which NKX2-5 has been associated with a broad spectrum of congenital heart disease, from atrial septal

defects to severe hypoplastic left heart syndrome[173]. Recent identification of a loss-of-function mutation in *NKX2-5* as responsible for an autosomal dominant form of BAV[206] has confirmed its role in human BAV. Functional characterisation of this defect again highlights the complex interactions of molecular pathways in pathogenesis, as it inhibited the normal synergism between NKX2-5 and GATA5.

Extensive functional analysis of a complex patient with a translocation affecting two genes implicated a mutation affecting Matrin 3 (*MATR3*) expression in pathogenesis of the cardiovascular phenotype, which included BAV[212]. The potential for mutations in this gene to cause BAV was confirmed in a mouse model. However, the general applicability of this finding in the wider context of BAV has yet to be established.

More recent family segregation studies have identified variants in *SMAD6*[213, 214], an effector of the BMP signalling pathway, and *MAT2A*, a component of the methionine pathway, as causative of BAV with thoracic aortic aneurysm (TAA). The penetrance of BAV and of TAA in the family with *MAT2A* variant is low. It is hypothesized that both these defects might lead to aberrant differentiation of VSMCs, which might predispose to TAA formation.

10.8.3.5 BAV: Genome-wide association studies

The phenotypic heterogeneity of BAV and the challenge of case identification (requiring echocardiography or other imaging, and / or long-term follow-up to assess for development of aortic pathology) makes it extremely difficult to conduct such a project in BAV. Results will also be confounded by the undoubted presence of significant numbers of subjects with unrecognised BAV in the control cohort, and by the reduced penetrance of the condition. These difficulties mean that this technique has not yet been published. The establishment of large, international consortia such as BAVCon[185] and AortaGen[215] however, make this a real possibility in the near future.

10.8.3.6 Thoracic aortic aneurysm (TAA) in BAV: genetics

TAA is common in patients with BAV, reported in around a half of affected individuals[182]. Both inherent genetic and structural abnormalities of the aortic wall, as well as aberrant haemodynamics, are likely to contribute to development of this phenotype. There has been some controversy over the relative importance of these contributing factors and, indeed, the

contribution from each may vary from patient to patient[193, 216-218]. The common embryological origin of part of the valve structure and ascending aorta from neural crest cells, may explain the aortopathy seen in BAV, with a key role for eNOS identified in the normal maturation of these structures.

Development of TAA sometimes perfectly cosegregates with BAV in families[185, 219-221], and in these cases, it may represent an inherent part of the "BAV phenotype." But there is emerging evidence that it may in some cases result from distinct genetic influences. First-degree relatives of patients with BAV have a greatly increased prevalence of TAA over the general population[140, 222]. This increased susceptibility to TAA can occur even in relatives with a normally-functioning, tricuspid aortic valve[141, 142]. For example, in the original paper by Garg et al[208], a clear link was identified between inactivating mutations in *NOTCH1* and autosomal dominant valve disease. 4 out of 9 affected individuals had TAA. However, there was one individual who had TAA without valve disease – i.e. he had a normally-functioning, tricuspid aortic valve, and did not have the familial *NOTCH1* mutation. This finding, strengthened by subsequent genotype-phenotype correlations in BAV, implies two distinct mechanisms of expressivity: one (*NOTCH1*) for calcific valve disease, and the second for highly penetrant TAA with uncommon valve calcification [208, 223].

Aortic elastic function has also been shown to be diminished in relatives with normal valve phenotype[142]. These studies are not merely of academic interest; the findings have considerable implications for screening of family members, as even those with a structurally normal valve may require surveillance for development of aortic complications.

In patients with BAV, there is much evidence to suggest an inherent structural abnormality of the aorta. Non-invasive imaging suggests impairment of the elastic properties of the aorta, even before any dilatation becomes apparent[120, 224-229]. Histologically, a pattern of cystic medial necrosis may be seen, analogous to that seen in the aorta of patients with Marfan syndrome. Indeed, reduced expression of *fibrillin-1* has been shown in the BAV aorta[196] and, recently, *fibrillin-1* gene mutations have been seen in patients with BAV without overt Marfanoid features[230].

In addition, factors involved in proteolysis, such as MMP2 and MMP9, along with their tissue inhibitors, display distinct expression patterns in the aortas of BAV patients, even in the

absence of significant aortic dilatation[197, 231-234]. One transcriptomic study found that the differential gene expression in dilated vs non-dilated aortas in patients with bicuspid valve was different to that seen in patients with tricuspid valve[235]. This suggests a different mechanism of aortic dilatation in BAV.

It may be that common, or indeed moderately rare, variants in genes classically associated with aortopathy may modulate TAA risk in the context of BAV. Candidate genes for TAA development in BAV include *FBN1*, as described above. Another candidate is *MYH11*, which encodes myosin heavy chain, one of the key proteins in the vascular smooth muscle cell contractile apparatus. This has been linked to familial TAA associated with persistent ductus arteriosus and, less frequently, BAV[71, 72]. *MYH11* mutations such as R247C seem also to act as phenotype modifiers, representing a susceptibility trait where this particular variant increases risk of aortic abnormalities, but is not sufficient, in itself, to cause major aortic complications[70, 236].

Common genetic variants may have much more subtle effects on phenotype -but their cumulative effect may be very great. Identification of these variants is problematic, and requires large-scale GWAS analysis of quantitative traits or sub-phenotypes. A target gene approach has, however, identified SNPs in *eNOS, ACE, MMP2* and *MMP9* which influence aortic aneurysm risk in BAV patients[237-240].

10.8.3.7 Thoracic aortic aneurysm (TAA) in BAV: haemodynamics

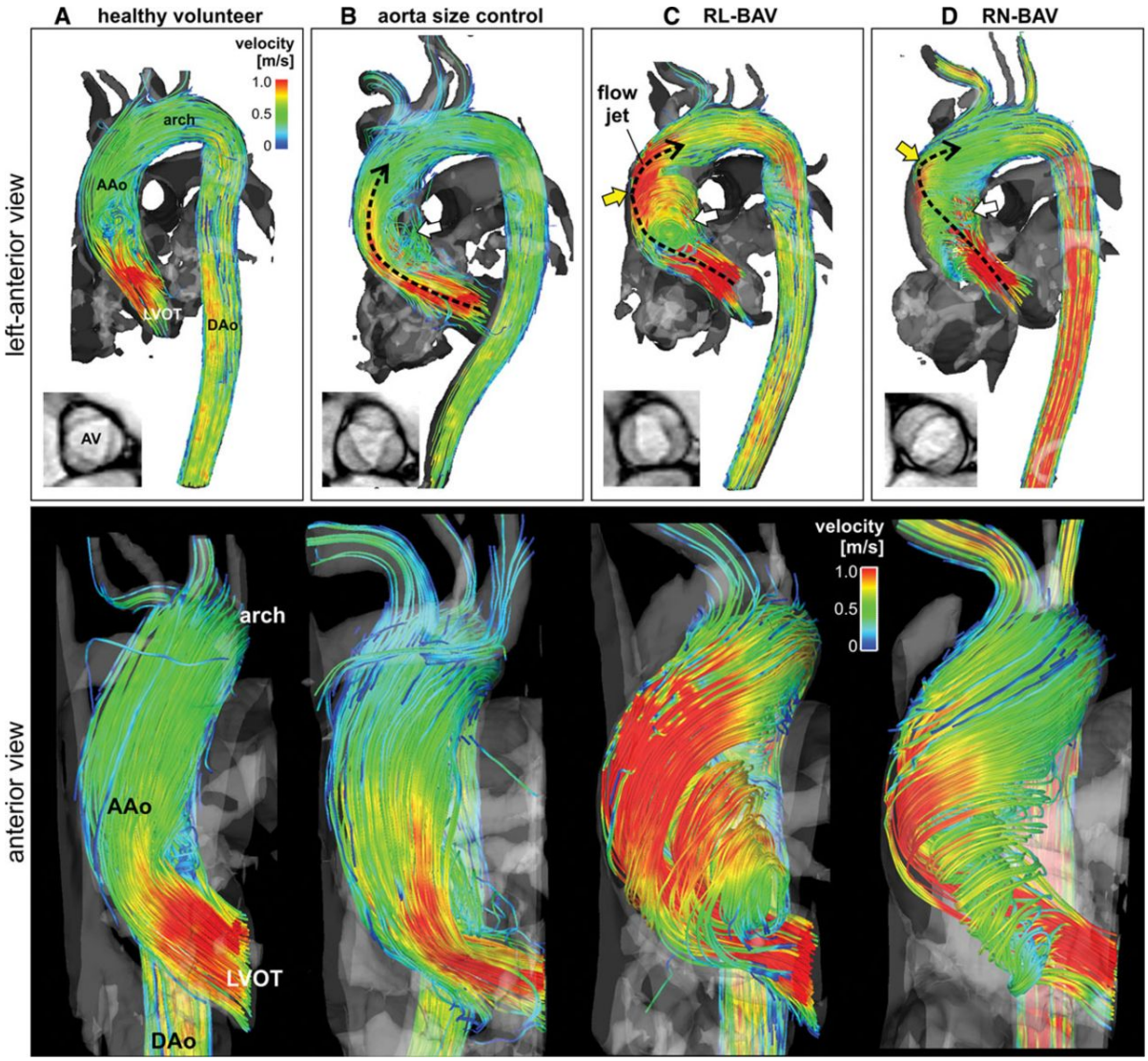
Even in the absence of classically defined "valve dysfunction," flow patterns in the aortic root and ascending aorta are abnormal in patients with BAV[241]. Furthermore, these flow patterns are distinct between morphological subtypes of BAV[242] (see Figure 10.7), exposing the aortic wall to differing patterns of shear stress and stretch[242-245]. A localized response of the aortic wall to this haemodynamic environment can provide a mechanistic explanation for the different natural history seen in these subtypes (with R-L coronary cusp fusion associated with dilatation of the aortic root at a younger age and R-N cusp fusion associated with increased aortic arch dimensions).

It is also proposed that haemodynamics play a crucial role in embryonic development, with flow patterns through the primitive valve structures altering phenotypic expression[246-249]. For

example, in the case of twins described above [192], it seems that different haemodynamic conditions in utero may be a possible explanation for the divergent phenotypes seen.

In keeping with this hypothesis, it has been shown that shear-responsive genes *KLF2, eNOS* and *ET-1* play a key role in the development of the cardiovascular system[246]. *In vitro*, human endothelial cells demonstrate different patterns of gene expression in response to differing shear stress[79, 250, 251]; a response which alters the underlying properties of the endothelium and the vascular smooth muscle. The vascular smooth muscle cells themselves also directly sense stretch. These both regulate the structure and elastic function of the aortic wall through synthetic, secretory, contractile and structural capabilities as discussed above. *In vivo,* alterations in gene expression in the BAV aorta have been shown for several flow-regulated genes, such as *eNOS*[252], MMPs[233, 237, 253], *PKD-2*[254] and genes in the TGF-β signalling pathway. Also, endothelial function has been shown to be impaired in patients with BAV[80]. These pieces of evidence together indicate that endothelial cell response to altered patterns of shear stress, coupled with a VSMC response to stretch in BAV might be a key component of aortopathy risk[182, 255, 256].

FIGURE 10.7: Blood flow patterns with different aortic valve morphologies, from Mahadevia et al[242]



Top: 3D streamline visualization of peak systolic blood flow in patients with BAV (C and D) in comparison with an aorta size–matched control subject (B) and a healthy volunteer (A). Note the presence of distinctly different 3D outflow flow jet patterns (black dashed arrows) in the ascending aorta (AAo) for patients C and D. Bottom: 3D flow patterns in the left ventricular outflow tract (LVOT) and AAo distal to the aortic valve. Note the different systolic aortic valve outflow flow jet patterns (red indicating high velocities > 1 m/s) and wall impingement zones that correspond to variable exertion of high wall shear forces between different valve groups (C and D) and aorta size–matched controls (B) and healthy volunteers (A). BAV indicates bicuspid aortic valve; RL, right and left coronary leaflet; and RN, right and noncoronary leaflet.[242]

Studies of haemodynamic abnormalities in BAV certainly provide compelling evidence for the involvement of flow-regulated mechanisms in disease; the "controversy" in the literature over haemodynamics versus genetics seems a rather false one; it is evident that both contribute to phenotype, and interact in complex and profound ways, both during

development and throughout adult life. The real questions lie in unravelling the mechanisms underlying the aortic response to the particular haemodynamic challenges of BAV.

10.8.3.8 Clinical implications of research in BAV

One of the major challenges in clinical management is selecting patients for surgical replacement of the aortic root or ascending aorta, which may be undertaken either alongside or independently of, valve replacement or repair. Current guidelines suggest thresholds of 5.5cm for surgical intervention in patients with BAV[257] and no additional risk factors. However, these "one-size-fits-all" guidelines don't capture the subtleties of risk differentiation between patients with different genetic susceptibilities and haemodynamics. If it were possible to build a more sensitive risk assessment, based on the cumulative contribution of individual genetic risks and haemodynamics, this would prevent unnecessary invasive intervention. Current guidelines also recommend screening of all first-degree relatives of probands with BAV[257]. A single imaging study to detect the presence of a BAV seems inadequate to truly determine a relative's risk of complications. The high prevalence of BAV makes it impractical to undertake regular repeated screening of all relatives. Therefore, improved risk stratification and family screening is one of the main goals of research into the genetic and haemodynamic origins of this condition and its complications.

10.9 GENETICS AND GENOMICS

10.9.1 Complexity from simplicity

The human genome is an extraordinary testament to how incredible complexity can be generated from great simplicity. The DNA in every nucleated cell in our bodies contains the full instruction manual for human development and function. As Francis Collins put it shortly after the publication of the first human genome, "*It's a history book – a narrative of the journey of our species through time. It's a shop manual, with an incredibly detailed blueprint for building every human cell. And it's a transformative textbook of medicine, with insights that will give health care providers immense new powers to treat, prevent and cure disease.*" And yet that detail is encoded by a string of just 4 different nucleotide bases, or genetic

"letters" – A (adenine), C (cytosine), G (guanine) and T (thymine), covalently linked together by alternating sugars and phosphates. Whilst in the late 20th century, the major scientific hurdle was the technological ability to sequence genomes, now the frontier has shifted to the interpretation of the information contained within. And in fact, it has become apparent that it is somewhat disingenuous to suggest that the genetic code is as simple as a string of 4 letters – there are countless modifications, both in the macromolecular conformation of the chromatin which "packages" the DNA strands, and also in the regulatory biochemical modifications of the histone proteins which control that 3D structure, as well as the bases themselves. Areas of the genome which were once thought to be non-functional, inert and unimportant ("junk" DNA), are now known to have important cis- and trans-regulatory roles and to play a key part in the expression of RNA and protein gene products. In the Encyclopedia of DNA Elements (ENCODE) project, around 80.4% of the human genome participates in at least one biochemical RNA and/or chromatin associated event in at least one cell type[258]. It has become apparent that, when trying to interpret the human genome, we are, in the words of a recent review[259], "little better than a two-year-old trying to make sense of the *Encyclopaedia Britannica*." The field of genome interpretation is still undergoing huge evolution, at an exhilarating pace.

Advances in the last 20 years have revolutionised the fields of genetics and genomics; starting with the Human Genome Project which was completed in 2003[260]. This detailed reference data, along with the advent of next-generation, massively-parallel sequencing technologies in the mid 2000s, has been truly transformative. Now the cost of sequencing an individual's entire genome has broken through the $1000 dollar barrier for the first time, with major companies such as Illumina promising to push towards $100[261]. This makes accessible the use of routine genome sequencing in healthcare, and also in research. This drop in sequencing costs has stimulated an explosion in genomic research and data-gathering. The 100,000 Genomes Project in the UK aimed to sequence the genomes from rare disease and cancer patients to stimulate the creation of an integrated Genomic Medicine service in the UK. Last year, the UK Health Secretary announced even more ambitious plans to sequence 5 million genomes within 5 years[262].

10.9.2 Genetic variation

Genetics is the study of heredity, broken down into single units or genes; genomics integrates information from across the genome, often at population-scale. Whilst there is some controversy about the exact definition of a gene, this is usually taken to mean a stretch of DNA which contributes to phenotype.

As genome sequencing has been applied to larger populations and more patients with disease, we have come to understand more about the natural variation in the human genome between individuals. Around 0.1% of the bases in the genome, equating to approximately 10 million nucleotides, exist in two or more common forms[263]. These variations in the base code which occur frequently in a population are termed single nucleotide polymorphisms (SNPs). Technically, the term "SNP" refers to a single base difference which occurs frequently (usually >1%) in the population. Of course, variants may come in different forms such as short insertions or deletions. These "indels" are much more challenging to call accurately from sequencing or even array data. The frequent occurrence of SNPs and common indels in the population makes them at first glance unlikely candidates for key determinants of pathogenesis; disease-causing variants are expected to be under negative selection. However, many SNPs and common indels play key roles in determining penetrance and expressivity of disease traits. Each person harbours around 4-5 million genetic variants[263]; interpreting all these and their possible effects on phenotype is a major challenge.

10.9.2.1 Open-access data and understanding human genetic variation

Over the last few years, the funding, accumulation and release of other publicly available huge population-level genomic and functional datasets such as has made possible much more ambitious strategies to define, filter and annotate the genome in varied populations. One example of this is gnomAD – the genome aggregation database[26, 27]. The latest release incorporates sequenced genomic data from 141,456 individuals (125,748 exomes and 15,708 genomes), aggregated from numerous different disease-specific and population studies. The resource is open-access and easily queried via a guided user interface or programmatically. It has enabled accumulation of insights such as reference variant frequencies across different populations, evaluation of constraint metrics (measures of a gene's tolerance to loss-of-function or missense mutations)[27], amongst many others.

A key example of open-access functional data is the Genotype-Tissue Expression (GTEx) project. This database contains RNA-sequencing and microarray data from many different tissue types, alongside dense genotyping of the same donor individuals. This enables querying of expression levels across different tissues. By treating gene expression levels in different tissues as quantitative traits, the effects of genetic variation on gene expression can also be examined. Variants which are highly correlated with gene expression can be identified as expression quantitative trait loci, or eQTLs[264, 265].

These types of freely accessible reference data have facilitated the interpretation and refinement of results from clinical genetic sequencing, as well as studies such as genome-wide association studies or segregation analyses.

10.9.2.2 The roles of genetic variation in human disease

For common diseases, and for quantitative phenotypes such as blood pressure or aortic pulse wave velocity, much of the heritability is expected to be due to common genetic variants – the so-called "common disease-common variant" hypothesis. Each individual variant exerts a very small effect size, but the collective effect on phenotype may be considerable. An example of this is in hypertension, where more than 900 different loci have been identified with effects on the underlying blood pressure distribution. Collectively, the significant SNPs explain 5.7% of variance in SBP[266]. With denser genotyping and imputation, the heritability due to common variants can account for around two-thirds of the heritability estimates from twin studies[267, 268]. This has led to much debate as to the origins of this "missing heritability". It is likely that twin studies might overestimate heritability, and that GWASs underestimate the heritability due to common variants due to imperfect tagging of causal SNPs[269]. However, much of the missing heritability is thought to be explained by rare genetic variation with a more profound effect on phenotype; the so-called "common disease-rare variant" theory. This hypothesis has been strengthened by a recent study which used whole genome sequencing data to examine the heritability of height and body mass index (BMI). The authors found that, by accounting for rare variants using whole genome sequencing data, they could "recover" almost all of the missing heritability[270]. Finally, rare cardiovascular and aortic diseases can be due to rare, disease-causing variants – the rare disease-rare variant hypothesis. These rare genetic influences may be discovered by linkage studies, candidate gene sequencing,

case:control or family segregation analysis. A variant which is rare enough to “cause” disease must be less common than the disease itself (significantly more so if the disease is genetically heterogeneous). However, the interpretation of these rare variants is not a trivial task. Each individual has around 550 “private” variants per exome – in other words, they have not been previously identified in population datasets, of which around 20-30 may be protein-altering and predicted to be functional[271]. Determining their effect on complex clinically-relevant traits or on disease causation is not easy. The interpretation of rare variants in disease is discussed below in section 10.12.4.

10.9.2.3 Assaying genetic variation

The techniques used to define genetic variants in an individual are chosen according to the question being asked. Next-generation sequencing of a particular gene, panel, or whole exome or genome sequencing, is selected when unknown rare variants are under study – for example when looking for rare pathogenic variants in a patient with suspected Mendelian disease. It provides single base resolution for most genomic regions, and high sensitivity. However, the cost of this technique, whilst falling, remains prohibitively high for most population studies. Specific known variants or common variants can usually be cheaply and efficiently assayed using SNP array genotyping, combined if necessary with imputation (described further in Chapter 13).

10.9.2.4 Next generation sequencing, whole exome and whole genome sequencing

Genetic sequencing can be used to assay all variation across a single gene, a panel of genes thought to be associated with disease, every protein-coding gene in the genome (whole exome sequencing) or the whole genome (including non-coding bases). Each of these apart from whole genome sequencing requires a step of target enrichment. There are different methods for this target enrichment step: amplicon-based methods which rely on multiplexed polymerase chain reaction (PCR) and hybridisation enrichment. In PCR enrichment, primer pairs are designed to bind to sequences either side of the target regions and introduce platform-specific adaptor sequences. These target regions are then amplified by PCR and pooled to form a library ready for NGS sequencing. In hybridisation enrichment, the genomic DNA is first fragmented into multiple short "reads” and then attached to platform-specific

adaptors. The genomic library fragments are then hybridised to baits – short DNA or RNA oligonucleotides complementary to the regions of interest, in the form of a capture array or in-solution probes. Targeted library fragments are pulled down and eluted, while off-target fragments are washed away. These sequences are then amplified via PCR.

Next generation sequencing enables millions of library fragments to be read in parallel, with each platform employing different techniques to achieve this.

Life Technologies 5500xl sequencer employs sequencing-by-ligation. Library fragments are annealed to beads and amplified using emulsion PCR, resulting in a bead being enriched with multiple copies of a single fragment of library. Beads are deposited onto a glass slide, and fluorescent complementary probes are ligated to the library fragment. The probe which ligates, and so the fluorescence detected by the imager, is determined by the sequence of the library fragment. Illumina platforms use sequencing-by-synthesis. First, library fragments are attached to a flow cell, annealing to complementary adaptor sequences coating the glass. These are copied multiple times to form clusters of identical library fragments. Fluorescent deoxynucleotides (dNTPs) with reversible terminators are incorporated into a newly synthesised complementary strand, images of the fluorescent label are captured, and terminators are removed, allowing the next base to be incorporated.

In each case, the sequences generated from the reads are reconstructed from the sequence of fluorophore signals at each location. The reads are then aligned bioinformatically against a reference genome. This requires huge computing power and careful quality control. “Calling” of variants is then undertaken by a series of sophisticated filtering tools which assess the alignment of reads at each position, and identify what proportion of them have an alternative base to the reference. Base calls are assigned a quality value, or “Q score” which enables high-confidence variants to be taken forward for interpretation.

This sequencing approach is very flexible and can be applied to single genes, panels of genes, exomes or even whole genomes according to the question being addressed.

For patients with a particular monogenic disorder, it may be appropriate to sequence a single gene. However, most clinical genetic testing is performed with panels of genes or exome sequencing. Whole genome sequencing employs very similar methods, but can exclude the target capture and enrichment step. This solves problems of unequal coverage and uneven

target amplification. However, it remains prohibitively expensive and extremely computationally intensive for large cohort studies. Recent projects designed to progress the generation and interpretation of whole genome sequencing data such as the 100,000 Genomes Project, have begun to address these problems, and the cost of genome sequencing is falling dramatically, with Illumina hinting at a sub-$100 genome available over the next several years[261].

In the modern era, the generation of sequence data, whilst not a trivial task, is perhaps the simplest part of the puzzle: calling, validating and interpreting variants is infinitely complex.

10.9.3 Identifying genetic variation that contributes to disease

The expected genetic architecture of disease will determine the most appropriate study models for that disease, and the types of variants which will be assayed. Broadly, studies can focus on related individuals by segregation or linkage analysis, or unrelated individuals by association analysis.

10.9.3.1 Family studies: linkage and segregation analysis

Linkage studies can identify chromosomal regions or individual candidate variants which cosegregate with disease or phenotype in families. Parametric linkage analysis uses a logarithm of odds (LOD score) to assesses the probability that a marker and putative disease gene cosegregating is due to the existence of linkage rather than due to chance. This method assumes a specific genetic model – for example, a dominant, fully-penetrant trait caused by one disease gene with two alleles, with the disease allele frequency known. For single-point (aka two-point) analysis, we consider one marker and the putative disease gene, and θ = recombination fraction between the marker and disease gene. We test the null hypothesis that the recombination fraction is 0.5 – in other words, if an individual has that marker, then there is an equal likelihood of having or not having the disease gene. This would be the case if the marker and disease gene were very distant from one another – for example, on separate chromosomes. If the marker and putative disease gene are situated close together, then the likelihood of recombination between them is reduced, and so the alternative hypothesis is that the recombination fraction is <0.5. Simplistically, the lower the chances of recombination

between the marker and putative disease gene, then closer they must be situated. In multipoint analysis, we can consider θ= the location of the disease gene on the chromosome. Here, we test the hypothesis of the gene being unlinked (θ= ∞) to the hypothesis of θ= particular chromosomal position. Non-parametric models can be used for complex diseases or traits where the inheritance pattern is unknown. Large pedigrees or multiple pedigrees sharing the same disease or trait are required for linkage analysis to be properly powered.

10.9.3.2 Case: control analyses & rare variant burden analysis

The role of rare variants in common disease is often difficult to unravel at population level, as the power to detect variant:disease association relies heavily on the variant frequency, as well as the effect size. Therefore, for an unknown low-frequency variant, huge sample sizes would be required to detect even a considerable increase in disease risk. Instead of testing rare variants individually, these are often grouped together into those which are likely to have similar functions. Perhaps the most common method is to group all rare, protein-altering variants in genes of interest, and then compare the "burden" of rare, protein-altering variants in a cohort of cases versus healthy control subjects. This can identify genes in which rare variants may increase disease risk, but adds little to the analysis of individual rare variants in that gene[272].

10.9.4 Interpreting rare variants in disease

In Mendelian disease with known gene associations, many causal variants can be confidently identified as pathogenic. However, even where disease genes are well-defined, it remains a significant challenge to assign pathogenicity to novel variants or those not well-described in previous literature. The American College of Medical Genetics and Genomics (ACMG) released variant classification guidelines in 2015[273], which provided a framework by which varied data can be assimilated to provide reproducible and reliable classification of genetic variants. The five-level classification (pathogenic, likely pathogenic, variant of uncertain significance (VUS), likely benign, and benign) allows easy communication of the weight of evidence for a particular variant causing disease. This is important for assessing the clinical actionability of genetic findings. An important message from these guidelines is that no one piece of information is sufficient, by itself, to assign pathogenicity. Instead, evidence must be

collated from multiple lines of enquiry to allow interpretation. This includes evidence from population-based studies (is the variant absent, or below an acceptable threshold for disease prevalence, in control populations?), family-based segregation analysis, in silico prediction of amino acid change and comparison with known pathogenic variants, in silico prediction of deleterious effect on gene or gene product[274], and functional laboratory studies demonstrating effect on gene product function.

Where inheritance follows a well-described Mendelian pattern, segregation analysis within a pedigree is one of the most robust ways to demonstrate pathogenicity. If a rare, protein-altering variant clearly follows the patterns of disease inheritance (for example, a *de novo* variant from a classical trio [affected proband plus two unaffected parents] in an autosomal dominant condition), then it is highly likely to be pathogenic if it falls within a plausible disease gene. However, it is rare that family structures are so tractable for variant interpretation.

One major limitation of these guidelines is that they deal mainly with straightforward patterns of Mendelian inheritance. Unfortunately, many diseases do not display this consistency of genotype-phenotype correlation. Instead, incomplete penetrance (individuals carrying a variant but not having a disease phenotype), variable expressivity (individuals carrying the same variant as one another but displaying very different phenotypes), or more complex inheritance patterns are very common. Even diseases traditionally thought of as clearly Mendelian such as the heritable cardiomyopathies display these features of more complex inheritance – for example with some titin truncating variants more likely to manifest with dilated cardiomyopathy[275] if combined with a second insult. In the context of aortic disease, bicuspid aortic valve and its associated aortopathy display features of more complex inheritance[182]. Often, a single variant in a particular gene is not sufficient, in itself, to cause disease, but a combination of variants in different genes may be sufficient to separate patients with the disease from control subjects without it. This is the notion of oligogenic inheritance. Bioinformatically, this is challenging to address without very large cohorts or very large family structures.

10.9.4.1: *In silico* tools for prediction of variant pathogenicity

A huge variety of *in silico* tools have arisen to predict pathogenicity of variants. The accuracy of these is patchy at best, and highly dependent upon the biological function and context of

the variant[274]. For coding variants, there are good tools to predict biochemical and conformational changes in the encoded proteins or protein domains. Other tools assess conservation across species or across homologous proteins. Yet more tools can model the effects of potential splice site mutations on splicing and transcript formation. And meta-tools such as M-CAP (Mendelian Clinically Applicable Pathogenicity score) can assimilate information from several of these and establish a score which reflects the likelihood of a particular variant being "damaging". However, the final common flaw in all these models is that there is still no good way to predict whether something damaging at a molecular level translates to damage at a whole organism, or even cellular level – or whether that particular variant is sufficient, in itself, to cause disease without requiring a further genetic or environmental insult to be expressed. These limitations prevent the use of *in silico* algorithms from providing anything more than "supportive" evidence for variant interpretation[273, 274].

The picture becomes even more complex when one considers non-coding variants. Again, a variety of tools can be used to assess whether a particular variant is likely to be a key site of regulation – for example, by modelling likely modifications to transcription factor binding, incorporating evidence from CHiP-seq (chromatin immunoprecipitation combined with DNA-sequencing), or modelling chromatin conformational changes or using data from 4C (chromatin conformation capture on chip) /Hi-C (high-throughput sequencing of ligated fragments from chromatin conformation capture) experiments to examine the likelihood of the variant affecting histone modifications or chromatin folding, or the availability of the site for active transcription / regulation. However, these data are hugely variable across tissues. If one considers a complex disease such as hypertension, how is it possible to decide which tissues are particularly relevant to this trait? Genetic influences may be exerted via tissues as diverse as the aorta, fibroblasts, brain, different endocrine organs, renal tissue, adipose tissue, reproductive tissue (and resulting endocrine changes) – the list goes on. There are also many tissues which are not well-represented in the major databases of functional and regulatory data: even the aorta is not as well-covered as more accessible tissues such as skin or cultured fibroblasts. Nevertheless, the open availability of such rich datasets to researchers enables one to gain at least an idea as to the likelihood of a particular variant being situated in a functionally important region. If the tools are not yet quite sufficiently accurate to make clinical judgements on, they are at least useful to narrow down a field of potential candidates

in either an individual for pathogenicity or in results from experimental studies to prioritise variants for further follow-up.

10.9.5 Common variant analysis: genome-wide association studies

Since the first large-scale GWAS was published in 2005[276], there has been a huge expansion in the use of these studies to investigate both diseases and traits of interest. The great promise of GWAS is that it enables testing of genotype: phenotype association across the genome for common variants with no *a priori* assumptions about the biological pathways or genes involved. The simplest model is case: control analysis. Here, there is a simple premise: if a variant is causative of higher risk of a disease, then the frequency of that variant should be higher in those with the disease than in those without. Similarly, for quantitative trait analysis, the underlying assumption is that if a variant has significant impact on phenotype, then individuals carrying that variant, will, on average across a large population, have a different distribution of measurements than those without the variant. The effect size of each individual common SNP may be very small, but the cumulative effect of many common SNPs can be very large.

This power in combination has been harnessed in polygenic risk scores derived from GWAS. Here, a polygenic model of disease risk seeks to finely predict an individual disease risk from several, or all genotyped markers. The hypothesis is that these polygenic risk scores may be used to inform future management and risk modification decisions for individual patients, on a personalised basis. Some developed polygenic risk scores (PRS) can differentiate risk well; for coronary artery disease, PRS can identify 8% of the population with a threefold increased risk; equivalent to the increase in risk conferred by a monogenic disorder: familial hypercholesterolaemia[277]. This allows targeting of preventative measures. However, detractors say that their clinical utility has yet to be shown; they remain relatively expensive to test at the population levels required to identify at-risk individuals. PRS represent much smaller percentages of overall risk than, say, exercise and lifestyle factors. Whilst monogenic disorders give a clear focus for risk reduction, PRS are an overall measure of multifactorial risk and so do not necessarily guide specific interventions beyond standard risk factor

management – indeed it is not yet clear to what extent meticulous risk factor control can mitigate genomic risk. Future work by groups such as Kathiresan's will undoubtedly aim to answer these questions in the near future.

Aside from PRS, GWASs have been hugely successful at identifying variants and genes associated with diseases and quantitative traits across a huge range of pathology and biology. To date, over 50,000 significant associations have been discovered[278]. New disease genes and pathways have been identified, such as the role of autophagy in Crohn's disease[279, 280], or the role of *Complement Factor H* SNPs (*CFH*), implicating complement activation in age-related macular degeneration[276]. The latter observation has made possible the identification of high-risk individuals who might benefit from surveillance and early treatment. Follow-up functional studies guided by GWAS may also identify drug targets, and there are many drugs currently in development based on pathways and genes identified by GWAS[281]. Additionally, GWAS can help to narrow down the search for genes in which variants can cause Mendelian forms of disease. Nearly one fifth of loci discovered by GWAS include a gene which is implicated in a related monogenic or oligogenic disorder[282]. This approach has proved fruitful in diseases such as obesity, where associations with *SH2B1*, *NPC1* and *ADCY3* have been discovered through GWAS, and these genes have later been implicated in monogenic forms of the disorder[283-285].

However, with the hypothesis-free nature of GWAS comes major challenges. One major limitation is the need for very large cohorts to deliver the statistical power to detect significant associations, given the need for multiple testing corrections across the millions of single nucleotide polymorphisms (SNPs) under investigation[125]. With large cohorts, and with no prior hypotheses, the precision of both the genotyping strategy and the phenotype becomes of vital importance. Meticulous quality control at each step of the process is mandatory, to avoid false positive associations, and, where cohorts are combined, to ensure that phenotyping, genotyping and imputation have been matched as closely as possible across cohorts. Minor violations of statistical models can produce great alterations in type I errors, and so it is really important that models are well-specified, that careful individual, SNP and cohort-level quality control checks are all performed, and that results are interpreted with an awareness of the complexities and limitations of the study design.

10.9.5.1 GWAS functional follow-up

The other major challenge of GWAS is in the interpretation and application of the results. More often than not, significant associations fall outside the coding region of the genome and therefore functional studies are required to identify the genes through which phenotype is affected at these loci. Even where associated SNPs fall within a coding region, or are intronic, it often takes painstaking work to untangle the underlying biology. One clear example of this is the mis-named "*FTO* locus" association with obesity. This was identified early in the history of GWAS, in 2007, and it took a further 8 years to appreciate that the causative variant worked not through *FTO* (*Fat Mass And Obesity Associated Alpha-Ketoglutarate Dependent Dioxygenase*) – the gene in which associated variants were found, but by regulation of *ARID5B* repression on expression of *IRX3* and *IRX5*[286]. The intronic variant in *FTO* actually disrupted a key motif for the *ARID5B* repressor. This took a hugely complex programme of research, integrating epigenomics, comparative genomics, human genetics, genome editing, and directed perturbations in samples from patients and mice. Very careful annotation and interpretation of GWAS results is necessary to identify the true drivers of phenotypic variation. Epigenetic modifications, expression data, chromatin conformation capture, and proteomic data can all be integrated to allow this careful dissection of results. The development over recent years of large-scale data repositories such as GTEx (Genotype-Tissue Expression Project)[29, 265], ENCODE (Encyclopaedia of DNA Elements)[258], Roadmap epigenomics[287], as well as integrative databases such as RegulomeDB[32] or Haploreg[33, 34], has enabled much of this to be done in silico, to narrow down likely causative genes or variants linked with the top hits from GWASs. Some of these resources are discussed further in Chapter 13.

Once a narrower range of candidates is identified, then a return to the lab, using knock-out or knock-in models, CRISPR-Cas9 editing or other cellular models according to phenotype, is still considered best practice for proof of causative effect. This approach relies on a high-confidence intermediate cellular or molecular phenotype being identified, or on the existence of reliable animal models of the trait / disease in question.

10.9.5.2 GWAS – not all it's cracked up to be?

Some believe that GWAS has become a victim of its own success; it has not failed to identify new loci, rather, it has identified too many. Over the last decade of GWAS, larger and larger sample sizes have provided the ability to detect variants with smaller and smaller effect sizes on a trait. The most extreme example of this to date is the recent meta-analysis of blood pressure indices, combining 4 studies to give a final sample size of >1 million individuals[266]. This study reported association of BP with >900 separate genomic loci, representing almost 1/20th of all known protein-coding loci in the human genome. Hence, for blood pressure, GWAS has identified loci which might influence renal biology, adrenal function, vascular tone, oxidative stress, cardiovascular development – the list goes on. Essentially, the message is that blood pressure is influenced by a bewildering variety of systems and processes in the body. And yet, in the >1million subject study cited above, this common variant influence at these >900 loci accounts for just 5% of the trait's variance overall. Other complex traits have proven equally polygenic. In other words, functional variation in almost any part of the genome is likely to have a measurable association with any given trait, even if that effect is infinitesimally small. The larger the cohort, the smaller the phenotypic effects one picks up – and therefore even SNPs which have a minor effect on, say, alcohol intake, might appear as statistically significant hits in a GWAS of blood pressure.

To what extent our understanding of complex trait biology can be enhanced by the discovery of loci with such infinitesimally small effect sizes remains uncertain. However, the detection of so many associations gives rise to the distinct possibility of omni-genic traits, whereby all genetic variants (or at least all variants in genes expressed in relevant tissues) could be associated with phenotype if the sample were large enough. The interpretation of such results is therefore increasingly difficult, and the biological or clinical applicability becomes rather more distant. As effect sizes become smaller, the routes of causative influence of SNPs on phenotype become more circuitous. These observations are of course no less valid than larger effect sizes, but certainly much more challenging to interpret, study, apply to clinical situations or modify.

The major challenge of genomics researchers is to prove that the next steps, beyond the massive landmark papers, can yield significant advances in understanding of biology and

disease, and yield new, clinically meaningful risk markers and treatment targets. With this in mind, it is vitally important to acknowledge that the "end-point" of associations derived from a GWAS study should be seen as the *start* of a programme of research, and not as an end in their own right.

10.9.6 Genetic architecture: the challenges

There is often a Linnaean instinct of researchers to classify genetic influences and studies into "rare" and "common" variant work. However, this is a false distinction – there is clearly a continuum of variant frequencies which have a spectrum of effect sizes on disease risk or phenotype. And different study designs must be used to get a true sense of the overall genetic architecture of any complex condition – no GWAS will be able to assess a novel, or private variant's effect on phenotype, whereas a family segregation study will never be able to identify common variants contributing to disease risk.

The challenge to researchers today is not solely or even mainly one of data-gathering, but of data interpretation – carefully dissecting common variant associations from large GWASs, using the wide array of functional datasets to annotate them, assessing the role of rare variants in related pathology, performing functional validation, and putting all the evidence together to create a coherent picture of the genetic and environmental factors which combine to influence phenotype.

This thesis therefore aims to begin this process: to examine aortic biology and pathology with a range of genetic and genomic tools at our disposal, understanding in depth the challenges and limitations of each technique, but trying always to keep an eye on the bigger picture.

11: DETERMINANTS OF HEALTHY AORTIC SIZE AND SHAPE

"Form follows function – that has been misunderstood. Form and function should be one, joined in a spiritual union." – Frank Lloyd Wright

11.1 INTRODUCTION

The modern vision of the aorta as a dynamic organ has led to renewed interest in the inextricable relationship of aortic form and aortic function[1]. Aortic size and shape affect blood flow patterns, and the resultant changes in wall strain and shear stress can cause aortic wall remodelling, in turn altering aortic form and elastic function.

11.1.1 Aortic dimensions

In the clinical setting, aortic diameter remains the key criterion for decision-making and diagnosis of aortic disease. Aortic diameter has the benefit of being an easily acquired, simple metric. Laplace's law reminds us that the wall tension is proportional to radius of the vessel, so that a larger diameter will place the wall under greater tension, regardless of any underlying structural abnormalities. Aortic diameter also is a strong predictor of dissection rate[288, 289]. However, the flaws of relying on this simplistic measure for aortic risk prediction are widely documented[290, 291], particularly in the setting of heritable thoracic aortic disease, where there may be some intrinsic abnormality of aortic structure. The majority of aortic dissections occur at aortic diameters below the traditional surgical cut-offs[290], and these basic numeric measures do not take account of anthropometric effects on dimensions: gender, ethnicity, age and body composition being key influences. A more nuanced approach to use of aortic dimensions takes account of these additional variables to define normal values[292, 293]. However, limited series using this subtler approach have been published to date. Guidelines still refer to absolute values for surgical decision-making[191, 294, 295], although in practice, and in more up-to-date guidance[295], indexed values are widely used.

Above and beyond their importance in thoracic aortic disease, aortic dimensions have also been shown to correlate with cardiovascular risk. The PAMELA, Jackson Heart and

Framingham studies have demonstrated that aortic root size is a predictor of cardiovascular events in a range of populations[14-16].

Many of the normal ranges for aortic dimensions have been derived from echocardiographic or CT series. There have been surprisingly few large-scale studies of normal aortic dimensions measured by MRI, and still fewer where body size has been accounted for in the published normal ranges[102, 296]. Several studies have examined what biometric variables affect aortic size. There are well-established effects of age, ethnicity, body surface area and other cardiovascular risk factors on aortic dimensions[102]. However, there are still some discrepancies between the different reports of blood pressure indices associated with aortic dimensions.

11.1.2 Aortic morphology

Whilst aortic dimensions are relatively well-documented (with some important caveats), the importance of aortic form has only recently been acknowledged. Aortic length, width, diameter, angulation and tortuosity are not static, but fluid and changing through life. All these parameters change with vascular ageing, can have profound effects on aortic flow patterns[17, 297, 298], and may contribute to, or at least correlate with, cardiovascular risk. However, the morphology of the healthy aorta is relatively under-studied. As Figure 10.3 (see Chapter 10) demonstrates, even in healthy volunteers, there is a huge variation in the shape, width, angulation and even branching pattern of the thoracic aorta. There is increasing interest in the importance of measurable morphological variables such as arch angle, aortic taper and arch height:width ratio, in risk of aortic and cardiovascular events, and there is evidence that these variables are not static throughout life. Instead, it has become apparent that the aortic arch undergoes remodelling and re-shaping throughout life, with the arch broadening, lengthening and the taper reducing with vascular ageing[17, 18, 299]. In mice, atherosclerosis extent correlates with degree of aortic curvature[300], and in humans, aortic geometry may contribute to stroke and dissection risk[132, 297], and is associated with the degree of calcification in the aortic arch [301]. It is uncertain whether differences in flow patterns due to the altered arch geometry, or intrinsic underlying developmental mechanisms are responsible for this increased risk.

This chapter will present aortic phenotypes from of a cohort of 1299 healthy Caucasian volunteers, recruited as part of the Digital Heart Project (DHP). DHP was originally conceived to enable detailed cardiovascular imaging of a carefully-selected group of *healthy* adults, and to examine genotype:phenotype correlations for a wide range of clinically-relevant variables. Several aortic phenotypes were measured using cardiovascular MRI, and the anthropometric and biometric variables which influence them were examined. Simple mixed regression models were then developed to contribute to genome-wide association studies of some of these aortic traits (presented in Chapters 13 and 14).

11.1.3 Studying a healthy population

Whilst larger, biobank-scale cohort studies have recruited adults for similar phenotyping since the conception of this project (e.g. UK Biobank;[302] see chapter 17), our own cohort has a number of particular features which set it apart. Perhaps most importantly, this was a cohort screened to be healthy, in contrast to the "population" cohorts recruited by other studies. These individuals were also relatively young (mean age: 40), enabling assessment of phenotype over the whole adult lifespan (from 18 to 80 years). In other large population studies such as MESA[102] and UK Biobank[302], recruitment started at 45 years of age.

Studying phenotype:genotype correlation in an exclusively healthy population has a number of advantages:

- Normal healthy ranges of phenotype can easily be defined (usually as values within the 95% confidence intervals for the cohort)
- Normal relationships between biometric variables can be discovered without confounding from pathological outliers or medication.
- Theoretically, a younger and healthier cohort will consist of individuals for whom phenotype is less confounded by environmental factors such as medications, smoking etc. Therefore a more robust relationship may be detected between phenotype and genotype.

In essence, a young, healthy cohort can give a clearer insight into the "pure" biology of the phenotypes and genomic influences under study. It may be a less directly translatable way to

research the basis of cardiovascular disease, but it will tell us about the normal biology of the aorta, without or before the influence of accelerated ageing, disease and risk factors.

11.1.4 Aortic phenotyping: CMR

In large studies such as the DHP, deciding what and how to measure, is one of the most challenging steps. As many different subprojects have different aims, it is important to ensure that each study has sufficient information, without making the task of data-gathering hugely time-consuming and unwieldy.

Cross-sectional imaging techniques have become the gold standard for clinical evaluation of cardiac function and the aorta. Cardiovascular MRI (CMR) can be used for accurate and comprehensive evaluation of aortic and cardiac phenotype, without exposure to ionizing radiation, and without the limitations imposed by acoustic windows with echocardiography. CMR also allows integration of the assessments of aortic structure, shape and function, with cine images and flow velocity mapping allowing rapid measurement of distensibility and pulse wave velocity, with minimal extra scanning time. CMR-derived values for aortic dimensions have been reasonably well-defined, although there remains some debate about the best way to report these. In clinical reporting, it is common practice to report raw dimensions as well as values indexed to body size. However, there is little consensus on the best metric for indexation, with body surface area widely used in clinical practice and in guidelines, but height being increasingly studied also. There is increasing interest in using 3D models and indeed 4D flow CMR for more detailed clinical assessment and phenotyping, and these will form the basis of future work.

11.2 AIMS

- To measure aortic dimensions and morphology in a healthy cohort using cardiovascular magnetic resonance imaging
- To define healthy aortic size and shape
- To define the influence of anthropometric and basic biometric measures on aortic size and shape

11.3 METHODS

11.3.1 Digital Heart Project: recruitment

In total 1776 healthy adult volunteers were recruited prospectively to the Digital Heart Project at Imperial College London (for cohort characteristics, see Table 11.1 below). At screening, exclusion criteria included known cardiovascular disease, treatment for cardiovascular risk factors (hypertension, hyperlipidaemia or diabetes), regular prescription medication (simple analgaesics, antihistamines and oral contraceptives were acceptable). Female subjects were excluded if pregnant or breast-feeding. Standard published safety contraindications to MRI were applied, including an absolute weight limit of 120kg. Two patients were removed from the study after imaging revealed significant cardiovascular pathology. This left 1774 healthy individuals, of whom 1299 were Caucasian. All subjects provided informed written consent for participation in the study, which was approved by the local research ethics committee.

11.3.2 Digital Heart Project: biometric assessment

All measurements were performed by specially-trained cardiac nurses at Hammersmith Hospital (the study centre). Participants' height and weight were measured without shoes, whilst wearing scrubs. Body Mass Index (BMI) was calculated as the total weight(kg) divided by height (in metres) squared. Body Surface Area (BSA) was calculated according to the Mosteller formula as follows: BSA (m^2) = (Height(cm) x Weight(kg) / 3600)^{½} [303]. Each subject was fasted for 4 hours prior to the visit. Total body fat mass was measured with multi-frequency bioelectrical impedance analysis (InBody 230, BioSpace, Los Angeles, CA) [304] and expressed as a percentage of the participant's total body weight.

Brachial BP measurement was performed according to European Society of Hypertension guidelines[305] after 5 minutes' rest, using a calibrated oscillometric device (Omron M7). Three measurements were taken. The first of these was discarded and the second two values were averaged. Mean arterial pressure (MAP) was calculated as [(2×diastolic pressure)+systolic pressure]/3. An electrocardiogram (ECG) was undertaken during the visit.

Volunteers self-reported their ethnic background and completed a brief questionnaire which asked questions about activity levels, prematurity, socioeconomic background, rural versus urban upbringing and smoking / alcohol history.

Blood samples were taken during the visit for DNA extraction and genotyping as described in Chapter 13, and serum was stored at -80^{o}C for lipid assays, performed by the Clinical Biochemistry laboratory at Royal Brompton Hospital.

Cohort characteristics are given in Table 11.1 below:

TABLE 11.1: Anthropometric characteristics of Digital Heart Project Cohort

Total Number	1774
Ethnicity n, (%)	
Afro-Caribbean	58 (3.3)
Asian Subcontinent	250 (13.5)
Caucasian	1299 (73.2)
Chinese	29 (1.6)
Japanese	3 (0.2)
Malaysian	55 (3.1)
African	59 (3.3)
Other Unspecified	21 (1.2)
Age (SD, range)	40 yrs (13.1, 18-80)
Female gender n,(%)	976 (55)
Mean height (SD)	170cm (9.7)
Mean weight (SD)	71.1kg (13.7)

A further description of the biometrics of the cohort is given in 11.4.3 (Table 11.4), comparing values by gender for the cohort.

11.3.3 Digital Heart Project: Cardiovascular Magnetic Resonance imaging (CMR)

CMR was performed on a 1.5T Philips Achieva system (Best, Netherlands). The maximum gradient strength was 33 mT/m and the maximum slew rate 160 mT/m/ms. A 32-element cardiac phased-array coil was used for signal reception. Scout images were obtained and used to plan 2D cine balanced steady-state free precession (b-SSFP) images in the left ventricular short axis (LVSA) plane from base to apex using the following parameters: field-of-view 370 mm×370 mm, repetition time/echo time 3.0/1.5 msec; flip angle 60°; bandwidth 1250

Hz/pixel; acquired pixel size 2.0 × 2.2 mm; section thickness 8 mm with a 2 mm gap; reconstructed voxel size 1.2 × 1.2 × 8 mm; number of sections 10 – 12; cardiac phases 30. A single breath-hold 3D LVSA b-SSFP sequence was acquired in the same orientation using the following parameters: repetition time/echo time 3.0/1.5 msec; flip angle 50°; bandwidth 1250 Hz/pixel; pixel size 2.0 × 2.0 mm; section thickness 2 mm overlapping; reconstructed voxel size 1.2 × 1.2 × 2 mm; number of sections 50 – 60; cardiac phases 20; sensitivity encoding (SENSE) factor 2.0 anterior-posterior and 2.0 right-left direction. The LVOT view was acquired using a retrospectively ECG-gated breath-hold b-SSFP sequence. The imaging slice was prescribed to bisect the aortic valve and aortic root on the basal LVSA image then angulated to pass through the LV apex on the 2 chamber image, using the following parameters: field-of-view: 380x380mm; matrix: 256 x 256; repetition time: 3ms; echo time: 1.5ms; flip angle: 60°; bandwidth: 962; slice thickness: 8 mm; reconstructed voxel size 1.5 x 1.5 x 8 mm; cardiac phases 50.

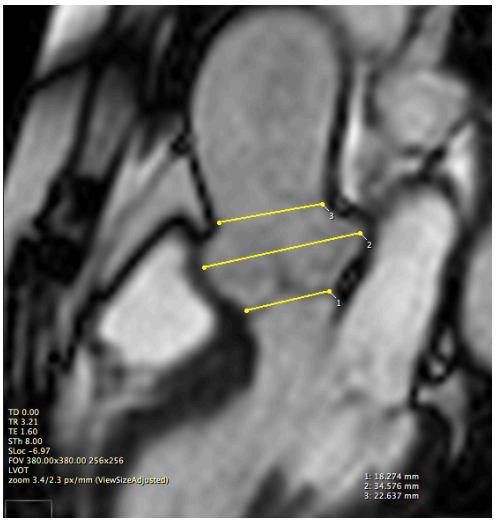
Phase-contrast sequences were acquired at the level of the pulmonary bifurcation, perpendicular to both the ascending and the descending thoracic aorta, enabling simultaneous study of both vessels. The phase-contrast data were acquired using a retrospectively ECG-gated breath-hold sequence with a through-plane velocity-encoding gradient of 200 cm/s. The sequence parameters were as follows: field-of-view 370 mm×370 mm, repetition time 2.8 ms, echo time 1.4 ms, flip angle 15°, and voxel size 1.65 mm×1.92 mm×10 mm, with a temporal resolution of approximately 20 ms (with interpolation to 10ms). For the calculation of aortic length, ECG-gated balanced steady state–free precession images were acquired through the thoracic aorta using the following parameters: field-of-view 320 mm×320 mm, repetition time 3.4 ms, echo time 1.7 ms, flip angle 60°, and voxel size 1.65 mm×1.92 mm×10 mm. Images were curated on an open-source image database (MRIdb, Imperial College London, UK) [306].

11.3.4 Derived measures

The Digital Heart project was originally conceived to collect comprehensive cardiac phenotyping data from healthy individuals, so the imaging protocol was not specifically tailored to measure aortic phenotype. Basic aortic images were obtained. These enabled assessment of aortic root diameters only in a single plane. The LVOT view was used for the

aortic root diameters as there was no short axis view of the root for cross-sectional measurements (see Figure 11.1 below). This LVOT / aortic root measurement is not consistent with SCMR or ESC guidelines on CMR measurement of the aortic root[307]. Therefore care needs to be taken in the interpretation of these measurements, as they are dependent upon the exact orientation of the long-axis plane used. Nevertheless, this view correlates with measurements commonly taken during echocardiographic evaluation of the aortic root, and so it is reasonable to use it for within-cohort evaluation of this phenotype.

FIGURE 11.1: Measurement of aortic root diameters

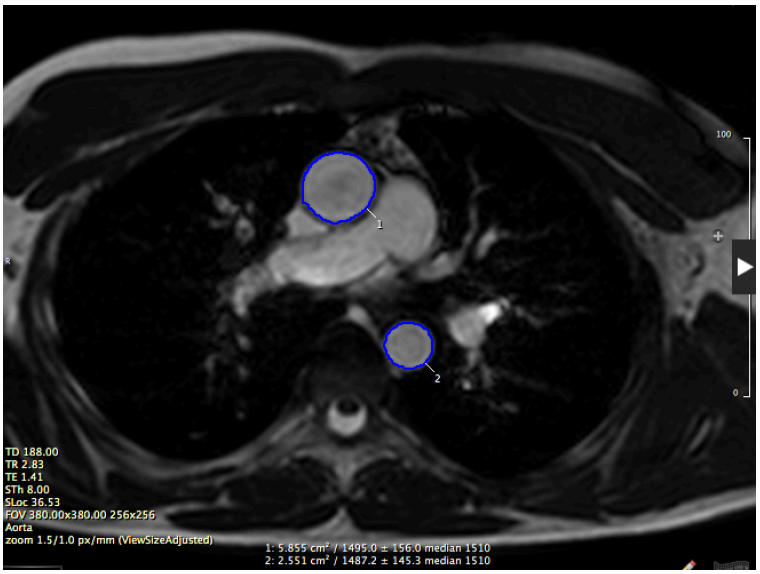


Screenshot from CVI42 of aortic root diameter measurements in end diastole in the left ventricular outflow tract (LVOT) view. 1: Aortic valve annulus; 2: Sinuses of Valsalva; 3: Sino-tubular junction

The diameters and areas derived from the ascending and descending aorta were obtained in line with ESC / SCMR recommendations. These measurements were taken using axial cine images obtained orthogonal to the sagittal oblique scout for the ascending aorta at the level of the pulmonary bifurcation. CVI42 software (Circle Cardiovascular Imaging Inc.) was used for automated edge detection, and the contour was manually checked and refined for each individual. Areas were automatically calculated from these contours. Diameters were derived using the formula below. Minimum (diastolic) dimensions were reported unless otherwise indicated.

$$\text{Diameter(mm)} = 2x\sqrt{\frac{Area(mm^2)}{\pi}}$$

FIGURE 11.2: Still image from axial cine demonstrating measurement of ascending and descending aortic areas



Screenshot from CVI42 showing a still from axial cine images taken at the level of the pulmonary bifurcation. Blue "circles" denote automated contouring (checked manually) of ascending (1) and descending (2) aorta.

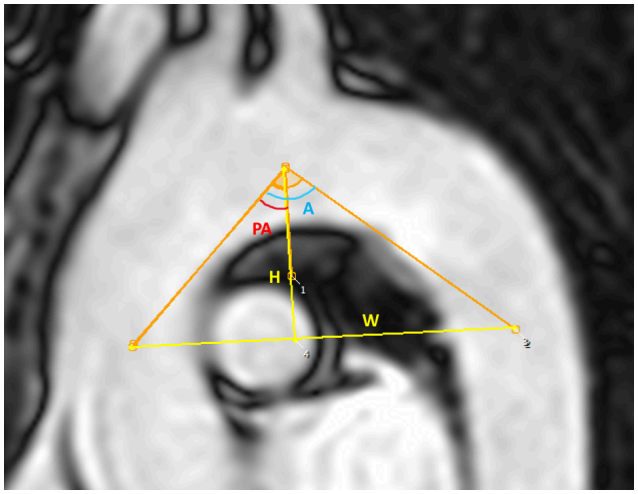
Taper was defined as the ratio of ascending aortic area to descending aortic area.

Views of the aortic arch anatomy were also limited. The arch view was limited to a single "candy-cane" view in a single plane. This view was used in CVI42 to examine arch morphology, including the height:width ratio, aortic arch angle and aortic arch symmetry, defined as shown in the diagram and formulae below (Figure 11.3) and in line with previous studies[18, 299, 308].

$$\text{Height:Width ratio} = \frac{H}{W} \quad \text{(see Figure 11.3)}$$

$$\text{Arch symmetry} = \frac{PA}{A} \quad \text{(see Figure 11.3)}$$

FIGURE 11.3: Aortic arch "candy-cane" view demonstrating aortic morphological variables



Aortic width (W) was defined as the distance in millimetres between the mid-point of the ascending and descending aorta, at the level of the right pulmonary artery, in a plane orthogonal to the ascending and descending aortas, corresponding to the plane of the axial aortic cine and phase contrast images. Aortic height (H) was measured in an orthogonal plane to the width, between this line and the midpoint of the highest point of the aortic arch. Aortic arch angle (A) was defined as the angle between the midpoint of ascending aorta, the midpoint of the peak of the aorta and the midpoint of descending aorta. The proximal arch angle (PA) was defined by dividing this aortic arch angle into proximal and distal components by the line defining aortic height, to enable assessment of aortic arch symmetry.

11.3.5 Reliability

A second observer measured aortic variables from 50 subjects, following the same protocol as observer 1. Observer 2 was blind to the results obtained by observer 1. Observer 1 also repeated 50 measurements, blinded to previous measurements. Interobserver and intra-observer reliability was calculated using the irr and Bland-Altman-Leh packages in R.

Intra-class correlation coefficients (twoway, agreement) with 95% confidence intervals are presented in the results sections below. The coefficient of variation was calculated by dividing the standard deviation of the differences between the two sets of measurements by the mean value of the measurement.

11.3.6 Statistical analysis

Normal values were defined as those falling between the 5th and 95th centile for our cohort. These were plotted as nomograms using the VGAM package in R, which fits a vector-generalised additive model to the data, using a vector (cubic smoothing spline) smoother to define the smoothed lines representing 5th, 50th and 95th centiles of the data.

Multiple linear regression analysis was performed in R (R version 3.5.1; R studio version 1.1.456). Distributions of variables were manually inspected using histograms, density plots and quantile-quantile (QQ plots) using the qqnorm function in R. For each measured covariate and phenotype, significant outliers were removed, and appropriate cut-offs for this (either >3SDs, 4SDs or 5SDs from the mean) were defined with reference to the accuracy of the measurement itself and the distribution of the variable.

Multiple regression models were constructed using the lm() function from the basic statistical package in R for each measured phenotype. Where phenotypes were not normally distributed (ascending and descending aortic dimensions), they were log-transformed prior to construction of the regression model. Normality of residuals was checked using histograms and QQ plots.

For raw dimensions, the following models were constructed:

Basic model:

Variable ~ Age + Gender + Height

Extended model:

Variable ~ Age + Gender + Height + single additional covariate

Selected model:

Variable ~ Age + Gender + selected significant covariates

The “selected model” was defined using a supervised stepwise model selection, in full awareness of the limitations of this technique. A single “best” blood pressure variable was chosen where appropriate. The stepAIC() function in R was then used to define a model based on stepwise iteration of the different model parameters, derived from the “full” model

containing all of the possible covariates. Any significantly collinear variables were removed from the model (using, for example, only one measure of body size), and covariates added if there were good biological reasons to include any which had been omitted. The use of this model is purely to define an approximately "maximum" valid adjusted R^2 and to examine approximately what proportion of variance could be predicted by a "best" selection of covariates; it will not be used for genetic studies, nor for prediction.

Height was used as a proxy for body size in the basic model, rather than body surface area, as this is a non-modifiable variable and avoids attenuation of the expected relationships with fat mass and body composition.

Also presented are multiple regression models for aortic dimensions indexed to body surface area: here, the regression models exclude height or other body size variables.

Effect sizes are presented as a raw effect size (unit change in the dependent variable per unit change in the independent variable) and as a partially-standardised effect size (unit change in the dependent variable per standard deviation change in the independent variable). Where variables were log-transformed, the effect size derived from the model created using the log-transformed variable is presented, which is not directly translatable into a "raw" effect size.

R^2 measures are presented to approximate the proportion of variance in the phenotype which is predicted by the measured covariates included in the model, split into non-modifiable (in the basic model) and modifiable (in the selected model).

For between-gender effects, a Kruskal-Wallis test of variance was performed to determine the significance of the difference between the two groups.

11.4 RESULTS: aortic dimensions

To determine the normal healthy range of aortic dimensions and the biometric variables which affect them, 1299 healthy Caucasian volunteers underwent aortic phenotyping using cardiovascular magnetic resonance imaging (CMR).

11.4.1 Reliability

Inter- and intra-observer reliability for aortic dimensions was in general excellent. The exception to this was the aortic valve annulus, which had an intra-class correlation of 0.81 between observers and wide confidence intervals. This is likely reflective of the sub-optimal method used to derive this measure, using a single plane at the LVOT. The aortic valve is often obliquely situated, and its complex 3-dimensional shape and structure means that a 2D value captures limited information which is variable according to the exact plane used for measurement.

TABLE 11.2. Good inter-observer reliability of aortic dimensions

	Inter-observer		Intra-observer	
Phenotype	**ICC (SD)**	**Coefficient of variability**	**ICC (SD)**	**Coefficient of variability**
Valve annulus diameter	0.81 (0.52-0.92)	4.87	0.84 (0.72-0.91)	5.10
SoV diameter	0.97 (0.93-0.98)	2.51	0.96 (0.92-0.98)	3.21
STJ diameter	0.91 (0.83 - 0.96)	5.13	0.96 (0.92-0.98)	4.88
AA area	0.98 (0.97 - 0.99)	4.04	0.97 (0.95-0.99)	4.26
DA area	0.96 (0.92 - 0.98)	6.03	0.94 (0.91-0.98)	6.59

ICC: intra-class correlation coefficient; SD: standard deviation; SoV: Sinuses of Valsalva; STJ: Sino-tubular junction; AA: Ascending aorta at level of pulmonary bifurcation; DA: Descending aorta at level of pulmonary bifurcation.

11.4.2 Defining normal values

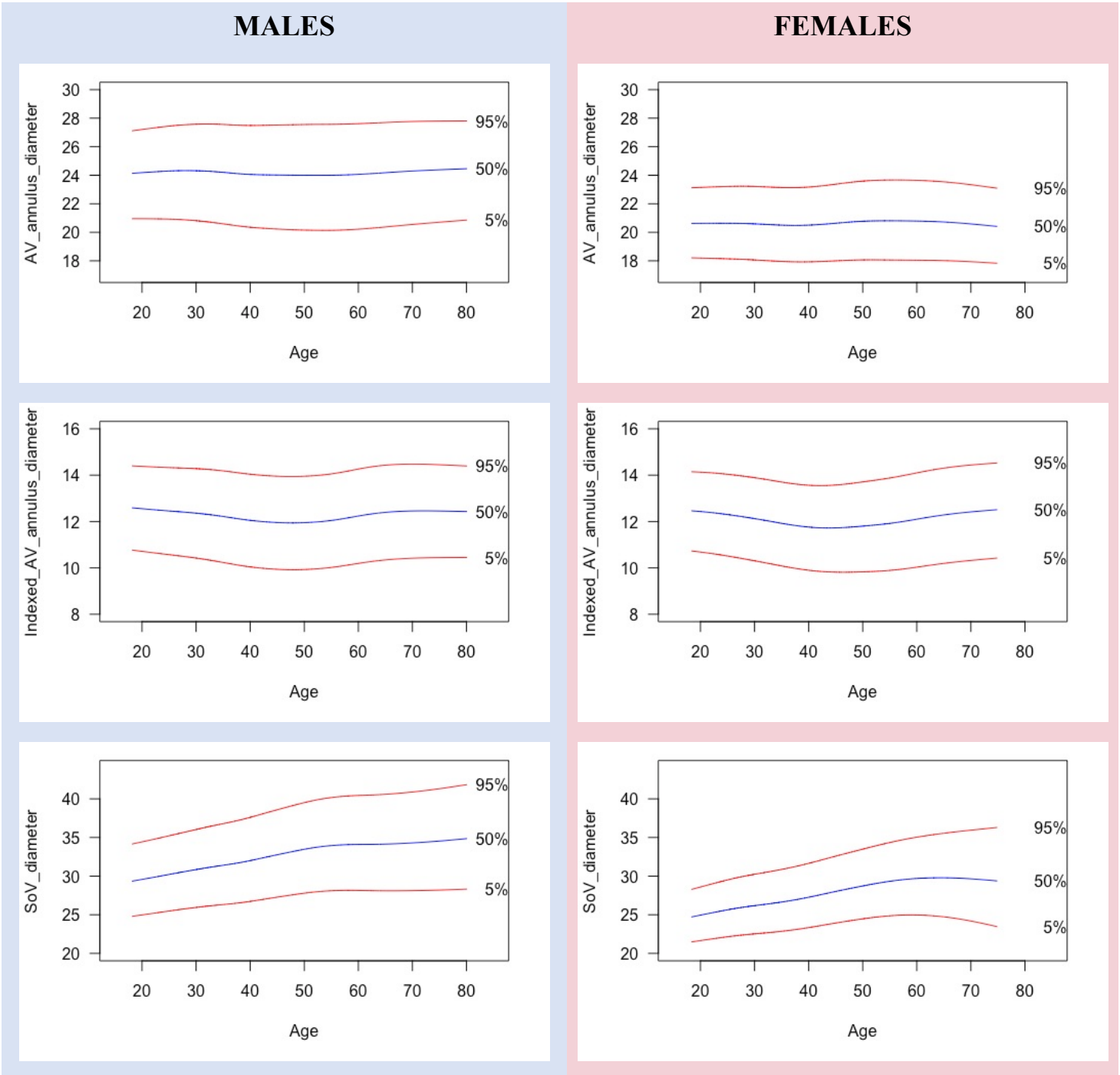
The ranges of our aortic measurements were broadly in line with previously published figures (see Table 11.3). There are surprisingly few detailed studies looking at the normal range for aortic measurements derived from CMR, and those that exist differ in measurement technique. 4 main papers present normal values; 2 of these, the MESA study[102], and the study by Redheuil et al[18] used mean aortic diameter (systolic-diastolic diameter /2), rather than diastolic, and looked exclusively at ascending aortic dimensions. Mean aortic diameter is not routinely used in clinical reporting. These studies are therefore excluded from Table 11.3 below. The Oxford data[296] provided the most comprehensive diastolic measurements of aorta at many levels – however, it did not provide any indexed reference ranges. The UK Biobank is due to report shortly on reference ranges for aortic dimensions in a population cohort[302]. In the meantime, the Digital Heart Project provides the largest healthy UK Caucasian reference cohort to date, and produces data in keeping with previously published dimensions. The nomograms below (Figure 11.4) allow us to define normal healthy aortic dimensions in a Caucasian population, and could be used as reference data for cardiovascular MRI reporting. It is the first time that such nomograms have been produced for BSA-indexed values derived from MRI.

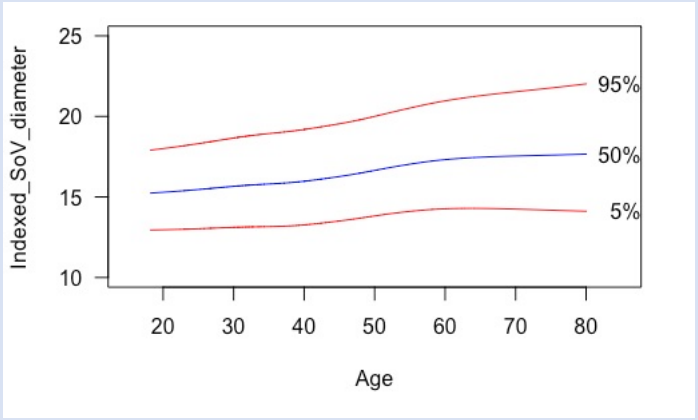
TABLE 11.3: Current data ranges are in line with previously-published studies

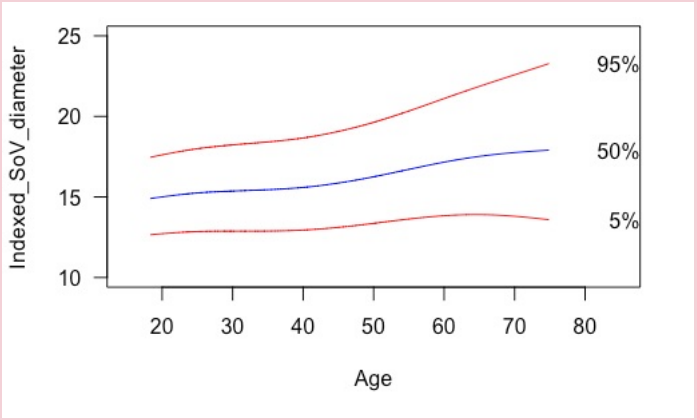
		Digital Heart Project n=1299		Oxford data[296] n=447		Burman et al[309] n=120	
		M	F	M	F	M	F
Valve annulus diameter mean (SD)	Raw	24.1 (2.2)	20.7 (1.6)	24.4 (2.7)	21.0 (1.8)	-	-
	Indexed	12.2 (1.2)	12.0 (1.2)	-	-	-	-
Sinus of Valsalva diameter mean (SD)	Raw	32.1 (3.7)	27.5 (3.0)	32.4 (3.9)	27.6 (2.9)	32.4 (4.2)	29.0 (3.3)
	Indexed	16.3 (1.9)	16.0 (2.0)	-	-	-	-
Sino-Tubular Junction diameter mean (SD)	Raw	25.2 (3.4)	21.9 (3.0)	25.0 (3.7)	21.8 (2.7)	-	-
	Indexed	12.8 (1.7)	12.8 (1.8)	-	-	-	-
Ascending aortic diameter mean (SD)	Raw	27.9 (3.8)	25.2 (3.4)	26.7 (3.9)	25.5 (3.7)	-	-
	Indexed	14.1 (1.9)	14.6 (2.0)	-	-	-	-
Descending aortic diameter mean (SD)	Raw	21.4 (2.5)	18.9 (2.4)	20.6 (2.8)	18.9 (2.0)	-	-
	Indexed	10.8 (1.3)	11.0 (1.4)	-	-	-	-
Ascending aortic area mean (SD)	Raw	626.1 (174.2)	507.2 (139.0)	-	-	-	-
	Indexed	315.3 (82.5)	293.3 (77.5)	-	-	-	-
Descending aortic area mean (SD)	Raw	320.4 (89.9)	284.4 (73.2)	-	-	-	-
	Indexed	183.4 (41.2)	164.4 (39.9)	-	-	-	-

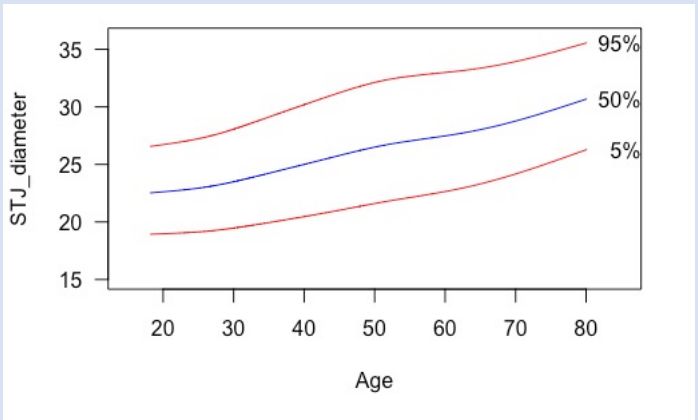
Mean and standard deviation of measured aortic dimensions, presented both raw and indexed to body surface area in line with common clinical practice. Comparison is made with equivalent measurements from the Oxford data[296] and the Burman paper[309] (only directly comparable measurement; aortic root size measured in LVOT view). All dimensions are in mm (diameters), mm/m^2 (indexed diameters), mm^2 (areas) or mm^2/m^2 (indexed areas).

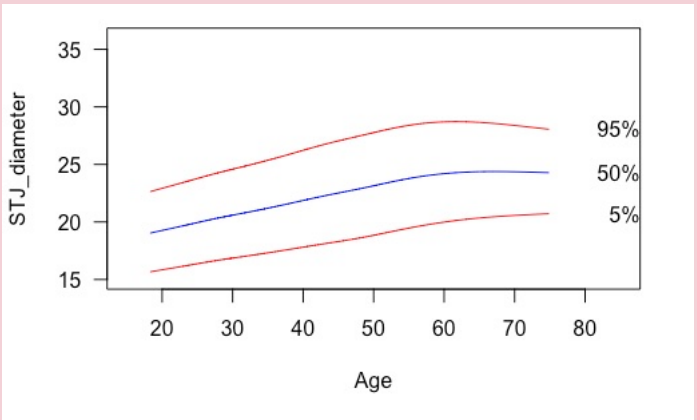
FIGURE 11.4: Nomograms displaying raw and BSA-indexed aortic dimensions

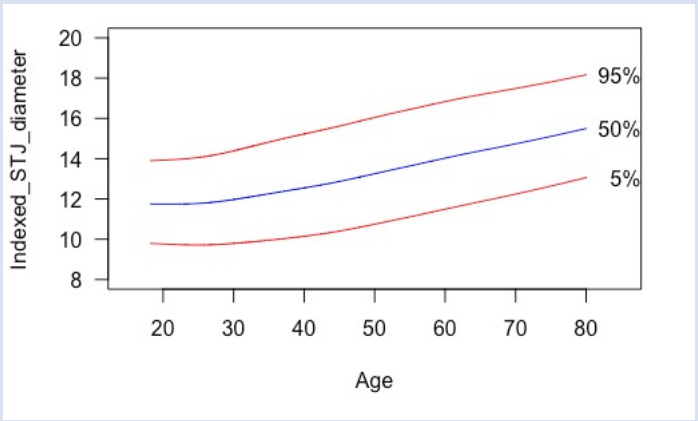


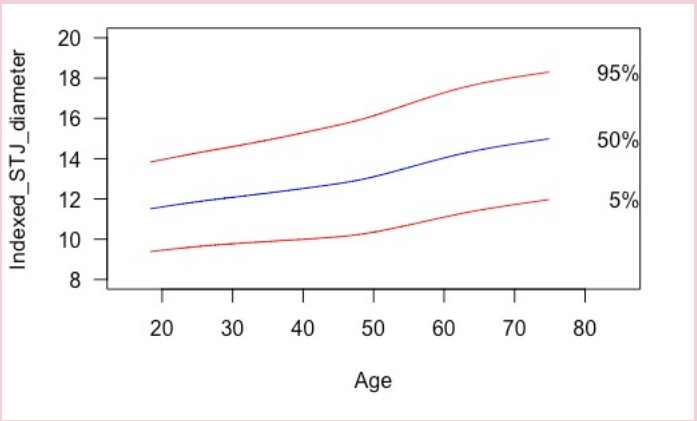

Indexed_SoV_diameter
25
20
15
10
95%
50%
5%
20 30 40 50 60 70 80
Age

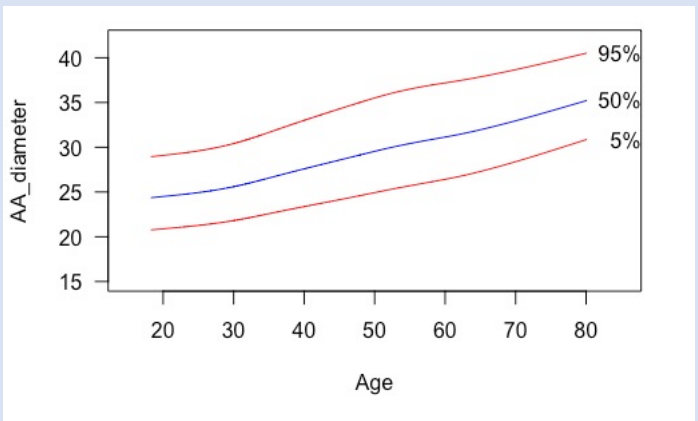

Indexed_SoV_diameter
25
20
15
10
95%
50%
5%
20 30 40 50 60 70 80
Age

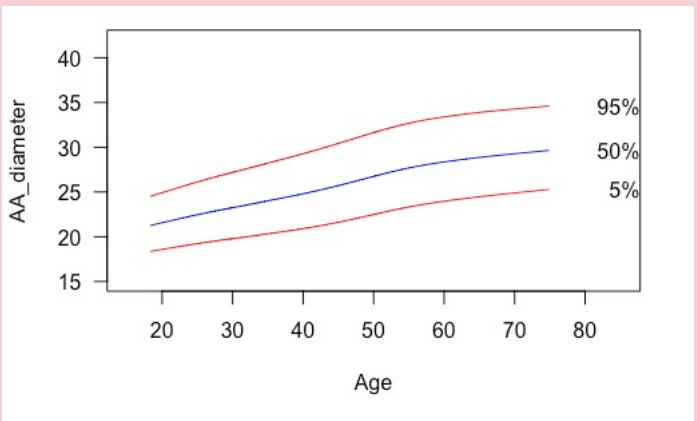

STJ_diameter
35
30
25
20
15
95%
50%
5%
20 30 40 50 60 70 80
Age

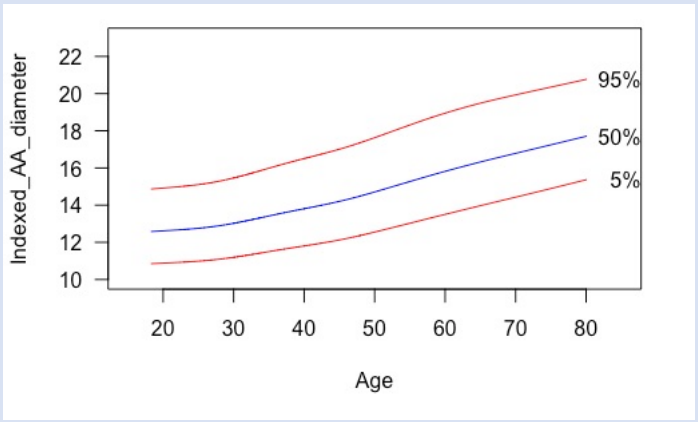

STJ_diameter
35
30
25
20
15
95%
50%
5%
20 30 40 50 60 70 80
Age

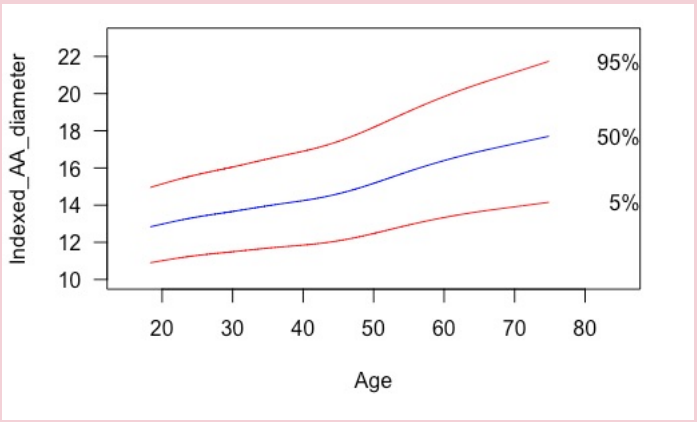

Indexed_STJ_diameter
20
18
16
14
12
10
8
95%
50%
5%
20 30 40 50 60 70 80
Age

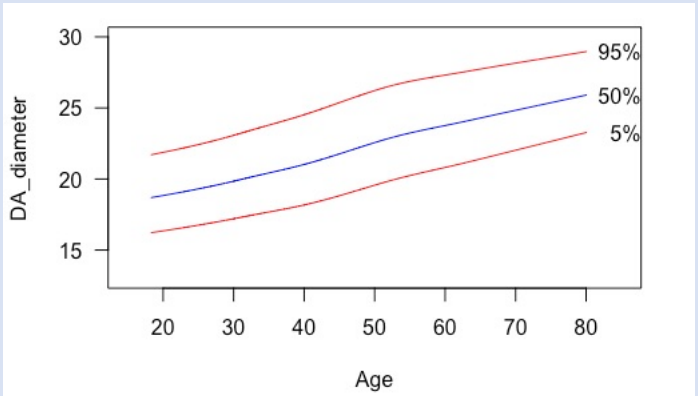

Indexed_STJ_diameter
20
18
16
14
12
10
8
95%
50%
5%
20 30 40 50 60 70 80
Age

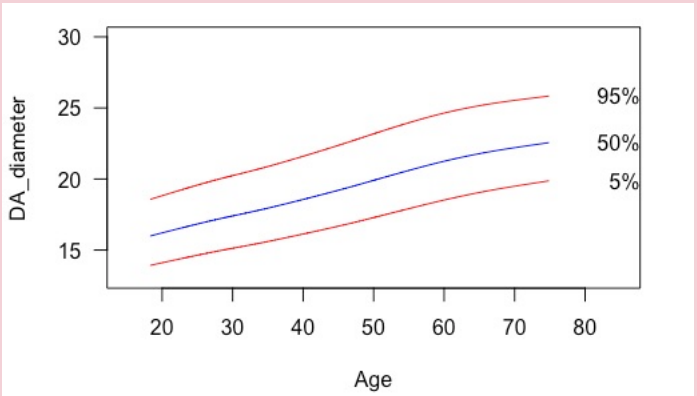

AA_diameter
40
35
30
25
20
15
95%
50%
5%
20 30 40 50 60 70 80
Age

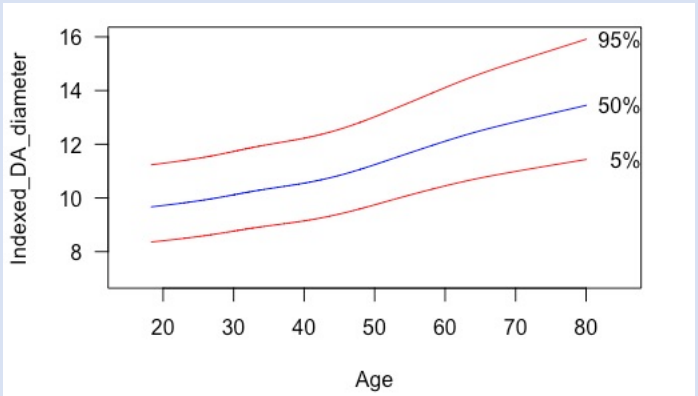

AA_diameter
40
35
30
25
20
15
95%
50%
5%
20 30 40 50 60 70 80
Age

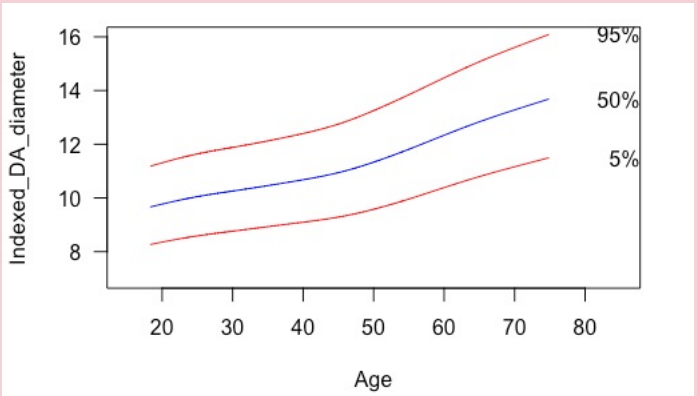

Indexed_AA_diameter
22
20
18
16
14
12
10
95%
50%
5%
20
30
40
50
60
70
80
Age

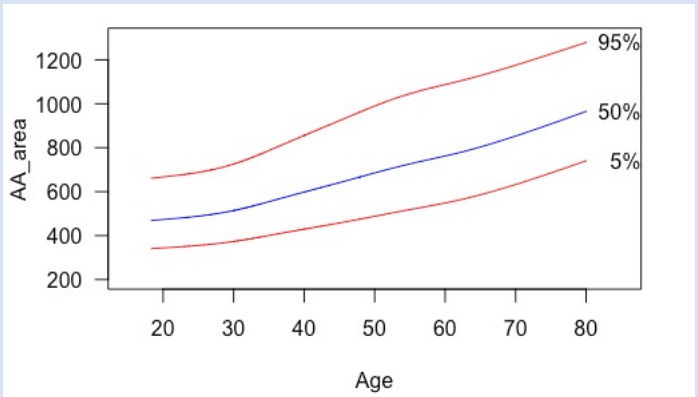

Indexed_AA_diameter
22
20
18
16
14
12
10
95%
50%
5%
20
30
40
50
60
70
80
Age

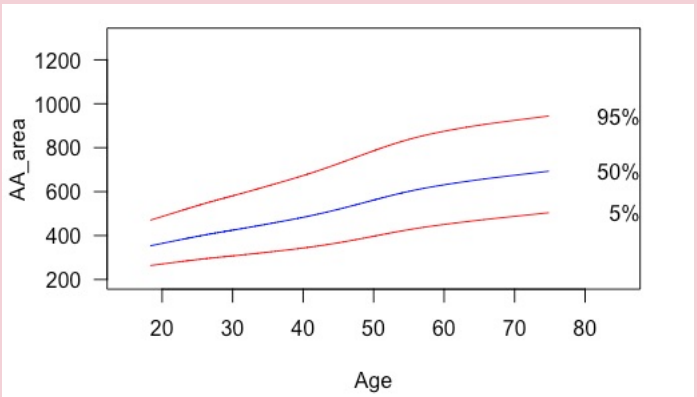

DA_diameter
30
25
20
15
95%
50%
5%
20
30
40
50
60
70
80
Age

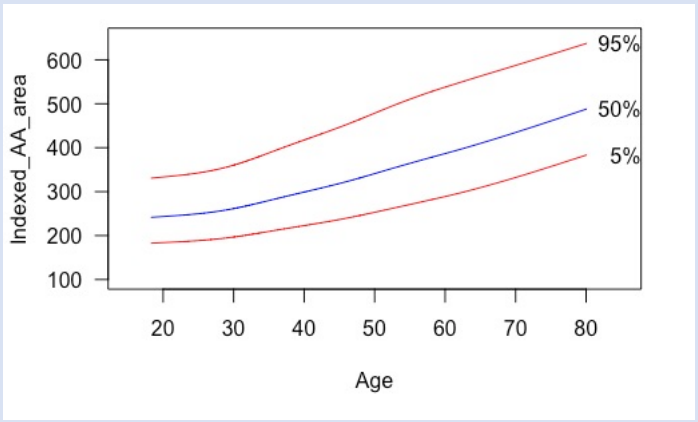

DA_diameter
30
25
20
15
95%
50%
5%
20
30
40
50
60
70
80
Age

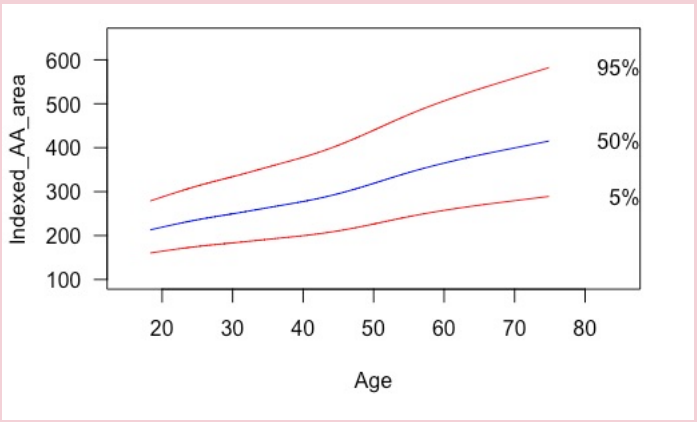

Indexed_DA_diameter
16
14
12
10
8
95%
50%
5%
20
30
40
50
60
70
80
Age

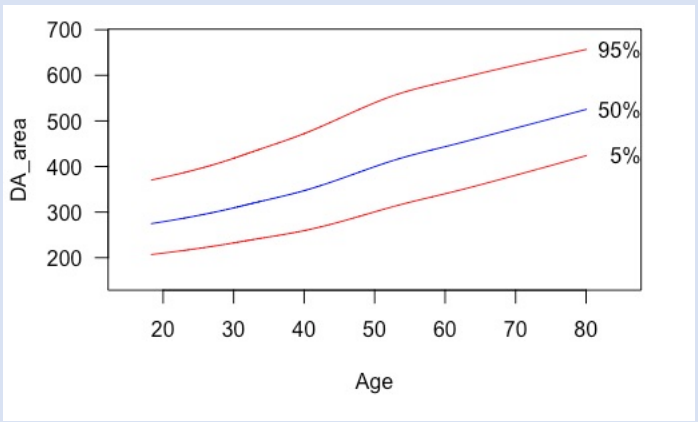

Indexed_DA_diameter
16
14
12
10
8
95%
50%
5%
20
30
40
50
60
70
80
Age

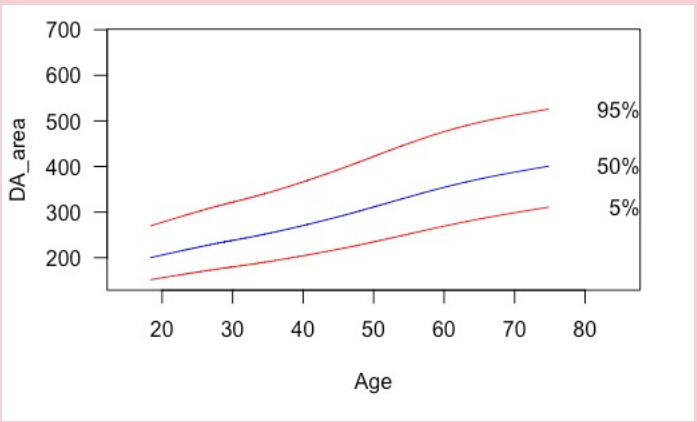

AA_area
1200
1000
800
600
400
200
95%
50%
5%
20
30
40
50
60
70
80
Age

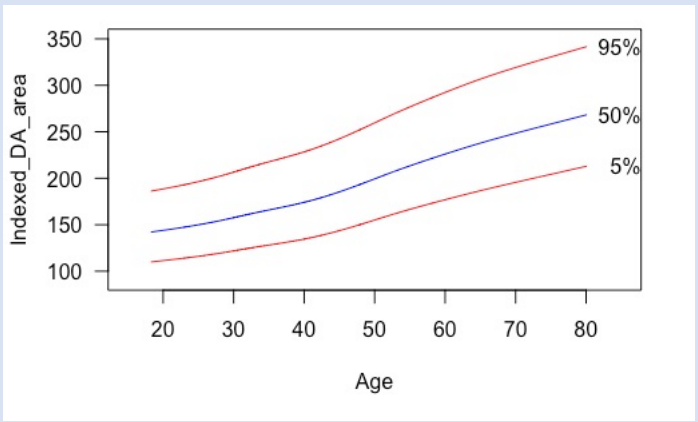

AA_area
1200
1000
800
600
400
200
95%
50%
5%
20
30
40
50
60
70
80
Age

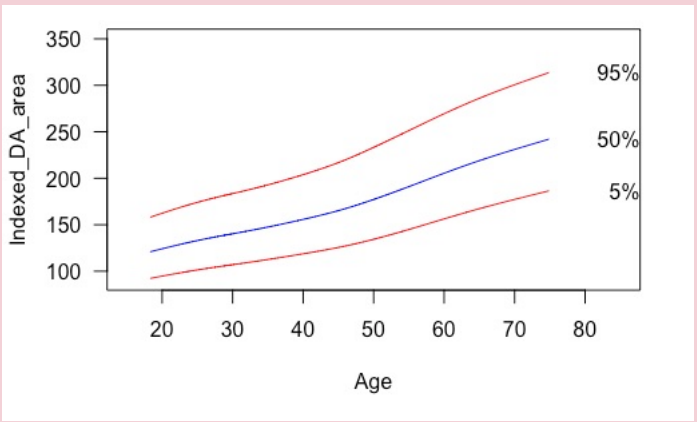












Nomograms displaying raw and Body Surface Area-indexed values for aortic dimensions in our healthy Caucasian population. *The 5th, 50th and 95th centiles are displayed across age groups, divided into males (left column; blue background) and females (right column; pink background). All dimensions are in mm (diameters), mm/m^2 (indexed diameters), mm^2 (areas) or mm^2/m^2 (indexed areas). AV_annulus = aortic valve annulus (measured in LVOT view); SoV = sinuses of Valsalva (measured in LVOT view); STJ = Sino-Tubular Junction (measured in LVOT view); AA = Ascending Aorta (measured at the level of the pulmonary bifurcation); DA = Descending Aorta (measured at the level of the pulmonary bifurcation).*

These normal values allow us to define a healthy range for aortic dimensions in the UK Caucasian population, and are a useful reference tool. The nomograms also enable us to visualize some of the key differences in dimensions between genders and with aging, explored further below.

11.4.3 Effect of gender on measured covariates

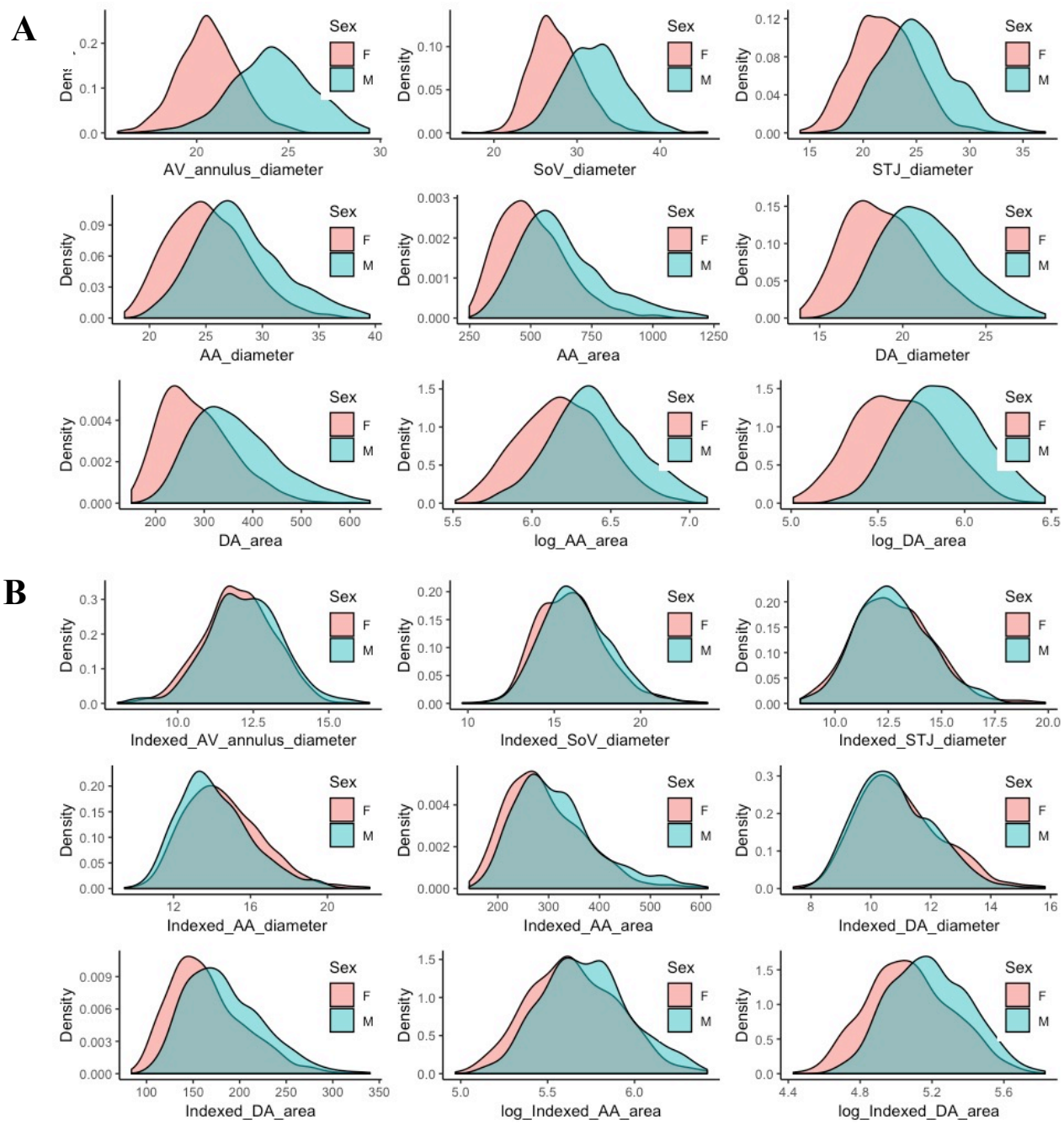
As expected, male gender was associated with higher height and weight, whilst females had a relatively higher fat mass component of weight. In line with previous studies, blood pressure indices were lower in women. However, once adjusted for age and body weight, diastolic blood pressure (DBP) did not differ significantly between the genders (ß=1.0; p=0.10). Gender remained a significant predictor of all other blood pressure indices. Total cholesterol did not differ between genders, but women had higher HDL levels and marginally lower LDL levels overall.

TABLE 11.4: Gender effects on measured covariates

	Male (n=585)		Female (n=714)		P value
	Mean	SD	Mean	SD	
Age (yrs)	40.6	13.3	41.1	13.6	0.56
Height (cm)	178.1	6.8	164.8	6.4	<0.001
Weight (kg)	79.5	11.4	65.5	10.6	<0.001
BSA (m^2)	2.0	0.2	1.7	0.1	<0.001
BMI (kg/m^2)	25.1	3.3	24.1	3.8	<0.001
Fat mass (%)	15.6	7.2	19.4	7.8	<0.001
SBP (mmHg)	124.4	12.2	114.6	13.2	<0.001
DBP (mmHg)	80.2	9.3	77.2	8.9	<0.001
MAP (mmHg)	94.9	9.5	89.7	9.9	<0.001
PP (mmHg)	44.3	8.8	37.4	8.2	<0.001
HR (bpm)	62.6	10.8	65.1	9.6	<0.001
Total cholesterol (mmol/l)	5.6	1.2	5.6	1.2	0.79
HDL	1.5	0.3	1.8	0.4	<0.001
LDL	3.4	0.9	3.2	0.9	0.015
Triglycerides	1.4	0.7	1.2	0.5	<0.001
Smoking	31%		32%		-
Activity level	Mean 2.8	Median 3	Mean 2.5	Median 2	-

Gender differences in measured covariates. BMI=Body Mass Index (kg/m2); BSA=Body Surface Area (m2); Height (cm); Weight (kg); Fatmass= percentage fat mass (%); SBP=Systolic Blood Pressure (mmHg); DBP=Diastolic Blood Pressure (mmHg); PP=Pulse Pressure (mmHg); MAP=Mean Arterial Pressure (mmHg); HR=Heart rate (bpm); Total cholesterol (mmol/l); HDL= High Density Lipoprotein (mmol/l); LDL=Low Density Lipoprotein (mmol/l); Triglycerides = serum triglycerides (mmol/l); Smoking= percentage ever smoked regularly; Activity level=self-reported, with 0 totally sedentary and 4 high-level exercise on most days. P values derived from Kruskal-Wallis tests.

FIGURE 11.5: Aortic root dimensions approximate normal distribution; more distal measures have skewed distributions which are approximately normalised by log-transformation.



Kernel density plots for aortic dimension, split by gender. A: "Raw" aortic dimensions. B: "Indexed" aortic dimensions adjusted for body surface area. Diameters in mm; areas in mm2; indexed values are mm/m2 (diameters) or mm2/m2 (areas). F=Female; M=Male; AV = Aortic Valve; SoV= Sinuses of Valsalva; STJ= Sino-Tubular Junction; AA or DA = Ascending or Descending Aorta at level of the pulmonary bifurcation

Raw aortic dimensions were, as expected, lower in females. Once adjusted for body surface area, the gender differences become much less marked, although remain significant for all measures except for indexed STJ diameter (see tables 11.5).

11.4.4 Regression modelling: significant predictors of aortic root dimensions

Mixed regression models were constructed to investigate the relative effects and effect directions of basic anthropometric and biometric covariates on aortic root dimensions. Summary data from these are presented in Tables 11.5 below.

TABLES 11.5: Regression model summaries for aortic root diameters

	AV annulus diameter				Indexed AV annulus diameter			
	Regression coefficient (ß)	mm change in diameter per SD change in covariate	P value	Adjusted R^2	Regression coefficient (ß)	mm/m2 change in indexed diameter per SD change in covariate	P value	Adjusted R^2
Basic model								
Age	0.01	0.13	**0.016**		-0.01	-0.08	**0.027**	
Gender (M)	2.28	-	**<0.001**		0.19	-	**0.006**	
Height	0.09	0.83	**<0.001**		**-**	**-**	**-**	
R^2 for basic model = 0.50					**R^2 for basic model =0.01**			
Basic model + 1 additional covariate								
Weight	0.03	0.35	**<0.001**	0.51	-	-	-	-
BSA	2.15	0.43	**<0.001**	0.51	-	-	-	-
BMI	0.08	0.29	**<0.001**	0.51	-	-	-	-
SBP	0.01	0.11	0.06	0.50	-0.01	-0.13	**0.001**	0.02
DBP	-0.01	-0.06	0.25	0.50	-0.02	-0.19	**<0.001**	0.03
MAP	0.00	0.01	0.87	0.50	-0.01	-0.18	**<0.001**	0.03
PP	0.02	0.21	**<0.001**	0.51	0.00	0.02	0.55	0.01
TGs	-0.10	-0.06	0.27	0.51	-0.27	-0.17	**<0.001**	0.02
Tchol	-0.03	-0.03	0.59	0.50	-0.04	-0.04	0.31	0.01
HDL	0.40	0.16	**0.008**	0.50	0.66	0.26	**<0.001**	0.05
LDL	-0.03	-0.02	0.68	0.50	-0.09	-0.08	**0.038**	0.01
Fatmass	0.00	0.01	0.85	0.50	-0.07	-0.64	**<0.001**	0.18
Smoking	-0.05	-0.02	0.65	0.50	-0.11	-0.05	0.16	0.01
Activity	0.22	0.19	**<0.001**	0.51	0.21	0.18	**<0.001**	0.03
HR	-0.01	-0.12	**0.026**	0.50	-0.01	-0.10	**0.004**	0.01
Selected model: AV annulus ~ Age + Sex + BSA + PP + HDL + Fatmass + Smoking + HR **Adjusted R^2=0.54**					**Selected model:** Indexed AV annulus ~ Age + Sex + DBP + HDL + Fatmass **Adjusted R^2=0.20**			

	SoV diameter				Indexed SoV diameter			
	Regression coefficient (ß)	**mm change in diameter per SD change in covariate**	**P value**	**Adjusted R^2**	**Regression coefficient (ß)**	**mm/m2 change in indexed diameter per SD change in covariate**	**P value**	**Adjusted R^2**
	Basic model				***Basic model***			
Age	0.12	1.65	**<0.001**		0.05	0.72	**<0.001**	
Gender (M)	3.23	3.23	**<0.001**		0.28	0.28	**0.009**	
Height	0.11	0.99	**<0.001**		**-**	**-**	**-**	
	R^2 for basic model = 0.483				**R^2 for basic model =0.130**			
	Basic model + 1 additional covariate				***Basic model + 1 additional covariate***			
Weight	0.04	0.50	**<0.001**	0.49	-	-	-	-
BSA	3.03	0.60	**<0.001**	0.49	-	-	-	-
BMI	0.10	0.38	**<0.001**	0.49	-	-	-	-
SBP	0.01	0.15	0.11	0.48	-0.01	-0.14	**0.020**	0.13
DBP	0.03	0.27	**0.003**	0.49	-0.01	-0.05	0.34	0.13
MAP	0.02	0.24	**0.010**	0.49	-0.01	-0.09	0.12	0.13
PP	-0.01	-0.09	0.32	0.49	-0.02	-0.15	**0.014**	0.14
TGs	-0.14	-0.09	0.38	0.48	-0.40	-0.25	**<0.001**	0.15
Tchol	0.07	0.08	0.42	0.48	0.00	0.00	0.97	0.14
HDL	0.33	0.13	0.19	0.48	0.77	0.31	**<0.001**	0.15
LDL	0.12	0.11	0.28	0.48	-0.04	-0.04	0.58	0.14
Fatmass	0.00	0.00	0.99	0.48	-0.10	-0.86	**<0.001**	0.25
Smoking	-0.29	-0.13	0.12	0.48	-0.27	-0.13	**0.021**	0.13
Activity	0.19	0.16	0.06	0.48	0.22	0.19	**<0.001**	0.14
HR	-0.01	-0.13	0.16	0.48	-0.01	-0.12	**0.038**	0.13
	Selected model: SoV ~ Age + Sex + Weight + DBP +HDL+ Fatmass + Smoking **Adjusted R^2=0.51**				***Selected model:*** Indexed SoV ~ Age + Sex + PP + HDL + Fatmass + Smoking **Adjusted R^2=0.27**			

	STJ diameter				Indexed STJ diameter			
	Regression coefficient (ß)	mm change in diameter per SD change in covariate	P value	AdjustedR²	Regression coefficient (ß)	mm/m² change in diameter per SD change in covariate	P value	Adjusted R²
	Basic model				***Basic model***			
Age	0.14	1.83	**<0.001**		0.06	0.84	**<0.001**	
Gender (M)	1.92	1.92	**<0.001**		-0.00	-0.00	0.98	
Height	0.1	0.94	**<0.001**		-	-	-	
	Adjusted R^2 for basic model = 0.45				**Adjusted R^2 for basic model =0.22**			
	Basic model + 1 additional covariate				***Basic model + 1 additional covariate***			
Weight	0.04	0.47	**<0.001**	0.46	-	-	-	-
BSA	2.88	0.57	**<0.001**	0.46	-	-	-	-
BMI	0.10	0.36	**<0.001**	0.46	-	-	-	-
SBP	0.01	0.14	0.11	0.46	-0.01	-0.11	**0.033**	0.22
DBP	0.04	0.33	**<0.001**	0.46	0.00	0.01	0.86	0.21
MAP	0.03	0.28	**<0.001**	0.46	-0.00	-0.04	0.46	0.22
PP	-0.02	-0.19	**0.024**	0.46	-0.02	-0.18	**<0.001**	0.22
TGs	-0.10	-0.06	0.49	0.46	-0.29	-0.18	**<0.001**	0.23
Tchol	0.01	0.01	0.91	0.45	-0.02	-0.03	0.64	0.22
HDL	0.44	0.18	*0.05*	0.45	0.73	0.29	**<0.001**	0.23
LDL	-0.01	-0.01	0.96	0.45	-0.09	-0.08	0.12	0.22
Fatmass	0.00	0.02	0.82	0.45	-0.08	-0.70	**<0.001**	0.31
Smoking	-0.34	-0.16	**0.042**	0.46	-0.26	-0.12	**0.011**	0.22
Activity	0.19	0.16	**0.041**	0.45	0.21	0.18	**<0.001**	0.22
HR	-0.02	-0.17	**0.030**	0.45	-0.01	-0.14	**0.004**	0.22
	Selected model: STJ ~ Age + Sex + BSA + DBP + HDL + Smoking+ HR **Adjusted R^2=0.50**				***Selected model:*** STJi ~ Age+Sex+PP+HDL+Fatmass + HR **Adjusted R^2=0.33**			

These summary statistics are derived from the basic, extended and selected models described in Methods. Normal effect sizes (mm or mm/m2 change in dependent variable per unit change in predictor) and partially-standardised effect sizes are presented (mm or mm/m2 change in dependent variable per SD change in independent variable) to allow comparison. R2 values are adjusted for the number of predictors in the model. Abbreviations: AV (aortic valve); SoV (sinuses of Valsalva); STJ (Sino-Tubular Junction); BSA (body surface area, m2); BMI (body mass index); SBP (systolic blood pressure, mmHg); DBP (diastolic blood pressure, mmHg); MAP (mean arterial pressure, mmHg); PP (pulse pressure, mmHg); TGs (serum triglycerides, mmol/l); Tchol (serum total cholesterol, mmol/l); HDL (serum high-density lipoprotein, mmol/l); LDL (serum low-density lipoprotein, mmol/l); Fatmass (body fat percentage); Smoking (ever smoked versus never-smoked); Activity (activity score on numeric index from 0 [sedentary] to 4 [high-level activity most days]); HR (heart rate, bpm).

11.4.4.1 Aortic root diameters: effect of anthropometric variables

A key predictor of all aortic root dimensions is age, as confirmed by the nomograms above. Age particularly predicts SoV and STJ diameters, where each decade is associated with a 1.2 and 1.4mm increase in diameter respectively, or a 0.72 and 0.84 mm/m^2 increase in indexed diameter. Height alone does not capture all of the variation in measurements due to body size and body composition; the addition of weight, body mass index (BMI) or body surface area (BSA) to the model explained more of the variability in non-indexed measurements.

Gender remains a significant predictor of all the non-indexed diameters, with male gender being associated with a 2.3, 3.2 and 1.9mm greater diameter at the AV annulus, SoV and STJ respectively. However, this gender difference was greatly lessened by indexation to BSA, with male gender predicting significantly larger indexed diameters only at valve annulus (0.2 mm/m^2) and SoV (0.3 mm/m^2) levels. There was no significant gender difference for indexed STJ diameter.

For each of the non-indexed measurements, the non-modifiable factors (age, gender and height), predicted around 45-50% of the variance (R^2 0.50, 0.48 and 0.45 for aortic valve annulus, sinuses of Valsalva and sinotubular junction respectively). For indexed measures, these figures were, unsurprisingly, much lower, with non-modifiable factors predicting less than 1% of the variance in indexed AV annulus diameter, but up to 21% of the variance in indexed sinotubular junction diameter.

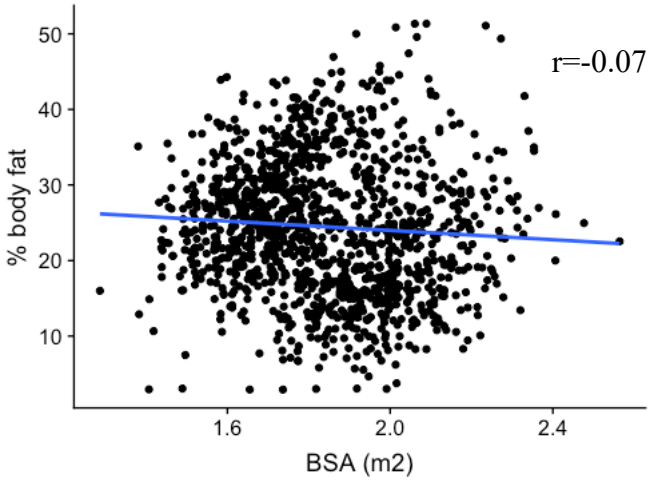
11.4.4.2: Aortic root dimensions: effect of modifiable risk factors

Addition of single additional covariates can predict at most a further 1% of the variance in raw diameters, compared with the basic (non-modifiable) model alone. The selected model which combines anthropometric variables and selected additional covariates predicts 50-54% of the variance. This demonstrates that measured modifiable cardiovascular risk factors only have a very small influence on root diameters (up to 4%), which is dwarfed by the effects of anthropometric variables.

For BSA-indexed measures, individual covariates explained a greater degree of variance, with the full model explaining 20 ,27 and 33% of the variance for aortic valve annulus, sinuses of Valsalva and sinotubular junction respectively.

For BSA-indexed diameters, percentage body fat (fatmass) was the single best predictor, explaining an additional 17% of variance of indexed AV annulus diameter beyond the basic model, and up to 9% of the variance of indexed STJ diameter. This may of course partially reflect a correlation with BSA, although this is not particularly strong (see Figure 11.6).

FIGURE 11.6: BSA does not strongly correlate with body fat percentage



Scatterplot showing BSA vs % body fat demonstrating minimal correlation between the variables. Line is linear regression fit (lm() function in R); r is Pearson's correlation coefficient calculated using cor.test() function in R.

As noted in previous studies, the association of blood pressure parameters with aortic root diameters is mixed. Blood pressure metrics were generally positively associated with raw, non-indexed diameters. For both SoV and STJ diameters, DBP was the most strongly associated BP variable, with a 10mmHg increase in DBP corresponding to a 0.3mm increase in SoV diameter and a 0.4mm increase in STJ diameter.

However, interestingly, once we look at indexed diameters, we see a reversal of blood pressure effect. Here, all significant blood pressure associations were negative – i.e. an increase in blood pressure corresponded to a decrease in indexed dimension. A 10mm increase in SBP corresponded to a 0.1mm/m^2 decrease in all indexed dimensions, and was the only BP variable significantly associated with all indexed root dimensions.

11.4.5 Regression modelling: significant predictors of ascending and descending aortic dimensions

Mixed regression models were constructed to investigate the relative effects and effect directions of basic anthropometric and biometric covariates on ascending and descending aortic areas. Of note, these variables had a skewed distribution, resulting in non-normal distribution of residuals and standard errors. These were normalised by a log-transformation of the area variable. Therefore, effect sizes presented in the tables below apply on a log scale and are difficult to interpret in terms of absolute effect size.

TABLES 11.6: Regression model summaries for aortic area variables

	AA area			**Indexed AA area**		
	Regression coefficient (ß; note log scale)	**P value**	**Adjusted R^2**	**Regression coefficient (ß; note log scale)**	**P value**	**Adjusted R^2**
Basic model (log(dimension)~ age + gender + height) or (log(indexed dimension) ~ age + gender)						
Age	0.014	**<0.001**		0.012	**0.027**	
Gender (M)	0.103	**<0.001**		0.073	**0.006**	
Height	0.008	**<0.001**		-	-	
	R^2 for basic model = 0.53			**R^2 for basic model =0.44**		
Basic model + 1 additional covariate						
Weight	0.005	**<0.001**	0.56	-	-	-
BSA	0.396	**<0.001**	0.56	-	-	-
BMI	0.014	**<0.001**	0.56	-	-	-
SBP	0.003	**<0.001**	0.54	0.001	**0.003**	0.44
DBP	0.005	**<0.001**	0.55	0.003	**<0.001**	0.45
MAP	0.005	**<0.001**	0.55	0.003	**<0.001**	0.45
PP	0.000	0.97	0.53	-0.001	0.298	0.44
TGs	0.01	0.37	0.53	-0.01	0.350	0.44
Tchol	0.002	0.72	0.53	-0.003	0.649	0.44
HDL	-0.005	0.78	0.53	0.034	**0.035**	0.44
LDL	0.002	0.75	0.53	-0.008	0.272	0.44
Fatmass	0.003	**0.001**	0.53	-0.004	**<0.001**	0.45
Smoking	-0.027	**0.032**	0.53	-0.034	**0.007**	0.44
Activity	0.011	0.12	0.53	0.019	**0.005**	0.44
HR	-0.001	0.32	0.53	-0.001	0.072	0.44
	Selected model: log(AA area) ~Age + Gender + BSA + DBP + Fatmass + Smoking + HR **Adjusted R^2=0.59**			***Selected model:*** log(Indexed AA area) ~ Age + Gender + DBP + Fatmass + Smoking + HR **Adjusted R^2=0.47**		

	DA area			Indexed DA area		
	Regression coefficient (ß; note log scale)	**P value**	**AdjustedR²**	**Regression coefficient (ß; note log scale)**	**P value**	**Adjusted R²**
Basic model (log(area)~ age + gender + height) or log((indexed area) ~ age + gender)						
Age	0.013	**<0.001**		0.012	**<0.001**	
Gender (M)	0.144	**<0.001**		0.115	**<0.001**	
Height	0.008	**<0.001**		-	-	
	Adjusted R² for basic model = 0.62			**Adjusted R² for basic model =0.52**		
Basic model + 1 additional covariate						
Weight	0.005	**<0.001**	0.65	-	-	-
BSA	0.385	**<0.001**	0.65	-	-	-
BMI	0.014	**<0.001**	0.65	-	-	-
SBP	0.002	**<0.001**	0.63	0.000	0.29	0.52
DBP	0.003	**<0.001**	0.63	0.001	0.12	0.52
MAP	0.002	**<0.001**	0.62	0.001	0.11	0.52
PP	0.001	0.38	0.61	0.000	0.99	0.52
TGs	0.006	0.56	0.61	-0.018	0.05	0.53
Tchol	-0.002	0.68	0.61	-0.007	0.16	0.52
HDL	-0.012	0.43	0.61	0.030	**0.034**	0.52
LDL	0.001	0.90	0.61	-0.009	0.13	0.52
Fatmass	0.003	**0.001**	0.62	-0.004	**<0.001**	0.53
Smoking	0.002	0.82	0.62	-0.004	0.69	0.52
Activity	0.013	**0.026**	0.62	0.021	**<0.001**	0.52
HR	-0.001	0.24	0.62	-0.001	**0.039**	0.53
Selected model: log(DA area) ~ Age+ Sex + BSA + DBP + Fatmass + Activity **Adjusted R²=0.66**				***Selected model:*** log(DA areai)~ Age+Sex+DBP+Fatmass+Activity **Adjusted R²=0.54**		

*These summary statistics are derived from the basic, extended and selected models described in Methods, with the dependent variable log-transformed. *Effect sizes are therefore on a log scale. R^2 values are adjusted for the number of predictors in the model. Abbreviations: AA or DA area: ascending or descending aortic area (mm^2) at level of pulmonary bifurcation; BSA (body surface area, m^2); BMI (body mass index); SBP (systolic blood pressure, mmHg); DBP (diastolic blood pressure, mmHg); MAP (mean arterial pressure, mmHg); PP (pulse pressure, mmHg); TGs (serum triglycerides, mmol/l); Tchol(serum total cholesterol, mmol/l); HDL (serum high-density lipoprotein, mmol/l); LD (serum low-density lipoprotein, mmol/l); Fatmass (body fat percentage); Smoking (ever smoked vs never-smoked); Activity (activity on numeric index from 0 [sedentary] to 4 [high-level activity most days]); HR (heart rate, bpm).*

11.4.5.1 Ascending & descending aortic dimensions: effect of anthropometrics

In ascending and descending aorta, age, gender and height were significantly associated with dimensions. These fixed factors explained 53 and 62% of the variance in raw ascending and descending aortic diameters respectively. Even for indexed diameters, age and gender explained 44 and 52% of the variation, demonstrating a much greater influence of these fixed anthropometric variables in this more distal section of the aorta than in the aortic root. Per decade of age, there was an area change equating to a diameter change of approximately 2mm per decade in ascending aorta, and a 1.4mm increase per decade in descending aorta (by geometric calculation).

11.4.5.2 Ascending & descending aortic dimensions: effect of cardiovascular risk factors

In the ascending and descending aorta, the association with blood pressure was much more consistent than in the aortic root, with SBP, DBP and MAP positively associated with raw and indexed areas. DBP was the most strongly predictive blood pressure metric. The effect of blood pressure variables was, however, small, accounting for at most 1-2% of the variance. There was no significant association of blood pressure with indexed descending aortic area in the extended model, although in the final selected model, DBP did demonstrate a modest but significant positive association with indexed DA area (ß=0.472; p=0.01).

Percentage body fat (fatmass) again was a significant predictor of aortic dimensions, once more correlated with an increase in absolute diameter, but a decrease in indexed dimensions. There was no significant association of lipid measures with AA or DA areas, with the exception of HDL, which showed a modest positive correlation with indexed AA and DA areas. Where significant associations were found, cardioprotective variables such as HDL and activity level were correlated with increased indexed areas, whereas cardiovascular risk factors such as smoking and increasing heart rate, were correlated with decreased areas. This is a similar pattern to that found in the aortic root. The exception to this rule in the AA and DA is blood pressure, which is tightly linked with aortic distensibility and elastic function.

11.5 RESULTS: aortic morphology

11.5.1 Aortic morphology: reliability

Having defined different methods of quantifying simple measures of aortic arch morphology using cardiovascular MRI, we tested the inter- and intra-observer reliability. The inter-observer reliability was, as expected, somewhat less for morphological measures than for the standard dimensions. Nevertheless, there was good agreement for basic measures such as arch height, width and arch angle.

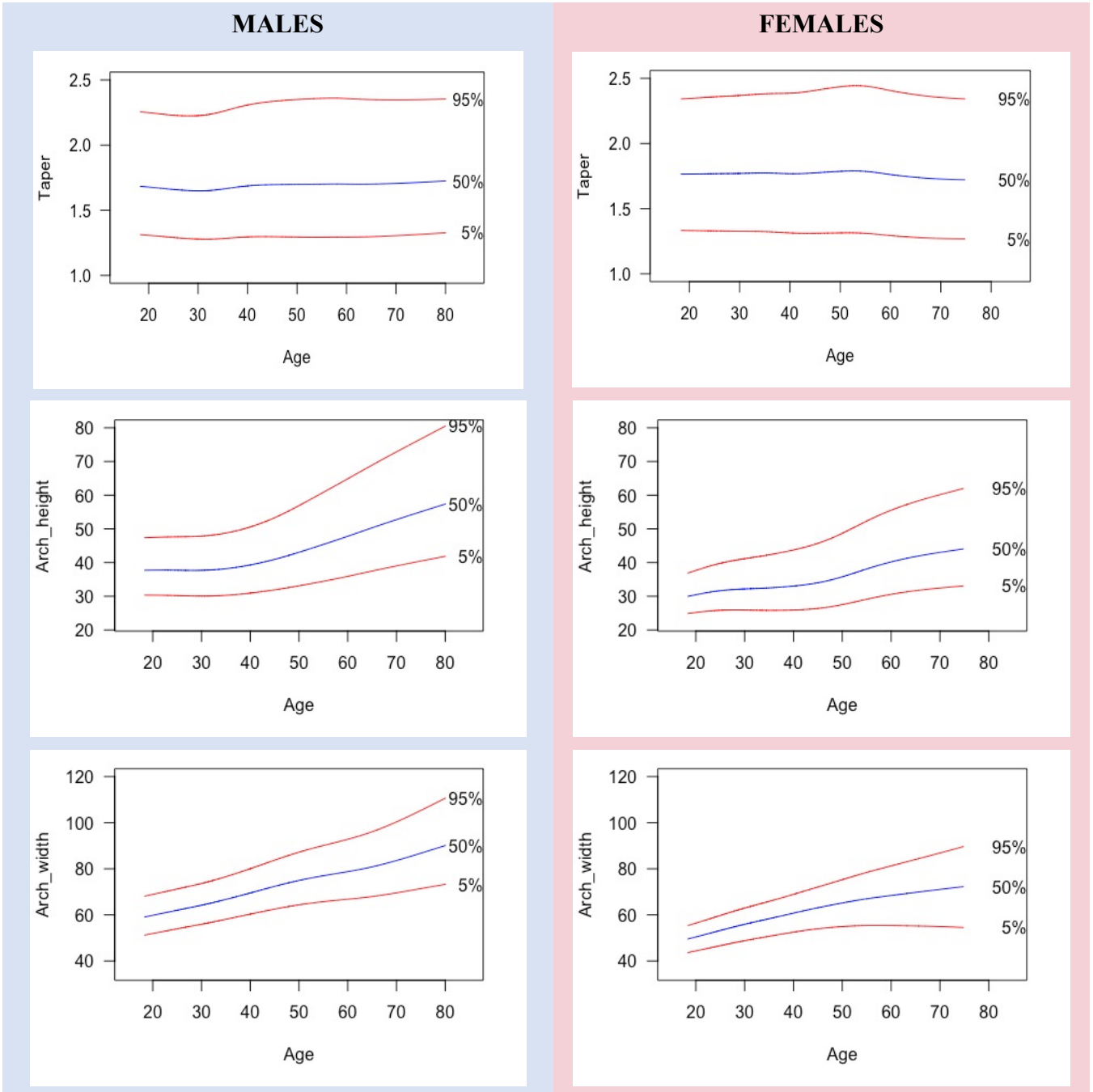
TABLE 11.7: Reliability of aortic morphology measurements is generally good

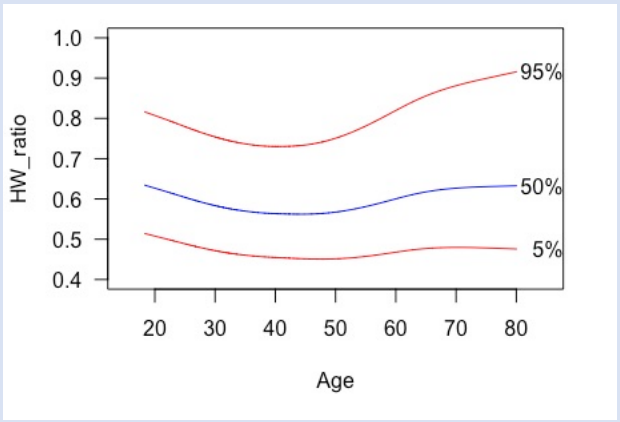
	Inter-observer		Intra-observer	
Phenotype	**ICC**	**Coefficient of variability**	**ICC**	**Coefficient of variability**
Arch height	0.91 (0.81-0.95)	8.02	0.93 (0.86-0.98)	7.13
Arch width	0.93 (0.87-0.97)	4.40	0.91 (0.83-0.97)	5.22
Arch length	0.97 (0.95-0.98)	4.07	0.97 (0.95-0.98)	4.15
Taper	0.96 (0.92-0.98)	5.70	0.97 (0.94-0.98)	4.99
Arch H:W ratio	0.86 (0.74 - 0.93)	10.56	0.90 (0.82-0.95)	9.50
Arch angle	0.81 (0.65 - 0.90)	8.20	0.85 (0.71-0.94)	7.78
Arch asymmetry	0.66 (0.31-0.83)	6.20	0.81 (0.68-0.89)	6.30

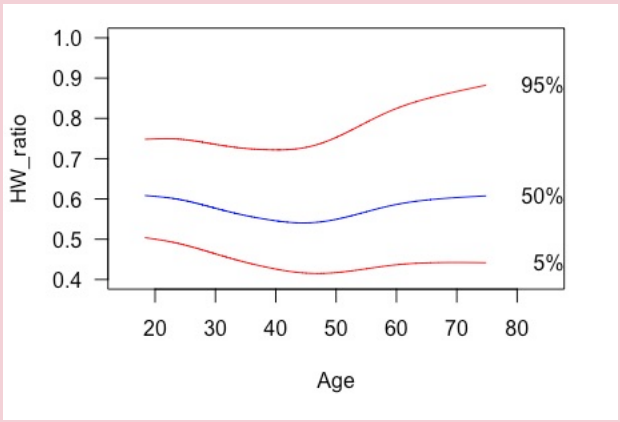
11.5.2 Aortic morphology: Defining normal values

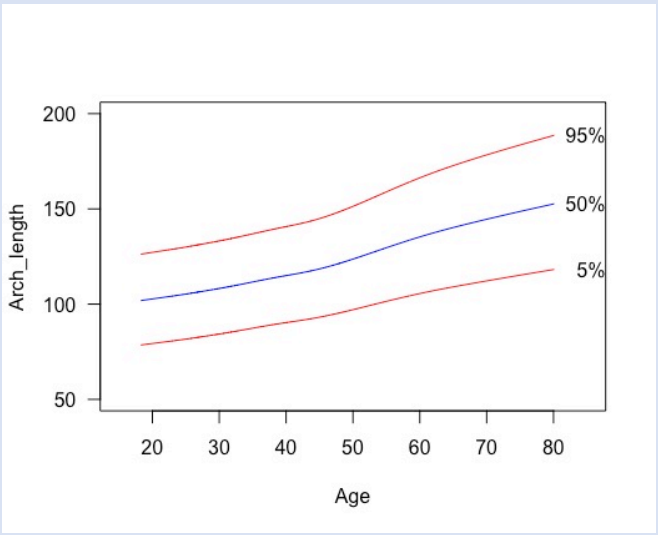
The nomograms below define centile plots according to age for each morphological measure.

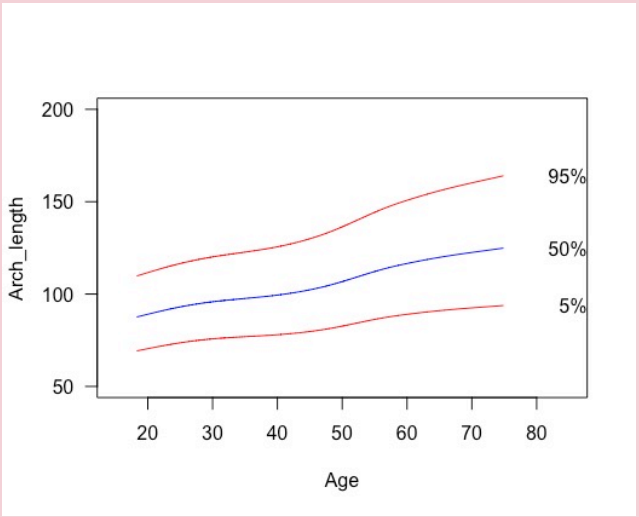
FIGURE 11.7: Nomograms demonstrating normal ranges of aortic morphology

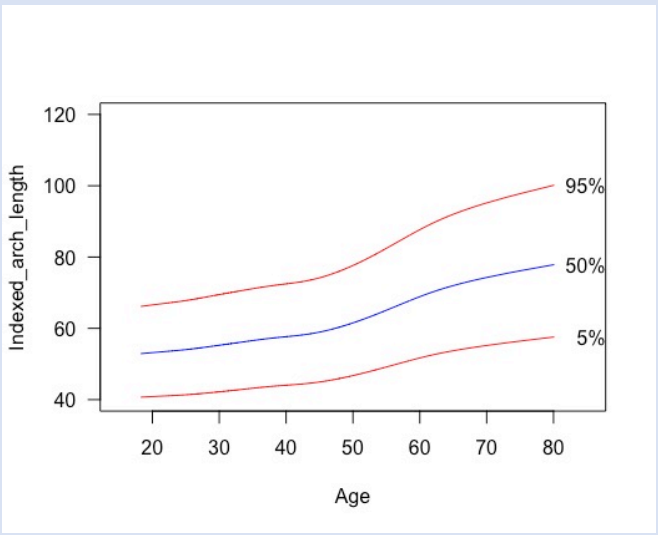


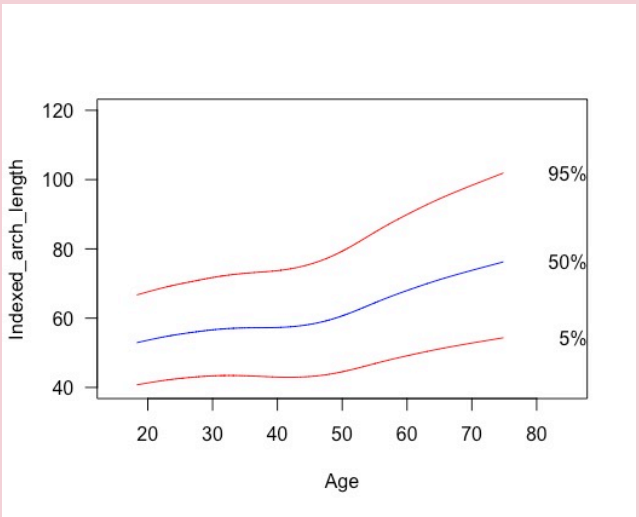

HW_ratio
1.0
0.9
0.8
0.7
0.6
0.5
0.4
95%
50%
5%
20
30
40
50
60
70
80
Age

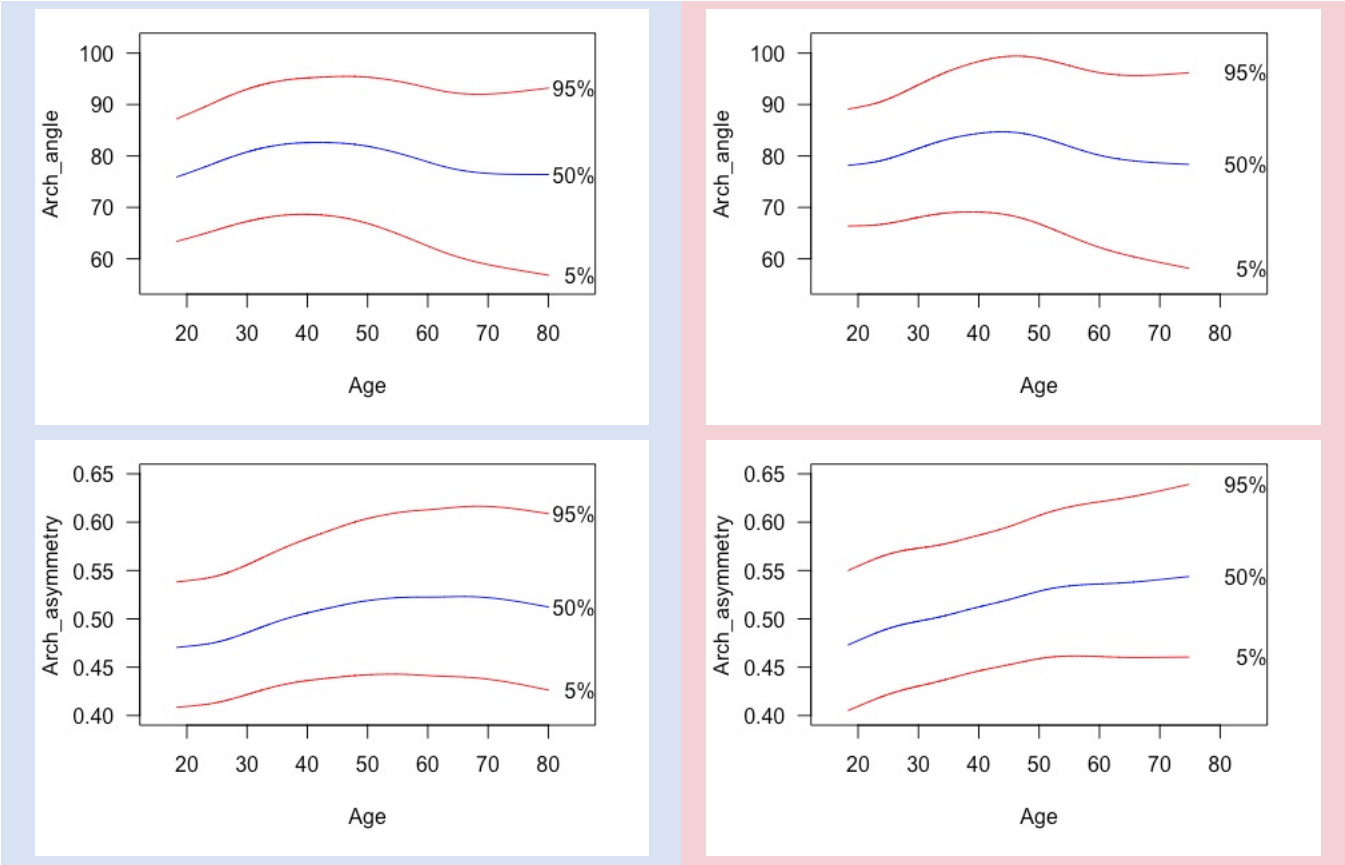

HW_ratio
1.0
0.9
0.8
0.7
0.6
0.5
0.4
95%
50%
5%
20
30
40
50
60
70
80
Age


Arch_length
200
150
100
50
95%
50%
5%
20
30
40
50
60
70
80
Age


Arch_length
200
150
100
50
95%
50%
5%
20
30
40
50
60
70
80
Age

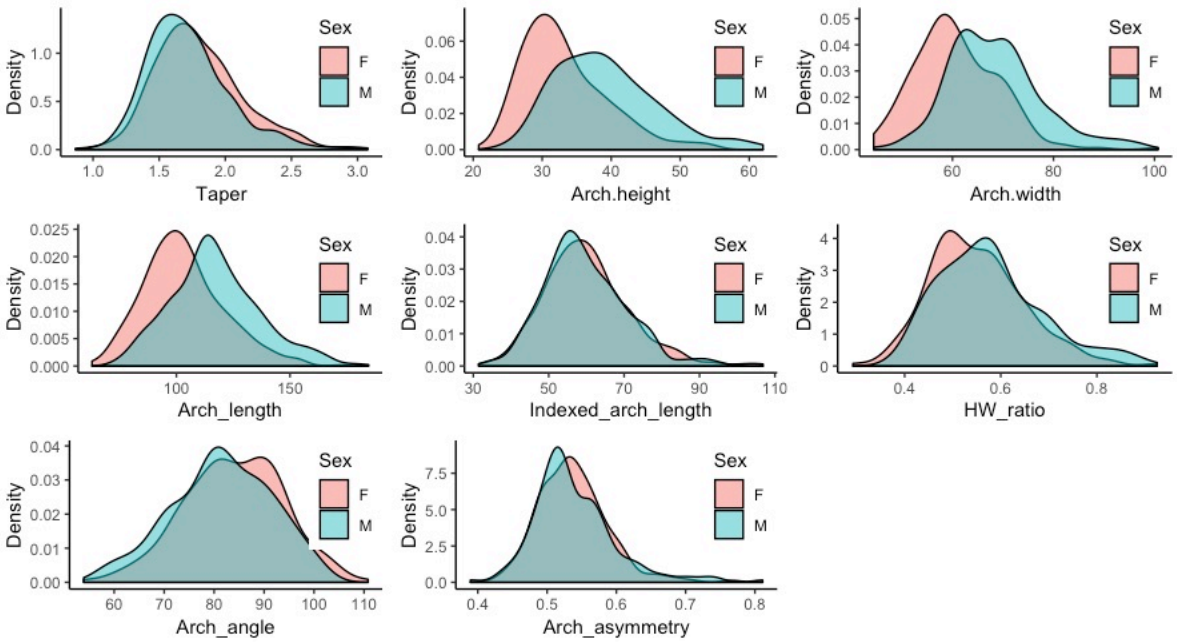

Indexed_arch_length
120
100
80
60
40
95%
50%
5%
20
30
40
50
60
70
80
Age


Indexed_arch_length
120
100
80
60
40
95%
50%
5%
20
30
40
50
60
70
80
Age



Nomograms displaying aortic morphological variables in our healthy Caucasian population. *The 5th, 50th and 95th centiles are displayed across age groups, divided into males (left column) and females (right column). All dimensions are in mm (arch height, arch width and arch length. Indexed arch length is indexed to body surface area. Taper is defined as the ratio of ascending to descending aortic area in diastole. HWratio is the ratio of arch height to arch width as described in Methods. Arch angle is measured in degrees and represented by A in Figure 11.3 (Methods) and arch asymmetry is defined as the ratio of proximal arch angle to total arch angle – in other words, an increase in this measure corresponds to a posterior "tipping" of the arch and a relative increase in proximal component of arch angle.*

FIGURE 11.8: Most aortic morphological variables are approximately normally-distributed



Kernel density plots for aortic morphology, split by gender. Dimensions are presented in mm. F=Female; M=Male; HW ratio = arch height:width ratio. Other variables are as defined above.

The nomograms show a widening, lengthening and increasing height of the aortic arch with age. With the exception of the variables directly related to body size (arch height, width and length), distributions of aortic morphological variables are approximately equal between genders.

11.5.3 Regression modelling: significant predictors of aortic morphology

TABLES 11.8: Regression models for aortic morphology

	Aortic height				Aortic width			
	Regression coefficient (ß)	**mm change in height per SD change in covariate**	**P value**	**Adjusted R^2**	**Regression coefficient (ß)**	**mm change in width per SD change in covariate**	**P value**	**Adjusted R^2**
Basic model (dimension ~ age + gender + height)								
Age	0.27	3.59	**<0.001**		0.48	6.50	**<0.001**	
Gender (M)	2.93	2.93	**<0.001**		7.01	7.01	**<0.001**	
Height	0.27	2.51	**<0.001**		0.10	0.90	**0.039**	
	Adjusted R^2 for basic model = 0.41				**Adjusted R^2 for basic model =0.60**			
Basic model + 1 additional covariate								
Weight	-0.13	-1.64	**<0.001**	0.44	0.22	2.88	**<0.001**	0.66
BSA	-10.14	-2.02	**<0.001**	0.44	17.81	3.54	**<0.001**	0.66
BMI	-0.39	-1.41	**<0.001**	0.44	0.69	2.51	**<0.001**	0.66
SBP	-0.01	-0.10	0.77	0.41	0.08	1.14	**<0.001**	0.60
DBP	0.00	-0.03	0.93	0.41	0.13	1.16	**<0.001**	0.61
MAP	-0.01	-0.11	0.74	0.41	0.13	1.32	**<0.001**	0.61
PP	-0.02	-0.19	0.53	0.41	0.05	0.50	0.12	0.60
TGs	-0.49	-0.30	0.39	0.44	1.21	0.75	0.06	0.55
Tchol	-0.06	-0.07	0.85	0.44	0.29	0.34	0.38	0.55
HDL	2.02	0.81	**0.01**	0.45	-0.89	-0.36	0.34	0.56
LDL	-0.27	-0.25	0.47	0.43	0.03	0.02	0.95	0.54
Fatmass	-0.17	-1.49	**<0.001**	0.43	0.21	1.87	**<0.001**	0.62
Smoking	-0.40	-0.19	0.50	0.41	1.51	0.70	0.02	0.61
Activity	0.62	0.54	0.06	0.42	0.43	0.37	0.22	0.60
HR	0.02	0.23	0.42	0.41	-0.05	-0.46	0.13	0.60
	Selected model: Aortic height ~ Age + Sex + Height + Weight + DBP + HDL+ Fatmass **Adjusted R^2=0.48**				***Selected model:*** Aortic width ~ Age + Sex + Weight + DBP + Smoking + Fatmass **Adjusted R^2=0.66**			

	Aortic Length				Indexed aortic length			
	Regression coefficient (ß)	**mm change in length per SD change in covariate**	**P value**	**Adjusted R^2**	**Regression coefficient (ß)**	**mm/m^2 change in indexed length per SD change in covariate**	**P value**	**Adjusted R^2**
Basic model (dimension ~ age + gender + height) or (indexed dimension ~ age + gender)								
Age	0.80	10.80	**<0.001**		0.36	4.65	**<0.001**	
Gender (M)	7.76	7.80	**<0.001**		1.8	-0.36	0.55	
Height	0.50	4.70	**<0.001**		-	-	**-**	
	Adjusted R^2 for basic model = 0.39				**Adjusted R^2 for basic model = 0.19**			
Basic model + 1 additional covariate								
Weight	-0.17	-2.24	**<0.001**	0.40	-	-	-	-
BSA	-13.39	-2.66	**<0.001**	0.40	-	-	-	-
BMI	-0.48	-1.74	**0.001**	0.40	-	-	-	-
SBP	0.05	0.74	0.18	0.39	-0.03	-0.38	0.28	0.21
DBP	0.10	0.96	0.07	0.39	-0.02	-0.21	0.54	0.21
MAP	0.09	0.87	0.11	0.39	-0.03	-0.34	0.31	0.21
PP	0.02	0.20	0.72	0.39	-0.03	-0.24	0.49	0.21
TGs	-0.22	-0.14	0.81	0.39	-1.16	-0.72	**0.04**	0.23
Tchol	0.39	0.46	0.43	0.39	0.08	0.10	0.79	0.22
HDL	4.63	1.86	**0.001**	0.39	4.68	1.87	**<0.001**	0.23
LDL	-0.31	-0.29	0.62	0.39	-0.56	-0.52	0.14	0.22
Fatmass	-0.22	-1.97	**0.003**	0.39	-0.48	-4.27	**<0.001**	0.30
Smoking	0.74	0.34	0.50	0.39	0.30	0.14	0.66	0.21
Activity	0.15	0.13	0.80	0.39	0.48	0.42	0.19	0.21
HR	0.03	0.35	0.50	0.39	0.00	0.03	0.94	0.21
	Selected model: Aortic length ~ Age + Sex + Height + Fatmass + DBP + HDL **Adjusted R^2=0.41**				***Selected model:*** Indexed arch length ~ Age + Sex + Fatmass + HDL + TGs **Adjusted R^2=0.33**			

	H:W ratio				Taper			
	Regression coefficient (ß)	**Unit change in ratio per SD change in covariate**	**P value**	**Adjusted R^2**	**Regression coefficient (ß)**	**Unit change in taper per SD change in covariate**	**P value**	**Adjusted R^2**
Basic model (phenotype ~ age + gender + height)								
Age	0.00	-0.002	0.77		0.001	0.009	0.35	
Gender (M)	-0.01	-0.01	0.41		-0.09	-0.09	**0.002**	
Height	0.003	0.03	**<0.001**		0.00	0.004	0.79	
	Adjusted R^2 for basic model = 0.05				**Adjusted R^2 for basic model =0.01**			
Basic model + 1 additional covariate								
Weight	-0.004	-0.051	**<0.001**	0.20	0.00	0.003	0.84	0.01
BSA	-0.316	-0.063	**<0.001**	0.21	0.02	0.003	0.83	0.01
BMI	-0.012	-0.044	**<0.001**	0.21	0.00	0.004	0.71	0.01
SBP	-0.001	-0.012	**0.04**	0.06	0.00	0.026	**0.02**	0.02
DBP	-0.001	-0.012	**0.03**	0.06	0.00	0.040	**<0.001**	0.03
MAP	-0.001	-0.014	**0.01**	0.06	0.00	0.037	**<0.001**	0.02
PP	-0.001	-0.006	0.25	0.05	0.00	-0.008	0.45	0.01
TGs	-0.016	-0.010	0.11	0.07	0.01	0.005	0.68	0.02
Tchol	-0.003	-0.003	0.59	0.06	0.01	0.007	0.51	0.02
HDL	0.039	0.016	**0.01**	0.08	0.01	0.003	0.80	0.02
LDL	-0.005	-0.004	0.49	0.06	0.00	0.001	0.94	0.02
Fatmass	-0.004	-0.040	**<0.001**	0.13	0.00	0.002	0.90	0.01
Smoking	-0.017	-0.008	0.11	0.05	-0.05	-0.024	**0.01**	0.02
Activity	0.005	0.004	0.41	0.05	0.00	-0.001	0.96	0.01
HR	0.001	0.007	0.15	0.05	0.00	0.000	0.97	0.01
	Selected model: H:W ratio~ Age + Height + Weight + Smoking + HR **Adjusted R^2=0.23**				***Selected model:*** Taper ~ Sex + Weight + SBP + DBP + TGs + Fatmass + Smoking **Adjusted R^2=0.04**			

	Arch angle				Arch asymmetry			
	Regression coefficient (ß)	° change in angle per SD change in covariate	P value	Adjusted R^2	Regression coefficient (ß)	Unit change in asymmetry per SD change in covariate	P value	Adjusted R^2
Basic model (dimension ~ age + gender + height) or (indexed dimension ~ age + gender)								
Age	-0.009	-0.12	0.81		0.002	0.02	**<0.001**	
Gender (M)	1.72	1.72	0.21		-0.009	-0.009	0.20	
Height	-0.33	-3.10	**<0.001**		0.001	0.01	**0.01**	
	Adjusted R^2 for basic model = 0.03				**Adjusted R^2 for basic model =0.14**			
Basic model + 1 additional covariate								
Weight	0.40	5.16	**<0.001**	0.23	0.000	-0.005	0.10	0.15
BSA	32.02	6.37	**<0.001**	0.23	-0.031	-0.006	0.10	0.15
BMI	1.23	4.45	**<0.001**	0.23	-0.001	-0.004	0.12	0.15
SBP	0.10	1.30	**0.02**	0.07	0.000	-0.004	0.12	0.15
DBP	0.14	1.25	**0.01**	0.07	0.000	0.000	0.96	0.14
MAP	0.15	1.51	**0.005**	0.08	0.000	-0.003	0.31	0.15
PP	0.05	0.44	0.40	0.06	0.000	-0.003	0.25	0.15
TGs	1.82	1.12	0.06	0.08	0.000	0.000	0.93	0.11
Tchol	0.41	0.48	0.43	0.07	-0.003	-0.003	0.37	0.12
HDL	-4.30	-1.72	**0.003**	0.09	0.010	0.004	0.20	0.13
LDL	0.67	0.63	0.29	0.07	-0.004	-0.004	0.25	0.12
Fatmass	0.44	3.88	**<0.001**	0.14	0.000	-0.001	0.66	0.15
Smoking	1.48	0.69	0.14	0.06	-0.007	-0.003	0.20	0.15
Activity	-0.50	-0.43	0.37	0.05	0.000	0.000	0.99	0.14
HR	-0.08	-0.83	0.09	0.06	0.000	0.003	0.30	0.15
	Selected model: Arch angle ~ Age + Height + Weight + Fatmass **Adjusted R^2=0.23**				***Selected model:*** Arch asymmetry ~ Age+Sex+Height **Adjusted R^2=0.14**			

These summary statistics are derived from the basic, extended and selected models described in Methods. R^2 values are adjusted for the number of predictors in the model. Abbreviations: H:Wratio (Height:width ratio of the aortic arch); BSA (body surface area, m^2); BMI (body mass index); SBP (systolic blood pressure, mmHg); DBP (diastolic blood pressure, mmHg); MAP (mean arterial pressure, mmHg); PP (pulse pressure, mmHg); TGs (serum triglycerides, mmol/l); Tchol(serum total cholesterol, mmol/l); HDL (serum high-density lipoprotein, mmol/l); LD (serum low-density lipoprotein, mmol/l); Fatmass (body fat percentage); Smoking (ever smoked vs never-smoked); Activity (activity on numeric index from 0 [sedentary] to 4 [high-level activity most days]); HR (heart rate, bpm).

11.5.3.1 Aortic morphology: effect of anthropometric variables

Aortic height, width and length are reasonably well-predicted by anthropometric variables (R^2 0.41, 0.60 and 0.39 respectively), whereas the more complex morphological variables (arch angle, taper, H:W ratio and asymmetry), are very poorly predicted.

Both arch height and arch width increase significantly with age, with each decade associated with approximately a 2.7mm increase in aortic arch height and a 4.8mm increase in aortic arch width. This goes hand-in-hand with an increase in aortic length of approximately 8mm per decade. These observations are in keeping with previously reported lengthening and unfolding of the aorta as previously described. We did not replicate previous observations of a decrease in H:W ratio with age[18].

Age is also associated with a significant increase in arch asymmetry, which equates to an increase in the proximal component of the arch angle and a "posterior tilting" of the aortic arch. This is in keeping with previous observations of aortic lengthening being most prominent in the ascending aorta. The arch angle itself was not significantly associated with age, but was associated with body size measures including height in particular. Taper was also not significantly predicted by age.

Male gender was, as expected, associated with increased aortic dimensions – i.e. increased arch height, width and length. It was not associated with increased indexed aortic length, nor with any changes in H:W ratio, arch angle or asymmetry. Male gender did predict a small reduction in aortic taper.

11.5.3.2 Aortic morphology: effect of modifiable risk factors

In general, cardiovascular risk factors explain a very modest proportion of the variance in aortic morphological variables.

An exception is body fat percentage, which is a significant predictor of several morphological variables, independently of age, height and gender. An increase of 10% fat mass correlates with a 2.2mm decrease in arch length, a 1.7 mm decrease in arch height and a 2.1mm increase in arch width, with a corresponding decrease in height:width ratio (H:W ratio) and increased arch angle.

HDL is another significant predictor of several variables, with increased levels associated with longer aortas, increased arch height and H:W ratio, and a corresponding decrease in arch angle.

As with some of the previously-reported aortic dimensions, these results show cardiac risk factors such as percentage body fat being associated with aortic variables in the opposite direction to the changes seen with ageing. The classically cardioprotective factor, HDL, is associated with aortic morphology traits with the same direction of effect as age. Again, this supports the idea that size and morphological change in the aorta is an adaptive mechanism which occurs as part of "healthy" cardiovascular ageing.

However, the picture is a little more complex with blood pressure metrics – which are almost certainly confounded by a bidirectional relationship between arch morphology and BP. Blood pressure measures seem to predict aortic arch width more strongly than they predict arch height. SBP, DBP and MAP all correlate with an increase in arch width, but have no significant association with aortic height, nor with arch length (although DBP has a borderline significant association in the selected model, with ß=0.12 and p=0.058). This predominant effect on arch width correlates with a decrease in H:W ratio and an increase in arch angle with increasing blood pressure. BP metrics also predict increased taper, with DBP and MAP being the most significantly associated measures. This is not surprising given the previously noted association with ascending, but not descending, aortic area (see section 11.3.4.2).

Aortic taper is remarkably poorly predicted by anthropometric and cardiovascular risk factor variables, with just 4% of the variance accounted for by our predictors. The major contributor to this phenotype is blood pressure, with SBP, DBP and MAP all associated with increasing taper (albeit at a very modest level). Smoking is correlated with a very small reduction in taper.

14% of the variance in arch asymmetry is explained by anthropometric factors (in particular, age); there are no other significant influences from cardiovascular risk factors at all.

11.6 DISCUSSION

11.6.1 Defining normal values

This is the largest UK-based cohort to report healthy reference values for raw and indexed aortic dimensions, and the first to report healthy reference ranges for aortic morphological parameters varying by age and gender. These figures may be used to interpret imaging findings in a clinical setting, and, for aortic morphology, provide a framework around which future research may be based.

11.6.2 The impact of anthropometric factors on aortic size and morphology

Gender remained a significant predictor of all indexed and non-indexed aortic dimensions as well as several of the aortic morphological variables. The size and morphology of the aorta changes with ageing in a manner which appears independent of gender. This involves a dilatation at all levels of the thoracic aorta, lengthening and widening of the aortic arch, and an increase in asymmetry – a posterior "tipping" of the aortic arch. These changes will all have a profound effect on flow and function of the aorta, and reflect the dynamic remodelling capacity of the aorta with age. The change in ascending and descending aortic areas per decade of age in our cohort was a little higher than previous studies[101], possibly reflecting the healthier nature of our cohort – supporting the idea that this expansion represents an adaptive mechanism (see below for further discussion).

11.6.3 The impact of cardiovascular risk factors on aortic size and morphology

We demonstrated opposite influences of blood pressure metrics on raw and indexed aortic dimensions in the aortic root. Where significant associations exist, the relationship was positive with raw dimensions, but negative with indexed dimensions. Notably, the direction of association of BP indices with ascending and descending aortic areas was positive for both raw and indexed values.

This somewhat mixed picture is in keeping with previous studies[310-312] which have consistently demonstrated a negative relationship of pulse pressure to aortic root diameter, but have been more variable for the other BP indices according to the exact specification of the regression model. None have examined the influence of BP on indexed aortic dimensions. These findings in the aortic root are perhaps most likely to be a statistical "quirk", reflecting

the modest correlation between these blood pressure indices and BSA. Associations between aortic traits and blood pressure variables are also always confounded by the bidirectional nature of the relationship: narrower, stiffer aortas require a greater pulse pressure to maintain organ perfusion; this increase in pressure stimulates aortic wall remodelling which increases diameter and increases wall stiffness. It is difficult, then, to make assertions about the "effect of blood pressure on aortic traits" as they are inextricably linked.

Nevertheless, looking at the other biometric variables reveals the possibility that cardiovascular risk factors in some way impair the "normal" age-related remodelling of the aorta. All the variables thought of as modifiable cardiovascular "risk factors" – LDL cholesterol, triglycerides, smoking, body fat mass percentage, heart rate, as well as blood pressure indices, are negatively associated with indexed aortic root dimensions and indexed ascending and descending aortic dimensions where significant associations exist. Cardio-protective factors such as HDL and activity levels are positively associated with indexed aortic root or indexed ascending or descending aortic dimensions. This has been noted in previous studies, but not further explored.[313]

It is tempting to speculate that the age-related increase in indexed aortic diameters is, within a predominantly "healthy" range of values, a protective mechanism, and that cardiovascular risk factors might act to attenuate this adaptive expansion in some way. The failure of this adaptive remodelling might therefore represent a more pathological state, rather than a beneficial "delay" in vascular ageing.

Again, it is difficult to tease out the possibility of these observations being due to correlation between these cardiovascular risk factors and BSA. There remains some signal in support of the "protective expansion" hypothesis however, from risk factors which correlate with BSA and aortic traits in opposite directions. For example, increasing activity levels correlate with a slight increase in BSA in our data (perhaps surprisingly). If activity levels predict aortic root dimensions predominantly by correlation with BSA, then we might expect increased activity to correlate inversely with indexed root diameters. However, activity levels correlate positively with both raw and indexed aortic root dimensions. Similarly, for heart rate, there is (perhaps somewhat surprisingly) a negative association with BSA, but additionally a negative association with aortic root diameters.

The effect of cardiovascular risk factors on aortic morphological variables is more limited, but again we see “protective” factors such as HDL exerting effects in the same direction as age – for example, associating with an increase in aortic length. Conversely, cardiovascular risk factors, such as percentage fat mass, associate in the opposite direction to age – for example, associating with a decrease in aortic length.

Vascular ageing is usually, and perhaps rightly, seen as a negative process. However, these observations suggest that the healthy aorta adapts well to changes with age, by undergoing remodelling, lengthening and widening. Cardiovascular risk factors might attenuate this adaptive process, perhaps exposing the aorta, and indeed left ventricle, to increased haemodynamic stresses.

Whatever the explanation for the association of risk factors with aortic size and shape changes, we have clearly demonstrated that vascular aging in large elastic arteries like the aorta is subject to significant influence from traditional atherosclerotic risk factors, and that the size and shape of the aorta are significantly associated with these covariates. This emphasises the dynamic nature of the aorta as a responsive organ which can undergo complex remodelling and re-shaping throughout life.

What is important to note however, is that all these effects remain very small, with the full models explaining a perhaps surprisingly limited proportion of the variance. It might therefore be the case that genetics, alongside other environmental variables not captured in our dataset – perhaps notably blood glucose and other metabolic factors – could explain a greater degree of the inter-individual differences in aortic root size.

11.7 LIMITATIONS

The assessment of aortic dimensions and, particularly, aortic morphology, was limited by the available CMR sequences from the Digital Heart Project (DHP). It would be an exciting addition to this work to collect 3-dimensional imaging from the whole of the thoracic aorta, from which one could derive 3-D reconstructions, to allow a much more comprehensive and refined evaluation of aortic morphology. However, the scanning time and processing time to interpret these images limits their use in routine clinical practice and in large-scale research projects like DHP; the benefit of "collapsing" a complex 3D structure to single unitary measurements is that they are easy to present, understand and interpret, both for patients, in statistical models and in clinical decision-making. Clearly a lot of information is "lost" by doing this, but creating easily-defined metrics is a good start to introduce aortic morphological assessment to the clinical setting.

The DHP also was limited in the ethnic mix of the participants. There were insufficient numbers of non-Caucasians to allow accurate modelling of normal ranges across the age spectrum, and therefore this analysis is limited to the UK Caucasian population. Clearly it would be desirable to expand the ethnic composition of the participants.

There are certainly environmental factors which have not been measured in the participants which might be expected to have some impact on aortic dimensions and morphology – in particular blood glucose and insulin resistance, as well as other metabolic factors such as CRP.

Finally, whilst we have demonstrated an association of variables with age, we have not examined longitudinal data to prove that in an individual, changes would occur at the rates we have described. The DHP is undertaking recall of selected participants, and we hope to gather some data to investigate the longitudinal changes in aortic traits.

11.8 CONCLUSIONS

We have presented normal values for MRI-derived aortic dimensions and morphology in a healthy Caucasian population. We have demonstrated significant changes in both dimensions and morphology with age, and shown that these are modified by the presence of additional cardiovascular risk factors, which have individually relatively small effects on the measurements. We have shown that traditional atherosclerotic risk factors are largely associated with a reduced "remodelling" of the aorta within these normal ranges; a fact worthy of further investigation.

11.9 ACKNOWLEDGEMENTS

The Digital Heart Project was conceived and coordinated by Professor Stuart Cook and Dr Declan O'Regan. CMR analysis was performed with help from Ben Statton, Wareed Alenaini and Marina Quinlan for inter-observer statistics and quality control.

12: DETERMINANTS OF HEALTHY AORTIC ELASTIC FUNCTION

"The willow which bends to the tempest often escapes better than the oak which resists it"

– Albert Schweitzer

12.1 INTRODUCTION

12.1.1 Overview

The aorta's elastic function is key to its role in buffering the pulsatile flow from left ventricular ejection[314]. The distensibility of the aorta allows it to accommodate the stroke volume, and in combination with the elastic recoil, smooths out the flow profile and maintains diastolic flow to the peripheries and, importantly, the coronary arteries[1].

Aortic function is an important determinant of cardiovascular mortality, both in the general population[2, 6, 7, 315, 316] and in specific aortic diseases[12, 120]. Aortic elastic function, as measured by pulse wave velocity (PWV) or ascending aortic distensibility, is a marker of cardiovascular risk[2, 5, 6]. Importantly, it exerts its effects on risk largely independently of known cardiovascular risk factors[2]. A 1m/s increase in PWV corresponds to a 7% increase in risk of cardiovascular events for an otherwise healthy 60 year old man[2]. Similarly, reduced ascending aortic distensibility has been shown in several studies to predict increased cardiovascular risk and overall mortality, with a hazard ratio of 2.7 for the lowest quintile of distensibility versus the highest[5].

This implies that there are novel mechanisms of vascular risk that are, thus far, poorly understood, and which could provide more accurate risk assessment and identify new treatment targets.

12.1.2 Measuring aortic function

There are many different techniques described to assess aortic elastic function; each subtly different in the precise component of aortic stretching and recoil that it measures. Most have been associated with coronary artery disease risk, with the strongest evidence coming from studies of PWV and ascending aortic distensibility[2, 5, 6].

Most studies of PWV to date have used the carotid-femoral technique for measuring PWV (CF-PWV)[317]. However, this method introduces significant error – both by incorporating abdominal aorta, iliac and femoral vessel properties into the measurements – and also by introducing significant inaccuracy into the path length measurement[136]. In addition, CF-PWV does not account for the elastic function of the proximal ascending aorta and proximal aortic arch - areas which might account for up to 50% of total arterial compliance. These sources of confounding are avoided by the direct measurement of path length and transit time in the aortic arch using cardiovascular MRI (CMR-PWV; see below for description of method). This has been validated in several studies[318]. There are few large series defining normal CMR-PWV, despite significant differences being reported in mean PWV in population cohorts using the two methods (3.6m/s using CMR-PWV vs 6-10m/s using CF-PWV)[121]. Those that exist, report values from population studies, with high prevalence of cardiovascular risk factors, making it hard to define a "healthy" range of PWV.

Similarly, reports of normal ranges of aortic distensibility come largely from population cohorts.

12.1.3 Impact of anthropometrics and cardiovascular risk factors on aortic function

Whilst the relationship between aortic elastic function and cardiovascular risk has been well-established, the influences on this relationship and the roles of additional cardiovascular risk factors in modifying this risk are poorly understood. Multivariate modelling has been undertaken using a variety of techniques, covariates and transformations of the dependent variables, making it difficult to compare results across studies. What is clear is that the relationship between aortic elastic function and simple biometric variables is far from simple itself. Gender, age, blood pressure, heart rate, lipids, blood glucose and other risk factors all interact in their influence on aortic function.

12.2 AIMS

- To measure aortic elastic function in a healthy cohort using cardiovascular magnetic resonance imaging
- To define healthy aortic elastic function
- To define the influence of anthropometric and basic biometric measures on aortic elastic function

12.3 METHODS

Pulse wave velocity and ascending and descending aortic distensibility were measured by cardiovascular magnetic resonance imaging (CMR) in the Digital Heart Project cohort, with the characteristics and CMR protocol as described in Chapter 11.3.

12.3.1 Distensibility measurement

Distensibility was quantified from the aortic cine images in a plane perpendicular to both ascending and descending aorta, at the level of the pulmonary bifurcation (see Chapter 11.3 for details of MRI protocol). Commercially available semi-automated software (CVI42) was used to measure the minimum and maximum luminal areas of the AA and DA. Pulse pressure was defined as systolic BP – diastolic BP, measured in accordance with ESC guidelines as described in Chapter 11.

Distensibility was then calculated using the following equation:

$$Distensibility = \frac{maximum\ area - minimum\ area}{minimum\ area\ x\ pulse\ pressure}\ x\ 1000$$

12.3.2 Pulse wave velocity measurement

Aortic arch pulse wave velocity was measured according to established methodology[319] using the open-access ArtFun software (LIB, INSERM 1146, France). Firstly, a contour was drawn around the ascending and descending aorta in the axial images, to define the regions of interest (ROIs). This ROI was propagated to the phase contrast images. Pixel intensity was

averaged over the aortic area, and thus a flow velocity curve was drawn up representing the velocity of flow through the axial image plane. The two curves from ascending and descending aorta were normalized and overlaid, with the transit time calculated from sigmoid curves fitted to the systolic upstroke of the two normalised flow velocity curves. Path length was calculated using spline interpolation, from user-defined points placed along the midline of the candycane aortic image. PWV was then calculated:

$$PWV(m/s) = \frac{path\ length\ (mm)}{transit\ time\ (s)}$$

FIGURE 12.1: Screenshot from ArtFun software, showing regions of interest defined around the ascending and descending aorta in the axial plane (cine and phase contrast images)

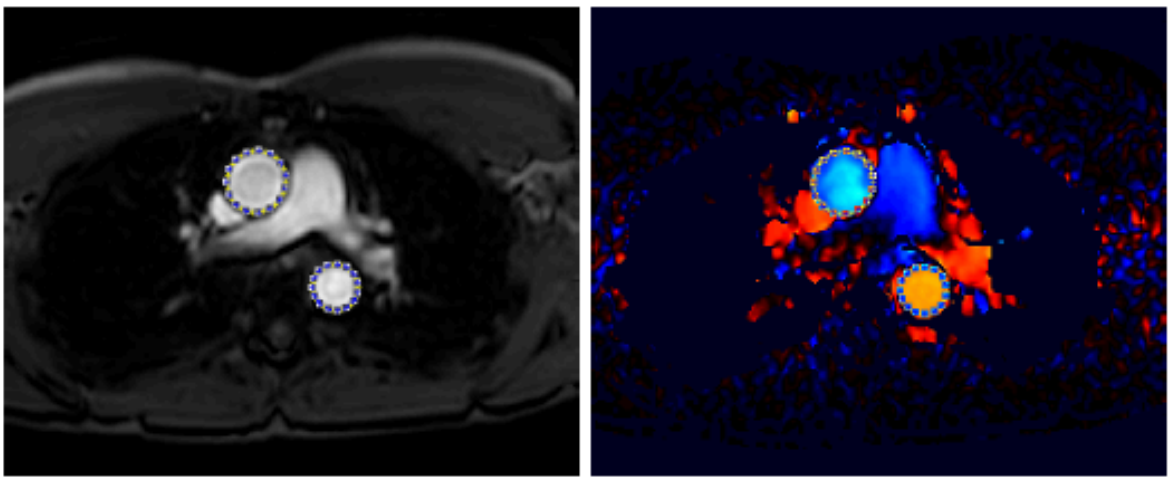
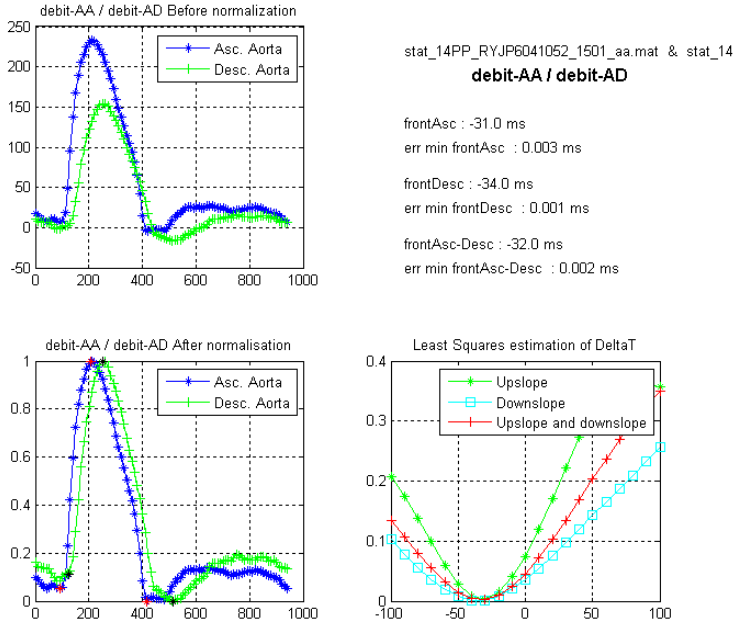
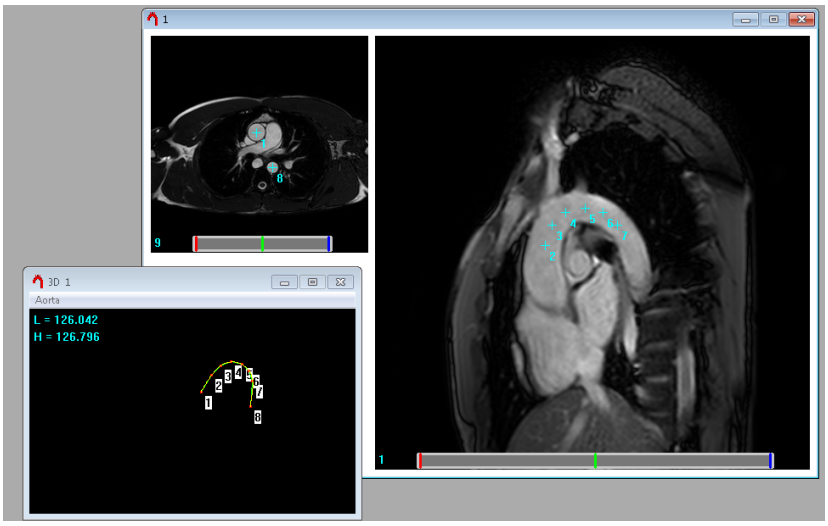


FIGURE 12.2: Matlab figure showing normalised flow profiles from regions of interest selected in ascending and descending aorta.



The curves in the left-hand panels show the flow velocities through the ascending and descending aorta over time, with the top panel showing the raw data, and the lower panel the normalized data. The label "front Asc" refers to the transit time calculated from sigmoid curves fitted to the systolic upstroke of the two generated flow curves.

FIGURE 12.3: Measurement of path length using aortic candy-cane view.



The path length is measured from a derived curve, created from user-defined points placed along the midline of the aortic arch, in the candy-cane view and the orthogonal aortic cine imaging. The image below shows the derived curve (bottom left pane) and the points placed in the mid-line of aorta.

12.3.3 LV parameter quantification (with thanks to Antonio de Marvao)

Analysis of the 2D cine sequences was performed by a trained cardiologist with 3 years of experience in CMR, using commercially available semi-automated software (CMRtools, Cardiovascular Imaging Solutions, London, UK) and using a standard methodology[320]. End-systole (ES) and end-diastole (ED) were identified as the smallest and largest ventricular frames at mid-ventricular level. For quantification of left ventricular (LV) function and volumes, endocardial and epicardial contours were delineated in all slices, in ED and ES. Measurements from each slice were summed using the method of disks. The LV base was identified using the long-axis images (four chamber and ventricular long axis) where the mitral valve position was traced in ES and ED. The systolic descent and twist of the mitral valve was accounted for by tracking the valve motion on the long axis cines, therefore automatically correcting for loss of systolic LV volume due to atrioventricular (AV) ring descent. The papillary muscles were included in LV mass (LVM) and excluded from the blood pool using the signal-intensity driven, semi-automated threshold function. Thresholding level was manually adjusted until endocardial appearances correlated with visual assessment. LVM was calculated from the total myocardial volume multiplied by the specific gravity of the myocardium (1.05 g/mL). LVM and LV blood volumes were indexed when indicated by dividing the parameter by body surface area (e.g. LVM indexed = LVM / BSA). Cardiac output (CO) was calculated using the measured stroke volume (EDV – ESV) and the subject's heart rate during the scan. Concentricity index (CI) was calculated by dividing LVM by end-diastolic volume (EDV).

12.2.4 Statistical analysis

Normal values were defined as those falling between the 5th and 95th centile for our cohort. These were plotted as nomograms using the VGAM package in R, which fits a vector-generalised additive model to the data, using a vector (cubic smoothing spline) smoother to define the smoothed lines representing 5th, 50th and 95th centiles of the data.

Multivariate linear regression analysis was performed in R (R version 3.5.1; R studio version 1.1.456), using the lm() function. For aortic function variables, neither the raw phenotype nor the log-transformed phenotype meets assumptions of normality – in particular, the residuals

from regression models using this "raw" phenotype are not normally-distributed. Standard transformations (log, 1/square root, etc) do not significantly "improve" the distribution, and therefore a rank-based inverse normal transformation was used (hereafter referred to as RN-"variable"). This was achieved by applying the qnorm() function of the "stats" package in R to the ranked variable. Normality of residuals was checked using histograms and quantile-quantile (QQ) plots. This approach to non-normal distribution was preferred to the use of robust (non-parametric) regression, due to the need to use the developed models to take forward for genome-wide association analysis. Whilst it would be possible to use non-parametric regression with robust errors for the subsequent genome-wide associations, transformation of the dependent variable is preferred due to the computational time efficiencies of a standard mixed model regression method. In addition, a rank normalisation facilitates the use of this dataset for meta-analysis with other cohorts.

For PWV, the influence of age is known to be very great, and this influence follows a non-linear relationship, usually modelled as a quadratic relationship (Age + Age^2). This therefore formed part of the basic model for PWV.

Basic models:

RN-distensibility ~ Age + Gender + Height

RN-PWV ~ Age + Age^2 + Gender + Height

Extended models:

RN-distensibility ~ Age + Gender + Height + single additional covariate

RN-PWV ~ Age + Age^2 + Gender + Height + single additional covariate

Selected model:

RN-distensibility ~ Age + Gender + selected significant covariates

RN-PWV ~ Age + Age^2 + Gender + selected significant covariates

Model selection was carried out using the same methods as described in Chapter 2. As the aortic function variables were rank-normalised, only standardised ßs are presented (standard

deviation change in the dependent variable per standard deviation change in the independent variable).

In order to define the effects of distensibility and PWV on indexed LV parameters, the effect of adding elastic function to a basic model including age, gender and SBP was examined.

Reliability estimates for PWV were calculated as described in Chapter 11.3.

12.4 RESULTS

12.4.1 Reliability

Reproducibility of PWV measurement was good for inter-observer variability, with correlation R^2 of 0.97 (95% confidence intervals 0.95-0.98). The coefficient of variability was 7.93. Distensibility calculations were derived from the area measures with reproducibility described in Chapter 11. Inter-test reproducibility was not able to be assessed in the current study, as recalled subjects did not have the relevant MRI sequences repeated. There is a current study ongoing which will recall subjects and assess this. Previous reports have demonstrated that there is good inter-study reproducibility for mPWV using the same method[136].

12.4.2 Defining normal values

This is the largest study to date reporting CMR-derived PWV in the UK population.

Summary statistics are presented and compared with previously-published data (where papers presented distensibility in kPa-1, this was approximated to our values (mmHg-1) by multiplication by 0.133:

TABLE 12.1: Comparison of mean aortic function values with other cohorts

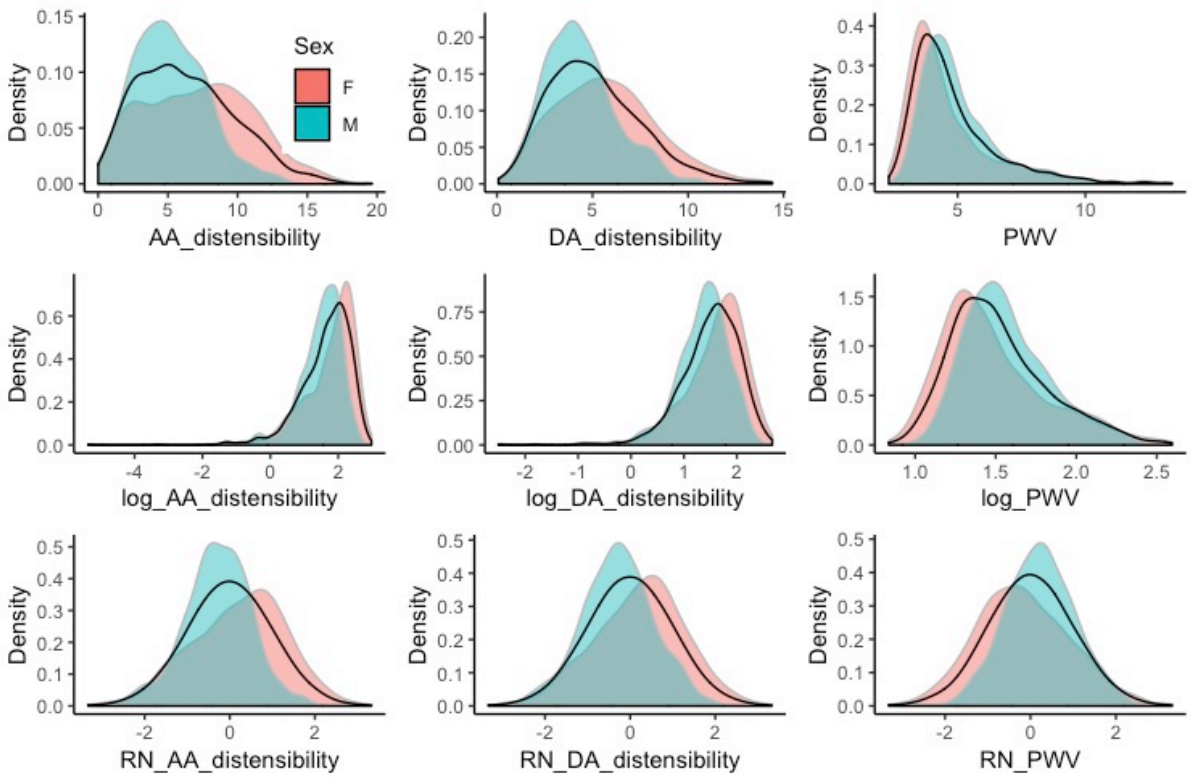
	Digital Heart Project n=1299, mean age = 40		Redheuil et al[101] n=111, mean age 47	Nethononda et al[321] n=777 mean age~ 47		Redheuil et al[18] n=100, mean age 46		Voges et al[322] n=71 mean age 17		MESA[6,5] n=3527; *min* age 45
	M	**F**	**All**	**M (40-49)**	**F (40-49)**	**M**	**F**	**M**	**F**	
Ascending aortic distensibility (10^{-3}mmHg^{-1})	5.2 (2.6)	7.2 (3.9)	4.8 (approx.)	3.8 (1.3)	4.0 (1.6)	3.8 (2.7)	3.8 (3.1)	8.5 (4.2)	9.2 (3.0)	Median 1.6
Descending aortic distensibility (10^{-3}mmHg^{-1})	4.4 (1.9)	5.7 (2.6)	5.6 (approx.)	3.9 (1.4)	4.5 (1.7)	4.4 (3.0)	4.5 (2.4)	7.7 (2.7)	8.8 (3.1)	
Aortic arch PWV (m/s)	5.0 (1.5)	4.6 (1.8)	6.4	6.8 (2.3)	6.5 (3.1)			3.7 (0.9)	3.5 (0.6)	Median 7.4 (IQR 5.6 - 10.2)

Summary of the different studies, including our own (Digital Heart Project) which have reported population values for aortic elastic function. Note the clear relationship with the age of the cohort – the MESA study has a much lower distensibility and higher PWV than the others, likely due to the older age of participants.

Of note, the mean PWV in our cohort was somewhat lower than in most previous studies (see discussion); this is likely to reflect the healthy characteristics and younger age of our cohort. If we consider the ages of the participants in each study, we can see that our data "fits" between the study of children and younger adults by Voges et al and the remaining population-based studies with older and less healthy participants. The same is true for our distensibility data – the mean distensibility in our cohort is higher than the previously-reported population studies, but lower than the paediatric study.

Overall, this data is in line with previous reports.

FIGURE 12.4: Aortic function variables are not normally-distributed



Density plots showing the distribution of different aortic function variables in our cohort. PWV= Aortic arch pulse wave velocity, in m/s; AA distensibility= ascending aortic distensibility at level of pulmonary bifurcation, in $10^{-3}mmHg^{-1}$; DA distensibility= descending aortic distensibility at level of pulmonary bifurcation, in $10^{-3}mmHg^{-1}$. The log-transformed variables are presented in the second row and considerably "worsen" the distribution. Rank-normalised values are shown in the third row (RN-AA distensibility etc.) and have a mean and standard deviation of 0 and 1 respectively.

The distribution shown in the density plots is particularly skewed for pulse wave velocity, and there is a very long upper tail. This suggests that several individuals have high pulse wave velocities which might be thought to be pathological, or at least might place them at higher risk of cardiovascular events.

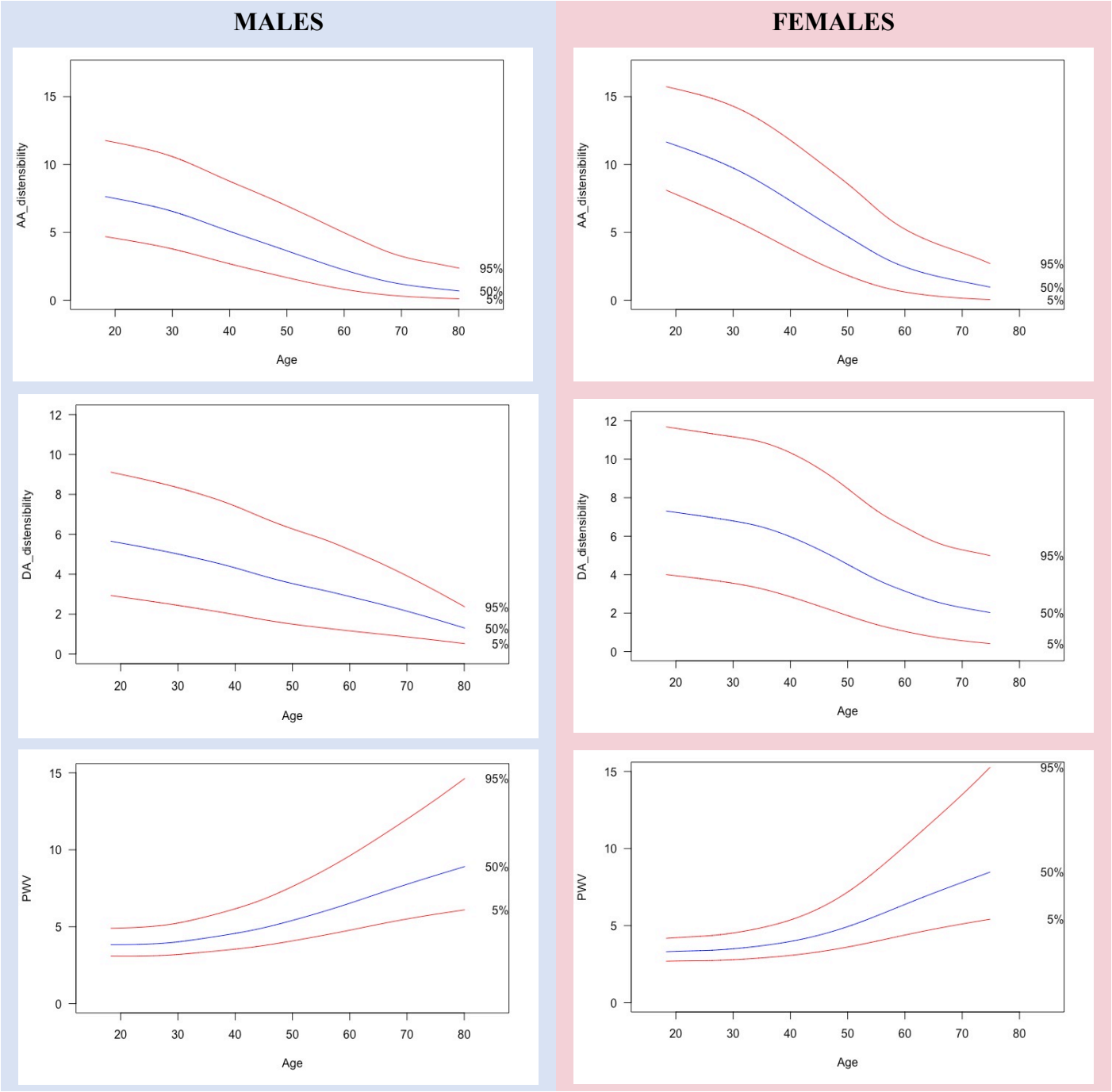
Distensibility follows a different pattern of distribution as there is a physical limit at zero, and the "pathological" values which are associated with increased cardiovascular risk occur at the lower end of the measurement range. It is possible that very high values of distensibility also represent abnormalities, for example in conditions such as Ehlers Danlos where there is a hyper-extensibility-type phenotype, and laxity of joints and tissues. What implications an

abnormally increased distensibility would have for more general cardiovascular risk phenotypes are uncertain.

These distributions are markedly non-normal. “Standard” log transformations did not improve the normality of the distribution (or that of the standard errors of regression models) and therefore a rank-normal distribution, as displayed in the third row of Figure 12.4, was used for regression modelling.

We next defined a normal healthy range of values for AA and DA distensibility and aortic arch PWV, as depicted in the nomograms below, split by gender and plotted against age, the major determinant of the phenotype.

FIGURE 12.5 Nomograms of aortic elastic function defining normal values by age and sex



These nomograms represent centile plots of aortic elastic function variables, derived from smoothing splines fit to the distributions in our cohort. These are split by gender. The red lines denote the 5th and 95th centiles; the blue line is the 50th centile. PWV= Aortic arch pulse wave velocity, in m/s; AA distensibility= ascending aortic distensibility at level of pulmonary bifurcation, in $10^{-3}mmHg^{-1}$; DA distensibility= descending aortic distensibility at level of pulmonary bifurcation, in $10^{-3}mmHg^{-1}$. Age is presented in years. The female plot is slightly curtailed just below 80 years by comparison with the male plots, due to the small number of subjects in the upper age group making a spline fitting inaccurate.

12.4.3 Regression modelling: anthropometric and biometric influences on aortic elastic function

The summary statistics from multiple regression models developed from each of the rank-normalised phenotypes are presented in Table 12.2 below:

TABLES 12.2: Summary statistics from multiple regression models of aortic elastic function

	AA distensibility (rank-normalised)			**DA distensibility (rank-normalised)**		
	Standardised regression coefficient (ß)	**P value**	**Adjusted R^2**	**Standardised regression coefficient (ß)**	**P value**	**Adjusted R^2**
			Basic model			
Age	-0.765	**<0.001**		-0.558	**<0.001**	
Gender (M)	-0.403	**<0.001**		-0.422	**<0.001**	
Height	-0.078	**0.004**		-0.077	**0.027**	
	R^2 for basic model = 0.62			**R^2 for basic model =0.36**		
			Extended model (basic + 1 covariate)			
Weight	0.02	0.34	0.62	-0.05	0.09	0.36
BSA	0.03	0.38	0.62	-0.07	0.08	0.36
BMI	0.02	0.26	0.62	-0.05	0.08	0.36
SBP	-0.28	**<0.001**	0.68	-0.30	**<0.001**	0.43
DBP	-0.09	**<0.001**	0.63	-0.08	**0.003**	0.37
MAP	-0.18	**<0.001**	0.65	-0.18	**<0.001**	0.39
PP	-0.29	**<0.001**	0.69	-0.35	**<0.001**	0.46
TGs	0.00	0.90	0.62	-0.05	**0.05**	0.36
Tchol	0.00	0.94	0.63	0.00	0.93	0.36
HDL	-0.03	0.14	0.63	0.01	0.74	0.36
LDL	0.01	0.77	0.62	-0.01	0.71	0.36
Fatmass	0.03	0.19	0.62	-0.04	0.25	0.36
Smoking	0.02	0.27	0.62	0.02	0.38	0.36
Activity	0.00	0.83	0.62	0.04	0.07	0.36
HR	-0.11	**<0.001**	0.64	-0.20	**<0.001**	0.40
	Selected model: RN-AA distensibility ~ Age + Gender +Height + DBP + PP + HR **Adjusted R^2=0.71**			***Selected model:*** RN-DA distensibility ~ Age + Sex + PP + HR **Adjusted R^2=0.49**		

PWV (rank-normalised)			
	Standardised regression coefficient (ß)	**P value**	**Adjusted R^2**
	Basic model		
Age	0.45	**<0.001**	-
Age^2	0.30	**0.02**	-
Gender (M)	0.39	**<0.001**	-
Height	0.03	0.34	-
R^2 for basic model (PWV ~ age + age^2 + gender + height) = 0.58			
	Extended model (basic + 1 covariate)		
Weight	-0.08	**0.002**	0.59
BSA	-0.10	**0.002**	0.59
BMI	-0.07	**0.002**	0.59
SBP	0.17	**<0.001**	0.61
DBP	0.16	**<0.001**	0.61
MAP	0.18	**<0.001**	0.61
PP	0.08	**<0.001**	0.59
TGs	0.03	0.17	0.58
Tchol	0.05	**0.03**	0.59
HDL	0.02	0.29	0.58
LDL	0.04	0.10	0.58
Fatmass	-0.09	**<0.001**	0.59
Smoking	0.00	0.93	0.58
Activity	0.01	0.74	0.58
HR	0.11	**<0.001**	0.60
Selected model: RNPWV ~ Age + Age2 + Gender + MAP + Fatmass + Tchol + HR **Adjusted R^2=0.62**			

These summary statistics are derived from the basic, extended and selected models described in Methods. The rank-normalised values of aortic function variables are used as the dependent variable. Standardised effect sizes are presented (SD change in dependent variable per SD change in independent variable) to allow comparison of effect sizes. One downside of using rank-normalised values is that effect sizes are difficult to translate into a specific change in the dependent variable. R^2 values are adjusted for the number of predictors in the model. The "selected" model was selected by a combination of Akaike's Information Criterion, using a step-wise method, and manual selection of variables to minimize collinearity and maximise R^2, without introducing too much noise from too many variables. Abbreviations: PWV (pulse wave velocity); BSA (body surface area); BMI (body mass index); SBP (systolic blood pressure); DBP (diastolic blood pressure); MAP (mean arterial pressure); PP (pulse pressure); TGs (serum triglycerides); Tchol (serum total cholesterol); HDL (serum high-density lipoprotein); LDL (serum low-density lipoprotein); Fatmass (body fat percentage); Smoking (ever smoked versus never-smoked); Activity (activity score on numeric index from 0 [sedentary] to 4 [high-level activity most days]); HR (heart rate).

12.4.3.1 Aortic elastic function is predicted by gender

Gender remained a significant predictor for all measures of aortic elastic function, even with correction for other variables. Male gender was associated with stiffer aortas by all measures (a decrease in ascending and descending aortic distensibility and an increase in PWV; standardised ß: -0.4, -0.4 and +0.4 respectively). This relationship persisted after addition of blood pressure variables to the model. The impact of gender was, however, most marked when we examine the change of aortic elasticity with age (see below).

Also of note are the different distributions of distensibility in men and women, with a much broader range of values evident in women (as seen in the density plots (Figure 12.4). This is explored a little further in the discussion, but briefly, this may be partially explained by the different distributions of aortic functional parameters in pre- versus post-menopausal women.

12.4.3.2 Aortic elastic function is strongly predicted by age

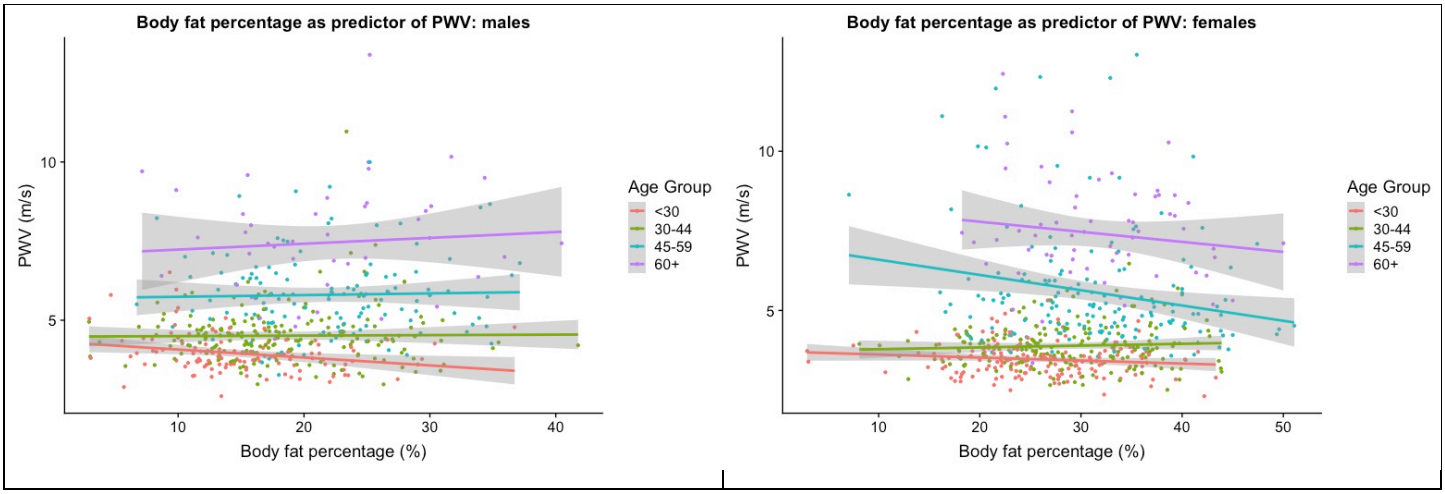
It is apparent from these nomograms (Figure 12.5) and multiple regression models (Table 12.2) that aortic elastic function declines dramatically with age. A notable feature is the different profile of this age-related decline in males and females. Males demonstrate a reasonably steady rate of decline with age. Younger women start off with much greater "elasticity" (i.e. higher distensibility and lower PWV) than men. However, around 40-50 years, there is a sudden steeper decline. It is tempting to speculate that this decline coincides with the menopause. Thereafter, older women show similar ranges of elastic function to men, although the range of measurements is greater in women.

12.4.3.3 Aortic elastic function: relationship with body size and body composition

Body size measures, conversely, are generally not very strongly related to these elastic function measures, with small effects of height on prediction of AA and DA distensibility, but no other significant associations. The impact of body composition is mixed, with increased body fat percentage associated with a modest reduction in PWV. When the cohort is split by gender, we see that this relationship is really only significant in women, and possibly younger men (see Figure 12.6, below). This relationship was strongest in older women (>45 years), although confidence intervals remain wide for the oldest age group due to much higher

variability in measurements. Local measures of elastic function (AA and DA distensibility) were not predicted by body fat percentage.

FIGURE 12.6 Body fat percentage is negatively associated with PWV in females but not males



Scatterplots showing the relationship of body fat percentage to PWV, split by gender and age group. Lines are derived from univariate linear regression per age group, with the standard errors shown in grey.

12.4.3.4 Aortic elastic function is predicted by blood pressure and heart rate

Haemodynamic covariates (BP indices and heart rate) had a greater impact on the local measures of aortic elastic function than on PWV, with blood pressure indices contributing up to 7% of the predictive value of the extended models. Pulse pressure (PP) was the best BP predictor of distensibility; a relationship one must treat with caution as it is used in the calculation of distensibility. Diastolic BP added further predictive power to the model. Mean arterial pressure was the most informative predictor of PWV. Heart rate also predicted elastic function by all measures, with higher heart rates associated with increased pulse wave velocity and reduced distensibility. The direction of these causal relationships is a matter of debate (see Discussion).

12.4.3.5 Aortic elastic function is poorly predicted by additional risk factors

Other cardiovascular risk factors – namely lipids and smoking - were largely not significantly associated with elastic function. Triglycerides modestly improved the basic model for DA distensibility; this relationship however became non-significant in the "Selected" model when BP indices were taken into account. Total cholesterol contributed in a modest way to the prediction of PWV, although individual lipid fractions were not significantly predictive.

12.4.4 Relationship between aortic elastic function and LV parameters

A very detailed assessment of the effect of aortic stiffness on LV parameters is beyond the scope of this thesis. However, the cohort and measurements undertaken afford an opportunity to examine this relationship in the healthy population. A preliminary analysis of the effect of aortic stiffness on indexed LV parameters is presented below:

TABLE 12.3 Effect of elastic function on models of LV dimensions and function

LVMi	**Standardised ß**	**P value**	**Adjusted R^2**
Age	-0.25	**<0.001**	
Gender	15.75	**<0.001**	
SBP	0.23	**<0.001**	
			Basic model: 0.44
+AA distens	0.10	0.51	0.44
+DA distens	0.42	**0.02**	0.45
+PWV	0.26	0.41	0.44

LVEDVi	**Standardised ß**	**P value**	**Adjusted R^2**
Age	-0.37	**<0.001**	
Gender	9.40	**<0.001**	
SBP	0.08	**<0.001**	
			Basic model: 0.27
+AA distens	-0.24	0.13	0.27
+DA distens	0.15	0.40	0.27
+PWV	-0.29	0.37	0.27

LVESVi			
	Standardised ß	**P value**	**Adjusted R^2**
Age	-0.19	**<0.001**	
Gender	5.87	**<0.001**	
SBP	-0.01	0.45	
			Basic model: 0.25
+AA distens	-0.28	**0.004**	0.26
+DA distens	-0.09	**0.037**	0.25
+PWV	0.16	0.40	0.25

LVSVi			
	Standardised ß	**P value**	**Adjusted R^2**
Age	-0.17	**<0.001**	
Gender	3.52	**<0.001**	
SBP	0.10	**<0.001**	
			Basic model: 0.15
+AA distens	0.02	0.80	0.15
+DA distens	0.25	**0.03**	0.16
+PWV	-0.44	**0.03**	0.16

LVEF			
	Standardised ß	**P value**	**Adjusted R^2**
Age	0.09	**<0.001**	
Gender	-3.21	**<0.001**	
SBP	0.05	**<0.001**	
			Basic model: 0.14
+AA distens	0.21	**0.004**	0.14
+DA distens	0.16	**0.047**	0.14
+PWV	-0.30	**0.036**	0.14

LVCI			
	Standardised ß	**P value**	**Adjusted R^2**
Age	0.001	0.053	
Gender	0.11	**<0.001**	
SBP	0.002	**<0.001**	
			Basic model: 0.23
+AA distens	0.004	**0.037**	0.24
+DA distens	0.003	0.09	0.23
+PWV	0.008	**0.041**	0.24

COi	Standardised ß	P value	Adjusted R^2
Age	-13.3	**<0.001**	
Gender	23.23	0.55	
SBP	10.94	**<0.001**	
			Basic model: 0.09
+AA distens	-37.80	**<0.001**	0.11
+DA distens	-42.67	**<0.001**	0.11
+PWV	21.39	0.229	0.09

These summary statistics are derived from basic and extended models of each left ventricular variable, as described in Methods. Where appropriate, the LV variables are indexed to body surface area (Indexed variable = variable/BSA). Standardised effect sizes are presented (SD change in dependent variable per SD change in independent variable) to allow comparison of effect sizes. Abbreviations: SBP (systolic blood pressure); AA distens (ascending aortic distensibility); DA distens (descending aortic distensibility); PWV (pulse wave velocity); BSA (body surface area); LVMi (BSA-indexed left ventricular mass); LVEDVi (BSA-indexed left ventricular end-diastolic volume); LVESVi (BSA-indexed left ventricular end-systolic volume); LVSVi (BSA-indexed left ventricular stroke volume); LVEF (left ventricular ejection fraction); LVCI (left ventricular concentricity index); COi (BSA-indexed cardiac output).

Ascending aortic distensibility was negatively associated with LVESVi and COi, but positively associated with LVEF and LVCI. In other words, stiffer aortas are associated with higher indexed LV end-systolic volumes and lower ejection fractions. The negative association with indexed cardiac output is more challenging to interpret: here, a stiffer aorta seems to predict a higher cardiac output. We can investigate this by adding heart rate – an important component of cardiac output – to the regression model. In a demonstration of the importance of the nature of the covariates in regression models, if heart rate is added into the basic model, this strengthens the association of AA distensibility with LVESVi and LVEF (ß = -0.44, $p<0.001$; ß = 0.27, $p<0.001$) but abolishes the statistical significance of its relationship with COi. Therefore, cardiac chronotropic response might maintain cardiac output in the face of a stiffened aorta.

DA distensibility was, perhaps surprisingly, the only aortic function metric which predicted left ventricular mass (LVMi), and was positively correlated. The lack of other associations could be due to the inclusion of SBP in our basic model. The tight relationship between blood pressure indices and aortic elastic function means that this might confound any effect of aortic stiffness on LVMi. This causal conundrum is one of the great challenges of studying aortic function.

Descending aortic distensibility was also negatively associated with LVESVi and COi but positively correlated with LVSVi and LVEF. Once again, addition of HR to the model abolished the associations with LVMi, COi and LVSVi but strengthened association with LVEF and LVESVi (ß=0.25, p=0.003; ß=-0.32, p=0.003).

Aortic arch pulse wave velocity was negatively associated with LVSVi and LVEF, and positively associated with LVCI. In other words, a stiffer aorta is associated with lower stroke volume, lower ejection fraction and increased concentricity of the left ventricle.

12.5 DISCUSSION

12.5.1 Normal values and healthy cohorts

The Digital Heart Project has provided the largest healthy UK-based cohort to date which has undergone MRI scanning for aortic function assessment. Whilst projects such as the UK Biobank[302] are poised to report distensibility values for a much larger population cohort, the age range is limited (minimum age of 45), the measurements are confounded by risk factors, and there is no full sequence for assessment of "summary" measures of aortic elastic function such as PWV.

Access to such a large cohort as the Digital Heart Project which spans the adult age range and is free of diagnosed cardiovascular risk factors allows us to define a healthy range of values for measures of aortic elastic function and the expected changes with "normal" aging. Our "normal ranges" are somewhat lower (PWV) or higher (distensibilities) than some previously published ranges[6, 321] – this is unsurprising given the selection bias in our cohort for healthy individuals. If such a thing as "healthy vascular aging" exists, this is the group which would represent that process.

12.5.2 Gender and aortic function

Female subjects have lower PWV and higher distensibility at all age points; observations which are not explained solely by the differences in BP or other anthropometrics between men and women. This has previously been demonstrated in multiple studies[6, 315, 321, 323]. Both aortic PWV and distensibility demonstrate steeper changes in women (a reduction in distensibility and increase in PWV) around the age of menopause, raising the probability that

these are under some degree of hormonal control[18, 323]. Indeed there is strong evidence that oestrogen and testosterone can modify vascular stiffness and endothelial function[323]. In the Multi-Ethnic Study of Atherosclerosis[324], lower testosterone in women and higher oestradiol in men were associated with increased distensibility. Conversely, in men, testosterone appears to play a protective role, with testosterone replacement in hypogonadal men reducing PWV[325] and LHRH analogue therapy to induce hypogonadism causes an increase in PWV[326]. In rabbits, testosterone appears to play a direct role in mediating aortic tone[327]. Oestrogens also play many different roles in regulating the vasculature; oestrogen is associated with increased nitric oxide bioavailability and reduced vascular tone[328]. It can also reduce arterial stiffening in post-menopausal women[329]. However, the mechanisms of its effects and the differential effects seen in male and female subjects are complex. Oestrogen receptors α and β seem to have opposite effects on mesenteric arterial stiffness in mice[330-332], whilst the G-protein coupled oestrogen receptor decreases glycosaminoglycan deposition in the medial layer of the aorta[333]. Additionally, knockout studies of oestrogen receptors in male and female vascular smooth muscle cells demonstrate a differential effect on proliferation according to gender. These observations raise the possibility that different genetic as well as environmental influences could mediate the effect of aging on aortic stiffness in men and women. In rats, the QTLs influencing aortic pulse wave velocity are sex-specific[334]; whether this also applies in humans is less certain.

12.5.3 Ageing and aortic function

Another intriguing observation from the nomograms presented above is that the variability of PWV greatly increases with ageing, whilst the variability of distensibility shrinks. This could imply that the local stiffening at the level of the ascending aorta is a “common final pathway” of vascular ageing, but is compensated for in some individuals by more global changes in the properties of the aortic arch – such as the dilatation, lengthening and widening described in the previous chapter. This expansion of the aortic arch, as the intrinsic material stiffens, mitigates the loss of elastic function, and retains a greater vascular compliance. Whether it is those individuals who fail to adapt in this way who exhibit the greatest vascular risk, remains an intriguing possibility. This increased variability also raises the question of whether the

effects of age on aortic stiffness are under genetic control, or whether it is environmental exposures throughout life which determine aortic function.

12.5.4 Haemodynamics and aortic function

I have also examined the effect of a number of different biometric parameters on elastic function. In particular, haemodynamic cardiovascular risk factors are related in a complex way to the measures of aortic elastic function presented here. The relationship between blood pressure (BP), heart rate (HR) and elastic function is a tricky one to untangle; local area changes captured by distensibility measurement are inherently sensitive to loading conditions, whereas PWV as a global measure of arch stiffness is more reflective of the overall regional elastic properties of the aortic wall. An increase in BP may both be an adaptation to, and causative of, deterioration in elastic function; the two parameters are very tightly linked. Several studies have attempted to determine whether increases in blood pressure precede increases in aortic stiffness or vice versa; these have produced conflicting results. A study of 414 treated hypertensives showed that increased pulse wave velocity at baseline was predictive of future increases in blood pressure ($\beta = 0.71 +/- 0.31$)[335]. In the Framingham Heart Study[336], a similar association was found in 1759 participants (Framingham Offspring cohort) with PWV associated with increased SBP and an increased risk of incident hypertension (OR 1.3 per SD PWV). However, in a younger age group (31-52) in the Bogalusa Heart Study, increases in blood pressure seem to precede an increase in pulse wave velocity[337], with the path coefficient from baseline BP to follow-up PWV being significantly greater than the path coefficient ($\rho 1$) from baseline PWV to follow-up BP ($\rho 2 = 0.19$ vs. $\rho 1 = 0.05$ ($P = 0.034$) for diastolic BP). These pieces of evidence all seem to suggest a bidirectional relationship between aortic stiffness and blood pressure, possibly with increased BP in early adulthood causing aortic remodelling and stiffening, which in turn causes an adaptive increase in blood pressure. This “negative spiral” has implications for the immediate challenge of constructing regression models which allow us to examine genetic effects on aortic stiffness: over-correction for blood pressure indices will attenuate the signals from aortic stiffness, whilst a failure to “correct” the model for BP will dilute the stiffness genetic signal with blood pressure loci.

Similarly, the clear positive association of heart rate with aortic function indices could reflect either an adaptive mechanism to compensate for increased aortic stiffness, or could additionally contribute to "adverse" remodelling in the aortic wall by exposing it to more frequent strain and shear stress.

12.5.5 Aortic function and LV size, shape and function

Similar arguments about cause and effect can be made to explain the relationship between aortic stiffness and left ventricular parameters. It is possible that these LV parameter observations represent an adaptation to a stiffer aorta, but equally possible that a pathological process of "cardiovascular ageing" has pleiotropic effects on both the LV and the aorta. The truth is probably a complex mixture of the two scenarios, along with confounding from changes in blood pressure and heart rate discussed above. The association of decreased distensibility with increased LVESVi can be explained by an increase in afterload "felt" by the LV, and a consequent remodelling. This increased LVESVi along with an unchanged LVEDVi would be expected to cause a decrease in stroke volume, as we saw in association with reduced DA distensibility (but not AA distensibility). We did, however, demonstrate an expected decrease in LVEF with reduced distensibility parameters in ascending and descending aorta. Somewhat surprisingly, we demonstrated a negative relationship of AA distensibility with indexed cardiac output. This could be explained by a reduction in heart rate with increased distensibility, and reflective of more efficient blood distribution to the peripheries in a healthy individual. Increasing pulse wave velocity was associated with reduced LV stroke volume and LVEF, along with a higher concentricity index; indicative of adverse LV remodelling to compensate for an increased afterload[18].

These data suggest that even within a healthy population, adverse LV remodelling could occur as a result of increased aortic stiffness. However, the complex causal relationship between haemodynamic parameters, aortic stiffness and LV measures remains to be untangled.

12.6 LIMITATIONS

The Digital Heart Project was limited in its assessment of aortic function – no central blood pressure measures were undertaken. This was mitigated by assessing brachial blood pressure using multiple readings and in similar conditions and as close to the CMR scan as possible.

Additionally, the recruited cohort was not ethnically diverse; the numbers of subjects with valid MRI scans from different ethnicities was too low to draw meaningful conclusions about ethnic differences in aortic parameters in an age- and gender-matched manner.

This study was designed to examine the relationship between biometric variables in a cross-sectional manner. Follow-up studies are planned to acquire some longitudinal data, but these data were not available at the time of write-up. Finally, the relationships between biometric variables and between these and aortic function measures and LV functional parameters are extremely complex. This is not intended to be a comprehensive investigation of the biometric influences on PWV but to form the basis of a good understanding of the relationships between these variables, and to construct valid regression models to take forward for genetic studies. There are additional biometric indices which might significantly affect aortic stiffness; blood glucose levels, insulin sensitivity, thyroid function and homocysteine levels as well as inflammatory markers have all been shown to affect aortic elastic function. However, addition of each covariate introduces noise and error into statistical models and it is therefore important to ensure that these models do not become too complex or noisy before application to genetic data.

12.7 CONCLUSIONS

Normal variations in aortic function with age have been defined in a healthy Caucasian cohort. I have shown that age, gender, body size and composition, blood pressure and heart rate, are all significant predictors of aortic function, and I have explored the nature of some of these relationships. In turn, I have demonstrated that aortic function, even in a healthy population, can predict left ventricular remodelling and function.

12.8 ACKNOWLEDGEMENTS

The Digital Heart Project was conceived and coordinated by Professor Stuart Cook and Dr. Declan O'Regan. CMR analysis was performed with help from Ben Statton, Wareed Alenaini and Marina Quinlan for inter-observer statistics. Dr Antonio de Marvao and Dr Tim Dawes undertook left ventricular phenotyping.

13: GENOME-WIDE ASSOCIATION STUDIES – METHODS, GENOTYPING, IMPUTATION AND QUALITY-CONTROL

"We think in generalities but we live in detail"- Alfred North Whitehead

13.1 INTRODUCTION

13.1.1 Genome-wide association studies (GWASs)

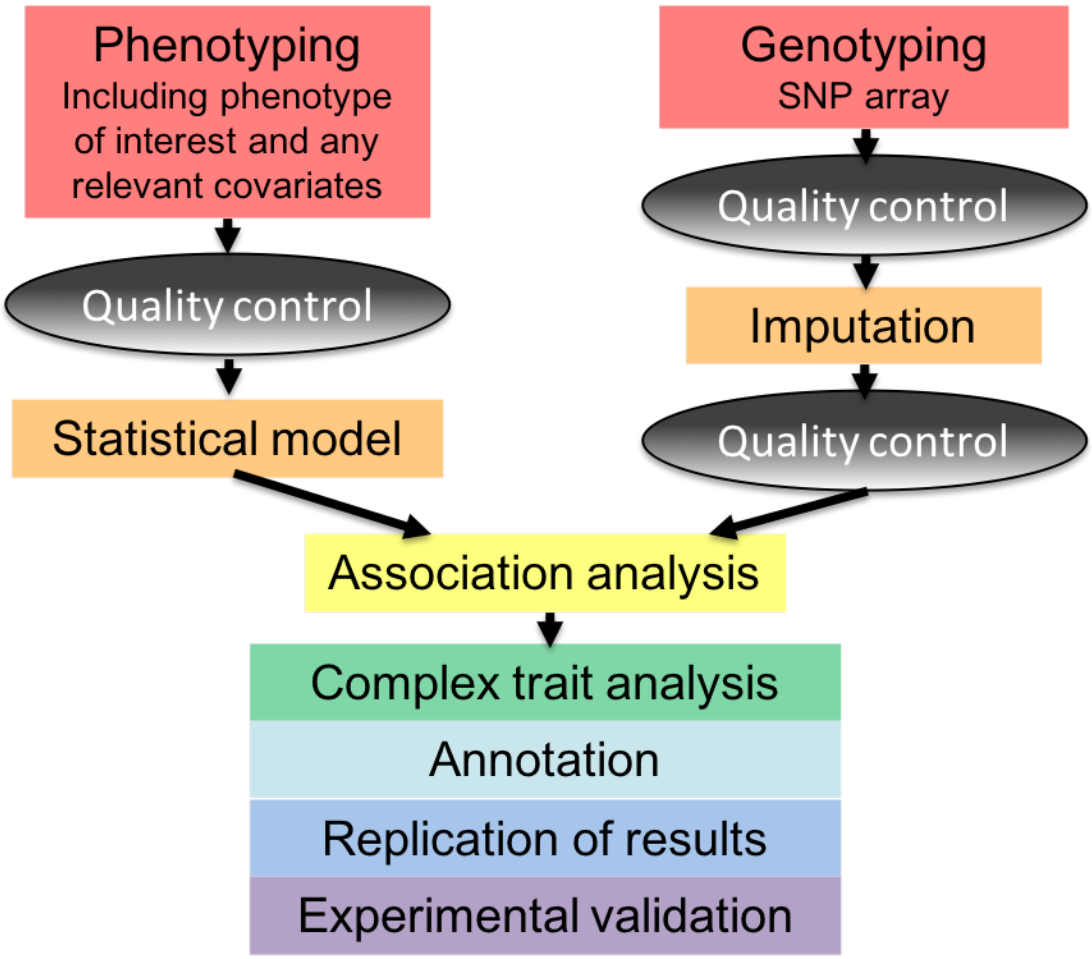
Genome-wide association allows hypothesis-free testing of the effect of common variants across the genome on phenotype. Individually, the effect of each variant associated with the trait in question may be small, but collectively, the combined influence of multiple common variants can be very great. GWASs have the major advantage of requiring no *a priori* assumptions about which genes or pathways are involved in the trait under investigation and have generated huge insights into the allelic architecture of common disease and biological traits. Multiple testing of association at single nucleotide polymorphisms (SNPs) across the genome and large, heterogeneous cohorts, mean that careful quality control at every step of the process is necessary to avoid false positive associations. The power to detect associations is limited in smaller cohorts such as the DHP; we have few subjects, but the advantage that they are a relatively homogeneous population of mostly young, healthy individuals without confounding from overt cardiovascular disease. This chapter will present methods for undertaking the genome-wide association analyses of the phenotypes described in Chapters 11 and 12 above.

13.1.2 Stages of a GWAS

Whilst the basic hypotheses of GWAS are very simple, the complexity of the input information and the potential for false positive associations is very great. Each step in the process of a GWAS is vitally important, and each stage requires meticulous quality control for control of Type I error rates.

Details of each step are described below; figure 13.1 gives a summary of the process:

FIGURE 13.1: Summary of the steps necessary for a genome-wide association study



The study genotypes then underwent detailed quality control, following established recommendations for QC of data prior to GWAS[338], and guidance on combining data from different genotyping batches/platforms[339]. Genotypes were pre-phased and imputed using IMPUTE2[340, 341]. Post-imputation QC was carried out as per established recommendations[338], and related samples were excluded from the data. SNPTEST version 2.5.2[342] was used for association analysis, and results were annotated using the Ensembl Variant Effect Predictor[343], GTEx[28, 29], dbSNP[25, 344] and RegulomeDB[32]. Complex trait analysis was performed using GCTA[345, 346].

Each of these stages requires careful planning and attention to detail.

13.2 AIMS

- To discuss and define the methods selected for quality control and imputation of the healthy cohort genotypes
- To create a high-quality genotype dataset for use in genome-wide association studies of healthy volunteers

13.3 GENOTYPING METHODS

13.3.1 General

Whilst whole genome sequencing might be the ultimate technique for detailed genomic coverage, the cost of genome sequencing at population scale has hitherto limited its use for GWAS. Instead, genotyping arrays have been carefully designed to maximise the coverage of the genome in a given population, and can rapidly, accurately and cheaply "type" up to several million SNPs per individual sample. These can be purchased as "off-the-shelf" options designed for general population use, or customised to maximise coverage of particular genomic regions of interest. The "off-the-shelf" options are optimised to cover coding regions and also provide good coverage of the whole genome at reasonable linkage disequilibrium. One limitation of these arrays is that they have been designed primarily to maximise imputation performance in Caucasian populations. Whilst some effort has gone into the development of arrays to cover different ethnicities, it is difficult within a single study to incorporate multiple different arrays designed for different populations. Cost also remains a factor. In our study, we selected the Illumina Human Omni-Express beadchip, as this provides good coverage across the genome in the Caucasian population, is widely available and (relatively) low-cost, and has been shown to enable high-quality imputation in Caucasian populations.

13.3.2 Beadchip array

The SNP array consists of a silicon wafer with silicon beads each binding a different "address" oligonucleotide. This is a SNP array which uses 3 separate oligonucleotides as capture sequences per variant (see Figure 13.2, below); 2 allele-specific oligonucleotides and one locus-specific oligonucleotide which acts as an "address" tag. Each allele-specific

oligonucleotide has a different universal PCR primer which is labelled with a different fluorescent dye (Cy3 and Cy5). After annealing and amplification, the labelled PCR products are directed to their complementary silicon bead by the address sequence. The relative fluorescence signals from the two different dyes are “read” by a scanner at each “address” on the beadchip; the ratio of the two fluorescent signals at each bead allows determination of genotype at one particular variant. Calling algorithms cluster the intensity of each colour signal at each address to determine whether the locus is homozygous for the reference allele, heterozygous or homozygous for the alternate allele.

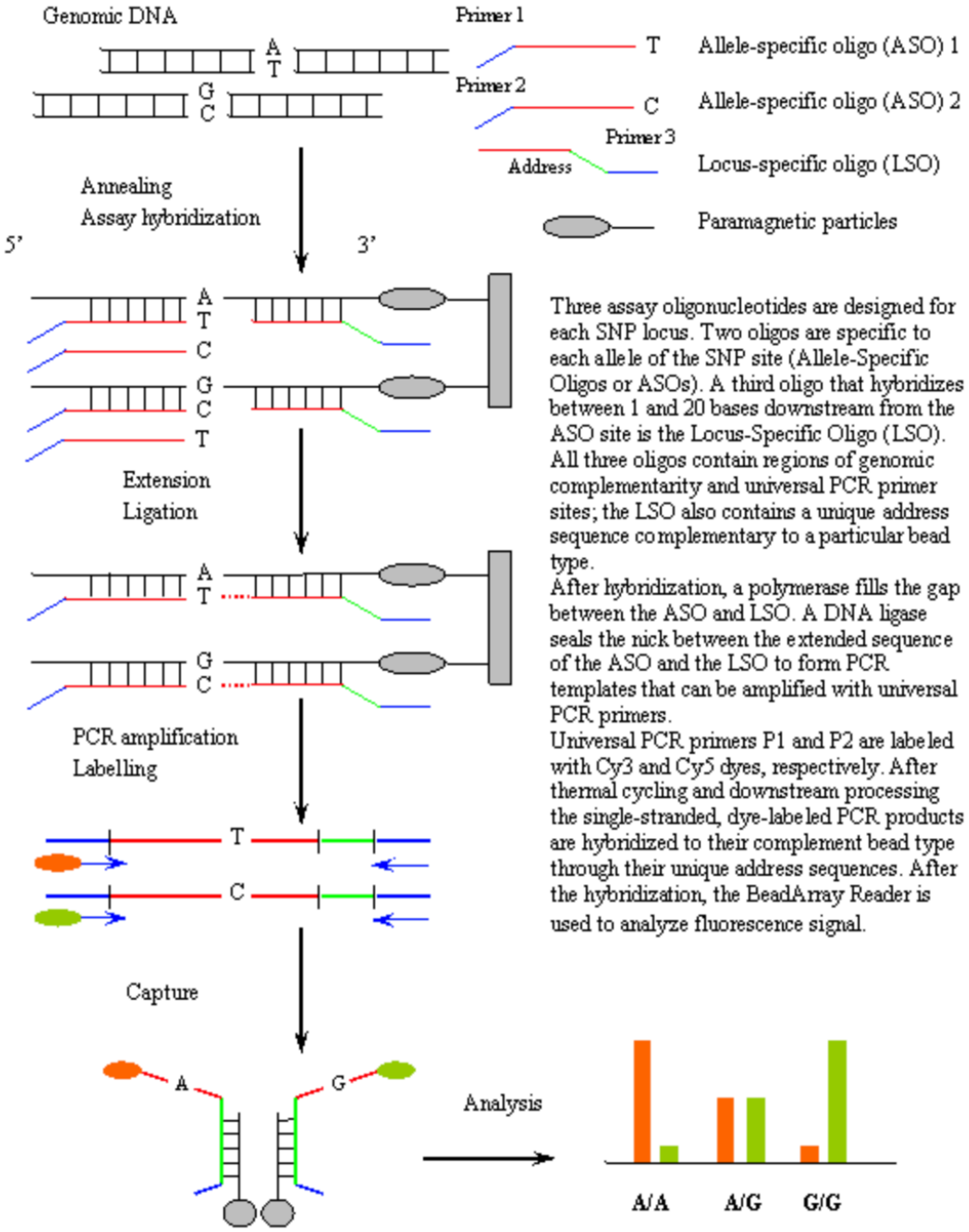
13.3.3 Genotyping in Digital Heart Project – different batches

The genotyping in our cohort was performed in three batches at two separate centres. Batch 1 was genotyped at the Sanger centre using the Illumina HumanOmniExpress-12v1-1 BeadChip ; batch 2 was genotyped in Singapore using the same beadchip and batch 3 was genotyped in Singapore using the updated version of this chip, the Illumina HumanOmniExpress 24 V1-1 BeadChip (v1-1; 713,599 SNPs). Gencall software[347] was used to call genotypes in all cases. Samples were genotyped alongside disease samples from different cohorts, and so the quality control steps, imputation and filtering included these disease cohort samples, and the healthy volunteer cohort was separated after imputation for association testing.

The genotyping in batches, using different chips and in different centres, precluded our re-calling the genotypes in one combined batch. Therefore, we had to undertake additional quality control measures to ensure that any batch effects did not skew our results.

FIGURE 13.2: How the beadchip array works

(from NCBI Probe DB at https://www.ncbi.nlm.nih.gov/probe/docs/techbeadarray/)



Three assay oligonucleotides are designed for each SNP locus. Two oligos are specific to each allele of the SNP site (Allele-Specific Oligos or ASOs). A third oligo that hybridizes between 1 and 20 bases downstream from the ASO site is the Locus-Specific Oligo (LSO). All three oligos contain regions of genomic complementarity and universal PCR primer sites; the LSO also contains a unique address sequence complementary to a particular bead type.

After hybridization, a polymerase fills the gap between the ASO and LSO. A DNA ligase seals the nick between the extended sequence of the ASO and the LSO to form PCR templates that can be amplified with universal PCR primers.

Universal PCR primers P1 and P2 are labeled with Cy3 and Cy5 dyes, respectively. After thermal cycling and downstream processing the single-stranded, dye-labeled PCR products are hybridized to their complement bead type through their unique address sequences. After the hybridization, the BeadArray Reader is used to analyze fluorescence signal.

13.4 PRE-IMPUTATION QUALITY CONTROL

Quality control (QC) of genetic data is probably the single most important determinant of the accuracy and viability of the results of association analysis. This begins with quality control at genotyping stage, with genotype calls being filtered by bioinformatics pipeline on the basis of signal intensity and the clustering of genotype signals in the cohort for calling. Once the genotype data has been called, there are a number of additional QC steps pre- and post-imputation. The QC pipeline we developed broadly follows steps in Anderson et al[338]. We called a mixed cohort of the healthy volunteers from the DHP alongside some disease cohorts (for cardiomyopathies primarily), as these were genotyped at the same time for different studies. There are benefits to retaining as many individuals as possible in the dataset for imputation, as the imputation accuracy can be improved by internal reference to other individuals within the dataset. Samples were genotyped and called in 3 separate batches, each with a mix of healthy subjects and disease cases. Initial per-subject and per-SNP QC was performed by batch, and then the batches were merged pre-imputation. The numbers below in the section on QC therefore reflect the mixed cohort. An overview of batches and numbers of subjects and SNPs excluded at each step is given in Table 13.1 below. I wrote code in bash, awk, perl and R and adapted some code written and kindly shared by Hannah Meyer (based in Ewan Birney's group at the European Bioinformatics Institute) for use with our dataset and on our computing cluster. PLINK (v1.9)[348] was used for some of the basic filtering steps and for genotype file data management.

13.4.1 Alignment with reference dataset

For imputation to work, the input data must match reference data as closely as possible. Aligning to the correct strand and ensuring that reference alleles match between the study and reference datasets is a crucial step. I used a combination of awk and PLINK commands to align the data to the positive strand, using strand files for each genotyping chip. Any SNPs where chromosomes, positions or reference and alternate alleles did not match reference genotype data (UK10K dataset; see section 13.3.1 below) or dbSNP after strand alignment were removed from the dataset. SNP numbers are shown at each step in the table below.

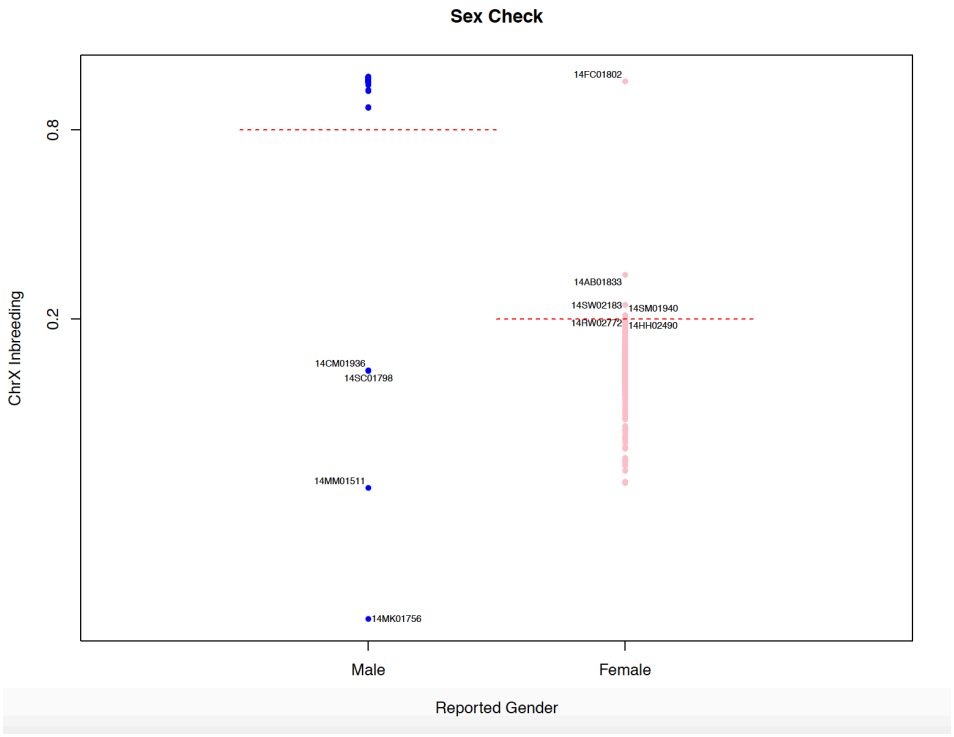
13.4.2 Per subject quality control

It is important to exclude from analysis any individual samples with poor quality of genotyping, or in which phenotype and genotype data does not match. In addition, due to the relatively small size of our cohort, it was important to select samples for inclusion on the basis of ethnicity, to avoid confounding by significant population stratification, and to allow accurate imputation. In larger, biobank-scale studies, it is possible to include individuals from different ethnicities; however, with just a handful of samples in each category of ethnic origin other than Caucasian, we do not have the power in our dataset to perform this analysis. Five quality control steps were generated to create lists of samples for exclusion. Table 13.1 (at the end of section 13.3) shows the numbers of samples excluded at each step.

13.4.2.1 Gender check.

Reported gender was compared with genotyped gender (using X chromosome homozygosity rates). Samples were deemed genotypically male with X homozygosity > 80% and female if X homozygosity was <20%. Samples in between these values were excluded from subsequent analysis. Subjects having a mismatch between reported and genotyped gender were manually checked and were either corrected or excluded from analysis as appropriate.

FIGURE 13.3: Gender cut-offs – example from batch1:



This graph shows X chromosome homozygosity for each sample, plotted by reported gender. There are a few far-outlying samples which represent in some cases sample mix-ups, or data entry errors.

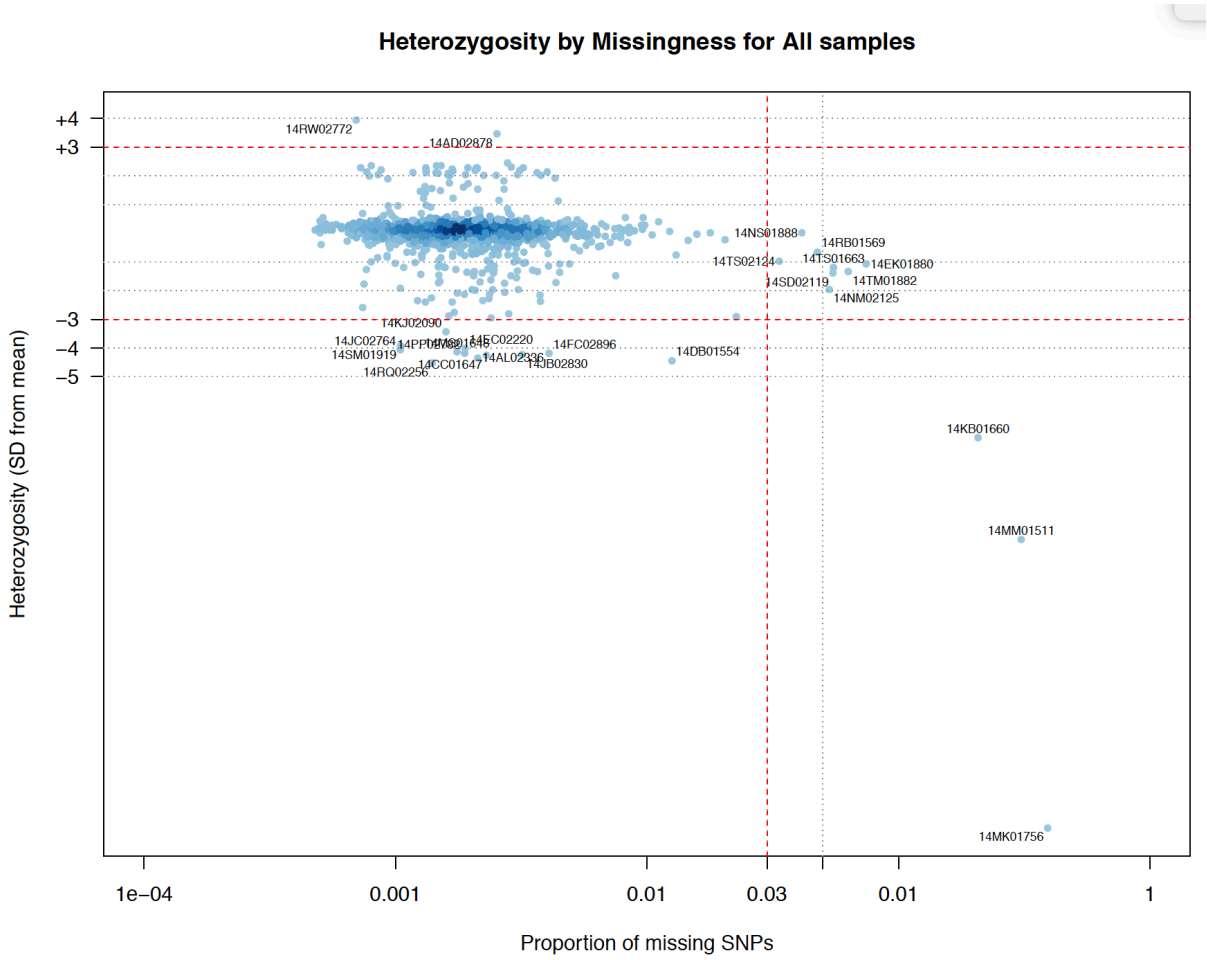
13.4.2.2 Missingness of genotype

The proportion of genotyped SNPs missing per sample was assessed. This reflects the accuracy and success of genotyping overall. Those with >3% missing data were excluded, in line with previous studies[349].

13.4.2.3 Heterozygosity rates

Samples >3 standard deviations away from the mean heterozygosity rate of our cohort were excluded, as these are likely to represent either DNA contamination (if the heterozygosity rate is high) or inbreeding if the rate is low.

FIGURE 13.4: heterozygosity by missingness in our cohort: example from batch 1



13.4.3 Per-SNP analysis

SNPs were filtered on the following basis to remove poorly-characterised markers which might result in false-positive associations:

13.4.3.1 Missing data

SNPs with a call rate <95% in any batch were filtered out pre-imputation. This is a fairly relaxed cut-off, to minimise information loss pre-imputation whilst ensuring that very low-quality SNPs are filtered out.

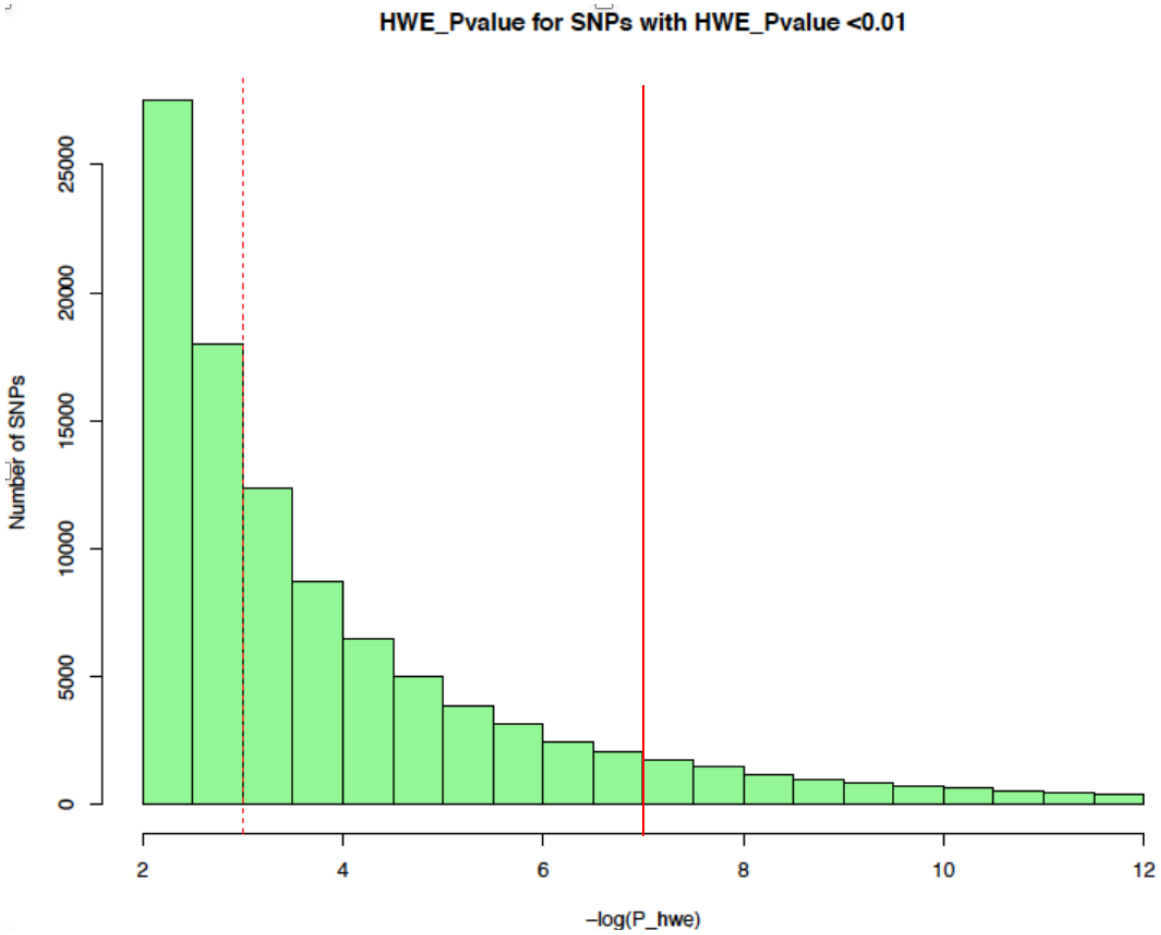
13.4.3.2 Hardy-Weinberg Equilibrium

The Hardy-Weinberg principle states that, in the absence of selection pressures, the proportion of homozygotes and heterozygotes will remain constant according to the

frequency of the SNP in the general population. If p is the frequency of allele a in the general population, and q is the frequency of allele A, then p2 + 2pq +q2 =1. The Hardy Weinberg principle can be used to explore whether observed genotype frequencies in the population differ from expected frequencies. Deviation from expected frequencies can imply genotyping errors or genotype calling errors. There is an argument for not applying HWE filtering, particularly where interest lies in genotype associations with traits which might exert a selection pressure (such as in case-control studies of early-onset disease). It is a matter of argument whether aortic traits could fulfil this criterion. However, as our study population is healthy, and common age of onset of cardiovascular disease is after reproductive age, it seems most likely that significant deviations from HWE are due to genotyping inaccuracies rather than genuine biological effects.

A relaxed cut-off value of $p < 10^{-7}$ is therefore used in this study, with SNPs deviating significantly from the assumptions of HWE in any batch being discarded from the dataset, due to concerns about differential genotyping quality. These SNPs were filtered using the --hwe command in PLINK.

FIGURE 13.5: Hardy-Weinberg Equilibrium p values for SNPs in batch 1



This histogram depicts the number of SNPs falling into each "bin" of p values, with the vertical red line denoting our cut-off value of p<1x10-7. For ease of presentation, only SNPs with a p value <0.01 are plotted.

13.4.4 Merging batches

The nature of the genotyping pipeline resulted in us being unable to re-call genotypes for the cohort as a whole. We therefore had to control for batch effects at each stage of QC. It was necessary to merge batches in as consistent a manner as possible, trying to avoid per-batch effects.

13.4.4.1 Restricting to SNPs genotyped in all batches

Our 3 genotyping batches were merged on the basis of only SNPs which passed initial QC checks (for HWE and missingness) in all 3 batches. This ensures consistency and avoids the problems of comparing genotyped with imputed SNPs at later stages. This is potentially an issue because the imputation process results in a probabilistic representation of the genotype

at each SNP. Therefore, if a SNP is imputed in some batches and directly genotyped in others, the probabilities will be 1 or 0 in the latter, but a fractional probability in the former. This can lead to statistical differences in the genotype associations with phenotype from each batch. Whilst restricting to only common SNPs means that some information is lost from batches which may have higher call rates due to larger sample sizes, this method is accepted as the best way to minimise the risk of false positives arising from batch effects[339]. I created a list of SNPs common to all batches, and then selected these from each batch using the –extract command in PLINK. The genotyping files were then merged.

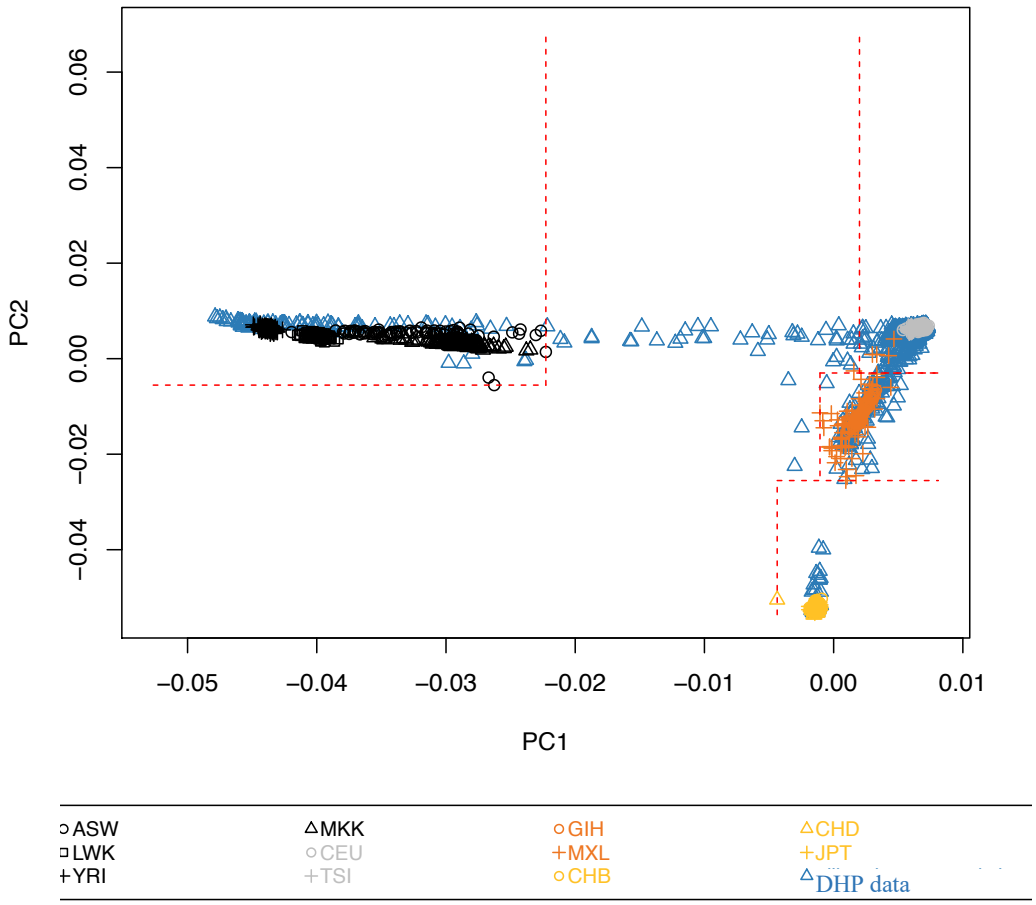
13.4.4.2 Differential missingness

Checks for differential missingness rates between batches were undertaken, using the –test-missing function in PLINK, using each batch in turn as cases. SNPs with p value for differential missingness of $<1x10^{-7}$ were excluded from further analysis. This excluded a further 5449 SNPs, leaving 670823 in the final pre-imputation dataset.

13.4.4.3 Final pre-imputation filtering - Ethnicity

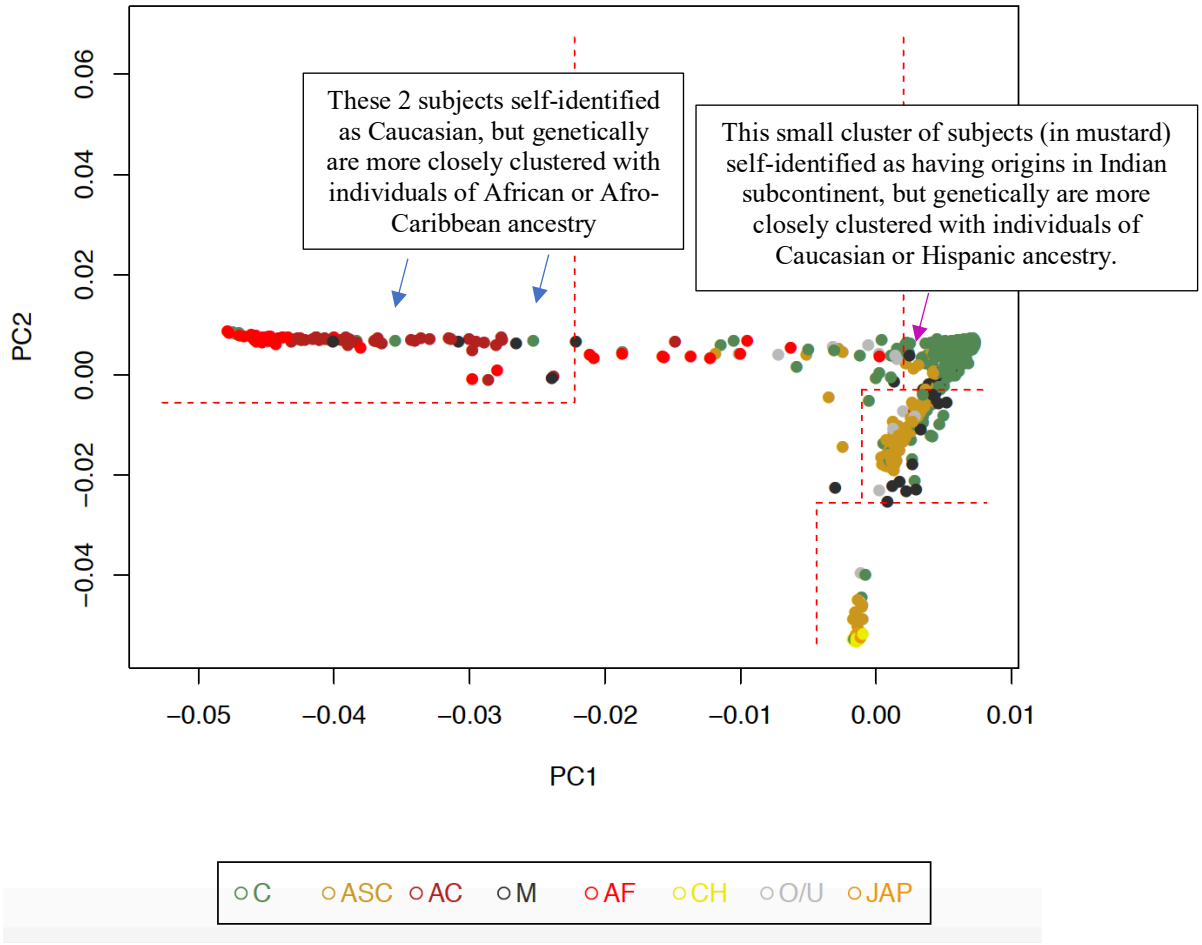
For accurate imputation, as close a match as possible needs to be made between the study and reference populations. We therefore restricted the association analysis to subjects of Caucasian ethnicity. We aligned study genotypes with those from the HapMap III cohort. We used PCA to separate different ethnicities into genotypically distinct clusters, and selected our study cohort on the basis of cut-offs defined by the mean and standard deviation of each HapMap ethnicity cohort (see figure 13.6). Only Caucasian samples were taken forward for imputation. Interestingly, the self-reported ethnicity did not match genetic ethnicity in several cases, with many subjects falling into different categories from those they specified (see figure 13.7). This is firstly testament to the rather over-simplistic nature of attempting to assign discrete ethnic categories to a complexly admixed population. On a more practical note, it has important implications for the accuracy of imputation in these subjects. We therefore used genotyped rather than reported ethnicity for defining subjects to include or exclude.

FIGURE 13.6: Ethnicity matching HapMap populations. Total cohort including disease samples shown



Principal components of genotype, matched to HapMap III as above. HapMap samples are shown in black, grey, orange and yellow. Our own samples are overlaid in blue. ASW: African ancestry in Southwest USA; CEU: Utah residents with Northern and Western European ancestry from the CEPH collection; CHB: Han Chinese in Beijing, China; CHD: Chinese in Metropolitan Denver, Colorado; GIH: Gujarati Indians in Houston, Texas; JPT: Japanese in Tokyo, Japan; LWK: Luhya in Webuye, Kenya; MXL: Mexican ancestry in Los Angeles, California; TSI: Toscani in Italia; YRI: Yoruba in Ibadan, Nigeria; DHP data: Digital Heart Project population (current study).

FIGURE 13.7: Genotyped ethnicity does not always match self-reported ethnicity



Principal components plot matched to HapMap III as above; total cohort (including some disease samples) shown. Colours represent self-reported ethnicity. C: Caucasian; ASC: Indian Sub-Continent; AC: Afro-Caribbean; M: Malaysian; AF: African; CH: Chinese; O/U: Other; JAP: Japanese.

TABLE 13.1: Summary of QC steps and numbers of SNPs and subjects at each step

	Batch 1		Batch 2		Batch 3	
	No. subjects	**No. SNPs**	**No. subjects**	**No. SNPs**	**No. subjects**	**No. SNPs**
Raw data	1346	719665	1216	716503	431	713014
Annotation and SNP alignment to + strand	1346	719387	1216	712411	431	701878
Per-subject qc						
Gender checks	-7	719387	-7	712411	-3	701878
Missingness checks	-11	719387	-23	712411	0	701878
Heterozygosity checks	-17	719387	-53	712411	-11	701878
Final numbers*	1316	719387	1152	712411	418	701878
Per-SNP qc						
Missingness	1316	-3624	1152	-702	418	-1437
Hardy-Weinberg	1316	-91	1152	-2769	418	-269
Final numbers*	1316	716417	1152	701472	418	692802

Merged data	**Number of subjects**	**Number of SNPs**
Differential missingness fails		-7205
Restricting to variants present in all batches	2886	670823
Ethnicity filtering for Caucasians only	2454	670823

FINAL PRE-IMPUTATION DATASET: 2454 mixed samples; 670823 SNPs

After these initial per-individual and per-SNP QC steps, imputation was carried out.

13.5 IMPUTATION

13.5.1 General considerations

Imputation of missing genotypes is a crucial step in generating a dataset for genome-wide association studies[342]. It allows prediction of genotypes at SNPs not directly assayed, and thus expansion of a set of approximately 670,000 SNPs to a dataset containing >9 million SNPs. The aim is primarily to increase power to detect significant associations, but also to enable fine-mapping of associations and facilitate meta-analysis.

Many factors can affect the accuracy of imputation – the major four being the properties of the initial genotyping panel, the rigour of the quality control procedures, the number of haplotypes in the reference panel and the closeness of the ethnicity match between the reference and study population. The initial genotyping panels (Illumina HumanOmniExpress-12v1-1 BeadChip; >710,000 SNPs and its updated version, the Illumina HumanOmniExpress 24 BeadChip) in our study are described above in Section 13.3.2, and have been shown to perform well as a basis for imputation for GWAS in UK study populations.

13.5.2 Reference panel

We selected the merged UK10K[350, 351] and 1000 genomes project[263] dataset as our reference panel as this has a proven record of accuracy of imputation for UK study cohorts, across a range of minor allele frequencies[351, 352]. It also includes a wide range of variants including some short indels and structural variants. For ease of presentation, the imputed variants are all referred to as "SNPs" in this thesis. The reference panel consists of 87,696,888 bi-allelic variants. I downloaded the merged UK10K-1000 genomes imputation dataset in 2017 (with permission) from European Genome-phenome Archive (EGA) with file prefix "_EGAZ00001225178_UK10K_1000GP3_MERGED.chr*.20160410". The initial dataset was annotated with dbSNP identifiers directly from dbSNP[25, 344, 353] version 150.

13.5.3 Pre-phasing

We pre-phased study genotypes using SHAPEIT[354]. The window size for phasing was set to 2Mb and the number of conditioning states per genotype to 200. This programme aligns the study dataset to the reference (in this case, the UK10K + 1KG imputation dataset), and then

resolves haplotype blocks using input from a fine-resolution recombination map of the genome (from HapMap phase III[355]). Recombination rates are estimated per region, and common haplotype blocks are identified within which recombination is a rare event. These haplotypes are then taken forward for imputation, reducing computational burden for the imputation step.

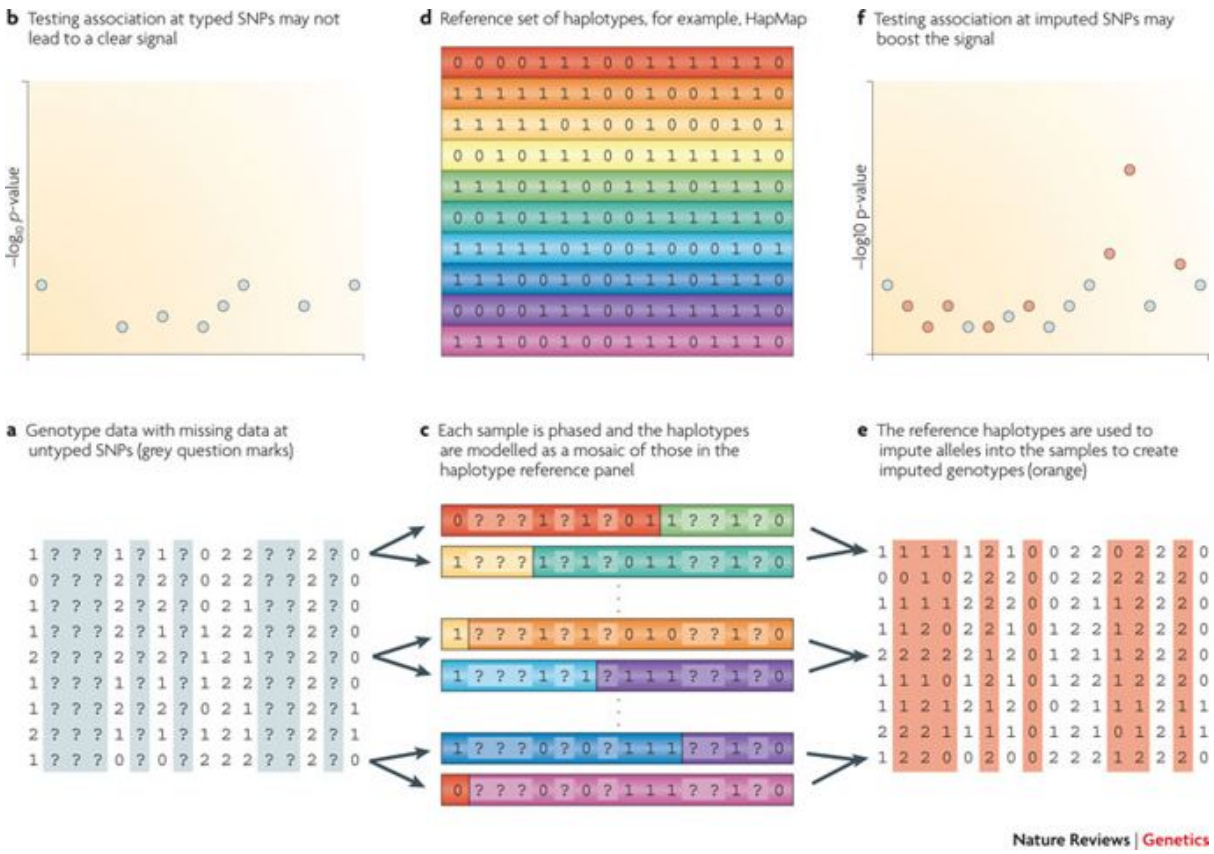
13.5.4 Imputation using IMPUTE 2

Imputation is the process by which missing genotypes may be "filled in" computationally. It allows testing of associations at ungenotyped markers and merging of different datasets (given appropriate quality control measures). It is now widely used in GWAS to enhance power to detect associations at loci which may not have been directly genotyped.

This process operates, in a simplistic way, like fitting together pieces of a jigsaw. For a given haplotype block, the software essentially finds the "best match" of the subject genotype with genotypes in the reference genome by iteration, and "tiles" the reference genomes to fill in the "missing" genotype data for the subject. IMPUTE2 uses a Markov chain Monte Carlo framework which incorporates uncertainty about phase, and fine-scale linkage disequilibrium data to assess the probabilities of recombination at particular points. The marginal probability distributions of missing genotype data are conditioned on the observed haplotype data. A forward-backward model allows checking of imputation accuracy by imputing genotype at known SNPs.

FIGURE 13.8: A graphical explanation of imputation

(from Marchini & Howie, *Nature Reviews Genetics* **volume 11**, pages 499–511 (2010)[342])



This process can only operate accurately for fairly common variation which is represented in the reference panel – rare, or private variation will not be captured by this method. I used IMPUTE 2 (v2.3.2)[341] for imputation. The imputation interval was set to 3Mb, with a buffer region of 250kb on either side of the analysis interval. The number of reference haplotypes was set to 1,000. For all other parameters, I used default settings. The output from IMPUTE 2 is a probabilistic score representing the likelihood of different genotypes at each SNP (in the form (0,1,2): 0.1, 0.80, 0.1. The first number is the probability of being homozygous for the reference allele; the second the probability of being heterozygous and the third the probability of being homozygous for the alternate allele. They therefore sum to 1, and can be combined with the dosage scores to derive an expected "genotype dosage" which represents the expected number of copies of the alternate allele.

This is an attractive model for association analysis, as the output score retains information about the uncertainty of genotype in the model for association, rather than losing this uncertainty by using hard thresholds to "call" particular genotypes.

The quality of imputation is captured both by the estimate of probability itself, but also by an imputation quality score. This information score has a value between 0 (complete uncertainty about genotypes) and 1 (total certainty of genotypes). It can be interpreted that an information score of α for a SNP in a sample of n individuals is equivalent to a set of perfectly observed genotypes in a sample of α*n individuals. Quality control of imputation is also carried out over "bins" of genotype, where concordance is checked using masking of known genotypes and reimputation.

13.6 POST-IMPUTATION QC

13.6.1 Per-SNP – whole cohort

The information info criterion score is generated by impute2 and gives SNP-level information about the accuracy of imputation as described above. We filtered out any SNPs imputed with an information score <0.4, in line with previous studies.

13.6.2 Per-SNP – separated by diagnosis

As the genotyping data contained samples recruited for several different studies, including disease datasets, the cohort was separated by diagnosis, and QC steps repeated as below, to ensure rigorous QC for all study cohorts.

13.6.2.1 Minor allele frequency

SNPs were excluded at this stage if the minor allele frequency in the cohort was <1%. SNPs occurring at lower frequencies are less likely to have been imputed with great accuracy, and association models are not likely to be statistically valid with minor allele counts of under 10 in the cohort.

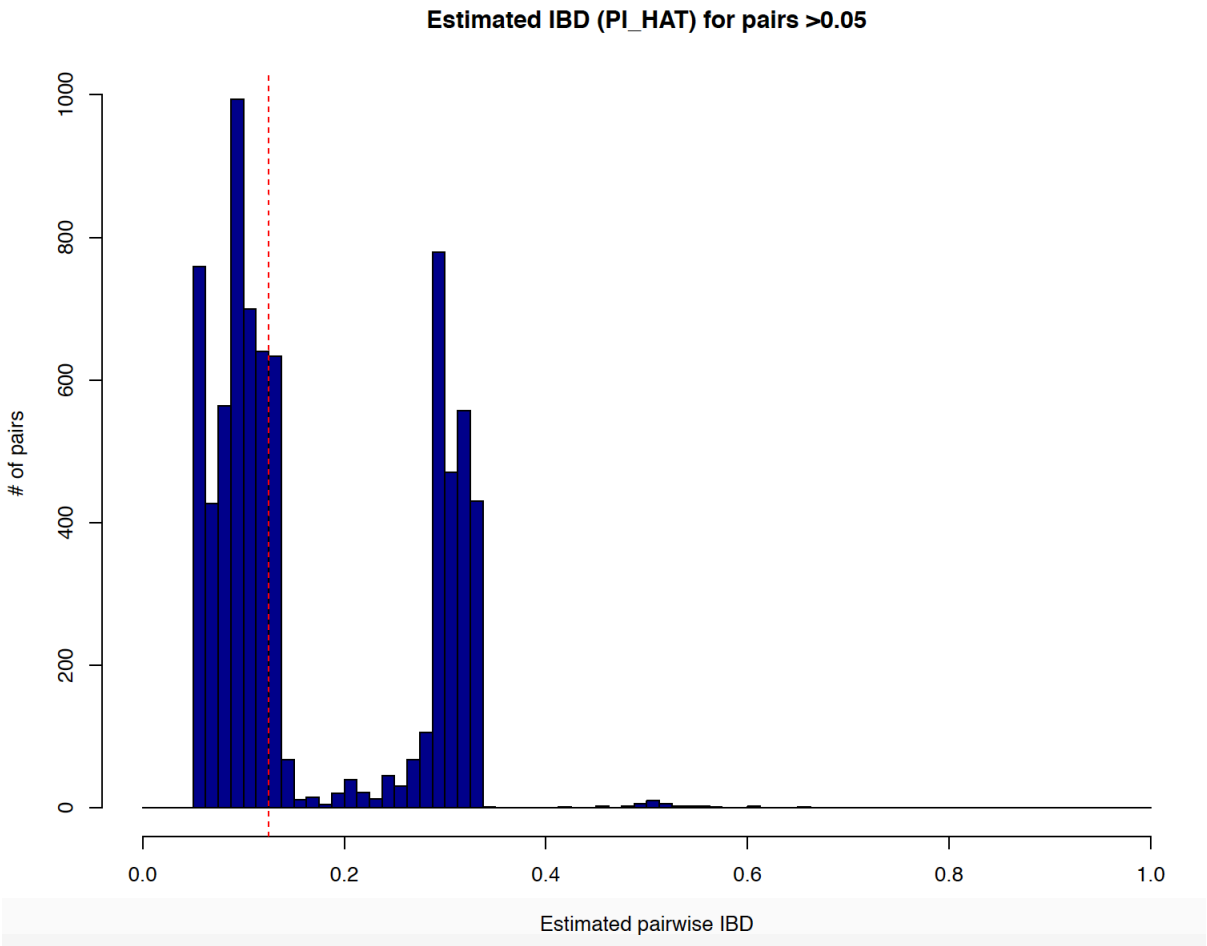
13.6.2.2 Diagnosis-specific HWE

SNPs were excluded if the cohort-specific HWE p value was $<1x10^{-7}$.

13.6.3 Per-individual: diagnosis-specific relatedness

I calculated relatedness between samples using a pair-wise estimation of IBD (Identity By Descent). This involves restricting the dataset to informative SNPs, by removing SNPs in LD (r2=0.05) or with MAF below 0.1, and then calculating the degree of similarity between genomes. Where this exceeded 0.187, samples were deemed to have a relationship closer than 1st cousins and therefore one of each pair was removed (post-imputation).

FIGURE 13.9: Calculating relatedness between individuals: estimated pairwise proportion Identical By Descent (IBD)



13.6.4 Final genotype dataset

For the healthy volunteer dataset, this left a total genome-wide dataset of 9,484,739 SNPs in 1218 individuals passing all quality checks.

13.7 GENOME-WIDE ASSOCIATION ANALYSIS

13.7.1 Statistical models

Simple models were selected to take forward for association analysis, in the form:

Phenotype ~ Age + Gender + Height + Weight + MAP + 3 principal components + genotype

This model was chosen firstly to avoid noise from incorporating too many measured variables. We felt it was important to maintain consistency of the model across phenotypes, to allow for complex trait and joint trait analysis, and to help with replication and meta-analysis. Mean arterial pressure was selected as the blood pressure trait of choice as it is a good summary measure of blood pressure exposure, without being used in the direct calculation of the phenotypes (such as distensibility).

For highly skewed variables with significant heteroskedasticity such as PWV and distensibility, after discussion with statisticians, I decided to transform the dependent variable by rank normalisation. Some results from log-transformed variable associations are presented where the comparison is informative – these are clearly differentiated from the primary results. A rank transformation has the major disadvantage of losing statistical power to detect significant associations, particularly those affecting the extremes of the distribution, but it ensures validity of the assumptions of a mixed regression model in this relatively small cohort, and ensures that the results are not skewed by less common variants found in individuals with values at the extremes of the distribution. There has been much discussion about the validity, desirability and methodology of variable transformation for GWAS in published literature, and I considered several different approaches (including log transformation, Box-Cox transformation, rank-based transformation) as described in the previous chapter. However, as the major variable of interest (PWV in particular) was so right-skewed, I elected to undertake a rank-based inverse normal transformation as this produced the closest approximation to a normal distribution of residuals to take forward for GWAS. This controls the type I error rate whilst conserving power as much as possible (clearly desirable given the relatively small cohort size in our study). This preserves the order of observations so that one can make inferences about a direction of effect of genotype; however, it does not allow one to assess the magnitude of the genetic effect directly from the mixed regression model. A

rank-based transformation also allows for easier meta-analysis in the future, with results from different cohorts easier to combine.

Principal components analysis was undertaken in PLINK v1.9 for the final healthy volunteer dataset using the –pca command. The first 3 principal components were used as covariates for association testing.

13.7.2 Association analysis: SNPTEST[356]

I used SNPTEST version 2.5.2 for association analysis. This software takes input directly from IMPUTE2 output, and constructs linear regression models as described above, using the phenotype as dependent variable, covariates as independent variables, and the expected genotype dosage as the "test" independent variable. An F-test of the association model with each SNP is conducted, and the p value of that statistical test returned, along with a β (estimate of SNP effect size) and standard errors. The models were constructed assuming additive effects of SNP on phenotype. This process is iterated across all 9+ million SNPs. To account for the very large number of significance tests carried out, a stringent p value threshold of $p<5x10^{-8}$ is applied to denote genome-wide significance. This has the benefit of controlling the false positive rate. However, it has the disadvantage of making it highly likely that true associations will be missed, particularly in the context of small effect sizes common in quantitative trait analysis. A p value of $p<5x10^{-5}$ is commonly used to describe "suggestive" associations[357] to define loci for follow-up.

13.7.3 Presentation of results

R package "qqman" was used to plot quantile:quantile plots, comparing expected rates of different significance levels (assuming no "true positive" significant associations with disease) with those observed. These plots were inspected to ensure there was no significant genomic inflation for each phenotype.

The same package was used to plot Manhattan plots for associations. These are presented in the results chapter below.

13.7.4 A note on validation / replication

It was hoped that the UK Biobank resource would be available for validation and replication of the results. However, unfortunately the imaging phenotypes and final genetic associations are not yet available. Additionally, imaging of the aortic arch was not so comprehensive as in this current study, so only basic measures such as dimensions and distensibility will be available (and not PWV or morphological measures). I therefore confined the detailed GWAS analysis here to phenotypes which we would be able to replicate in UK Biobank (with the exception of aortic PWV). Replication of associations is crucial for controlling Type II errors in GWAS, and therefore this will be an important component of follow-up work for this thesis.

13.8 COMPLEX TRAIT ANALYSIS

A number of techniques have been developed to enhance biological insights and to boost statistical power in GWAS like the current study where cohorts are relatively small. I have applied these to the aortic trait dataset to enable more comprehensive biological insights into the genetic architecture of these phenotypes.

13.8.1 Estimating SNP-based heritability

I used GCTA software to undertake analysis of SNP-based heritability using both a standard single-component genetic restricted maximum likelihood (GREML-SC) and the LD- and MAF-stratified (GREML-LDMS) method[345]. Rather than treating each SNP in isolation, this fits large groups of SNPs simultaneously, modelling their individual effects as random, conditioning on the joint effects of all the other SNPs in the model. This allows unbiased estimation of the proportion of variability in phenotype which is due to SNPs across the genome. The GREML-LDMS method further stratifies the analysis by MAF and accounts for heterogeneity in linkage disequilibrium across the genome. This technique uses groups of SNPs binned by MAF, and constructs a genetic relatedness matrix for each bin. The association of phenotype with each bin, accounting for genetic relatedness between samples, and for different patterns of linkage disequilibrium across the genome and between typed and causal SNPs, is combined in a multicomponent GREML analysis. However, with the relatively small size of our cohort, the GREML-LDMS method was not

robust and the standard errors were too great to produce meaningful estimates of SNP-based heritability. Therefore, the results presented here are from "standard" GREML-SC.

13.8.2 Gene-based analysis

This technique aggregates SNPs for association testing; in this case by genomic location of defined genes. I used the FastBAT method from GCTA[358], which applies a fast set-based association technique. This calculates the association for a set of SNPs from an approximated distribution of chi-squared statistics over the set of SNPs, using LD data from a reference genome with individual level genotypes, and summary data from GWAS. The gene list is derived from the UCSC table browser for all RefSeq genes (hg19). A window of 50kb upstream and downstream of the start and end positions of the genes was included.

13.8.3 Joint trait analysis

Many complex traits are related in the biological mechanisms underlying them. This is particularly true of the aorta, where shape and function are tightly linked. Joint trait analysis boosts power to find meaningful associations with aortic biology. There are many different proposed techniques to undertake joint trait analysis. I used SCOPA software[359] for joint trait analysis. This uses reverse regression methodology to assess the significance of each SNP's association with combined phenotypes – in other words, it treats genotype as the outcome, and constructs a regression model which incorporates each phenotype as the covariates. A Bayesian information score is calculated to derive the "best" model fit, and p values for the model at each SNP are derived. Using this method rather than some of the canonical correlation statistical methods for joint trait analysis allows joint trait analysis within the same population. The use of other methods such as GCTA and metaUSAT[360] or meta-CCA[360] with traits measured in the same individuals can lead to over-inflation of association statistics.

13.9 ANNOTATION AND EVALUATION OF ASSOCIATION RESULTS

The generation of significant associations with phenotype is, as discussed in Chapter 10, just the start of a process of evaluation, annotation and, where possible, functional validation. This process can be thought of as having two complementary functions: firstly, to find the SNP or variant which is responsible for driving the effect of a locus on phenotype and secondly, to

determine which gene or genes (or indeed other regulatory features) underlie the functional effect of that SNP. The generation and release of many different open-access databases helps to annotate significant "hits" with evidence for functional and regulatory roles.

13.9.1 The Genotype-Tissue Expression project (GTEx)[28, 265]

GTEx is an ongoing collaboration to analyse RNA expression within different human tissues. The open-access database incorporates RNA-sequencing data collected from 48 different tissues, in 620 individuals. By pairing this data with genotype, expression quantitative loci (eQTLs) can be identified. These are loci at which variants are associated with higher or lower expression of one or more genes. These loci can act as cis-eQTLs – i.e. they act locally on the expression of a nearby gene, or as trans-eQTLs – i.e. they act distantly. By allowing open access to this data across different tissues, this database allows researchers to determine whether a SNP of interest acts as an eQTL in tissues relevant to the trait in question. It is important to note that the multiple testing adjustments required again set the bar high for designation of eQTLs. It may well be that very small changes in expression levels of certain genes – such as those encoding transcription factors – may have very significant downstream effects, and yet these may not reach statistical significance. Nevertheless, this is a key resource for annotation of non-coding variants in particular.

13.9.2 Epigenomics

A number of resources have been developed to identify areas of the genome which are important for regulation, and to identify SNPs which appear to be particularly functionally active. These include methods to assess DNA methylation, histone modifications, chromatin accessibility and chromatin interactions. The Roadmap Epigenomics Consortium[287], and The Encyclopaedia of DNA Elements (ENCODE)[258], have made available genome-wide maps of these key features, along with mRNA expression across several hundred cell types and tissues.

Other databases, such as RegulomeDB[32] and Haploreg[33, 34] seek to amalgamate functional information obtained from multiple different sources. For example, RegulomeDB includes data from Roadmap and ENCODE as well as other published data, on chromatin states, DNAse sensitivity (a marker of how open, active and accessible the DNA strand is at a particular site), CHiP factors, transcription factor binding, as well as a limited selection of eQTLs and

differentially methylated regions. In the case of RegulomeDB, this is synthesised into a single score which reflects the likelihood of a particular SNP having an important regulatory function. There are also a number of cell-type-specific resources such as Hi-C data, and CHiP-seq datasets (discussed in section 10.9), which can be used for look-up of putative causal variants. However, in these data, aortic samples are to date rather sparse.

None of these genomic or epigenomic tools can fully replace functional validation in a laboratory setting – defining a cellular phenotype and using CRISPR-Cas 9 editing to introduce a polymorphism of choice, and assessing the effect on cellular phenotype. The limitation of this approach is that it is, in general, low-throughput, expensive, and depends on having a good intermediate cellular phenotype which can be modelled *in vitro*. The *in silico* tools above can be extremely useful for identifying candidate genes and variants which are responsible for observed associations, and for prioritising them for further functional follow-up.

For this study, I used GTEx, Haploreg v4.1 and RegulomeDB to assess the likely functional roles of variants in close linkage disequilibrium with the top SNP at each associated locus. I used this information, along with literature review, to filter down candidate genes for each locus. It is difficult to present all the detailed annotation for every locus; instead, I present a brief summary of the data supporting the involvement of a particular gene at each locus.

13.10. CONCLUSIONS

I have created a carefully quality-controlled, imputed genotype dataset for 1218 healthy volunteers, comprising genotype data at 9,484,739 SNPs.
I have described the methods used for generating the association statistics presented in the next chapter, and will discuss their limitations in Chapter 14.

13.11. ACKNOWLEDGEMENTS

Genotype arrays were processed by Sanger centre and the Singapore Genomics team. R, perl and awk code was generously shared by Dr Hannah Meyer, European Bioinformatics Institute, which I adapted and developed for use with this cohort on our cluster. Thanks to Dr James Ware and Dr Nicky Whiffin for fruitful discussion on techniques and approaches.

14: GENOME-WIDE ASSOCIATIONS WITH AORTIC TRAITS

"Science is beautiful when it makes…connections between different observations. Examples include the double helix in biology and the fundamental equations of physics." – Stephen Hawking

14.1 INTRODUCTION

14.1.1 Overview

In this chapter, I will present genome-wide association results from a number of different quantitative aortic traits in a healthy population, described in Chapters 11 and 12 above. Our study size is small (n=1218): this limits our power to detect associations reaching genome-wide significance ($p<5x10^{-8}$). However, the healthy, young nature of the population make it more likely that a significant proportion of the variance will be due to genetic factors.

Genome-wide association studies promise much: the ability to assess genetic influences on phenotype across the genome in an unbiased, hypothesis-free manner, and the possibility to aggregate results into genes or pathways to unravel the molecular biology underpinning phenotype. With modern methods of genotyping and imputation, large-scale population studies can be performed on thousands of individuals. Undoubtedly, great insights into disease biology, treatment targets and the biology of complex traits have been gained from these studies. For example, the role of PCSK9 in regulation of lipid levels and in mediating coronary artery disease risk was clearly demonstrated by GWAS[361]. However, GWAS presents a phenomenal statistical challenge, and must be undertaken with a keen awareness of the pitfalls of the techniques, its limitations and the need for detailed further research to follow-up, replicate and validate any associations found. The results presented here are from a discovery cohort; we will need to replicate these in an independent cohort to take further and hope to do so with the data from UK Biobank.

14.1.2 Genomics of aortic traits – previous studies

There have been remarkably few large-scale studies of aortic traits. Most research has focussed on aortic elastic function, and has used carotid-femoral pulse wave velocity (cfPWV) to assess this. The largest study to date has been the meta-analysis by the AortaGen consortium of GWAS of cfPWV[215]. This consolidated data from a total of 20,634 participants.

Despite this large cohort, only 1 locus was associated at genome-wide significance, with one other locus looking strongly suggestive of association. The significant locus was in the *BCL11B* gene desert. Functional studies have not revealed the mechanism for this association – although Maskari et al[362] suggest that it may be related to lymphocyte infiltration. The other strongly suggestive locus was in *C10orf112*, now known as *MALRD1*. Again, the mechanism of this association is unclear.

The other large study to address the genomics of aortic stiffness was the Framingham Heart Study[363]. However, none of the SNPs studied were associated at genome-wide significance. A few suggestive hits were identified, particularly in *MEF2C*, a regulator of cardiac morphogenesis. A genome-wide association scan reported by Tarasov et al[364] identified SNPs in *COL4A1* and *MAGI1* associated with cfPWV, although the latter did not replicate.

Other studies have examined the genetic underpinnings of aortic stiffness using comparison between subjects at the extremes of aortic stiffness. Yasmin et al[365] found associations with tag SNPs in *FBLN1* and *ACAN* when comparing those in the lowest quartile with those in the highest. A twin study also identified a SNP in the promoter region of *CIB2*[366] which was significantly associated with cfPWV.

Candidate gene studies or functional studies have found associations of particular SNPs or genes / gene products with aortic stiffness. These can be found in pathways as varied as inflammatory markers, matrix metalloproteinases, the renin/aldosterone/angiotensin pathways, glucose homeostasis, lipid and triglyceride metabolism as well as extracellular matrix components.

These varied studies have been extremely heterogeneous in their approach to measuring phenotype, to transforming the phenotype for statistical association, to the inclusion (or non-inclusion) of covariates, and to their selection of genes for study. There has been very little replication of results, and the number of participants lags way behind studies of easy-to-measure traits such as hypertension.

The lack of positive findings in these genome-wide studies may be attributable, at least in part, to the complexity of the phenotype in question. Particularly with cf-PWV, which incorporates properties along diverse regions of the vascular tree, it seems clear that influences such as hypertension, blood glucose, inflammation, lipid metabolism, vascular

calcification will be at play. Additionally, environmental confounders have immense impact on phenotype – the effect of these may well dwarf the subtle genomic effects that GWAS are designed to identify. In many of the studies, the participants were older – meaning that aortic stiffness is confounded by many additional cardiovascular risk factors. The cfPWV measurement itself can also be confounded by body composition of the participants, and by difficulties in measuring path length. Finally, cfPWV (and MRI-derived PWV) is a highly skewed variable – therefore, constructing valid regression models is challenging. Transforming the phenotype has been approached by many different groups in different ways – from a logarithmic transformation in some GWAS, to an inverse transformation in others, and no transformation in more. This makes comparisons between studies much more challenging.

These considerations led to the design of the current study: by selecting a study population which is relatively young, and free of overt cardiovascular disease or risk factors, and by using a more precise phenotyping strategy, we can enhance the chances of finding meaningful genetic associations. Additionally, we can use methods to enhance power – such as a gene-based association strategy and pathway-based analyses.

For aortic dimensions, there have been a few studies which included aortic root diameter as part of a more general examination of cardiac traits. The largest of these identified several loci significantly associated with echocardiographic root diameter[367], detailed in the results section for replication in section 14.5.1.2 below. At the time of writing, no studies have yet reported genome-wide associations with MRI-derived dimensions, although there are several such initiatives underway, including the UK Biobank imaging studies.

14.1.3 A note on power and replication

It is important to note from the start that our cohort is relatively underpowered to pick up small effect sizes. Nevertheless, it is possible to identify some patterns of suggestive associations in our data, and to combine the data in gene-based, pathways-based and joint trait analysis to enhance power. We plan to use this cohort to prioritise genes and pathways for further functional analysis, and to combine it with the UK Biobank imaging cohort for further aortic characterisation and replication of the results presented below.

14.2 HYPOTHESES

- Common genetic variants are significantly associated with aortic phenotypes
- Genome-wide association with aortic quantitative traits can identify genes important for aortic biology.

14.3 AIMS

- To identify common genetic variants which are significantly associated with aortic phenotypes
- To complete gene-based analysis of aortic traits
- To quantify the proportion of trait variance explained by the common variants tested
- To identify genes and pathways which are important for aortic biology
- To discuss the limitations of GWAS

14.4 METHODS

For detailed methods, see Chapter 13.

14.4.1 GWAS

The genome-wide association results presented in this chapter were generated from the Digital Heart Project cohort described in Chapters 11, 12 and 13 above. The study population was a healthy, adult population, free from known cardiovascular risk factors or cardiovascular disease. The mean age was 40 – a very young cohort compared with previous publications on aortic traits. Aortic traits were measured using cardiovascular magnetic resonance imaging (CMR), as described in chapters 11 and 12, above. Genotyping and association methods are presented in detail in chapter 13 above. The term "SNP" is used loosely throughout this chapter (and this thesis) for ease of presentation to denote a common variant – most are indeed single nucleotide polymorphisms; some will be small insertions or deletions. Briefly, genotyping was undertaken with the Illumina human OmniExpress beadchips v12.1 and v24.1. Stringent quality control at sample and SNP level was carried out prior to imputation with the 2016 combined UK10K and 1000 Genomes dataset. IMPUTE2 was used for imputation. Samples were excluded if sex checks were failed, missingness rates were > 3%,

relatedness checks were failed (IBD >0.187, ethnicity checks were failed (using PCA). SNPs were excluded if MAF was <1%, HWE p value was $<1x10^{-7}$, or missingness rates were > 3%.

Association analysis was performed with SNPTEST v2.1, using covariates age, gender, height, weight, mean arterial pressure and the first 3 principal components of genotype for all phenotypes except PWV, where an age^2 term was also included in the model (due to non-linearity of relationship). This simple model was selected to minimise data loss and to allow meaningful combination of phenotypes both within this study and across studies. Phenotypes were transformed where necessary to meet assumptions of regression models – these transformations are discussed in greater detail in chapter 13 above, and are noted in the results sections below where appropriate.

Results were visualised with the qqman package in R and using the web-based version of LocusZoom. Annotation was performed with dbSNP, LocusZoom and the Ensembl Variant Effect Predictor. A SNP was deemed to be associated at genome-wide significance if the p value for association was $p<5x10^{-8}$, in line with previous studies. Suggestive levels of association were taken as $p<5x10^{-5}$.

The significance thresholds described here do not account for the testing of multiple phenotypes for association with the same genotype dataset – if this data were to be used as a discovery dataset for publication, appropriate adjustments would need to be made to the significance thresholds. However, as this dataset is used for prioritisation of variants, genes and pathways for further follow-up studies and for replication of results from UK Biobank, for the purposes of this thesis, the "standard" significance thresholds are used.

Each locus with a p value for association $<1x10^{-5}$ was examined using LocusZoom to determine the robustness of the association peak. For genuine associations with common variants, we would expect nearby variants in linkage disequilibrium (LD) with the lead SNP to also display a degree of association with the trait. Therefore, on a Manhattan plot of chromosomal location versus $-\log_{10}$ p value, one expects to see a large "peak" composed of multiple points of association underlying the lead SNP. Where this does not exist, and there is a single SNP seemingly strongly associated with phenotype, this is likely to be (though not inevitably) a false positive association. Those lead SNPs which had robust association peaks are presented in the results below.

14.4.2 Annotation

For each association at suggestive levels of significance, I inspected the association peak visually using LocusZoom as described above. I examined the top SNP in Haploreg v4.1 and RegulomeDB for evidence of functional effect, and for any other SNPs in close linkage disequilibrium with the top SNP which might have functional effects. I looked for eQTLs using GTEx, noting particularly the tissue or cell type in which the eQTL occurred. I also gathered information about gene expression from GWAS. I also performed PubMed searches for each gene, or feature such as LINCs or microRNAs in close proximity to the association peak, and used GeneCards for an overview of gene function. The searches used the terms "Gene" plus "aortic" or "vascular" or "cardiac". This identified particular genes of interest at many of the loci, which are presented below. Only genes which seemed to have biologically plausible links with cardiac or vascular phenotype are presented below. None of the LINCs or microRNAs which I identified close to the association peaks had proven links with vascular phenotypes, so these were in general not considered further, for simplicity of results presentation.

14.5 RESULTS

The results below are genome-wide association results of quantitative aortic traits measured using MRI in healthy volunteers.

Ours is a small cohort, and unsurprisingly, there are very few genome-wide significant single-SNP results. This relative underpowering is evident from the Quantile:Quantile (QQ) plots of our data, which reveal deflation of test statistics. Nevertheless, there are some interesting signals for suggestive associations, and the gene-based and pathways analysis find some significant associations which represent important aortic biology.

For each trait, I present a Manhattan plot which displays the genomic co-ordinates on the x axis, plotted against the negative log of p value ($-\log_{10}$) on the y axis. The horizontal threshold lines represent genome-wide significance (at $p<5x10^{-8}$; not always shown if significance levels do not reach this) and suggestive significance ($p<1x10^{-5}$). QQ plots for each trait can be found in appendix 1. I also present a table of selected top associated SNPs. Only those SNPs which appear to represent robust associations are presented (i.e. SNPs in linkage disequilibrium also appear associated with phenotype as expected).

14.5.1 Aortic Root Dimensions

There were no genome-wide significant loci for any of the aortic root traits examined. However, there were several loci which were associated with dimensions at suggestive levels of significance ($p<1x10^{-5}$); these are presented below. Heritability estimates and association (Manhattan) plots for each of the aortic root dimensions (end-diastolic diameter measured in LVOT view at valve annulus, sinuses of Valsalva and sino-tubular junction) are shown below. Our data replicated effect direction at nominal significance levels for several previously-published associations. Gene and pathways- based analysis yielded further insights into the genetic underpinnings of these traits.

14.5.1.1 Aortic root: Narrow-sense heritability estimates

The narrow-sense heritability for each of the aortic root diameter traits was assessed using GREML (Genome-based Restricted Maximum Likelihood) as described in chapter 13. In brief, this enables an assessment of the proportion of variance in the phenotype which can be explained by all the SNPs in the dataset. In other words, this approximately equates to what proportion of the trait variance is explained by common genetic variants.

TABLE 14.1: Estimates of narrow-sense heritability

Trait	VG/Vp	SE	P value
Aortic valve annulus diastolic diameter	0.21	0.25	0.18
SoV diastolic diameter	0.51	0.29	***0.039***
STJ diastolic diameter	0.65	0.30	***0.019***

VG/Vp: trait variance explained by genotype as a proportion of the overall trait variance; SE: standard error; P value: significance value for estimate of VG/Vp. Note large SEs for these estimates.

The model for AV annulus diameter was not significant. This is probably due to the fact that the valve annulus is much trickier to measure in a single dimension than the other diameters, due to the 3-dimensional structure of the valve annulus and the difficulties in defining leaflet hinge points etc on CMR. This is reflected in the reproducibility statistics in chapter 11. It would appear that the genetic contribution to this measurement is perhaps rather less than the genetic contribution to the traditional “aortic root” measurements. For SoV and STJ diameters, the genetic contribution to phenotype variance from our measured variants was 51% and 65% respectively, although the standard errors for the estimates were large.

14.5.1.2 Aortic root dimensions: replication of previous results

We were able to replicate previous associations with aortic root diameter (with SoV diameter), but only at nominal levels of significance; these effects disappeared in our own data when corrected for multiple testing, but the effect direction was consistent with previous reports.

TABLE 14.2: Replication of previous GWAS associations with aortic root diameter (with sinuses of Valsalva diameter in our cohort)

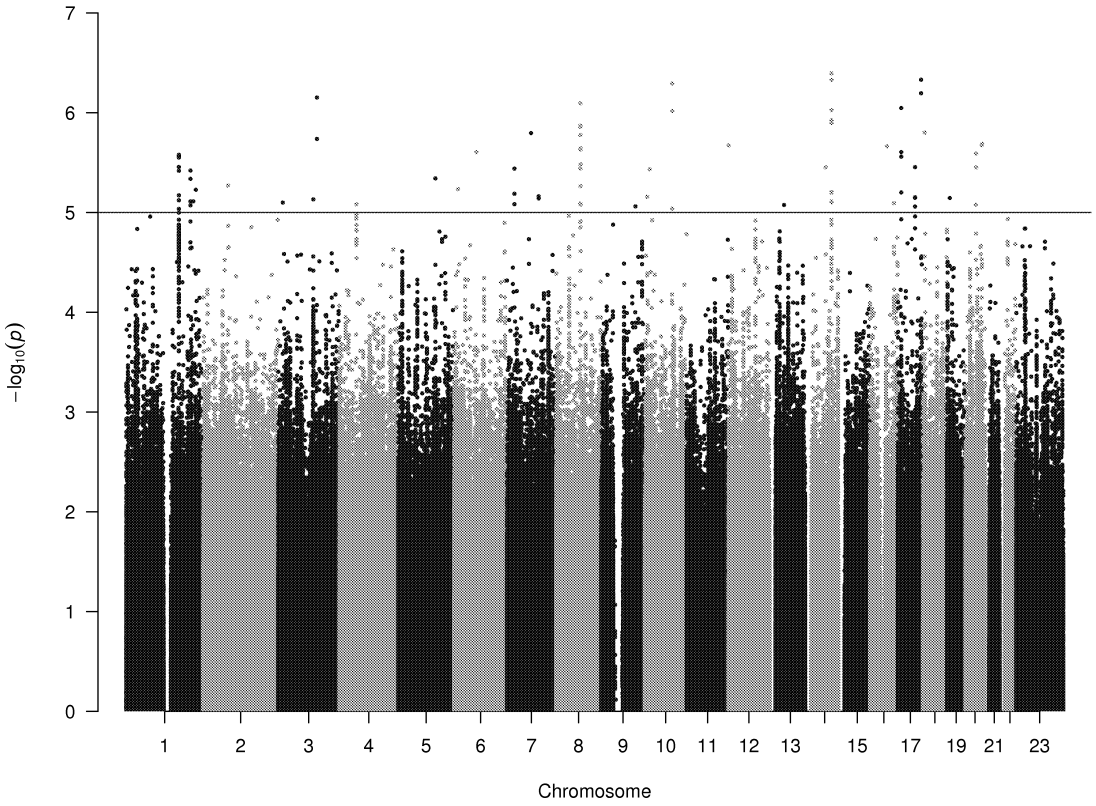
dbSNP ID	Chr:Position (GRCh37)	Near gene	Ref allele	Effect allele	Beta (effect size)	P value in our cohort	Previous beta (effect size)	Previous p value
rs6702619	**1:100046246**	***PALMD***	**T**	**G**	**0.25**	**0.032**	**0.021**	**<1.1x10^{-16}**
rs11207426	1: 59685919	*FGGY*	A	G	-0.21	0.10	-0.017	2.8x10^{-9}
rs17469907	5: 122528420	*CCDC100*	G	A	-0.03	0.84		1x10^{-11} (NR)
rs7127129	11:70027913	*TMEM16A*	G	A	-0.14	0.26	0.015	2.4x10^{-9}
rs4765663	12: 2178760	*CACNA1C*	C	G	0.18	0.26	0.02	4x10^{-9}
rs10878359	*12: 66404624*	*HMGA2*	*C*	*T*	*-0.42*	*0.0008*	*?*	*1.62x10^{-11} (NR)*
rs10770612	12: 20230639	*PDE3A*	G	A	-0.19	0.20		3.2x10^{-12} (NR)
rs806322	**13: 50841444**	***KCNRG***	**G**	**A**	**-0.39**	**0.002**	**-0.021**	**2.2x10^{-15}**
rs2649	15: 63886593	*USP3*	C	T	-0.15	0.39	-0.021	5.3x10^{-8}
rs17696696	**16: 75393352**	***CFDP1***	**T**	**G**	**-0.27**	**0.029**	**-0.016**	**2.7x10^{-10}**
rs1532292	*17: 2097483*	*SMG6*	*G*	*T*	*-0.31*	*0.013*		*1.3x10^{-11} (NR)*
rs17608766	**17: 45013271**	***GOSR2***	**C**	**T**	**-0.60**	**0.0008**	**-0.02**	**2.3x10^{-10}**

Previous p values and effect sizes obtained from Vasan et al[368] and Wild et al[367]. NR = not replicated in previous meta-analysis. In bold are SNPs replicated in our data with consistent effect direction and p value <0.05. In italics are SNPs which have an association p value in our data of <0.05, but for which the previous effect direction was not available for comparison. ***Gene names:*** *PALMD: Palmdelphin; FGGY: FGGY Carbohydrate Kinase Domain Containing; CCDC100 (aka CEP120): Centrosomal Protein 120; TMEM16A (aka ANO1): Anoctamin 1; CACNA1C: Calcium Voltage-Gated Channel Subunit Alpha1 C; HMGA2: High Mobility Group AT-Hook 2; PDE3A: Phosphodiesterase 3A; KCNRG: Potassium Channel Regulator; USP3: Ubiquitin Specific Peptidase 3; CFDP1: Craniofacial Development Protein 1; SMG6: SMG6, Nonsense Mediated MRNA Decay Factor; GOSR2: Golgi SNAP Receptor Complex Member 2*

14.5.1.3 Aortic root: aortic valve annulus single-SNP GWAS results

The Manhattan plot below (Figure 14.1) shows associations across the genome with aortic valve annulus diameter. Whilst no associations at genome-wide significance (p<5x10^{-8}) were found, there were suggestive associations (p<1x10^{-5}) at chromosomes 1, 4, 8, 10, 14 and 17. These loci are presented in Table 14.3 below and loci of interest are displayed in LocusZoom plots in Figure 14.2 below.

FIGURE 14.1: Manhattan plot of genome-wide associations with aortic valve annulus end diastolic diameter



The y axis displays -log_{10} p values of association – i.e. the "higher" the point on the graph, the more significant the association. The x axis is position on each chromosome. Each dot represents a single SNP; however, these often overlie one another. The horizontal threshold line represents a "suggestive significance" level threshold of p= 1x10^{-5}.

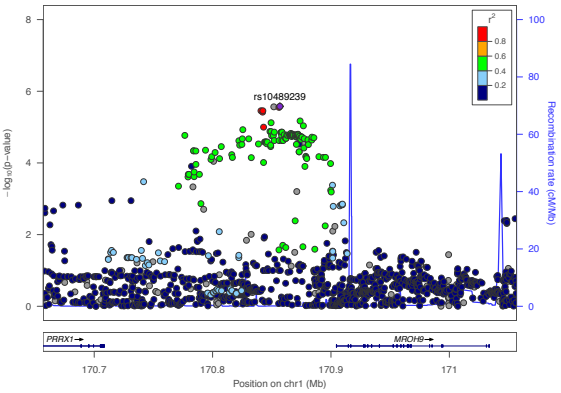
TABLE 14.3: Top SNPs associated with aortic valve annulus diameter

dbSNP	Chr	Position	Ref	Alt	MAF	Beta	p-value	Closest protein-coding gene	Other nearby candidate genes	Comments
rs10489239	1	170856852	T	C	0.215	0.428	$2.64x10^{-6}$	*MROH9*	*PRRX1*	*MROH9* linked to serum creatinine / LDL levels; *PRRX1* linked with AF, BP response to candesartan, serum creatinine
rs2703477	4	55472400	T	C	0.331	-0.354	$8.30 x10^{-6}$	*KIT*	*PDGFRA*	*KIT*: GWAS hit for coronary spasm in women. K/o mice have adverse cardiac remodelling post-MI and worse cardiac function. KIT is QTL for VEGFR2 concentration - a key mediator of vascular development
rs202093145	8	78964253	T	TTTA	0.020	1.301	$8.04 x10^{-7}$	*PKIA*		*PKIA* expressed in endothelial cells; no clear links with aortic traits
rs150929788	10	88372659	T	C	0.011	1.708	$5.07 x10^{-7}$	*OPN4*	*LDB3, WAPL, BMPR1A*	*LDB3* associated with DCM and k/o mice have abnormal cardiac morphology; *BMPR1A* close interaction with TGF-ß signalling pathway and conditional k/o mice have multiple cardiac and vascular abnormalities
rs143724322	14	85083598	T	C	0.012	1.743	$4.02 x10^{-7}$	*FLRT2*	very distant from any genes	*FLRT2* involved in cell-cell adhesion and prev. GWAS linked to heart rate. Mouse k/o have abnormal endocardium and epicardium.
rs4792081	17	10901306	A	G	0.388	0.391	$8.96 x10^{-7}$	*PIRT*	*TMEM220*	*PIRT*-deficient female mice have increased susceptibility to develop obesity and glucose intolerance.
rs71373400	17	55584594	A	G	0.096	0.594	$3.51 x10^{-6}$	*MSI2*	*AKAP1*	SNPs in MSI2 assoc. with coronary artery calcification by GWAS. AKAP1 k.o mice - increased LVH induced by pressure overload and worsened heart failure so may be important for remodelling.

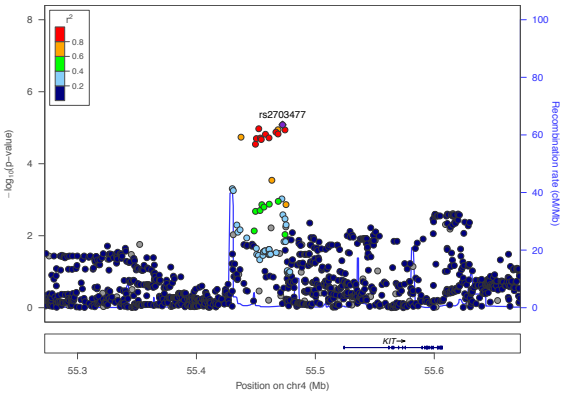
Chr: chromosome; Ref: Reference allele; Alt: effect allele; MAF minor allele frequency; Beta: effect size (change in diameter (mm) per copy of effect allele); Comments: annotations from multiple sources including GWAS catalog[36], GeneCards[369], Mouse Genome Informatics database[370], GTEx[265] and PubMed searches for gene function information; MROH9: Maestro Heat Like Repeat Family Member 9; PRRX1: Paired Related Homeobox 1; KIT: KIT Proto-Oncogene Receptor Tyrosine Kinase; PDGFRA: Platelet Derived Growth Factor Receptor Alpha; PKIA: CAMP-Dependent Protein Kinase Inhibitor Alpha; OPN4: Opsin 4; LDB3:LIM Domain Binding 3; WAPL: WAPL Cohesin Release Factor BMPR1A: Bone Morphogenetic Protein Receptor Type 1A; FLRT2: Fibronectin Leucine Rich Transmembrane Protein 2; PIRT: phosphoinositide interacting regulator of transient receptor potentials; TMEM220: Transmembrane Protein 220; MSI2: Musashi RNA Binding Protein 2; AKAP1: A-Kinase Anchoring Protein 1.

FIGURE 14.2: LocusZoom plots for selected loci associated at suggestive significance levels with aortic valve annulus diameter

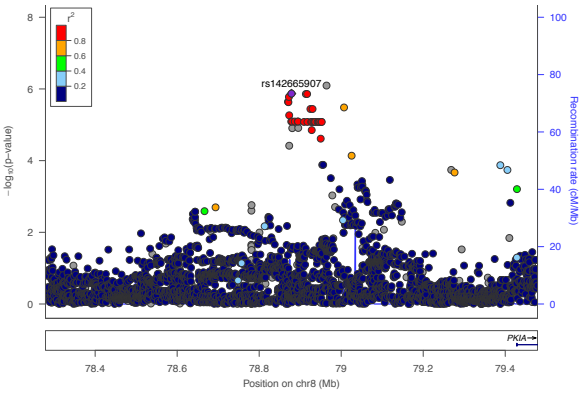
a. Chromosome 1 locus (MROH9 / PRRX1)



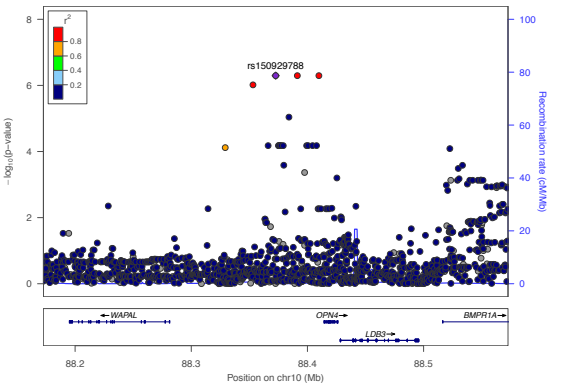
b. Chromosome 4 locus (KIT/PDGFRA)



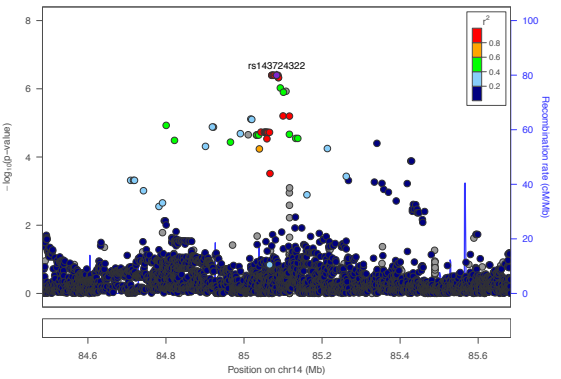
c. Chromosome 8 (PKIA)



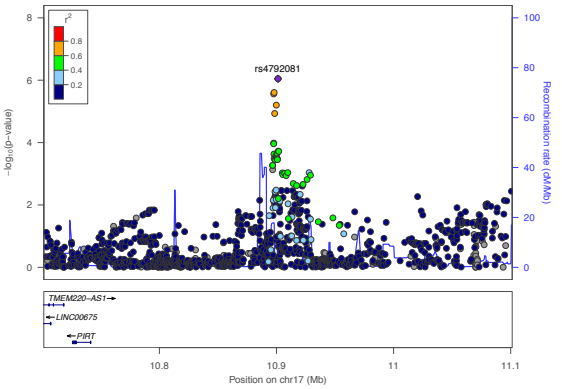
d. Chromosome 10 (LDB3 / BMPR1A)



e. Chromosome 14 (FLRT2)



f. Chromosome 17 (PIRT)



Suggestive loci are shown as examples above in LocusZoom. These plots are essentially a "zoomed in" version of a Manhattan plot, but also demonstrate the linkage disequilibrium structure of the association peak. They demonstrate the spatial relation of the associated SNPs with annotated genes, and with nearby SNPs in linkage disequilibrium (LD) with the lead variant. The plots selected here demonstrate the expected structure of a "true positive" association, with SNPs in LD with the lead variant also associated with phenotype, albeit at lower significance. The labels a-f denote the chromosomal location of the loci, and the top possibilities for candidate genes at the locus, based on biological plausibility and functional evidence, detailed in Table 14.3 and below.

The chromosome 14 locus is the most significantly associated with valve annulus diameter. Within the locus, there are some histone enhancer marks in vascular endothelial cells but no eQTLs or regulatory regions which are clearly linked with vascular traits. The closest protein coding gene, a huge 910kbp away, is *FLRT2 (Fibronectin Leucine Rich Transmembrane Protein 2),* which encodes a cell-cell adhesion molecule which regulates early embryonic vascular and neurological development. This regulates fibroblast growth factor signalling cascades. It is required for normal cardiac morphogenesis and structure of the cardiac basement membrane – knockout mice have abnormal cardiac development. *FLRT2* is therefore a potential candidate gene.

The top SNP at the chromosome 10 association peak is a non-coding SNP (rs150929788). This position has enhancer histone marks in foetal heart tissue. The closest gene, *OPN4 (opsin 4)* encodes a protein whose function is predominantly in the retina[371], and does not appear to have a plausible biological link with aortic traits. However, two further nearby genes, *LDB3* and *BMPR1A,* both have significant known impacts on cardiac traits. The first, *LDB3 (LIM domain binding 3),* approximately 30.5kbp from the lead SNP, is implicated as a cause of dilated cardiomyopathy and left ventricular non-compaction[372]*. BMPR1A (bone morphogenetic protein receptor type 1A)* is part of the BMP signalling pathway which has key roles in cardiovascular development and homeostasis[372]. Conditional knockout mice have multiple cardiac and vascular abnormalities[373]. Variants in *BMPR1A* are known to be causes of juvenile polyposis syndrome through their effects on SMAD signalling[374]. In another form, this syndrome (caused by variants in *SMAD4*), can also be linked to aortic dilatation and dissection, so it would make intuitive sense that *BMPR1A* may also contribute to aortic phenotype. The top SNP is not known to be part of a significant eQTL for either of these candidates.

The chromosome 8 locus tags *CAMP-Dependent Protein Kinase Inhibitor Alpha (PKIA)* – whose product is an inhibitor of chronic ß-adrenergic signalling through CAMP-Dependent Protein Kinase (PKA). This is a key regulator of the effects of sympathetic activation and research suggests that PKA inhibition might reduce pathological cardiac remodelling in response to ß-adrenergic stimulation[375].

The top SNP at chromosome 1, rs10489239, overlaps an eQTL for *MROH9 (Maestro Heat Like Repeat Family Member 9)* in transformed fibroblasts (NES 0.36, p $4.8x10^{-6}$)[29]. This gene is expressed in cardiac and arterial tissue, but is not known to be expressed at high level. By contrast, the other nearby gene at this locus, *PRRX1*, is highly expressed in arterial tissue. Knockouts in mice are known to have abnormal aortic arch, pulmonary trunk and foetal ductus morphology. This gene has also been associated with hypertension phenotypes and atrial fibrillation in previous GWAS. This information perhaps makes this the more plausible candidate to mediate effect on aortic size.

The chromosome 4 locus identifies *KIT* as an interesting candidate gene, being a key regulator of VEGF2 signalling. VEGF2 is an important molecule in angiogenesis and vascular development. KIT knockout mice exhibit adverse cardiac remodelling after myocardial infarction[376]. However, another key regulator of cardiac development, *PDGFRA*, lies approximately 300 kb from our association peak. *PDGFRA* regulates neural crest cell migration and development of the great arteries[377], and is also expressed in some aortic valve interstitial cells.

These results therefore implicate *FLRT2, LDB3/BMPR1A, PKIA, PRRX1* and *KIT/PDGFRA* as potential mediators of valve annulus diameter, and genes to follow up with functional studies.

14.5.1.4 Aortic root: aortic valve annulus gene-based GWAS results

Gene-based analysis allows the combination of association data across SNPs and can enhance the power of GWAS to identify biologically important signals. By amalgamating the association for each SNP annotated to a gene, genes in which multiple SNPs have small effect sizes on the trait can be identified. For the current analysis, the fast-BAT method in GCTA was used (see chapter 13 for details).

None of these gene-based associations were significant when multiple testing corrections were applied (for 20,000 comparisons). The top 10 genes associated with aortic valve annulus diameter are listed in Table 14.4 below:

TABLE 14.4: Gene-based associations with aortic valve annulus diastolic diameter

Gene	Chr	Gene start	Gene end	No. SNPs	Chi-sq	P value	P value most sig SNP	Most sig SNP
OTX1	2	63277191	63284966	79	262.719	9.95×10^{-5}	6.04×10^{-5}	rs72811598
QTRTD1	3	113775581	113807268	107	400.418	0.0001	7.38×10^{-6}	rs9823071
RNF150	4	141786724	142054616	467	1029.14	0.0002	3.33×10^{-5}	rs4317270
KCNG1	20	49620192	49639675	154	385.625	0.0002	2.16×10^{-5}	rs73119384
KIAA1407	3	113682983	113775460	160	452.995	0.0003	2.76×10^{-5}	rs1386478
STPG2	4	98480024	99064391	341	1129.11	0.0004	0.0001	rs35714120
FRY	13	32605436	32870776	426	874.046	0.0005	1.54×10^{-5}	rs17507316
LMAN2L	2	97371666	97405813	70	192.716	0.0005	0.0003	rs2314398
FGFRL1	4	1005609	1020686	221	551.048	0.0006	8.61×10^{-5}	rs71168806
KIAA0226L	13	46916134	46964177	123	371.775	0.0007	0.0001	rs61684147

Top 10 known protein-coding genes associated with aortic valve diastolic diameter. Chr: chromosome; Gene start/end: left and right-side boundaries of the gene (GRCh37); No.SNPs: number of SNPs occurring within gene contributing to the analysis; Chi-sq: sum of chi-squared test statistics of all SNPs within the gene region; P value: segment-based test p value; P value most sig SNP: smallest single-SNP p value in the segment. ***Gene names:*** *OTX1: Orthodenticle Homeobox 1; QTRTD1: aka QTRT2, Queuine TRNA-Ribosyltransferase Accessory Subunit 2; RNF150: Ring Finger Protein 150; KCNG1: Potassium Voltage-Gated Channel Modifier Subfamily G Member 1; KIAA1407 (aka CCDC191): Coiled-Coil Domain Containing 191; STPG2: Sperm Tail PG-Rich Repeat Containing 2; FRY: FRY Microtubule Binding Protein; LMAN2L: Lectin, Mannose Binding 2 Like; FGFRL1: Fibroblast Growth Factor Receptor Like 1; KIAA0226L (aka RUBCNL): RUN And Cysteine Rich Domain Containing Beclin 1 Interacting Protein Like*

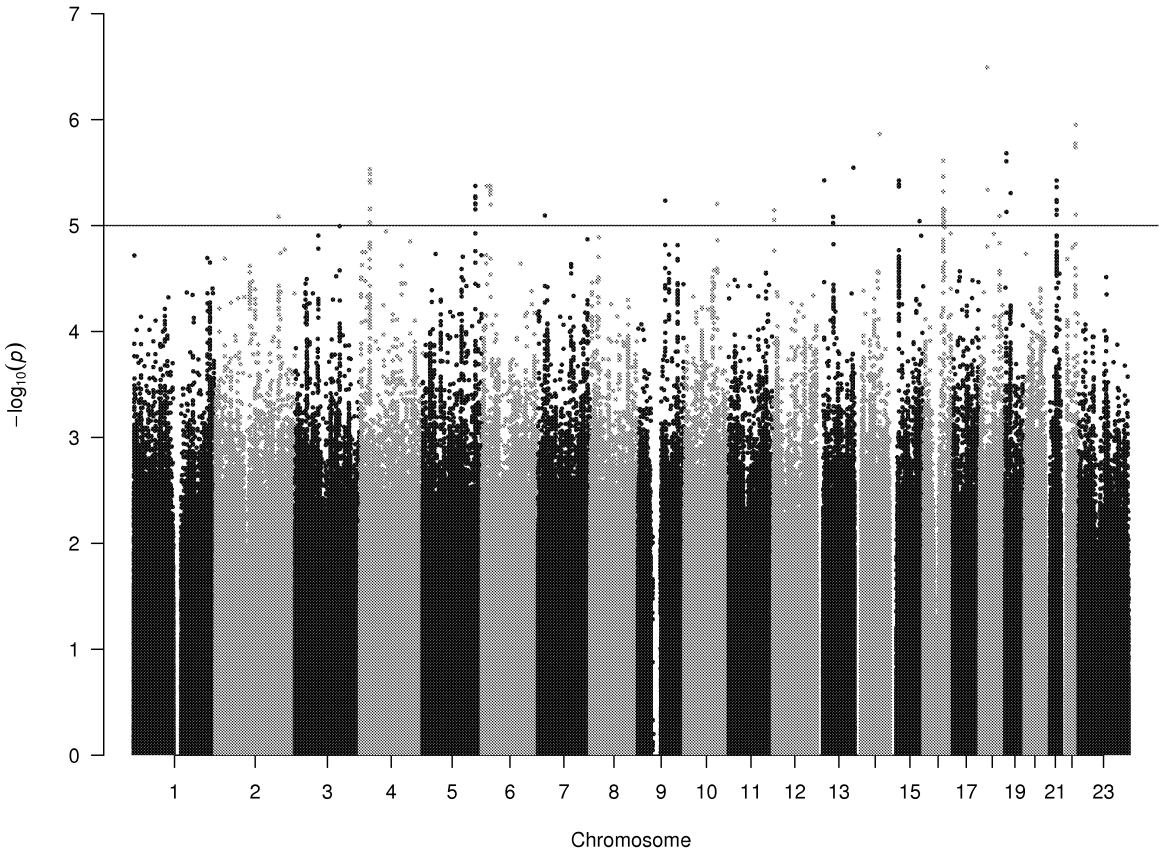
Of the genes listed here, perhaps *FGFRL1* stands out. A member of the fibroblast growth factor receptor family, this is highly conserved and with known roles in development[378]. Knockout mice die pre-natally due to significant developmental defects and exhibit, amongst other abnormalities, defects of the semilunar valves and cardiac outflow tract[376]. Others here seem not to have intuitive links with cardiovascular traits.

Due to the low confidence of the heritability estimates for aortic valve annular dimension, pathways analysis was not undertaken for this phenotype.

14.5.1.5 Aortic root: sinuses of Valsalva diastolic diameter single-SNP GWAS results

Similar to the aortic valve annular dimensions, no SNPs reached genome-wide significance for association. The Manhattan plot below (Figure 14.3) shows associations across the genome with aortic valve annulus diameter. Whilst no associations at genome-wide significance ($p<5x10^{-8}$) were found, there were suggestive associations ($p<1x10^{-5}$) with robust association peaks at chromosomes 4, 5, 6, 15, 16, 19, 21 and 22. These loci are presented in Table 14.5 below and loci of interest are displayed in LocusZoom plots in Figure 14.4 below.

FIGURE 14.3: Manhattan plot showing genome-wide associations for sinuses of Valsalva end-diastolic diameter



The y axis displays -log_{10} p values of association – i.e. the "higher" the point on the graph, the more significant the association. The x axis is the position on each chromosome. Each dot represents a single SNP; however, these often overlie one another. The horizontal threshold line represents a "suggestive significance" level threshold of p = 1x10^{-5}.

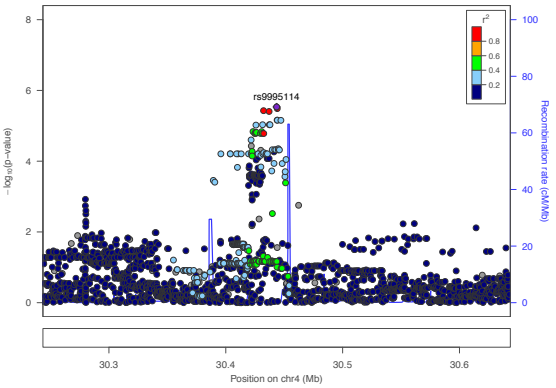
TABLE 14.5: Top SNPs associated with sinuses of Valsalva diastolic diameter

dbSNP	Chr	Position	Ref	Alt	MAF	Beta	p-value	Closest protein-coding gene	Other nearby candidate genes	Comments
rs9995114	4	30443522	G	T	0.045	-1.377	2.92×10^{-6}	*PCDH7*		*PCDH7* assoc. with Type I DM in GWAS
rs367900261	5	161004181	A	ATA ACT	0.343	-0.571	4.23×10^{-6}	*GABRB2*	*GABRA6*	Autonomic nervous system
rs16891477	6	26240752	T	C	0.040	-1.429	4.22×10^{-6}	*HIST1H4F*	multiple histone cluster genes	Associated with height and other body measures. Lead SNP eQTL for *ZNF322*
rs58368968	15	27869057	TCT ATC TTG CAA GA	T	0.332	-0.603	3.75×10^{-6}	*GABRG3*	*OCA2*	*GABRG3* has suggestive links to interaction of diabetes x coronary artery calcification by GWAS
rs13334148	16	60313728	T	C	0.220	-0.700	2.45×10^{-6}		None within 500kb	-
rs78617911	19	3772432	T	C	0.029	1.704	2.08×10^{-6}	*RAX2*	*MATK, APBA3, TJP3, MRPL54*	*MATK* role in IGF pathway
rs57996174	21	34837048	C	T	0.096	0.955	3.76×10^{-6}	*TMEM50B*	*IFNGR2, GART, DNAJC28, SON*	*SON* role in cardiovascular development
rs778312249	22	44218580	T	TG	0.419	0.605	1.12×10^{-6}	*SULT4A1*	*EFCAB6*	*EFCAB6* differentially expressed in LVH (rats)

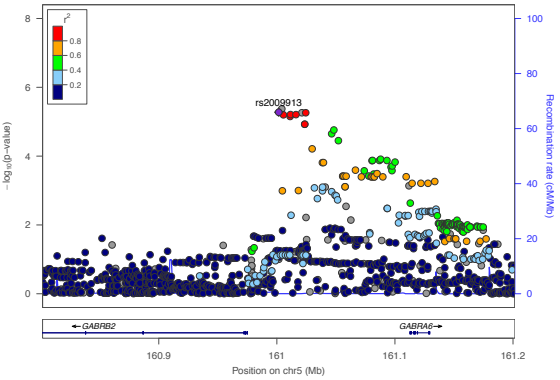
Chr: chromosome; Ref: Reference allele; Alt: effect allele; MAF minor allele frequency; Beta: effect size (change in diameter (mm) per copy of effect allele); Comments: annotations from multiple sources including GWAS catalog[36], GeneCards[369], Mouse Genome Informatics database[370], GTEx[265] and PubMed searches for gene function information. ***Gene names:*** *PCDH7: Protocadherin 7; GABRB2: Gamma-Aminobutyric Acid Type A Receptor Beta2 Subunit; GABRA6: Gamma-Aminobutyric Acid Type A Receptor Alpha6 Subunit; HIST1H4F: Histone Cluster 1 H4 Family Member F; GABRG3: Gamma-Aminobutyric Acid Type A Receptor Gamma3 Subunit; OCA2: OCA2 Melanosomal Transmembrane Protein; RAX2: Retina And Anterior Neural Fold Homeobox 2; MATK: Megakaryocyte-Associated Tyrosine Kinase; APBA3: Amyloid Beta Precursor Protein Binding Family A Member 3; TJP3: Tight Junction Protein 3; Mitochondrial Ribosomal Protein L54; TMEM50B: Transmembrane Protein 50B; IFNGR2: Interferon Gamma Receptor 2; GART: Phosphoribosylglycinamide Formyltransferase, Phosphoribosylglycinamide Synthetase, Phosphoribosylaminoimidazole Synthetase; DNAJC28: DnaJ Heat Shock Protein Family (Hsp40) Member C28; SON: SON DNA Binding Protein; SULT4A1: Sulfotransferase Family 4A Member 1; EFCAB6: EF-Hand Calcium Binding Domain 6; Type I DM: Type I diabetes mellitus; LVH: left ventricular hypertrophy.*

FIGURE 14.4: LocusZoom plots for selected loci associated at suggestive significance levels with sinuses of Valsalva diameter

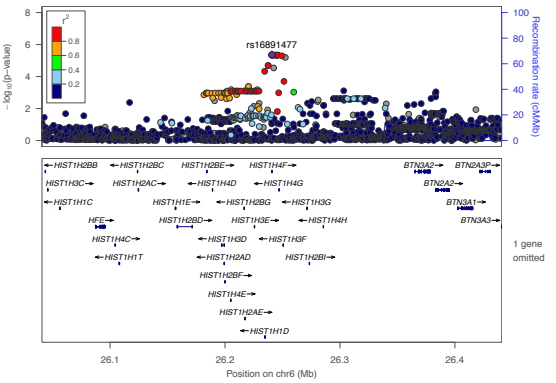
a. Chromosome 4 (PCDH7)



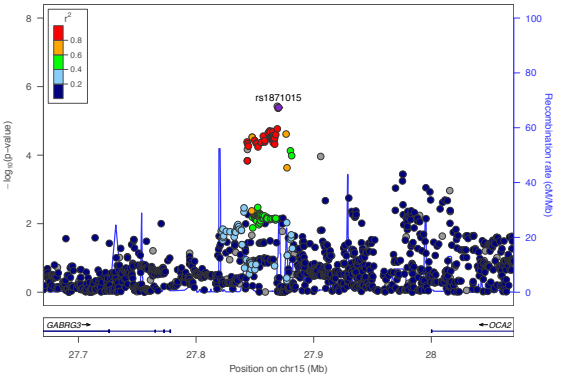
b. Chromosome 5 (GABRB2/GABRA6)



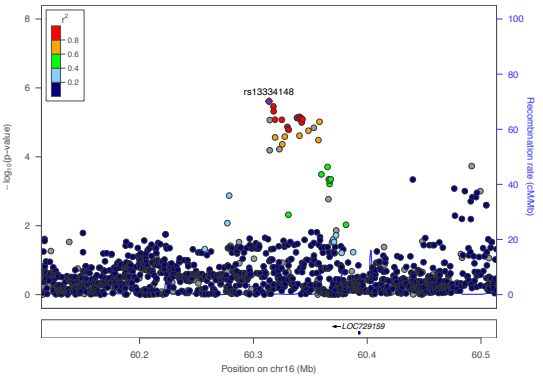
c. Chromosome 6 (histone cluster)



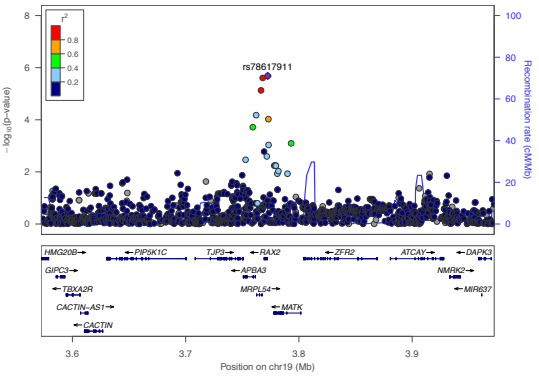
d. Chromosome 15 (GABRG3)



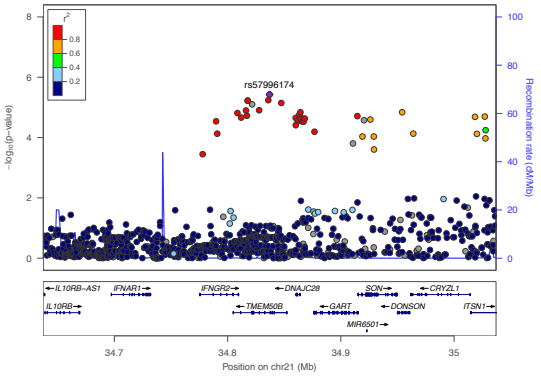
e. Chromosome 16 (nil)



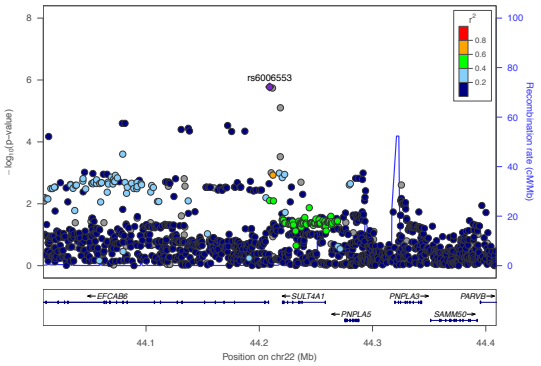
f. Chromosome 19 (MATK)



g. Chromosome 21 (SON?)



h. Chromosome 22 (EFCAB6)



Suggestive loci are shown as examples above in LocusZoom. These plots are essentially a "zoomed in" version of a Manhattan plot, but also demonstrate the linkage disequilibrium structure of the association peak. They demonstrate the spatial relation of the associated SNPs with annotated genes, and with nearby SNPs in linkage disequilibrium (LD) with the lead variant. The plots selected here demonstrate the expected structure of a "true positive" association, with SNPs in LD with the lead variant also associated with phenotype, albeit at lower significance. The labels a-h denote the chromosomal location of the loci, and the top possibilities for candidate genes at the locus, based on biological plausibility and functional evidence, detailed in Table 14.5 and below.

The chromosome 22 locus is the most significantly associated of all the robust association peaks. The top SNP is just downstream of *SULT4A1* and just upstream of *EFCAB6.* Neither of these two genes has strong links with aortic traits, although *EFCAB6* has been identified as a differentially expressed gene in left ventricular hypertrophy in rats[379]. This perhaps makes *EFCAB6* the more likely causal gene at this locus.

The chromosome 4 association peak is just upstream of *protocadherin 7 (PCDH7).* This encodes a cell adhesion molecule which is overexpressed in the aorta and the heart at all stages of development, and is known to be a key determinant of neural crest development – a process important for aortic root structure[380].

The chromosome 21 locus is a very broad peak spanning a large gene cluster. Most of the genes here are involved in inflammatory and immune system responses. The one gene with potential cardiovascular links is *SON (SON DNA binding protein),* a gene implicated in zttk syndrome which includes intellectual disability, and congenital heart conditions in some patients, implying a role in cardiovascular development.

The chromosome 16 locus is in a relative gene desert, with no protein-coding genes within 500kb, and no informative epigenomic data. The chromosome 19 lead SNP is just 2kb upstream of *Retina And Anterior Neural Fold Homeobox 2 (RAX2),* which is predominantly

involved in eye development with no clear vascular role. The association peak also tags a gene cluster which includes *MATK, megakaryocyte-associated tyrosine kinase.* The protein product of this gene plays an important role in IGF-1-stimulated proliferation of vascular smooth muscle cells.

The chromosome 5 locus is a particularly interesting one: this tags *GABRB2* and a cluster of $GABA_A$-receptor subunit genes. GABA (Gamma-Aminobutyric Acid) is known to play a key role in central cardiovascular regulation, mediating part of the central baroreceptor response and eliciting parasympathetic and sympathetic cardiovascular responses. It is notable that a separate subunit of the GABA type A receptor; *GABRG3* is tagged by an independently associated locus on chromosome 15. $GABA_A$ receptors are the major inhibitory neurotransmitter receptors in mammalian brain, and there is much data to suggest a role for these receptors in mediating central control of pressor responses to various stimuli[381], and that modulation of these responses could attenuate hypertension, at least in rat models[382].

So $GABA_A$ receptor subunit genes at two independent loci, *PCDH7, MATK* and *EFCAB6* seem to be the most convincing candidates for mediating control of SoV end-diastolic diameter. Of these, the $GABA_A$ receptor genes seem the most interesting for future follow-up, as the function of their protein product is potentially modifiable throughout life.

14.5.1.6 Aortic root: sinuses of Valsalva diameter: gene-based GWAS results

None of these gene-based associations were significant when multiple testing corrections were applied (for 20,000 comparisons). The top 10 genes associated with aortic root (sinuses of Valsalva) diameter are listed in Table 14.6 below:

TABLE 14.6: Gene-based associations with sinuses of Valsalva diastolic diameter

Gene	Chr	Gene start	Gene end	No. SNPs	Chi-sq	P value	P value most sig SNP	Most sig SNP
SYNDIG1	20	24449834	24647253	343	903.127	$5.04x10^{-5}$	$6.30x10^{-5}$	rs6132757
PENK	8	57353512	57359282	89	346.528	0.0001	0.0002	rs13265897
SLC38A8	16	84043388	84075762	517	1151.76	0.0002	$1.19x10^{-5}$	rs12445301
IGF2R	6	160390130	160527583	270	651.171	0.0004	0.0001	rs397887238
ERBB2IP	5	65222381	65376851	194	557.004	0.0005	0.0005	rs27701
IGF1R	15	99191767	99507759	703	1348.4	0.0005	0.0001	rs77320724
CRISP2	6	49660070	49681303	96	306.598	0.0007	0.0002	rs1535289
CRISP3	6	49695091	49712168	83	260.384	0.0007	0.0002	rs1535289
ESR1	6	152011630	152424408	603	1113.93	0.0008	$4.18x10^{-5}$	rs2982683
SAA2-SAA4	11	18252901	18270221	215	503.508	0.0008	$3.26x10^{-5}$	rs7128060

Top 10 protein-coding genes associated with sinuses of Valsalva (SoV) diameter. Chr: chromosome; Gene start/end: left and right-side boundaries of the gene (GRCh37); No.SNPs: number of SNPs occurring within gene contributing to the analysis; Chi-sq: sum of chi-squared test statistics of all SNPs within the gene region; P value: segment-based test p value; P value most sig SNP: smallest single-SNP p value in the segment; SYNDIG1: Synapse Differentiation Inducing 1; PENK: Proenkephalin; SLC38A8: Solute Carrier Family 38 Member 8; IGF2R: Insulin Like Growth Factor 2 Receptor; ERBB2IP (aka ERBIN): Erbb2 Interacting Protein; IGF1R: Insulin Like Growth Factor 1 Receptor; CRISP2: Cysteine Rich Secretory Protein 2; CRISP3: Cysteine Rich Secretory Protein 3; ESR1: Estrogen Receptor 1; SAA2-SAA4: Serum Amyloid A2- Serum Amyloid A4 Readthrough.

Here, two genes stand out as they are encoded on distant chromosomes, but both are within the top 10 protein coding genes associated with SoV diameter. These are the insulin-like growth factor receptor 1 and 2 genes. *IGF2R* has a strong association with coronary artery disease ($p = 5x10^{-15}$, top SNP rs688359)[383]. *IGF1R* is less obviously associated with coronary artery disease, but there has been some interest in the use of IGF-1 and its binding proteins as a serum biomarker for heart failure with preserved ejection fraction[384]. The IGF regulatory axis is a key determinant of multiple metabolic pathways, including glucose regulation. Of note, *MATK*, encoding a mediator of IGF-1 induced vascular smooth muscle cell proliferation, was independently identified from the single-SNP association peaks. This provides further evidence for the importance of the IGF signalling pathways in regulating aortic root diameter.

Another gene to note is the oestrogen receptor gene (*ESR1*). This has been the subject of some debate as to its effect on coronary artery disease risk; current limited evidence suggests that polymorphisms might affect coronary artery disease risk in Asian but not European populations[385].

PENK (Proenkephalin) is an interesting candidate. Proenkephalin levels are markers of cardiorenal outcomes following myocardial infarction, and correlate well with prognosis in acute and chronic heart failure[386-389] and in heart failure with preserved ejection fraction[390]. Proenkephalin is a stable endogenous opioid biomarker related to renal function. There has been some suggestion that it may be a determinant of, as well as a marker for, cardiovascular outcomes[389].

ERBB2IP is also a gene to note- the roles of the ERBB signalling pathways in cardiac development and function are further discussed below.

Finally, *SYNDIG1*, the most significantly-associated gene with SoV diameter, encodes a recently-identified AMPA receptor interacting protein that directly binds to the AMPA receptor subunit GluA2, and regulates excitatory synapse formation. Glutamate signalling via these synapses mediates central baroreceptor reflexes, and plays a key role in cardiovascular homeostasis and the autonomic nervous system. *SYNDIG1* has been shown to be a mediator of the number of excitatory synapses formed[391]. The role of glutamate in cardiovascular traits is discussed further in section 14.5.1.8 below, where *GRIA4*, a component of glutamatergic AMPA receptors, is identified as a potential determinant of sino-tubular junction diameter.

14.5.1.7 Aortic root: sinuses of Valsalva diameter: pathways analysis GWAS results

Pathways analysis was conducted using the summary statistics from the association tests. These were the input for iGSEA4GWAS v2, which returns significance values for pathways associated with phenotype, as described in Chapter 13. Table 14.7 below shows the top 5 KEGG pathways and GO biological processes associated with SoV diameter. These all reached significance, after FDR correction, apart from the fifth KEGG pathway (Dilated Cardiomyopathy) which has a borderline significance level of p=0.051.

TABLE 14.7: Top 5 KEGG pathways and top 5 Gene Ontology biological processes associated with SoV diameter

KEGG Pathway	FDR-corrected p value
KEGG: ERBB SIGNALING PATHWAY	0.010
KEGG: INSULIN SIGNALING PATHWAY	0.012
KEGG: PROXIMAL TUBULE BICARBONATE RECLAMATION	0.013
KEGG: LYSOSOME	0.016
KEGG: DILATED CARDIOMYOPATHY	0.051

GO Pathway: Biological Process	FDR-corrected p value
GO: ENZYME LINKED RECEPTOR PROTEIN SIGNALING PATHWAY	0.010
GO: REGULATION OF CELL ADHESION	0.013
GO: REGULATION OF G PROTEIN COUPLED RECEPTOR PROTEIN SIGNALING PATHWAY	0.021
GO: GROWTH FACTOR BINDING	0.022
GO: MESODERM DEVELOPMENT	0.033

The most significantly associated pathway is the ERBB signalling pathway. This plays a crucial role in development, but also in maintenance of cardiovascular function, making it a potential target for therapeutic intervention. This family of four receptor tyrosine kinases (EGFR/HER1, ErbB2/HER2, ErbB3/HER3, ErbBr/HER4) binds extracellular growth factor ligands and activate intracellular signalling pathways, including those regulating differentiation, proliferation, cell motility and cell survival. Knockouts of any of the four main receptor tyrosine kinases are lethal, causing major cardiovascular developmental defects[392]. Other mouse models, such as the waved-2 model, have reduced activity of ErbB1 / EGFR, (about 10-15% of normal), and lead to hyperplastic outflow tract valves. Trials of recombinant neuregulin-1, one of the major ligands of the ERBB receptors, are underway for treatment of heart failure[386] – perhaps, if successful, data could be collected to assess the effect of this drug on the aortic root too.

The insulin signalling pathway is also identified as a key pathway by this analysis, in keeping with the gene-based results which highlighted *IGFR1* and *IGFR2* as potentially associated genes, and with the single-SNP associations which also identified *MATK* (a mediator of IGF-1 induced vascular smooth muscle cell proliferation).

The significance of lysosome and bicarbonate reclamation in the kidneys is less easy to address – although glutamate and GABA signalling are highly dependent upon bicarbonate metabolism. The fifth KEGG pathway – Dilated Cardiomyopathy – is just the "wrong" side of significance thresholds. Nevertheless, it is of great interest to note that genetic influences which affect the structure and function of myocardial tissue in DCM might also be at play in determining aortic root structure.

The GO biological pathways reveal few surprises – there are clear roles of cell adhesion, growth factor binding and mesodermal development in aortic root biology, and their significance in this analysis is both biologically plausible and summarises nicely the different ways in which genetics might influence aortic root size.

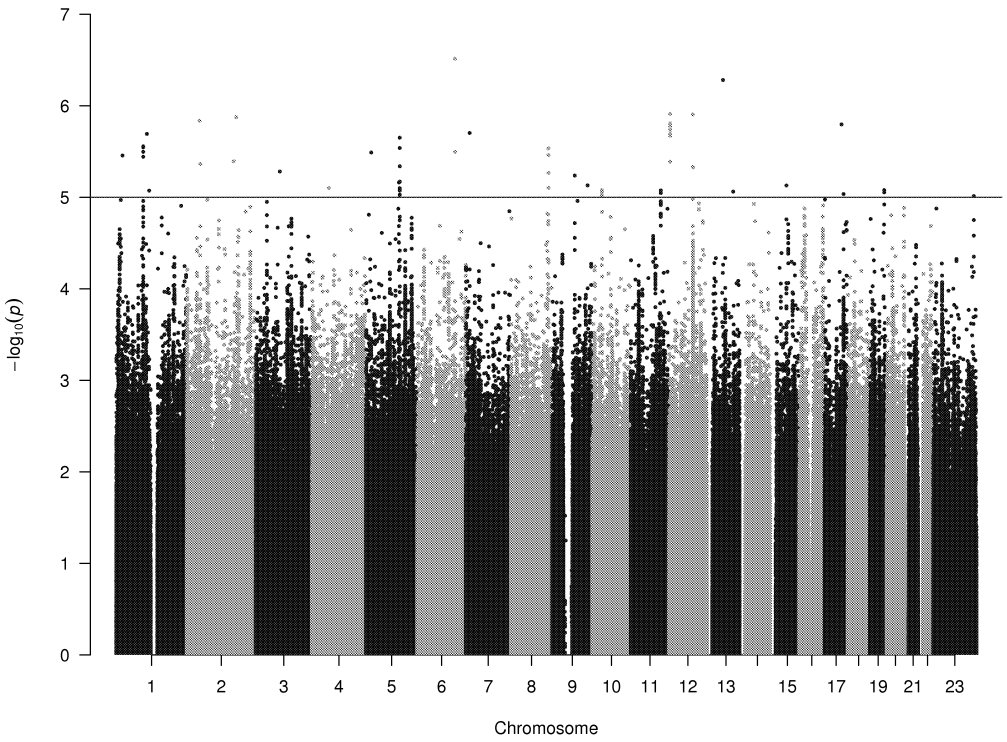
14.5.1.8 Aortic root: sinuses of Valsalva diameter: summary of GWAS results

Analysis of single-SNP, gene and pathways associations with SoV diameter has suggested the role of many different genes and pathways in determining aortic root size. These include the ERBB signalling pathway, insulin signalling, including the IGF-receptor genes *IGFR1* and *IGFR2, MATK*, the $GABA_A$ receptor subunit genes, *SYNDIG, PENK, PCDH7, PKIA* and *EFCAB6.*

14.5.1.9 Aortic root: sinotubular junction diameter: single-SNP GWAS results

The Manhattan plot below (Figure 14.5) shows associations across the genome with sinotubular junction diameter. Whilst no associations at genome-wide significance ($p<5x10^{-8}$) were found, there were suggestive associations ($p<1x10^{-5}$) at chromosomes 1, 5, 8, 11, 12 and 19. These loci are presented in Table 14.8 below and loci of interest are displayed in LocusZoom plots in Figure 14.6 below. It is of interest to note that the loci identified here are distinct from the loci associated with SoV diameter, suggesting that the genetic drivers of aortic dimension here could be different from those in the aortic root itself. As we move further along the aortic path, the genetic influences are likely to change, partly as a result of the different cells which contribute to the formation of the structure, and partly because different haemodynamic influences predominate.

FIGURE 14.5: Manhattan plot showing associations of genotype with sino-tubular junction (STJ) diastolic diameter.



The y axis displays -log_{10} p values of association – i.e. the "higher" the point on the graph, the more significant the association. The x axis is the position on each chromosome. Each dot represents a single SNP; however, these often overlie one another. The horizontal threshold line represents a "suggestive significance" level threshold of p = 1x10^{-5}.

TABLE 14.8: Top SNPs associated with sino-tubular junction diastolic diameter

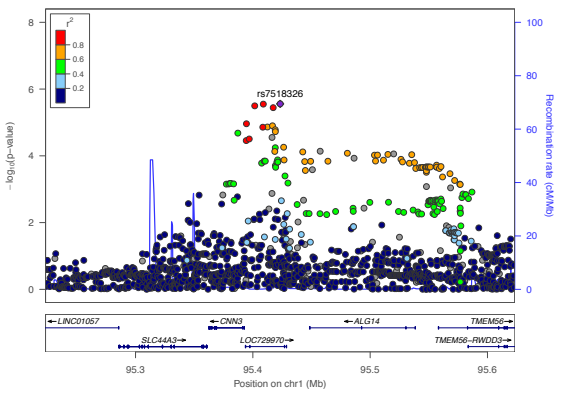
dbSNP	Chr	Position	Ref	Alt	MAF	Beta	p-value	Closest protein-coding gene	Other nearby candidate genes	Comments
rs7518326	1	95423291	A	G	0.462	0.509	2.77x10^{-6}	*CNN3*	*ALG14, SLC44A3, TMEM56*	*CNN3* implicated in regulation of smooth muscle contraction.
rs869031125	5	119073527	CT	C	0.018	-1.964	2.23 x10^{-6}	*FAM170A*		No evidence for cardiovascular function
rs10094464	8	132165751	A	G	0.089	-0.868	2.92 x10^{-6}	*ADCY8*		*ADCY8* gain-of function increases endurance in mice
rs56035506	11	105842387	A	G	0.074	0.922	8.41 x10^{-6}	*GRIA4*	*MSANTD4*	*GRIA4* part of glutamate receptor complex; other subunits known to mediate central baroreceptor reflexes

rs11829154	12	4077231	A	G	0.316	0.558	1.22 x10^{-6}	*PARP11*		No evidence for cardiovascular function
rs9943701	12	84318750	C	T	0.018	1.983	1.25 x10^{-6}	*None within 500kb*		-
rs3219340	19	50897887	G	T	0.087	-0.867	8.35 x10^{-6}	*POLD1*	*NRIH2, NAPSA, SPIB, MYBPC2*	*POLD1* known to be biomarker of ageing; no clear cardiovascular function. SPIB identified as driver of coronary artery disease by network analysis

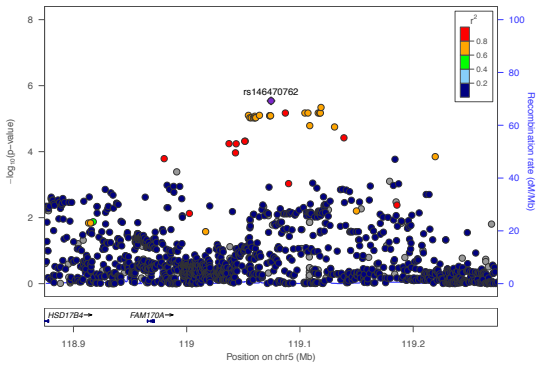
Chr: chromosome; Ref: Reference allele; Alt: effect allele; MAF minor allele frequency; Beta: effect size (change in diameter (mm) per copy of effect allele); Comments: annotations from multiple sources including GWAS catalog[36], GeneCards[369], Mouse Genome Informatics database[370], GTEx[265] and PubMed searches for gene function information. ***Gene names:*** *CNN3: Calponin 3; ALG14: ALG14, UDP-N-Acetylglucosaminyltransferase Subunit; SLC44A3: Solute Carrier Family 44 Member 3; TMEM56: Transmembrane Protein 56; FAM170A: Family With Sequence Similarity 170 Member A; ADCY8: Adenylate Cyclase 8; GRIA4: Glutamate Ionotropic Receptor AMPA Type Subunit 4; MSANTD4: Myb/SANT DNA Binding Domain Containing 4 With Coiled-Coils; PARP11: Poly(ADP-Ribose) Polymerase Family Member 11; CCND2: Cyclin D2; TIGAR: TP53 Induced Glycolysis Regulatory Phosphatase; CRACR2A: Calcium Release Activated Channel Regulator 2A; POLD1: DNA Polymerase Delta 1, Catalytic Subunit; NR1H2: Nuclear Receptor Subfamily 1 Group H Member 2; SPIB: Spi-B Transcription Factor; MYBPC2: Myosin Binding Protein C, Fast Type; NAPSA: Napsin A Aspartic Peptidase; cXorf51B: Chromosome X Open Reading Frame 51 B.*

FIGURE 14.6: LocusZoom plots for selected loci associated at suggestive significance levels with sino-tubular junction (STJ) diameter

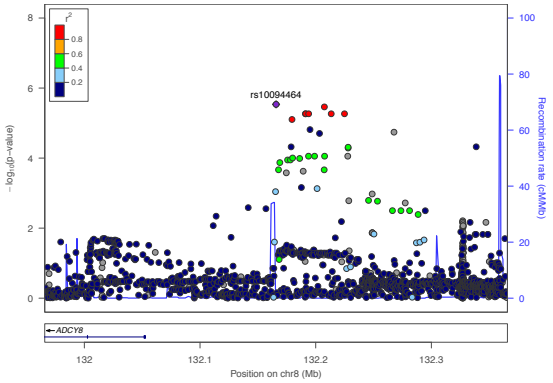
a. Chromosome 1 (CNN3/TMEM56)



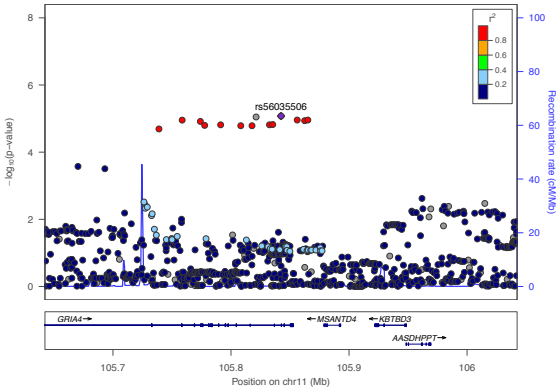
b. Chromosome 5 (FAM170A)



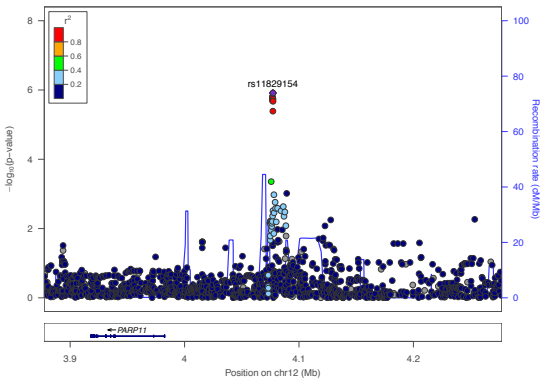
c. Chromosome 8 (ADCY8)



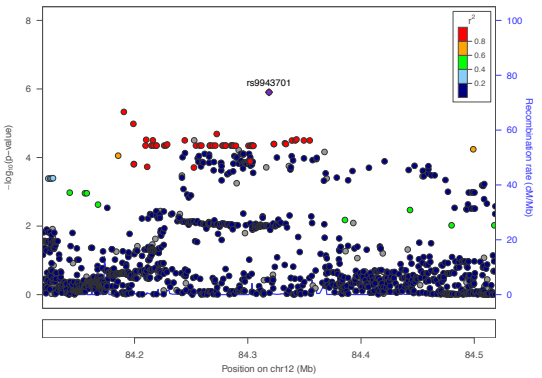
d. Chromosome 11 (GRIA4/MSANTD4)



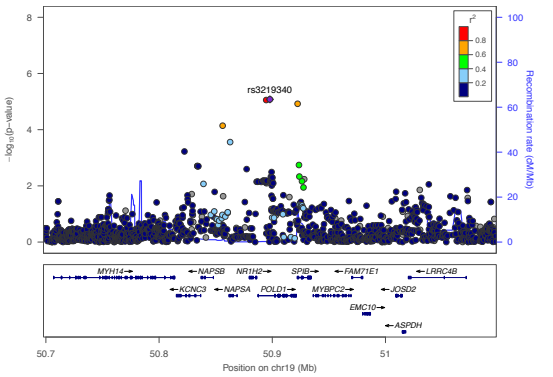
e. Chromsome 12 (PARP11)



f. Chromosome 12 (no candidates)



g. Chromosome 19 (SPIB/MYBPC2)



Suggestive loci are shown as examples above in LocusZoom. These plots are essentially a "zoomed in" version of a Manhattan plot, but also demonstrate the linkage disequilibrium structure of the association peak. They demonstrate the spatial relation of the associated SNPs with annotated genes, and with nearby SNPs in linkage disequilibrium (LD) with the lead variant. The plots selected here demonstrate the expected structure of a "true positive" association, with SNPs in LD with the lead variant also associated with phenotype, albeit at lower significance. The labels a-g denote the chromosomal location of the loci, and the top possibilities for candidate genes at the locus, based on proximity, biological plausibility and functional evidence, detailed in Table 14.8 and below.

Here, the most significantly associated locus is the chromosome 12 locus upstream of *PARP11 (Poly(ADP-Ribose) Polymerase Family Member 11)*. This does not have obvious links with cardiovascular traits.

The chromosome 1 locus tags a region upstream of *CNN3 (Calponin 3*), but also including *ALG14 (ALG14, UDP-N-Acetylglucosaminyltransferase Subunit)* and *TMEM56 (transmembrane protein 56). ALG14* has been associated with plasma fatty acids concentration[393], *TMEM56* is differentially expressed in whole blood of patients with coronary artery disease[394], and calponin 3 is associated with the thin filament and regulates smooth muscle contraction. Therefore, particularly *TMEM56* and *CNN3* have plausible biological links with phenotype. The top SNP, rs7518326, is also a significant eQTL for *RWDD3*, at least in oesophageal tissue[29] – this gene regulates multiple signalling pathways including HIF1A, VEGFA and NFKB. However, the lack of eQTL signal in relevant tissues to cardiovascular traits perhaps argues against this being a significant mechanism for phenotype association.

GRIA4 (Glutamate Ionotropic Receptor AMPA Type Subunit 4) on chromosome 11 encodes a subunit of AMPA glutamate receptors. Glutamate is the major excitatory neurotransmitter in the central nervous system, and plays a key role in centrally-mediated baroreceptor reflexes. It is, essentially, the counterpoint to GABA, the inhibitory neurotransmitter whose receptors are identified as likely mediators of aortic root dimensions above. Glutamate is activated as part of the renin-angiotensin system[395], and mediates pressor responses[396]. It is also important during cardiovascular development and adaptation, with glutamate signalling identified as a likely contributor to the closure of the ductus arteriosus in early infancy[397]. Little is published on the specific function of *GRIA4*, but it is highly expressed in the developing brain and in the adrenal glands.

At the chromosome 19 locus, there is s gene cluster containing a few potential candidate genes. The top SNP is intronic in *POLD1 (DNA Polymerase Delta 1, Catalytic Subunit)*, whose product is known to be a marker of ageing, with expression levels falling consistently with chronological age. Whether this might be relevant to a cardiovascular ageing phenotype is uncertain. *SPIB (Spi-B Transcription Factor)* is also near to this association peak. This gene encodes a transcription factor which has been identified as a key multi-tissue molecular driver of coronary artery disease by network analysis based on differential gene expression between

cases and controls[398], but little is known of its mechanism in this setting. *MYBPC2 (Myosin Binding Protein C, Fast Type)* is predominantly involved in skeletal muscle contraction. A closely related gene, *MYBPC3* is active in cardiac muscle and is implicated as a major cause of Hypertrophic Cardiomyopathy. Interestingly, the protein product of *MYBPC2* also appears to be expressed in the heart[399], so it is conceivable that it also plays a role in the cardiovascular system, although this is unproven.

The genes near the chromosome 5 and 8 association peaks have little data to support functional roles in the aortic root.

14.5.1.10 Aortic root: sinotubular junction diameter: gene-based GWAS results

Some of the genes identified by single-SNP associations were prominent in the gene-based association results. However, again, none of these gene-based associations were significant when multiple testing corrections were applied (for 20,000 comparisons). The top 10 genes associated with sinotubular junction diameter are listed in Table 14.9 below:

TABLE 14.9: Gene-based associations with sinotubular junction (STJ) diastolic diameter

Gene	Chr	Gene start	Gene end	No. SNPs	Chi-sq	P value	P value most sig SNP	Most sig SNP
CNN3	1	95362504	95392779	284	777.1	$8.33x10^{-5}$	$2.77x10^{-6}$	rs7518326
TRNP1	1	27320194	27327377	77	242.295	0.0002	0.0003	rs780567116
ALG14	1	95448278	95538507	244	741.011	0.0002	$2.77x10^{-6}$	rs7518326
GABRA1	5	161274196	161326965	132	422.157	0.0003	$7.29x10^{-5}$	rs10042696
PDCD2	6	170884659	170893780	63	265.611	0.0003	0.0002	rs734249
PRPF6	20	62612430	62664453	211	612.44	0.0003	$1.30x10^{-5}$	rs113017561
FAM46B	1	27331510	27339333	90	267.378	0.0003	0.0003	rs780567116
CD96	3	111260925	111371206	178	490.965	0.0003	0.0002	rs11716404
SCUBE1	22	43599228	43739394	522	1071.64	0.0004	$5.40x10^{-5}$	rs11090149
XIRP1	3	39224705	39234085	162	499.89	0.0004	$1.12x10^{-5}$	rs1274964

Top 10 protein-coding genes associated with sino-tubular junction (STJ) diameter. Chr: chromosome; Gene start/end: left and right-side boundaries of the gene (GRCh37); No.SNPs: number of SNPs occurring within gene contributing to the analysis; Chi-sq: sum of chi-squared test statistics of all SNPs within the gene region; P value: segment-based test p value; P value most sig SNP: smallest single-SNP p value in the segment; CNN3: Calponin 3; TRNP1: TMF1-Regulated Nuclear Protein 1; ALG14: ALG14, UDP-N-Acetylglucosaminyltransferase Subunit; GABRA1: Gamma-Aminobutyric Acid Type A Receptor Alpha1 Subunit; PDCD2: Programmed Cell Death 2; PRPF6: Pre-MRNA Processing Factor 6; FAM46B (aka TENT5B): Terminal Nucleotidyltransferase 5B; CD96: CD96 Molecule; SCUBE1: Signal Peptide, CUB Domain And EGF Like Domain Containing 1; XIRP1: Xin Actin Binding Repeat Containing 1

Notable genes here include *GABRA1 (Gamma-Aminobutyric Acid Type A Receptor Alpha1 Subunit)*. This gene and locus of association is close to, but distinct from, the locus implicating *GABRB2, GABRA6* and *GABRG3* in association with SoV diameter, discussed in section 14.5.1.5 above. This association of STJ diameter with *GABRA1* therefore lends further weight to the notion that centrally-mediated control of cardiovascular reflexes and the autonomic nervous system might be of vital importance in determining the dimensions and remodelling of the aortic root.

CNN3 was already identified as important by the single SNP results; the apparent gene association with *ALG14* is at the same locus, and is discussed above.

SCUBE1 was first identified in vascular endothelium, but is also a marker of platelet activation. Its expression levels are increased with acute ischaemic events[400] and in hypertension[401]. Variants could therefore plausibly affect vascular endothelial function and therefore development or remodelling of the aortic root.

Variants in *XIRP1 (Xin Actin Binding Repeat Containing 1)* are implicated in cardiomyopathies, with knockout mice developing cardiomyopathy and abnormal cardiomyocyte structure, as

well as conduction abnormalities[402]. Due to common developmental origins of the heart and aortic root, variants in this gene could also therefore have an effect on aortic root development and size.

Other genes listed here such as *TRNP1, PDCD2, PRPF6, FAM46B* and *CD96* do not have such obvious links with cardiovascular phenotypes.

14.5.1.11 Aortic root: sino-tubular junction diameter: pathways analysis GWAS results

Pathways analysis was conducted using the summary statistics from SNPTEST. These were input into iGSEA4GWAS v2, which returns significance values for pathways associated with phenotype, as described in Chapter 13. Table 14.10 below shows the top 5 KEGG pathways and GO biological processes associated with STJ diameter. These all reached significance, after FDR correction.

TABLE 14.10: Top 6 KEGG pathways and top 5 Gene Ontology biological processes associated with STJ diameter

KEGG Pathway	FDR-corrected p value
KEGG: LINOLEIC ACID METABOLISM	0.004
KEGG: CHEMOKINE SIGNALING PATHWAY	0.008
KEGG: CALCIUM SIGNALING PATHWAY	0.010
KEGG: VEGF SIGNALING PATHWAY	0.025
KEGG: DILATED CARDIOMYOPATHY	0.034
KEGG: VASCULAR SMOOTH MUSCLE CONTRACTION	0.044

GO pathway: Biological Process	FDR-corrected p value
GO: POSITIVE REGULATION OF CELLULAR PROTEIN METABOLIC PROCESS	0.010
GO: POSITIVE REGULATION OF PROTEIN METABOLIC PROCESS	0.011
GO: EXCRETION	0.022
GO: REGULATION OF PROTEIN MODIFICATION PROCESS	0.022
GO: POSITIVE REGULATION OF PHOSPHATE METABOLIC PROCESS	0.025

The Gene Ontology pathways are perhaps rather less informative in this analysis, being very general, broad pathways. Nevertheless, the KEGG pathways have highlighted some interesting associations.

Of these, linoleic acid metabolism comes as a bit of a surprise. This is a relatively small pathway with several clustered genes, and it may be that the association is a false positive one. Nevertheless, there is good evidence to associate omega-3 fatty acids with cardiovascular outcomes and it is possible that these could contribute to remodelling or development of the aortic root.

VEGF signalling is a much more plausible link – it is a key driver of angiogenesis. KIT, one of the mediators of this signalling pathway, was implicated in the valve diameter GWAS above.

The significance of association with the dilated cardiomyopathy pathway highlights the interconnection of ventricular and aortic phenotypes. It makes logical sense to suppose that genes which regulate myocardial structure and function might also regulate aortic root structure and function, based upon their common developmental origins, and yet the two structures are often studied independently.

A significant association with the vascular smooth muscle contraction pathway also highlights a set of genes which are implicated in aortopathies such as bicuspid aortic valve and familial aortopathies[166]. This makes apparent the functional overlap between genes in which common variants mediate normal variation in aortic structure, and those in which rare variants might cause pathological changes.

14.5.2 Joint trait analysis: aortic root dimensions

I performed a joint trait analysis of sinuses of Valsalva and sino-tubular junction diameters. Joint trait analysis can boost power to detect biologically relevant traits when there is correlation between phenotypes. SCOPA[359] uses a reverse regression methodology, treating the genotype as outcome and the residuals of both phenotypes (accounting for covariates) as predictors. This analysis is not necessarily designed to boost power overall (as in traditional independent sample meta-analysis), but to detect signals of association which might not be apparent from univariate analysis alone. An analysis of the relative significance values of association between the individual trait associations and the joint trait analysis may also be used to make inferences about the route by which a SNP might affect a trait – i.e. whether this represents biological pleiotropy (a SNP or gene having effects on multiple different phenotypes independently), or whether the effects on one phenotype are mediated via its

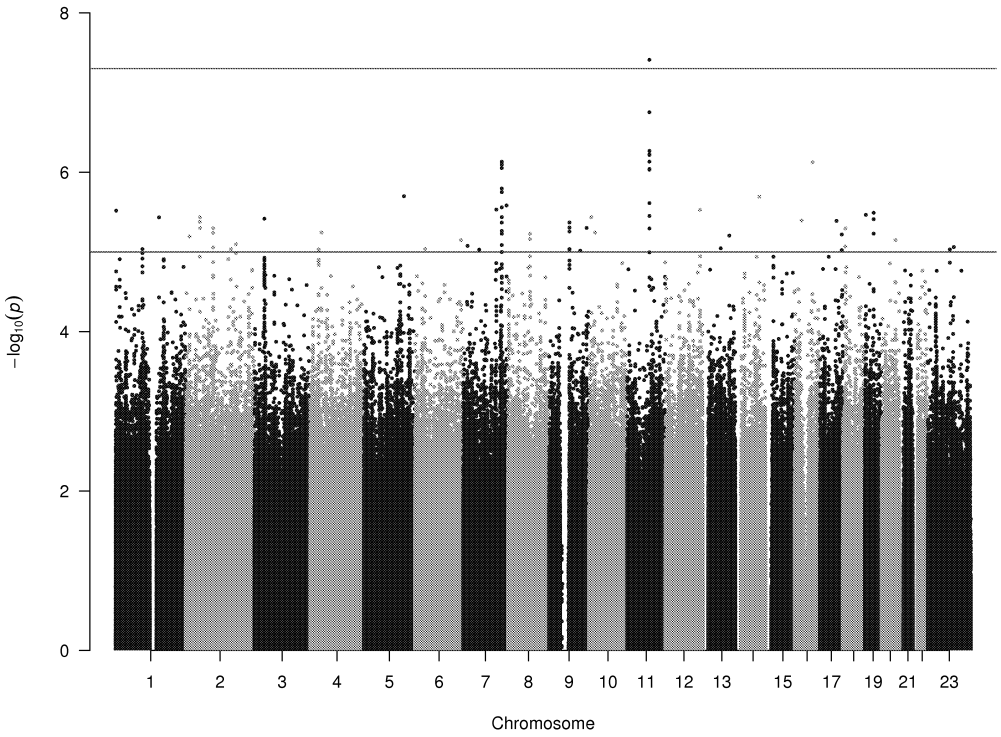
effects on the other. The summary statistics of the joint trait analysis can, as for "standard" GWAS output, be used for gene- and pathways-based analyses.

14.5.2.1 Joint trait analysis SoV + STJ diameter – single SNP analysis

When jointly assessing associations with SoV and STJ diameter, the joint trait analysis provides little boost in significance levels in general. As the two traits are so closely correlated, accounting for the correlation structure between them removes much of the additive benefit of the joint analysis, with statistical significance of each SNP's association with the joint traits being very similar to the single-trait levels. For example, the top SNP associated with STJ diameter, rs11829154, has a p value for association with joint traits of 5.4×10^{-6}. For SoV diameter, the effect size (ß) is 0.29 (p=0.021); for STJ diameter, the effect size (ß) is 0.56 ($p=1.2 \times 10^{-6}$). There is, in other words, an association with both traits, but when one accounts for correlation between them, the p value for association with the joint traits is less than that with the individual trait of STJ diameter. It is difficult to make assertions about pleiotropy in this context, as SoV diameter and STJ diameter cannot easily be separated in a causal relationship.

The joint trait analysis did, however, highlight some new loci significantly associated with aortic root diameter which were not apparent from single trait analysis. These are shown in Figure 14.7 below.

FIGURE 14.7: Manhattan plot showing associations of joint trait sino-tubular junction diameter + sinuses of Valsalva diameter with genotype.



The y axis displays -log_{10} p values of association – i.e. the "higher" the point on the graph, the more significant the association. The x axis is the position on each chromosome. Each dot represents a single SNP; however, these often overlie one another. The horizontal threshold lines represent a "suggestive significance" level threshold of p = $1x10^{-5}$ and a nominated genome-wide significance level of p=$5x10^{-8}$.

This joint trait analysis of SoV and STJ diameter identified one locus meeting genome-wide significance on chromosome 11, and a further interesting association peak at chromosome 7. Of great interest is the fact that these two independent loci may be functionally related. The chromosome 11 top SNP, rs485003, is an intronic variant within the *TENM4* gene. *TENM4 (Teneurin Transmembrane Protein 4)* is known to be a key driver of early mesodermal development[403], and therefore it is plausible that variants in this gene could affect the structure of mesodermally-derived tissues such as the aortic root. The top chromosome 7 SNP, rs1358447, is an intronic variant in *DGKI (Diacylglycerol Kinase Iota).* However, several variants in complete linkage disequilibrium (r^2=1) are significant eQTLs for *PTN (pleiotrophin)* in arterial tissue. *PTN* encodes a secreted growth factor which has a role in neurite outgrowth, and enhances proliferation of fibroblasts, epithelial, and endothelial cells, and plays a key role in angiogenesis. It is expressed in vascular atherosclerotic plaques, and can induce capillary-like sprouting from explanted aortic tissue[404]. Again, this represents a plausible biological link

between gene and phenotype. Of particular interest is the fact that a downstream effector of *PTN* appears to be *TENM4*, at least in the brain. *PTN* -/- mice had increased expression of *TENM4* (Z>3) in hippocampus, along with 49 other genes, confirmed by microarray and qPCR[405]. The expression of *TENM4* was not increased by silencing of *PTN* in an in vitro neuronal culture, indicating that other cells may be the main expressors of *TENM4*. It is tempting to speculate that these may be vascular cells. Whether *PTN* controls *TENM4* expression in the cardiovascular system remains untested, but the independent association signals with aortic root size in our analysis make this pathway an interesting target for further research.

14.5.2.2 Joint trait analysis SoV + STJ diameter – gene and pathways analysis

TABLE 14.11: Gene-based associations with joint trait of sinuses of Valsalva diastolic diameter and sinotubular junction diameter

Gene	Chr	Gene start	Gene end	No. SNPs	Chi-sq	P value	P value most sig SNP	Most sig SNP
FAM25C	*10*	*47177203*	*47181688*	*47*	*294.81*	*4.41×10^{-11}*	*0.0002*	*rs4415693*
FAM25G	*10*	*47177221*	*47181688*	*47*	*294.81*	*4.41×10^{-11}*	*0.0002*	*rs4415693*
ANXA8	*10*	*47157983*	*47174143*	*95*	*520.99*	*7.84×10^{-9}*	*6.96×10^{-5}*	*rs375838536*
NPY4R	*10*	*47083533*	*47088320*	*179*	*788.15*	*3.96×10^{-8}*	*3.93×10^{-5}*	*rs7914445*
GPRIN2	*10*	*46993545*	*47000568*	*133*	*532.50*	*7.34×10^{-7}*	*2.63×10^{-5}*	*rs58260169*
AGAP9	*10*	*47191843*	*47213626*	*25*	*141.84*	*2.06×10^{-6}*	*0.0010*	*rs76674881*
SYT15	*10*	*46955443*	*46970601*	*118*	*439.99*	*1.44×10^{-5}*	*2.63×10^{-5}*	*rs58260169*
PDGFD	11	103777913	104035027	403	887.27	0.0001	9.90×10^{-5}	rs17102144
NOMO3	16	16326388	16388668	107	389.45	0.0002	1.09×10^{-5}	rs541161755
NPIPA7	16	16473101	16487829	31	109.88	0.0003	0.0008	rs574094453

Top 10 protein-coding genes associated with joint trait analysis of sinuses of Valsalva (SoV) diameter with sinotubular junction diameter. Chr: chromosome; Gene start/end: left and right-side boundaries of the gene (GRCh37); No.SNPs: number of SNPs occurring within gene contributing to the analysis; Chi-sq: sum of chi-squared test statistics of all SNPs within the gene region; P value: segment-based test p value; P value most sig SNP: smallest single-SNP p value in the segment. ***Gene names:*** *FAM25C; Family member with sequence similarity 25C; FAM25G: Family member with sequence similarity 25G; ANXA8: Annexin A8; NPY4R: Neuropeptide Y receptor 4; GPRIN2: G Protein Regulated Inducer Of Neurite Outgrowth 2; AGAP9: ArfGAP With GTPase Domain, Ankyrin Repeat And PH Domain 9; SYT15: Synaptotagmin 15; PDGFD: Platelet Derived Growth Factor D; NOMO3: NODAL Modulator 3; NPIPIA7: Nuclear Pore Complex Interacting Protein Family Member A7. Genes located in the same chromosomal region on chromosome 10 are shown in grey.*

This list of associated genes derived from the joint trait association results throws up some interesting candidates. The first 7 genes in the list in fact are clustered in the same chromosomal region of association – so perhaps just one of these genes is truly associated with phenotype, although they all reach statistical significance. Strong candidates include *NPY4R – neuropeptide Y receptor 4*. Neuropeptide Y is known to mediate many metabolic processes including appetite and, insulin secretion. The specific functions of NPY receptor 4 are rather less studied, although knockout mice also have considerable cardiac autonomic defects[406], with reductions in both vagal inhibitory responses and adrenergic responses to stimuli, suggesting that *NPY4R* plays a central role in cardiovascular homeostasis and autonomic response. If this were the gene responsible for this association signal, it would fit nicely with the suggested roles of the GABAergic and glutamatergic signalling pathways, supporting a key role of the autonomic nervous system in mediating aortic root dimensions. *GPRIN2* is another candidate at this same locus and is very highly expressed in cardiac tissue. Little is known of its specific function.

PDGFD (Platelet-Derived Growth Factor D) is the second most highly associated gene, assuming that the first 7 genes all represent the same locus of association. *PDGFD* knockout mice demonstrate a clear but mild vascular phenotype. PDGFD has been shown to play a key role in angiogenesis[407]. It has been suggested that *PDGFD* variants increase cardiovascular mortality in elderly men[408], although the evidence for this is not strong. Additionally, PDGFD derived from perivascular adipose tissue has even been linked with aortic aneurysm formation in mice – perivascular adipose tissue expressed high levels of PDGFD in leptin-deficient obese mice, and inhibition of PDGFD led to a reduction in aortic aneurysm formation with angiotensin II infusion in these mice. Overexpression of *PDGFD* in adipose tissue alone, increased the incidence of angiotensin-II-induced aortic aneurysm and dissection significantly – in the abdominal and thoracic aorta. The adventitial layer of the aorta was markedly thicker with more fibrosis, and the TGF-ß pathway was activated[409]. These data reveal a key role for *PDGFD* in aortic remodelling. Of note is the fact that one of the PDGF receptor subunit genes, *PDGFRA*, is a possible association with aortic valve annulus diameter. This joint association further focusses attention on PDGF signalling.

NOMO3 is a modulator of the NODAL signalling pathway, which is important for mesodermal differentiation and development of the heart and cardiovascular system. Evidence for a specific role for *NOMO3* is, however, lacking.

The top 10 genes presented above have no overlaps with genes identified in association with the individual traits which contributed to the joint analysis. However, in positions 14 and 15 are *SYNDIG1* and *SCUBE1*, which are associated with STJ and SoV diameter respectively, above.

Joint trait analysis may also be used to identify pathways not significantly associated with either single trait – the top 5 are presented below.

TABLE 14.12: Top 5 KEGG pathways and top 5 Gene Ontology biological processes associated with joint trait analysis of SoV and STJ diameter.

KEGG Pathway	FDR-corrected p value
KEGG: TGF BETA SIGNALING PATHWAY	0.005
KEGG: FC EPSILON RI SIGNALING PATHWAY	0.014
KEGG: FOCAL ADHESION	0.017
KEGG: NEUROTROPHIN SIGNALING PATHWAY	0.018
KEGG: WNT SIGNALING PATHWAY	0.024

GO pathway: Biological Process	FDR-corrected p value
GO: REGULATION OF PHOSPHORYLATION	0.006
GO: REGULATION OF SMALL GTPASE MEDIATED SIGNAL TRANSDUCTION	0.006
GO: REGULATION OF PROTEIN AMINO ACID PHOSPHORYLATION	0.007
GO: POSITIVE REGULATION OF SIGNAL TRANSDUCTION	0.007
GO: SMALL GTPASE MEDIATED SIGNAL TRANSDUCTION	0.007

Here we see the TGF-beta pathway as a key mediator of the combined aortic root dimension phenotype. This comes as no surprise given the immense data on the role of this pathway in aortic diseases such as Marfan syndrome and Loeys-Dietz syndromes. The signalling pathway is known to be fundamental to aortic integrity, and there is some evidence that more common variants in TGF-ß receptor genes can mediate phenotypic differences in aortic pathology[410]. The top 5 genes driving this pathway association are presented in Table 14.12.a below:

TABLE 14.12.a: Top 5 genes within the TGF-ß pathway associated with joint trait analysis of SoV and STJ diameter

Gene	P value of association	Top SNP	Top SNP P value
ACVR2B	0.043	rs397701814	0.0075
BMPR2	0.044	rs4675283	0.0005
ID1	0.046	rs112380827	$5.68x10^{-5}$
MYC	0.048	rs3931651	0.0003
THBS2	0.062	rs72503851	0.0017

The FC epsilon ri signalling pathway is active predominantly in mast cell activation – an immune response to allergens. Whilst chronic inflammation can certainly have a great impact on vascular function and remodelling, this does not immediately appear to have clear links with aortic phenotype.

Cell-cell and cell-extracellular matrix adhesion is known to be of vital importance in the aorta. Physical attachment of vascular smooth muscle cells to different components of the extracellular matrix via integrins, clustered into focal adhesions, allows mechanical force transduction from extracellular matrix to intracellular actomyosin fibres. This mechanotransduction activates multiple signalling cascades involved in the homeostasis of aortic wall extracellular matrix composition and turnover, including TGF-B signalling. Genes and molecules in this pathway are known to be important determinants of vascular smooth muscle cell contractile and secretory phenotype[22], and are implicated in aortopathies. Genes in the focal adhesion pathway have not previously been associated with aortic root dimensions in a healthy population.

Neurotrophin signalling has been shown to be of importance in the early development of the left ventricular outflow tract, with high levels of expression of neurotrophin receptors in the aortic wall[411]. Neurotrophin plays a key role in the migration and development of neural crest cells – key components of the developing aorta. There has been much recent interest in this pathway as a mediator of aortic valve calcification; whether it also plays a role in aortic root development or remodelling has not previously been ascertained.

A similarly important pathway for aortic biology is the Wnt signalling pathway, also associated with aortic root dimensions in the current analysis. Wnt signalling is known to regulate

smooth muscle cell proliferation, migration and apoptosis, predominantly through non-canonical pathways.

Pathways analysis of joint trait results has therefore revealed several key pathways for aortic biology converging to mediate effects on aortic root diameter. These subtle effects of key signalling pathways on root dimensions have not been shown before, and could be key to phenotypic variation in both health and aortic disease.

14.5.3 Ascending and descending aortic areas association results

Genetic associations with ascending and descending aortic diastolic area at the level of the pulmonary bifurcation, measured by MRI, are presented in the following sections. For these skewed phenotypes, the regression model was log (area) ~ age + gender + height + weight + MAP + 3PCs+genotype. It should be noted that the effect sizes reported below are therefore on a logarithmic scale. No GWAS have previously reported associations with these traits.

Association (Manhattan) plots for ascending and descending aortic diastolic areas are shown below. There were no genome-wide significant loci for any of the traits examined. However, there were several loci which were associated with dimensions at suggestive levels of significance ($p<1x10-5$); these are presented below. Gene and pathways- based analysis yielded further insights into the genetic underpinnings of these traits.

14.5.3.1 Ascending and descending aortic areas: Narrow-sense heritability estimates

The narrow-sense heritability for each of the aortic area traits was assessed using GREML as described in chapter 13. In brief, this enables an assessment of the proportion of variance in the phenotype which can be explained by all the SNPs in the dataset.

TABLE 14.13: Estimates of narrow-sense heritability for ascending and descending aortic areas

Trait	VG/Vp	SE	P value
Ascending aortic area	0.461	0.297	0.053
Descending aortic area	0.845	0.290	0.0008

VG/Vp: trait variance explained by genotype as a proportion of the overall trait variance; SE: standard error; P value: significance value for estimate of VG/Vp.

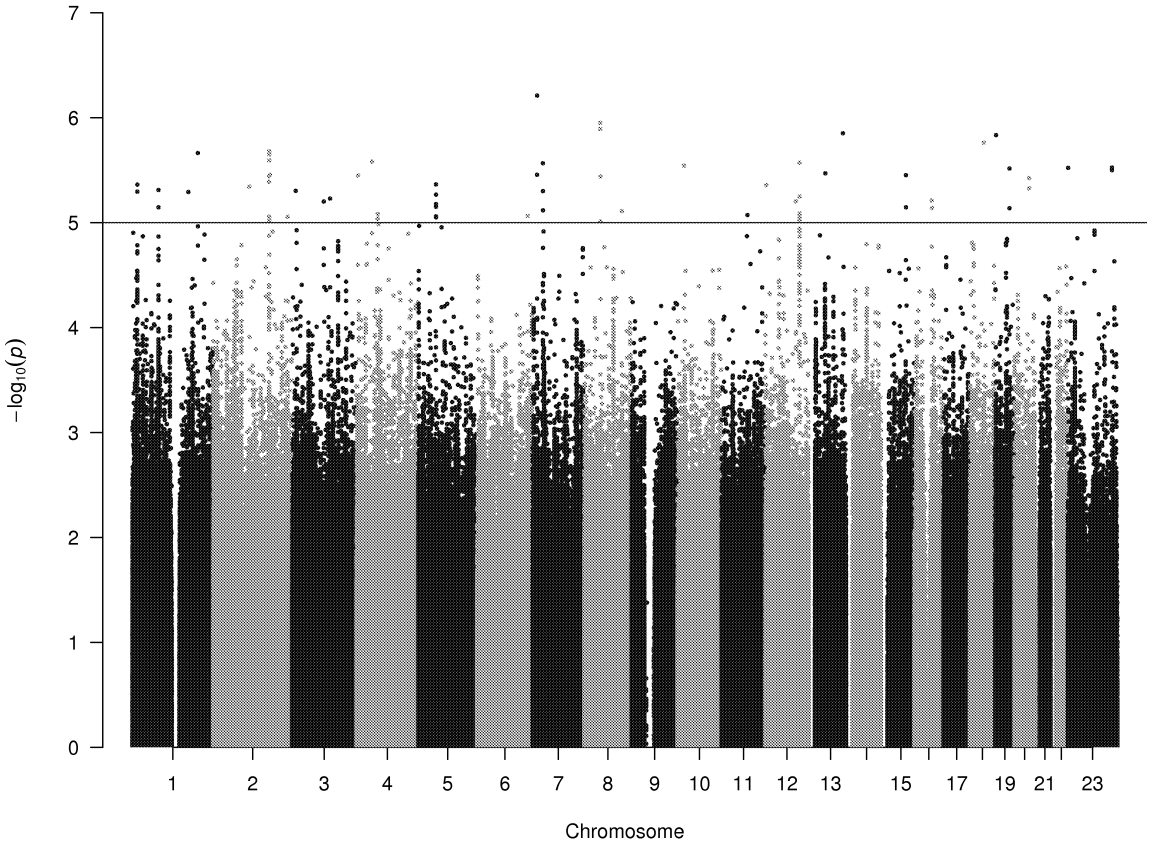
For ascending and descending aortic areas respectively, the genetic contribution to phenotype variance from our measured variants was 46% and 85% respectively, although the standard errors for the estimates were large, and the model for ascending aortic area did not quite meet statistical significance. The heritability estimate for descending aortic area is very large, suggesting that developmental and innate genetic mechanisms dominate over environmental influences in this section of the aorta. These figures suggest that cardiovascular risk factors and additional environmental variables might be of more

importance in determining ascending aortic phenotype, perhaps in keeping with the more marked changes seen with ageing in the proximal thoracic portion of the aorta.

14.5.3.2 Ascending aortic area: single SNP associations

As with other phenotypes, no single SNPs met the genome-wide significance threshold of $p<5x10^{-8}$. However, there were several loci which were associated with ascending aortic area at suggestive levels of significance.

FIGURE 14.8: Manhattan plot showing associations of ascending aortic diastolic area with genotype.



The y axis displays -log_{10} p values of association – i.e. the "higher" the point on the graph, the more significant the association. The x axis is the position on each chromosome. Each dot represents a single SNP; however, these often overlie one another. The horizontal threshold line represents a widely-accepted yet arbitrary "suggestive significance" level threshold of p = $1x10^{-5}$.

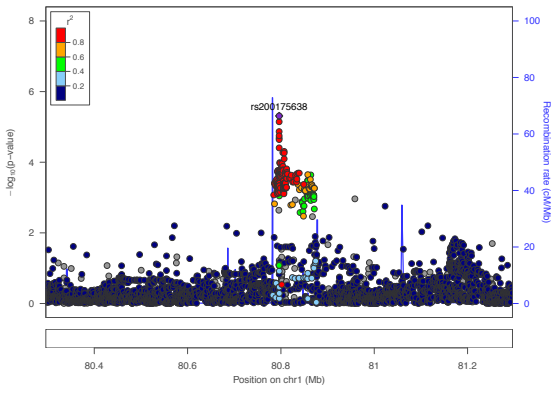
TABLE 14.14: Top SNPs associated with ascending aortic diastolic area

dbSNP	Chr	Position	Ref	Alt	MAF	Beta	p-value	Closest protein-coding gene	Other nearby candidate genes	Comments
rs200175638	1	80796241	G	A	0.214	0.047	4.89×10^{-6}			No clear candidates
rs2249572	2	172425464	G	A	0.128	-0.055	2.09×10^{-6}	*CYBRD1*		*CYBRD1* k/o mice have decreased heart weight. Role in iron homeostasis
rs201897330	4	65919755	C	T	0.393	-0.038	8.28×10^{-6}	*EPHA5*		No obvious role in cardiovascular system although important for brain development
rs2112935	5	54268487	T	C	0.128	0.055	4.32×10^{-6}	*ESM1*	*GZMK*	*ESM1* expressed highly in endothelium; involved in angiogenesis
rs116928433	7	31873333	G	A	0.017	-0.144	2.71×10^{-6}	*PDE1C*		*PDE1C* regulates proliferation and migration of VSMCs; overexpressed in heart and vascular system
rs11112418	12	105630463	G	A	0.084	0.068	2.67×10^{-6}	*APPL2*	*KIAA1033, C12orf75, ALDH1L2*	*KIAA1033* associated with short stature.
rs369840319	16	54402618	C	CAA	0.094	-0.065	6.15×10^{-6}	*IRX3*		*IRX3* known to be involved in obesity and adipogenesis.

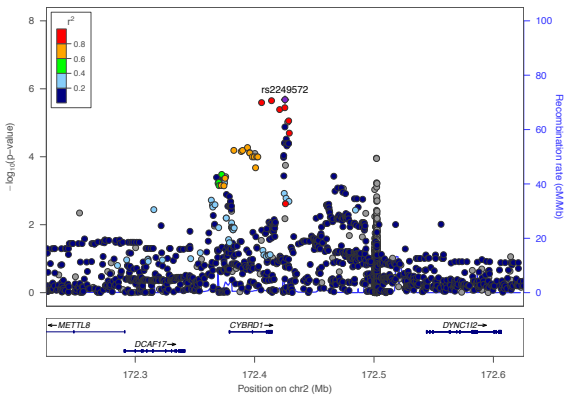
Chr: chromosome; Ref: Reference allele; Alt: effect allele; MAF minor allele frequency; Beta: effect size (change in diameter (mm) per copy of effect allele); Comments: annotations from multiple sources including GWAS catalog[36], GeneCards[369], Mouse Genome Informatics database[370], GTEx[265] and PubMed searches for gene function information; CYBRD1: Cytochrome B Reductase 1; EPHA5: Ephrin Receptor A5; ESM1: Endothelial Cell Specific Molecule 1; GZMK: Granzyme K; PDE1C: Phosphodiesterase 1C; APPL2: Adaptor Protein, Phosphotyrosine Interacting With PH Domain And Leucine Zipper 2; KIAA1033 aka WASHC4: WASH Complex Subunit 4; C12orf75: Chromosome 12 open reading frame 75; ALDH1L2: Aldehyde Dehydrogenase 1 Family Member L2; IRX3: Iroquois Homeobox 3.

FIGURE 14.9: LocusZoom plots for selected loci associated at suggestive significance levels with ascending aortic area

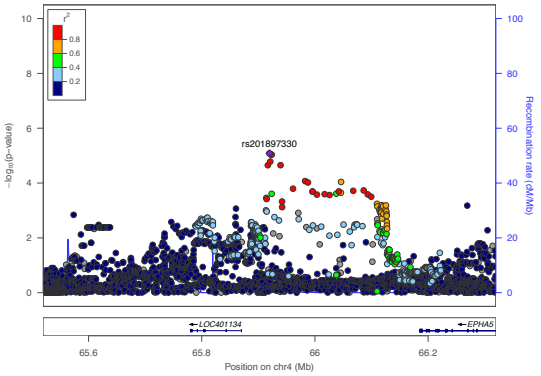
a. Chromosome 1: unknown



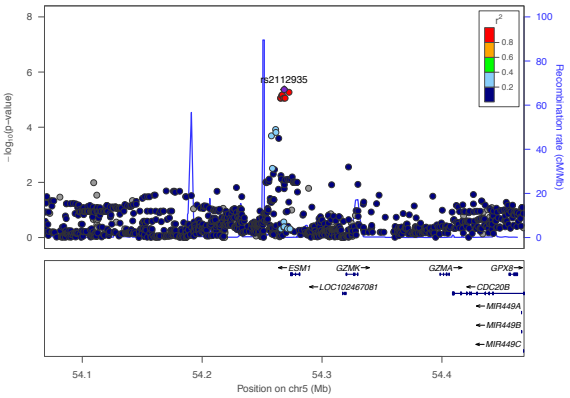
b. Chromosome 2: CYBRD1



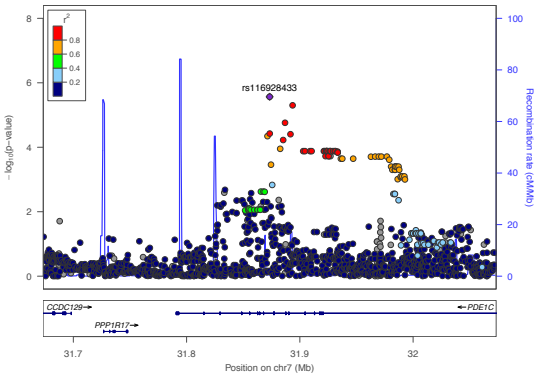
c. Chromosome 4: EPHA5



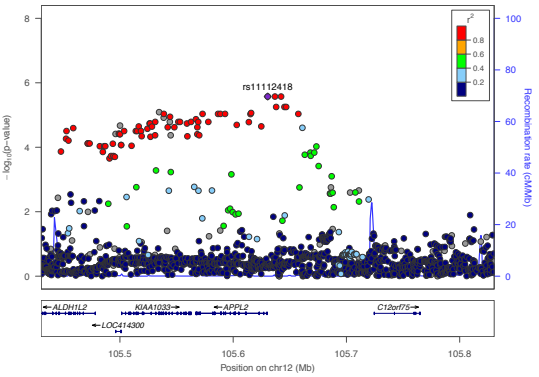
d. Chromosome 5: ESM1



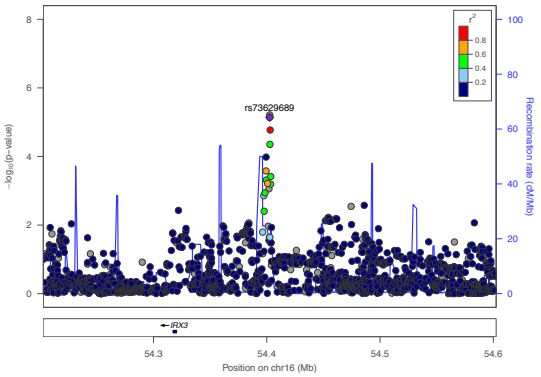
e. Chromosome 7: PDE1C



f. Chromosome 12: APPL2/KIAA1033



g. Chromosome 16: IRX3



Suggestive loci are shown as examples above in LocusZoom plots. These plots are essentially a "zoomed in" version of a Manhattan plot, but also demonstrate the linkage disequilibrium structure of the association peak. They demonstrate the spatial relation of the associated SNPs with annotated genes, and with nearby SNPs in linkage disequilibrium (LD) with the lead variant. The plots selected here demonstrate the expected structure of a "true positive" association, with SNPs in LD with the lead variant also associated with phenotype, albeit at lower significance. The labels a-g denote the chromosomal location of the loci, and the top possibilities for candidate genes at the locus, based on biological plausibility and functional evidence, detailed in Table 14.14 and below.

The most significant locus here is at chromosome 2, with the top SNP being approximately 10kb upstream of *CYBRD1 (cytochrome B reductase 1)*. The top SNP is a significant eQTL for *CYBRD1*, at least in peripheral blood. The protein product of this gene facilitates iron absorption in the small intestine, but is also very highly expressed in fibroblasts and in arterial tissues[29]. Knockout mice have reduced heart weight, so there is theoretical plausibility that *CYBRD1* might also play a role in the cardiovascular system. There is some evidence for macrophage iron being important in the development in atherosclerosis, and interest in this pathway for treatment. In a hyperlipidaemic mouse model, knockout of hepcidin (a hormonal regulator of iron homeostasis with impacts on intestinal absorption) reduced aortic macrophage inflammatory phenotype and protected from atherosclerosis[412].

Additional loci of interest here include the chromosome 7 locus, implicating *PDE1C (Phosphodiesterase 1C)*. This gene encodes a phosphodiesterase which has been implicated in pathological remodelling in the cardiovascular system[413], and is a potential key therapeutic target. Its expression is upregulated in failing hearts, and knockout myocytes are protected from hypertrophy or apoptosis in response to pressure overload or ISO-induced apoptosis. *PDE1C* knockout hearts exhibited less interstitial fibrosis in response to transverse aortic constriction-induced pressure overload[414]. *PDE1C* is not itself expressed at high levels in fibroblasts or myofibroblasts, and it seems that myocyte production of PDE1C might regulate fibrosis through a paracrine mechanism[414]. It is tempting to speculate that similar mechanisms could be at play in the aorta, and that PDE1C could regulate remodelling of the aorta. *PDE1C* is not constitutively expressed in human aortic smooth muscle cells; however its expression is induced when phenotype switching occurs in SMCs to a synthetic, secretory phenotype[415]; a step which occurs early in atherosclerosis and which is implicated also in genesis of thoracic aortic aneurysms. The association here with ascending aortic area bolsters the idea that PDE1C could be a potentially important therapeutic target in aortic remodelling and prevention of aortic aneurysms.

ESM1 is also an interesting candidate – it is highly expressed in endothelial cells and serum levels of its protein product, endocan, serve as biomarkers of endothelial dysfunction in a number of settings from erectile dysfunction, where levels predict cardiovascular risk, to endothelial dysfunction in renal disease or hypothyroidism. Its possible involvement in aortic

remodelling and determination of aortic size demonstrates the importance of the endothelium in aortic traits.

Other candidate genes, *EPHA5 (Ephrin Receptor A5)* and *IRX3 (Iroquois Homeobox 3)*, are known to have developmental roles. *IRX3* in particular is implicated in obesity and adipocyte function, and is functionally linked with a putative determinant of PWV, *ARID5B*. Direct links with aortic traits are less clear at other loci.

14.5.3.3 Ascending aortic area: gene-based GWAS results

None of these gene-based associations were significant when multiple testing corrections were applied (for 20,000 comparisons). The top 10 genes associated with ascending aortic area are listed in Table 14.15 below:

TABLE 14.15: Gene-based associations with ascending aortic area

Gene	Chr	Start	End	No. SNPs	Chi-sq	P value	P value most sig SNP	Most sig SNP
TMEM51	1	15479027	15546974	323	954.02	9.77×10^{-5}	4.34×10^{-5}	rs71000384
CYP26B1	2	72356366	72374991	130	424.14	0.0002	2.24×10^{-5}	rs397771158
OR6C68	12	55886161	55887100	87	280.32	0.0002	0.0003	rs796841689
HMCN1	1	185703682	186160085	311	800.70	0.0004	3.44×10^{-5}	rs2208711
ZNF860	3	32023265	32033228	140	477.21	0.0005	0.0004	rs9845645
TNFRSF8	1	12123433	12204264	292	612.43	0.0005	3.39×10^{-5}	rs75169069
C1ORF195	1	15490691	15498120	157	483.20	0.0006	1.88×10^{-5}	rs12118545
CYBRD1	2	172378756	172414643	179	526.01	0.0006	2.09×10^{-6}	rs2249572
OR6C70	12	55862983	55863922	83	249.06	0.0006	0.0003	rs796841689
NUDT19	19	33182866	33204702	130	362.59	0.0007	1.56×10^{-5}	rs7254931

Top 10 protein-coding genes associated with ascending aortic area. Chr: chromosome; Gene start/end: left and right-side boundaries of the gene (GRCh37); No.SNPs: number of SNPs occurring within gene contributing to the analysis; Chi-sq: sum of chi-squared test statistics of all SNPs within the gene region; P value: segment-based test p value; P value most sig SNP: smallest single-SNP p value in the segment; TMEM51: Transmembrane protein 51; CYP26B1: Cytochrome P450 Family 26 Subfamily B Member 1; OR6C68: Olfactory Receptor Family 6 Subfamily C Member 68; HMCN1: hemicentin1 (aka fibulin 6); ZNF860: Zinc Finger Protein 860; TNFRSF8: Tumour Necrosis Factor Receptor Superfamily Member 8; C1ORF195: Chromosome1 Open Reading Frame 195; CYBRD1: cytochrome B reductase 1;OR6C70: Olfactory Receptor Family 6 Subfamily C Member 70; NUDT19: Nudix Hydrolase 19

Of these genes, *CYP26B1* is involved in the metabolism of retinoic acid. SNPs in this gene have been associated with larger atherosclerotic lesions[416]. *HMCN1 - hemicentin 1*, also known as *fibulin-6*, is a regulator of TGF-ß -mediated fibrotic response in ventricular fibroblasts, and is

implicated in cardiac remodelling[417]. Increased expression after myocardial infarction stimulates TGF-ß-dependent fibroblast migration[418]. One could imagine that this remodelling could be important also in the aorta. Certainly, the key role in modulating the TGF-ß pathway makes this a very interesting candidate for future follow-up. *NUDT19* knockout mice have enlarged hearts[376], but there is little evidence to suggest a mechanism for this. Other listed genes including *TMEM51, ZNF860, TNFRSF8, C1orf195* and the two olfactory receptors do not have any clear links with cardiovascular phenotype.

14.5.3.4 Ascending aortic area: pathway association results

Table 14.16 below shows the top 5 associated KEGG and GO biological processes derived from the genome-wide association results for ascending aortic area.

TABLE 14.16: Top 5 KEGG pathways and top 5 Gene Ontology biological processes associated with ascending aortic area.

KEGG Pathway	FDR-corrected p value
KEGG: HISTIDINE METABOLISM	0.0810
KEGG: FATTY ACID METABOLISM	0.0814
KEGG: PROXIMAL TUBULE BICARBONATE RECLAMATION	0.0827
KEGG: ADIPOCYTOKINE SIGNALING PATHWAY	0.1107
KEGG: ALDOSTERONE REGULATED SODIUM REABSORPTION	0.1343

GO pathway: Biological Process	FDR-corrected p value
GO: GAMETE GENERATION	0.0097
GO: DNA PACKAGING	0.0173
GO: SEXUAL REPRODUCTION	0.0212
GO: RNA SPLICING	0.0309
GO: CHROMOSOME ORGANIZATION AND BIOGENESIS	0.0525

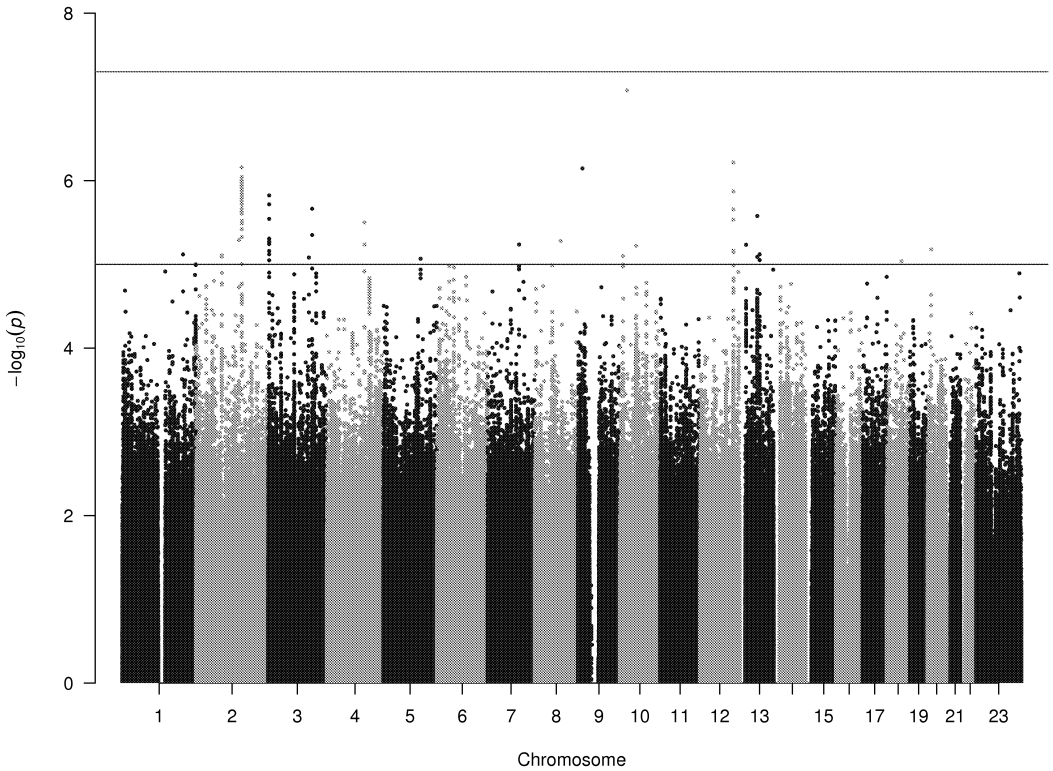
Interestingly, fewer pathways were identified as significantly associated with ascending aortic area than with the aortic root phenotypes. This may be because particular individual members of multiple pathways have effects on ascending aortic area, or because environmental influences are stronger here. No KEGG pathways reached statistical significance for association. However, the results are suggestive that pathways affecting cardiovascular risk factors - such as fatty acid metabolism and the aldosterone pathway which

regulates blood pressure- might be important determinants of aortic phenotype. The GO biological process terms which were significant were rather surprising, identifying processes important for sexual reproduction as well as general DNA organisation terms. Whether these are representative of hormonal effects on aortic biology, or are false positive associations, is uncertain.

14.5.3.4 Descending aortic area: single SNP association results

As with other phenotypes, no single SNPs met the genome-wide significance threshold of $p<5x10^{-8}$. However, there were several loci which were associated with descending aortic area at suggestive levels of significance.

FIGURE 14.10: Manhattan plot showing associations of descending aortic diastolic area with genotype.



The y axis displays -log_{10} p values of association – i.e. the "higher" the point on the graph, the more significant the association. The x axis is the position on each chromosome. Each dot represents a single SNP; however, these often overlie one another. The horizontal threshold line represents a widely-accepted yet arbitrary "suggestive significance" level threshold of p = $1x10^{-5}$.

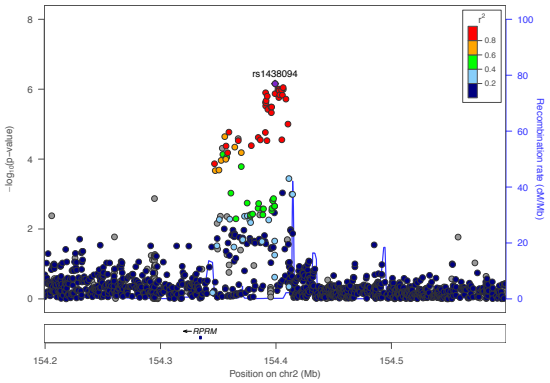
TABLE 14.17: Top SNPs associated with descending aortic diastolic area

dbSNP	Chr	Position	Ref	Alt	MAF	Beta	P-value	Closest protein-coding gene	Other nearby candidate genes	Comments
rs1438094	2	154399249	C	T	0.41	-0.03	6.96×10^{-7}	*RPRM*		No clear links with cardiovascular phenotype
rs764593	3	3712236	T	C	0.12	-0.05	1.50×10^{-6}	*LRRN1*		*LRRN1* differentially regulated by NKX2-5 during OFT development
rs116344051	3	148734342	A	G	0.02	0.11	2.17×10^{-6}	*GYG1*		Variants associated with limb girdle muscular dystrophy & cardiomyopathy
rs7724725	5	125774522	G	A	0.07	0.06	8.55×10^{-6}	*GRAMD3*		Assoc. with body fat distribution
rs42232	7	106243626	T	A	0.24	-0.04	5.79×10^{-6}	*CCDC71L*		Assoc. with atherosclerosis esp carotid
rs117288436	10	25109950	G	C	0.02	-0.14	8.35×10^{-8}	*PRTFDC1*		No clear cardiovascular link
rs12580178	12	113024793	G	A	0.31	-0.04	6.05×10^{-7}	*PTPN11*	*RPL6, ALDH2*	*PTPN11* variants cause Noonan syndrome, assoc. with BP; top SNP eQTL for *ALDH2* – assoc. with multiple cardiovascular phenotypes
rs17051454	13	21523628	G	T	0.08	-0.06	5.83×10^{-6}	*LATS2*	*XPO4*	
rs398023183	13	67561825	A	AT	0.38	0.03	7.57×10^{-6}	*PCDH9*		Adhesion molecule
rs6109718	20	13102026	A	G	0.30	0.03	2.29×10^{-5}	*SPTLC3*	*ISM1*	*SPTLC3* associated with LDL cholesterol levels in previous GWAS.

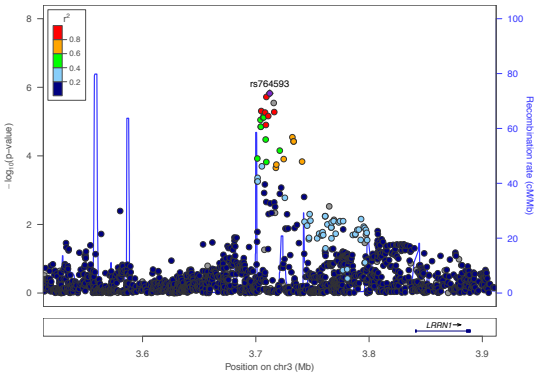
Chr: chromosome; Ref: Reference allele; Alt: effect allele; MAF minor allele frequency; Beta: effect size (change in diameter (mm) per copy of effect allele); Comments: annotations from multiple sources including GWAS catalog[36], GeneCards[369], Mouse Genome Informatics database[370], GTEx[265] and PubMed searches for gene function information; RPRM: Reprimo, TP53 Dependent G2 Arrest Mediator Homolog. LRRN1: Leucine Rich Repeat Neuronal 1; OFT: ventricular Outflow Tract; GYG1: Glycogenin 1; HLTF: Helicase Like Transcription Factor; HPS3: Hermansky-Pudlak Syndrome 3, Biogenesis Of Lysosomal Organelles Complex 2 Subunit 1; GRAMD3: GRAM Domain Containing 2B; CCDC71L: Coiled-Coil Domain Containing 71 Like; PRTFDC1: Phosphoribosyl Transferase Domain Containing 1; PTPN11: Protein Tyrosine Phosphatase, Non-Receptor Type 11; RPL6: Ribosomal Protein L6; ALDH2: Aldehyde Dehydrogenase 2; LATS2: Large Tumor Suppressor Kinase 2; XPO4: Exportin 4; PCDH9: Protocadherin 9; SPLTC3: Serine Palmitoyltransferase Long Chain Base Subunit 3; ISM1: Isthmin 1.

FIGURE 14.11: LocusZoom plots for selected loci associated with descending aortic diastolic area

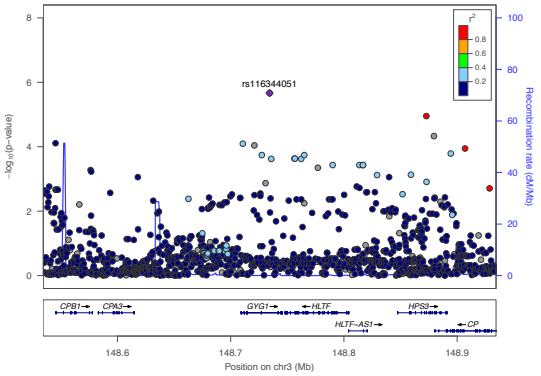
a. Chromosome 2: RPRM(?)



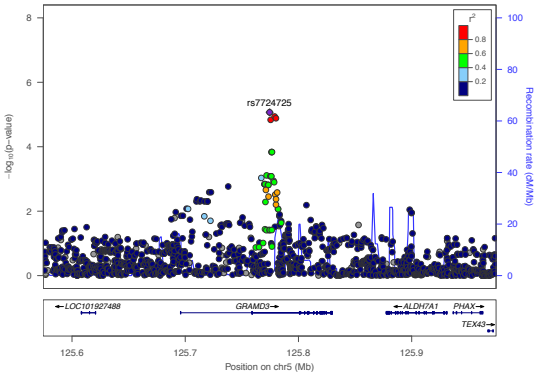
b. Chromosome 3: LRRN1



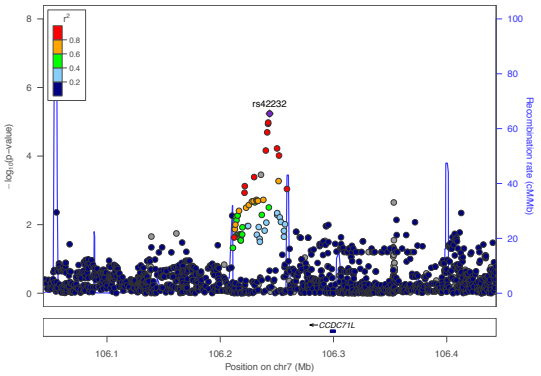
c. Chromosome 3: GYG1



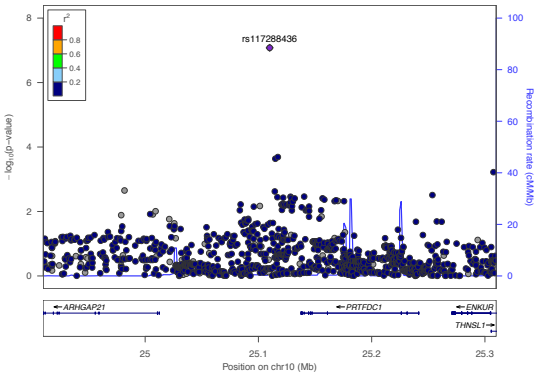
d. Chromosome 5: GRAMD3



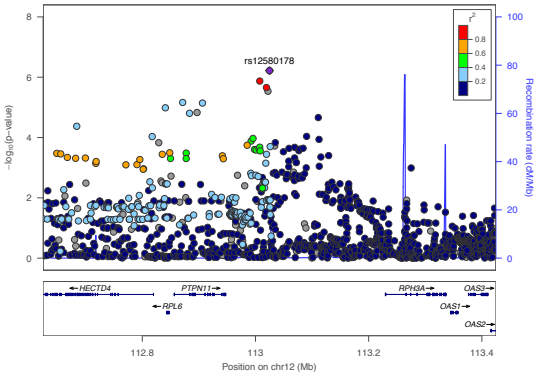
e. Chromosome 7: CCDC71L



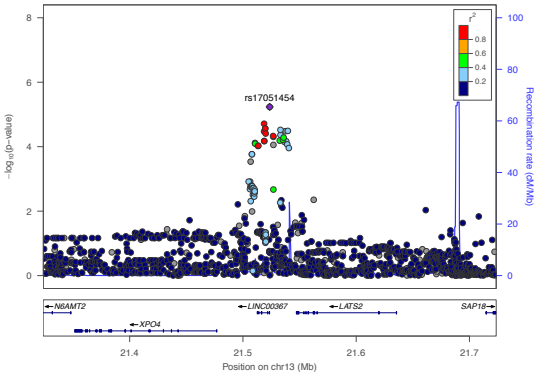
f. Chromosome 10: PRTFDC1(?)



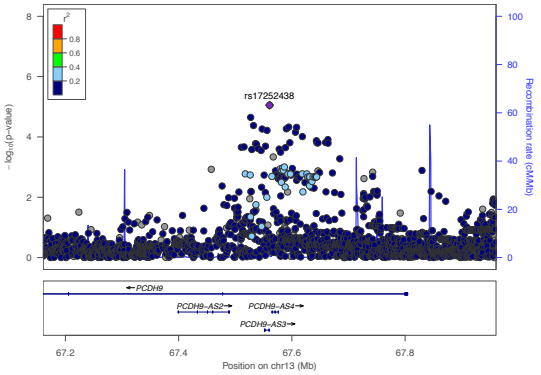
g. Chromosome 12: PTPN11 / ALDH2



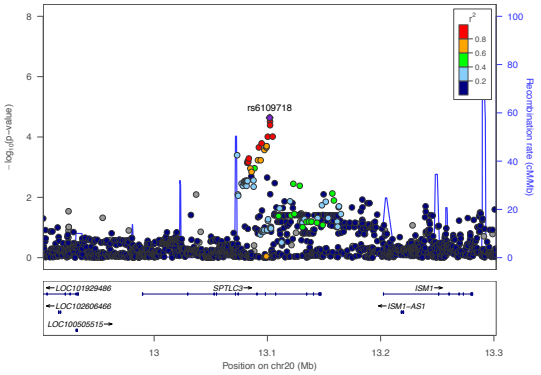
h. Chromosome 13: LATS2



i. Chromosome 13: PCDH9



j. Chromosome 20: SPTLC3/ISM1



Suggestive loci are shown as examples above in LocusZoom plots. These plots are essentially a "zoomed in" version of a Manhattan plot, but also demonstrate the linkage disequilibrium structure of the association peak. They demonstrate the spatial relation of the associated SNPs with annotated genes, and with nearby SNPs in linkage disequilibrium (LD) with the lead variant. The plots selected here demonstrate the expected structure of a "true positive" association, with SNPs in LD with the lead variant also associated with phenotype, albeit at lower significance. The labels a-j denote the chromosomal location of the loci, and the top possibilities for candidate genes at the locus, based on biological plausibility and functional evidence, detailed in Table 14.17 and below.

Here, the peaks seem to be more sharply defined than some of the other phenotypes. Whilst the SNP on chromosome 10 appears to have the greatest statistical association, the structure of the association peak, shown in Figure 14.11 (f) above, near *PRTFDC1 (Phosphoribosyl Transferase Domain Containing 1),* seems less robust, and the nearby gene has no obvious links with cardiovascular phenotypes. This may therefore be a false positive association. However, this gene appears again in the association results with aortic distensibility phenotypes, so is worth noting for further investigation.

Associations of interest here include the most significant locus on chromosome 12, close to *PTPN11 (Protein Tyrosine Phosphatase, Non-Receptor Type 11).* Variants in this gene have been shown to cause Noonan syndrome[419]; a RASopathy which causes a developmental syndrome frequently involving the heart and aorta (hypertrophic cardiomyopathy, pulmonary

stenosis and coarctation of the aorta being some of the more frequent manifestations). *PTPN11* has also been strongly associated with several blood pressure indices in multiple GWASs. It is therefore a strong candidate for affecting aortic phenotype. However, the top SNP at this locus is also an eQTL for *ALDH2 (Aldehyde Dehydrogenase 2*) in arterial tissue and other tissue types[29]. Common variants in *ALDH2* are associated with multiple cardiovascular phenotypes, including coronary artery disease, many blood pressure indices and the interaction of alcohol consumption with blood pressure, as well as body mass index. Mice with a single point mutation in *ALDH2* exhibit multiple cardiovascular abnormalities, particularly when diabetes is induced in these mice by feeding a high-fat diet. Here, heart failure with preserved ejection fraction is more frequently observed, and the mutant mice have blunted heart rate and cardiac contractility responses to exercise, as well as increased cardiac hypertrophy[376]. There are several papers which confirm a role for *ALDH2* in cardiac remodelling in response to pressure overload[420], and evidence also suggests it may be important in mediating beneficial remodelling in response to aerobic exercise training. The presence of eQTLs for *ALDH2* in intronic regions of *PTPN11* raises the question as to whether some of the previous GWAS associations of *PTPN11* with blood pressure are in fact mediated via effects on expression of *ALDH2*.

A number of the other genes at loci associated with descending aortic area have also been shown to affect cardiovascular risk factors. These include *GRAMD3 (GRAM Domain Containing 2B)*, in which common variants are associated with body fat distribution[421], and *CCDC71L (Coiled-Coil Domain Containing 71 Like),* in which common and rare variants have been associated with atherosclerosis and carotid intimal-medial thickness[422]. *SPTLC3 (Serine Palmitoyltransferase Long Chain Base Subunit 3)* has also been associated with lipid indices and with the blood pressure response to thiazide diuretics. These suggestive associations might imply that regulation of descending aortic diameter is mediated largely through effects on other risk factors.

Three further loci tag genes of particular interest. Glycogenin (encoded by *GYG1*) is a glycosyltransferase which mediates one of the first steps in glycogen synthesis, catalysing the formation of a short glucose polymer from uridine glyphosphate glucose. Variants in this gene are a cause of limb girdle muscular dystrophy and also are thought to be a cause of

cardiomyopathy. Whether defects in glycogen synthesis could also affect aortic function (an "aortomyopathy"?) is not established, but this possibility certainly has some theoretical plausibility.

PCDH9 (protocadherin 9) is a member of the protocadherin family, like *PCHD7* implicated in aortic root dimensions above. However, it is not known to play a role in the cardiovascular system.

LATS2 (Large Tumour Suppressor Kinase 2) on chromosome 13, encodes part of the Hippo signalling pathway – which is of great interest in vascular remodelling. LATS2-mediated phosphorylation of YAP is regulated by laminar flow sensed by vascular endothelium in the aorta, and inhibits downstream signalling cascade, allowing vascular endothelial cells to remain quiescent and atheroprotective[423]. LATS2 also negatively regulates cardiac hypertrophic responses to pressure overload[424].

Other loci yielded little evidence of proximity to, or functional regulation of genes with clear association with aortic traits.

14.5.3.5 Descending aortic area: gene-based association results

Again, none of the genes were significantly associated with phenotype after FDR correction.

TABLE 14.18: Gene-based associations with descending aortic area

Gene	Chr	Start	End	No. SNPs	Chi-sq	P value	P value most sig SNP	Most sig SNP
RPL6	12	112842993	112847443	51	217.5	$5.62x10^{-5}$	$6.87x10^{-6}$	rs11066301
POMT2	14	77741298	77787225	213	658.9	0.0001	$5.32x10^{-5}$	rs59623393
NGB	14	77731833	77737655	195	576.0	0.0001	$5.32x10^{-5}$	rs59623393
TMED8	14	77808113	77843396	150	471.8	0.0002	$7.99x10^{-5}$	rs56154515
OR2L3	1	248223983	248224922	78	302.2	0.0003	$1.00x10^{-5}$	rs74153022
USP38	4	144106069	144145027	61	255.4	0.0003	$1.46x10^{-5}$	rs2010767
SYDE1	19	15218213	15225789	200	518.6	0.0003	$7.70x10^{-5}$	rs60510905
PTPN11	12	112856535	112947717	105	375.0	0.0003	$6.87x10^{-6}$	rs11066301
TMEM63C	14	77648101	77725838	322	784.9	0.0004	$5.32x10^{-5}$	rs59623393
OR2M5	1	248308449	248309388	98	361.41	0.0004	$1.99x10^{-5}$	rs4614295

Top 10 protein-coding genes associated with ascending aortic area. Chr: chromosome; Gene start/end: left and right-side boundaries of the gene (GRCh37); No.SNPs: number of SNPs occurring within gene contributing to the analysis; Chi-sq: sum of chi-squared test statistics of all SNPs within the gene region; P value: segment-based test p value; P value most sig SNP: smallest single-SNP p value in the segment; RPL6: Ribosomal Protein L6; POMT2: Protein O-Mannosyltransferase 2; NGB: Neuroglobin; TMED8: Transmembrane P24 Trafficking Protein Family Member 8; OR2L3: Olfactory Receptor Family 2 Subfamily L Member 3; USP38: Ubiquitin Specific Peptidase 38; SYDE1: Synapse Defective Rho GTPase Homolog 1; PTPN11: Protein Tyrosine Phosphatase, Non-Receptor Type 11; TMEM63C: Transmembrane Protein 63C; OR2M5: Olfactory Receptor Family 2 Subfamily M Member 5

PTPN11 is among the top 15 most significantly associated genes – although the p value for its association does not remain significant after multiple testing. As described above, this has very plausible links with vascular phenotype, but this and other genes at the same locus (*RPL6*) may be significant due to tagging of eQTLs for *ALDH2* (as discussed above).

POMT2 (Protein O-Mannosyltransferase 2) is a gene associated with alpha-dystroglycanopathies (a form of congenital muscular dystrophy), and variants have been associated with cardiovascular abnormalities in these patients, including aortic root dilatation[425]. *NGB (Neuroglobin)* is predominantly expressed in neural tissue, but seems to exert vasomotor effects which may be protective in the cardiovascular system. In mice overexpressing neuroglobin, myocardial infarct size after ligation of the left coronary artery was significantly reduced[426]. *POMT2* and *NGB* are both situated on chromosome 14, and are likely to be tagged by the same locus of association. Both have plausible links with cardiovascular phenotype, but there is limited evidence of the specific roles of either gene. None of the other genes listed have obvious links with cardiovascular traits.

14.5.3.6 Descending aortic area: pathway association results

Table 14.19 below shows the top 6 associated KEGG and top 5 GO biological processes derived from the genome-wide association results for descending aortic area.

TABLE 14.19: Top 6 KEGG pathways and top 5 Gene Ontology biological processes associated with descending aortic area.

KEGG Pathway	FDR-corrected p value
KEGG: CELL ADHESION MOLECULES CAMS	0.009
KEGG: SNARE INTERACTIONS IN VESICULAR TRANSPORT	0.022
KEGG: DILATED CARDIOMYOPATHY	0.025
KEGG: AMINO SUGAR AND NUCLEOTIDE SUGAR METABOLISM	0.026
KEGG: HEDGEHOG SIGNALING PATHWAY	0.054
KEGG: TGF BETA SIGNALING PATHWAY	0.063

GO pathway: Biological Process	FDR-corrected p value
GO: REGULATION OF CYCLIN DEPENDENT PROTEIN KINASE ACTIVITY	0.004
GO: NEGATIVE REGULATION OF GROWTH	0.021
GO: CALCIUM ION TRANSPORT	0.022
GO: REGULATION OF DNA BINDING	0.023
GO: NEGATIVE REGULATION OF CELL CYCLE	0.023

The KEGG pathways associated with descending aortic diameter highlight the role of cell adhesion molecules – including the protocadherin family. The association with the KEGG pathway for dilated cardiomyopathy again highlights the overlap of genetic influences on aortic and cardiac traits.

Another pathway highlighted here is the Hedgehog signalling pathway. This pathway has complex interactions with the Notch1 and TGF-ß signalling pathways. *Sonic Hedgehog* (*SHh*) expression is reduced in abdominal aortic aneurysms, and Hedgehog pathway inhibition results in a reduction in Notch1 signalling and vascular smooth muscle cell differentiation, and an increase in pro-fibrotic TGFß1 expression[427]. SHh is an important regulator of endothelial cell growth, differentiation and formation of new blood vessels[428]. The Hedgehog signalling pathway is also important for cardiac repair and regeneration, and has been a target of research in this area[429]. It therefore appears poised as a regulator of multiple vascular

homeostatic pathways such as Notch, VEGFA and TGF-ß, and is well-placed to mediate aortic homeostatic mechanisms.

The TGF-ß pathway is also identified above as the 6th most significantly associated pathway with descending aortic area. This is a well-characterised regulator of aortic phenotypes, and interacts significantly with Hedgehog signalling.

14.5.4 Aortic elastic function association results

Genetic associations with different measures of aortic elastic function are presented below. These phenotypes include ascending and descending aortic distensibility, measured by cardiovascular magnetic resonance imaging (CMR) at the level of the pulmonary bifurcation. I also present genetic associations with aortic arch pulse wave-velocity, also measured using CMR as described in Chapter 12. These highly-skewed phenotypes were rank-normalised, as models built using untransformed phenotypes or log-transformed phenotypes (as previous studies have used), yielded unstable and erroneous results (discussed further below), with regression models which did not fulfil assumptions.

14.5.4.1 Replication of previous association findings

We did not replicate any of the previous associations with carotid-femoral pulse wave velocity, even at nominal significance. This may be due to differences in the measured phenotype and also differences in the population under study, as well as differences in the specified association models and covariates.

14.5.4.2: Aortic elasticity: Narrow-sense heritability estimates

The narrow-sense heritability for each of the aortic elasticity traits was assessed using GREML as described in chapter 13. In brief, this is an assessment of the proportion of variance in the phenotype which can be explained by all the SNPs in the dataset. In other words, this approximately equates to what proportion of the trait variance is explained by common genetic variants in a healthy population.

TABLE 14.20: Estimates of narrow-sense heritability for ascending and descending aortic distensibility and aortic arch Pulse Wave Velocity

Trait	VG/Vp	SE	P value
Ascending aortic distensibility	0.613	0.281	0.007
Descending aortic distensibility	0.086	0.278	0.373
PWV	0.128	0.295	0.327

VG/Vp: trait variance explained by genotype as a proportion of the overall trait variance; SE: standard error; P value: significance value for estimate of VG/Vp. PWV: aortic arch Pulse Wave Velocity.

The heritability estimates for both pulse wave velocity and descending aortic distensibility were low, and neither model reached statistical significance. This limits the conclusions one can draw from this (the true estimate might be higher than these figures would suggest). However, the low estimates imply that perhaps environmental variables or rare genetic variation not captured in this dataset are of more importance in determining these measurements than common genetic variation. This is in contrast to the narrow-sense heritability estimates for aortic areas, where the descending aortic area had a very high heritability.

Conversely, for ascending aortic distensibility, the genetic contribution to phenotypic variance is relatively high, at 63% (albeit with large standard errors).

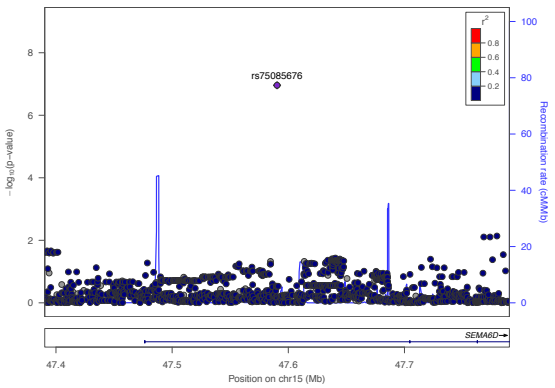
14.5.4.3 Ascending aortic distensibility: Model selection and single-SNP association results

The Manhattan plots below in Figures 14.13 a and b show genome-wide associations with ascending aortic distensibility; a phenotype associated with cardiovascular risk in many settings. It is interesting to note the varied results obtained with 2 different statistical approaches. The first, to log-transform the dependent variable, produced a somewhat inflated Q-Q distribution (see Figures 14.14 a and b), although this is not totally unreasonable for a complex trait with polygenic influences. By contrast, rank-normalisation prior to association analysis produces an under-inflated Q-Q plot, demonstrating again that this cohort is underpowered for discovery.

It is tempting to use a logarithmic transformation with the justification that the changes in phenotype which are of prognostic significance occur in the lower range of the phenotype which is bound by 0. A log-transformation will boost the significance of associations with this

"end" of the phenotype. It is tempting to speculate that at least some of the hits from the log-transformed phenotype will be valid determinants of variation at the lower end of the distensibility distribution. However, this statistical model is not technically valid, as the residuals from regression do not conform to a normal distribution. This will increase the likelihood of a Type I error in the results. Rank-normalising the phenotype brings the residuals to a much closer approximation of the normal distribution, fulfilling the assumptions of regression modelling. However, the power to discriminate between the fine changes in distensibility at the lower end of the scale is lost. To investigate this dilemma further, I examined the LocusZoom plots for loci identified as genome-wide significant from the log-transformed model. Very few "significant" loci from the log-transformed model have robust association peaks associated; several are single "floating" SNPs which are highly likely to represent false positive associations (see Figure 14.12 below for example) . Those which are robust might be more likely to represent significant biological associations.

FIGURE 14.12: Example of LocusZoom plot derived from the log-transformed distensibility association model, showing "orphan" significant SNP likely to represent a false-positive association



When we compare the two statistical approaches, the individual SNP associations are different. Several loci which appear genome-wide significant in association with log-transformed distensibility "disappear" in the rank-normalised analysis. Whether this simply reflects a higher type I error rate with log transformation, or illustrates the fact that considerable power is lost with more dramatic transformation of the dependent variable, is unclear. It is likely to be a mix of both. This data demonstrates the importance of careful model selection – reasonable arguments can be made for either form of transformation, and yet as the study is relatively under-powered, very different results are obtained from the two

approaches. Clearly, it is tempting for researchers to adopt the approach which is likely to produce the most "hits" and then justify this post-hoc. However, the association model risks high false positive rates if the residuals do not adhere to assumptions of normality. A rank-normalisation of phenotype also enables much easier meta-analysis with data from other studies. But this transformation comes at the expense of power and the potential loss of meaningful biological signals, particularly where changes of interest occur at the lower end of a variable's range. Where cohort sizes are massive, such as in the UK Biobank, it would be reasonable to have a more relaxed approach to model selection. However, where sample size and power are limited, and the statistical model makes such major differences to the results, one has to be more conservative in approach. For this reason, and after lengthy discussions with statisticians, I have concentrated on presenting results from association with rank-normalised phenotypes, as well as presenting alongside them results from the log-transformed association where the comparison is informative.

FIGURE 14.13a: Manhattan plot of association with log-transformed ascending aortic distensibility

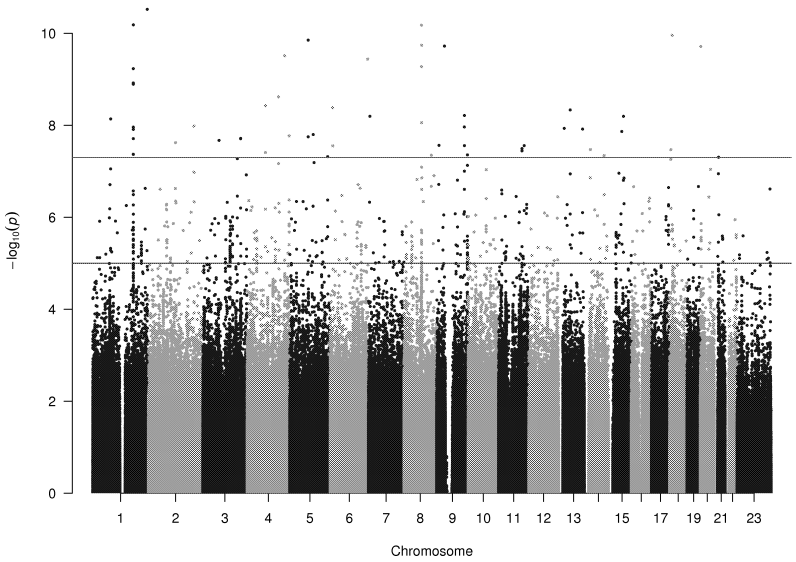
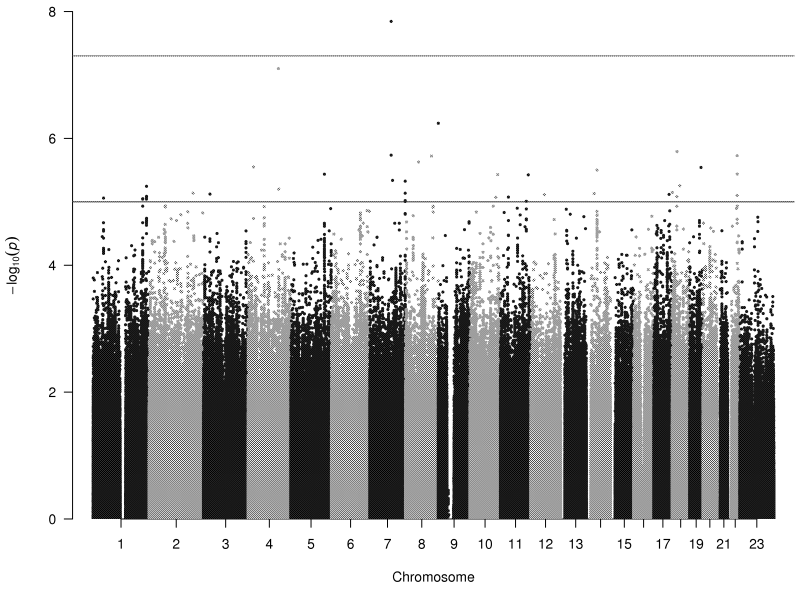


FIGURE 14.13 b: Manhattan plot of association with rank-normalised ascending aortic distensibility (note different y axis scale from 14.13a above)

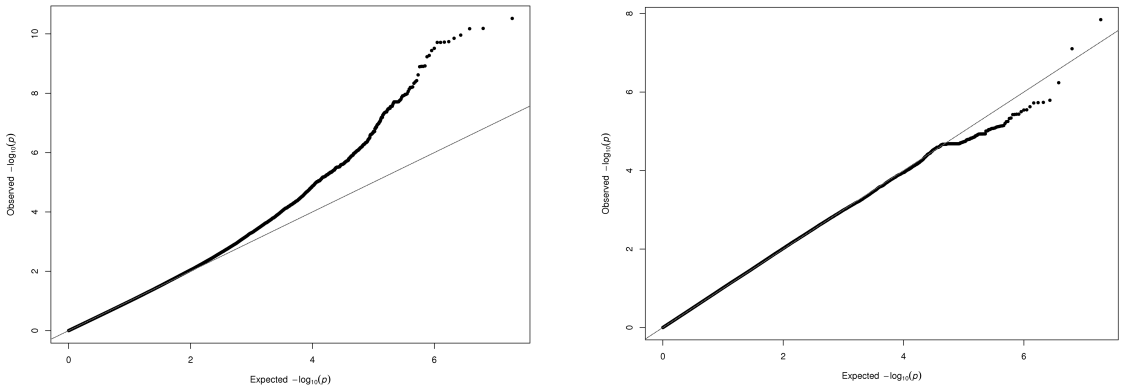


The y axis displays -log$_{10}$ p values of association – i.e. the "higher" the point on the graph, the more significant the association. The x axis is the position on each chromosome. Each dot represents a single SNP; however, these often overlie one another. The horizontal threshold lines represent a widely-accepted yet arbitrary genome-wide significance threshold at p=5x10^{-8} and a "suggestive significance" level threshold of p = 1x10^{-5}.

FIGURE 14.14: Quantile: quantile plot of p values from genome-wide associations with ascending aortic distensibility – effect of different phenotype transformations

a. Q-Q plot of p values from log-transformed model

b. Q-Q plot of p values from rank-normalised model



These plots show the quantile-quantile plots of the distribution of expected p values from genome-wide association against the observed values. The straight line on each plot represents x=y, i.e. the distribution of p values which would be generated by chance under a null hypothesis of no real associations with genotype. For an oligogenic trait, one would expect the plotted points to follow the x=y line closely until a low p value. A deviation from this pattern "above" the x=y line - an inflation of test statistics - implies that there is some population stratification causing the genotype overall to correlate with phenotype. However, for highly multigenic traits, one can expect to see the plotted points departing from the line "earlier", given high enough power in the study design, due to genuine associations. Deviations below the line mean that there is underinflation of the test statistics, most commonly due to underpowering.

None of the association peaks reach statistical significance for association with rank-normalised ascending aortic distensibility. There was one SNP on chromosome 7 with a p value for association below the genome-wide significance threshold. However, this SNP appeared to be a lone association without a robust peak underlying it, and is therefore likely to represent a false positive association. Other loci were associated at “suggestive” levels of significance. These are presented below in Table 14.21.

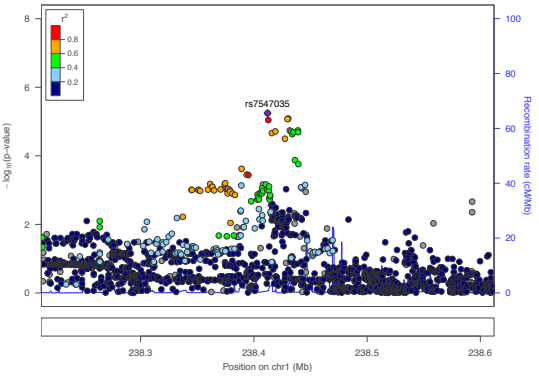
TABLE 14.21: Top SNPs associated with ascending aortic distensibility (rank normalized)

dbSNP	Chr	Position	Ref	Alt	MAF	Beta	p-value	Closest protein-coding gene	Other nearby candidate genes	Comments
rs7547035	1	238412097	A	G	0.443	-0.114	5.69×10^{-6}	*ZP4*	*RYR2*	*RYR2* variants known cause of ARVC, arrhythmias. Calcium channel which regulates intracellular calcium levels.
rs13374145	1	46930432	T	C	0.026	-0.349	8.72×10^{-6}	*FAAH*	*DMBX1*	*FAAH/DMBX1* locus has been associated with BMI.
rs9998160	4	137404436	T	A	0.190	-0.197	6.30×10^{-6}	*None within 500kb*		No clear links with CV phenotype
rs78044121	4	72644297	G	A	0.112	0.170	1.19×10^{-5}	*GC*		*GC* involved with vitamin D binding and transport. Associated in previous GWAS with coronary artery calcification.
rs398050845	5	149910119	A	ATT TTT TTT TTT	0.479	0.124	3.67×10^{-6}	*NDST1*	*RPS14, SYNAPO*	*NDST1* participates in synthesis of heparan sulphate. K/o mice have cardiovascular abnormalities including aortic malformations.
rs62481707	7	158132993	G	A	0.048	0.273	4.72×10^{-6}	*PTPRN2*		Major autoantigen for insulin dependent diabetes mellitus; may be involved in insulin secretion.
rs17810905	11	36105256	G	A	0.125	-0.176	8.43×10^{-6}	*LDLRAD3*		Low density lipoprotein receptor which plays a role in amyloid precursor protein trafficking.
rs10484197	14	47419743	T	C	0.023	-0.393	3.17×10^{-6}	*MDGA2*		Mainly role in central nervous system. No clear links with cardiovascular phenotype
rs7230451	18	22069393	G	T	0.331	-0.128	1.62×10^{-6}	*HRH4*	*IMPACT*	*HRH4* histamine receptor with links to asthma and autoimmune conditions. No evidence for cardiovascular role.

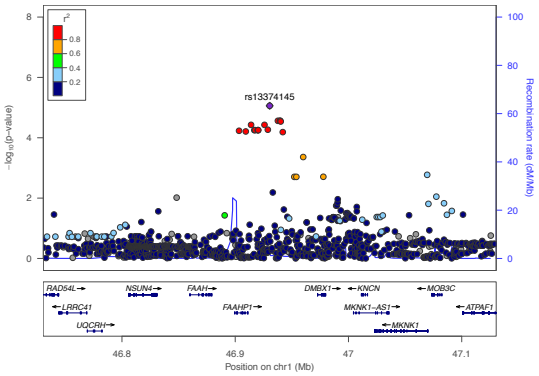
Chr: chromosome; Ref: Reference allele; Alt: effect allele; MAF minor allele frequency; Beta: effect size (change in diameter (mm) per copy of effect allele); Comments: annotations from multiple sources including GWAS catalog[36], GeneCards[369], Mouse Genome Informatics database[370], GTEx[265] and PubMed searches for gene function information; ZP4: Zona Pellucida Glycoprotein 4; RYR2: Ryanodine Receptor 2; FAAH: Fatty Acid Amide Hydrolase; DMBX1: Diencephalon/Mesencephalon Homeobox 1; GC: GC, Vitamin D Binding Protein; NDST1: N-Deacetylase And N-Sulfotransferase 1; RPS14: Ribosomal Protein S14; SYNAPO: Synaptopodin; PTPRN2: Protein Tyrosine Phosphatase, Receptor Type N2; LDLRAD3: Low Density Lipoprotein Receptor Class A Domain Containing 3; MDGA2: MAM Domain Containing Glycosylphosphatidylinositol Anchor 2; HRH4: Histamine Receptor H4; IMPACT: Impact RWD Domain Protein

FIGURE 14.15: LocusZoom plots showing loci associated with ascending aortic distensibility (rank-normalised phenotype)

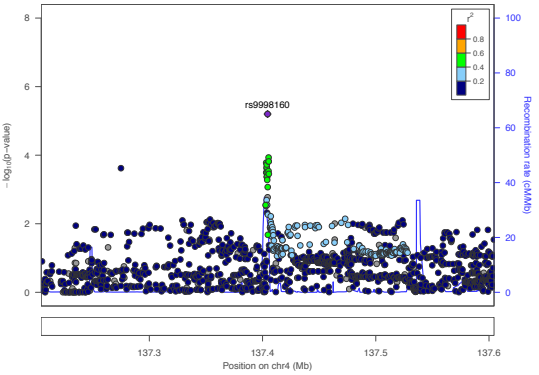
a. Chromosome 1: RYR2



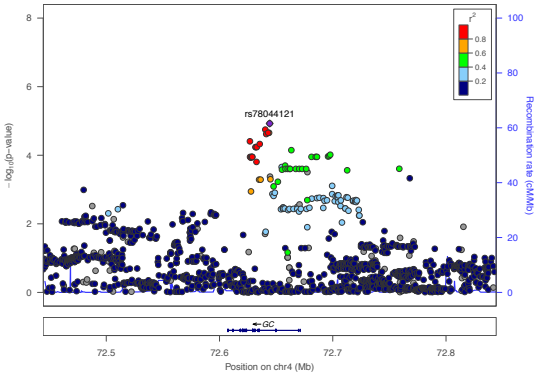
b. Chromosome 1: FAAH / DMBX1



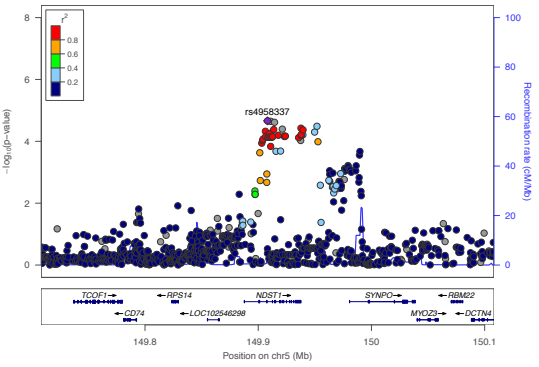
c. Chromosome 4: Unknown



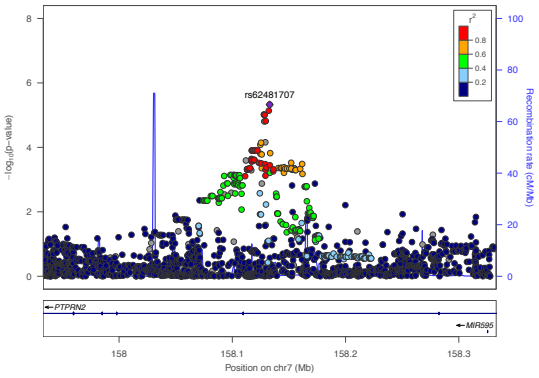
d. Chromosome 4: GC



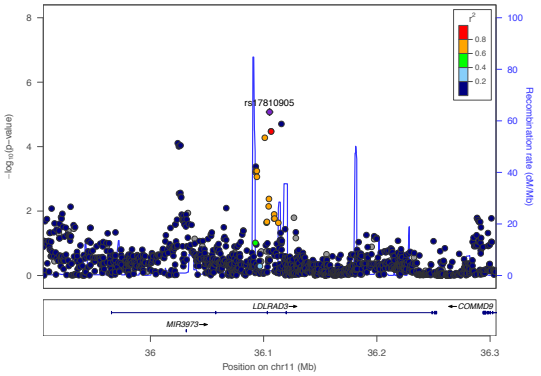
e. Chromosome 5: NDST1



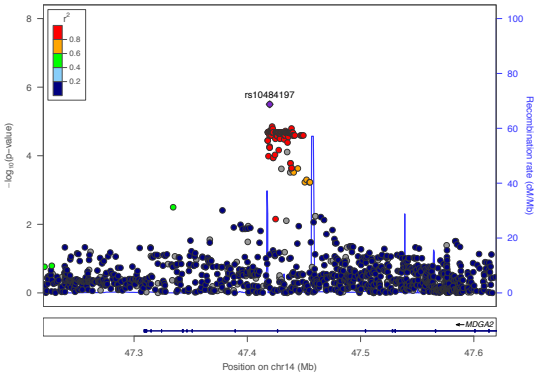
f. Chromosome 7: PTPRN2



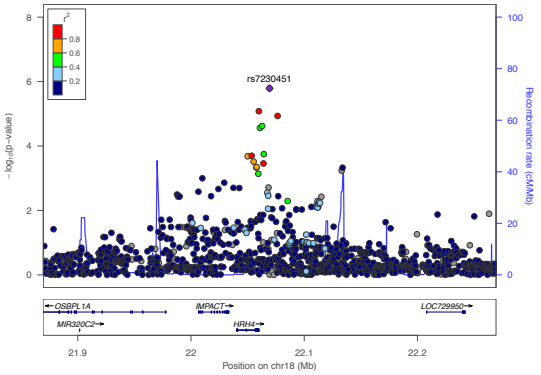
g. Chromosome 11: LDLRAD3



h. Chromosome 14: MDGA2



i. Chromosome 18: IMPACT



Suggestive loci are shown as examples above in LocusZoom plots. These plots are essentially a "zoomed in" version of a Manhattan plot, but also demonstrate the linkage disequilibrium structure of the association peak. They demonstrate the spatial relation of the associated SNPs with annotated genes, and with nearby SNPs in linkage disequilibrium (LD) with the lead variant. The plots selected here demonstrate the expected structure of a "true positive" association, with SNPs in LD with the lead variant also associated with phenotype, albeit at lower significance. The labels a-i denote the chromosomal location of the loci, and the top possibilities for candidate genes at the locus, based on biological plausibility and functional evidence, detailed in Table 14.22 and below.

The chromosome 1 locus is within the regulatory region of *ZP4 (Zona Pellucida 4)*, a gene involved in very early development. However, it is also within 400kbp of *RYR2 (Ryanodine Receptor 2)*. This encodes a calcium channel which regulates intracellular calcium flux. Variants in *RYR2* are associated with arrhythmogenic right ventricular cardiomyopathy (ARVC) and other cardiac arrhythmias. *RYR2* also seems important in LRP1 (Low Density Lipoprotein Receptor Related Protein 1) control of vessel wall structure and function[430], and there is an established key role of calcium homeostasis in aortic function.

Another locus at chromosome 4 tags *GC (GC vitamin D binding protein)*, whose protein product binds vitamin D and targets it to tissues. It is a key regulator of vitamin D levels, which

have been linked to many aspects of cardiovascular health, including blood pressure indices and atherosclerosis risk, as well as endothelial dysfunction. Vitamin D also is a key determinant of intracellular calcium levels, again in keeping with the established role of calcium homeostasis in aortic function.

The chromosome 5 locus tags *NDST1 (N-Deacetylase And N-Sulfotransferase 1)*, a key component of heparan sulphate biosynthesis. The top SNP with data available (rs4958337) is intronic and is a significant eQTL for *NDST1* in left ventricular tissue. Heparan sulphate is a crucial cofactor for many signalling pathways. Smooth muscle cell-specific knockout of *NDST1* has been shown to affect aortic stiffness *ex vivo*, and this was associated with alterations in myosin and actin isoforms expressed in the aorta and in isolated aortic vascular smooth muscle cells[431].Knockout mice exhibit cardiovascular malformations such as interrupted aortic arch.

Other loci here have little evidence for direct effects on aortic function, but might regulate cardiovascular risk factors. For example, *PTPRN2* in which the association peak is intronic, regulates insulin secretion, and *FAAH* is associated with body mass index.

14.5.4.4 Ascending aortic distensibility: Gene-based analysis

This analysis interestingly produced results with greater overlap between the two statistical approaches (log-transformation versus rank-normalisation). This is possibly because the amalgamation of information across multiple different SNPs reduces the probability of false positive associations arising from spurious single SNP results. The top 15 genes from each analysis are shown below, with the 3 genes common to both “top 15” lists highlighted in red.

TABLE 14.22: Gene associations with rank-normalised ascending aortic distensibility

Gene	Chr	Start	End	No. SNPs	Chisq (obs)	P value	Top SNP p value	Top SNP
AGTR1	**3**	**148415657**	**148460790**	**160**	**515.437**	**0.0001**	**$9.17x10^{-5}$**	**rs35773629**
C1ORF140	1	221503269	221509638	83	277.341	0.0001	$8.93\ x10^{-6}$	rs2996005
PRTFDC1	**10**	**25137535**	**25241573**	**258**	**632.54**	**0.0002**	**0.0003**	**rs533598728**
NDST1	5	149887673	149937773	166	459.489	0.0002	$3.67\ x10^{-6}$	rs398050845
DNMT3B	20	31350190	31397162	157	529.25	0.0004	$5.06\ x10^{-5}$	rs6058892
IMPACT	18	22006608	22033494	171	442.381	0.0005	$1.62\ x10^{-6}$	rs7230451
HRH4	18	22040592	22059921	183	490.657	0.0005	$1.62\ x10^{-6}$	rs7230451
DMRTC2	19	42349085	42356397	65	235.21	0.0006	0.0003	rs7251154
LYPD4	19	42341147	42348736	62	235.292	0.0006	0.0003	rs7251154
ZNF404	19	44376514	44384288	80	254.309	0.0007	$1.97\ x10^{-5}$	rs7248807
HEXIM1	17	43224683	43229468	159	450.333	0.0008	$3.38\ x10^{-5}$	rs61160218
TUBB6	18	12308256	12326568	159	434.556	0.0009	$6.60\ x10^{-5}$	rs62098961
COPG1	3	128968452	128996616	85	284.324	0.0009	0.0001	rs58063531
ACTC1	**15**	**35080296**	**35087927**	**227**	**460.956**	**0.0010**	**$8.07\ x10^{-5}$**	**rs752876**
HEXIM2	17	43238263	43247406	182	470.20	0.0012	3.38E-05	rs61160218

TABLE 14.23: Gene associations with log-transformed ascending aortic distensibility

Gene	Chr	Start	End	No.SNPs	Chisq(obs)	P value	Top SNP p value	Top SNP
AGTR1	**3**	**148415657**	**148460790**	**160**	**569.65**	**$2.93x10^{-5}$**	**$1.19x10^{-5}$**	**rs35773629**
ACTC1	**15**	**35080296**	**35087927**	**227**	**580.359**	**$2.98x10^{-5}$**	**$2.13x10^{-6}$**	**rs35706982**
GJD2	15	35044641	35046782	218	571.141	$5.57x10^{-5}$	$2.13x10^{-6}$	rs35706982
PRTFDC1	**10**	**25137535**	**25241573**	**258**	**621.041**	**0.0003**	**0.0001**	**rs533598728**
GTF2H2C	5	68856050	68888729	40	189.521	0.0004	0.0005	rs34962033
PYROXD2	10	100143321	100174978	191	537.465	0.0006	$4.01x10^{-6}$	rs10748726
THNSL1	10	25305507	25315593	91	320.505	0.0007	$2.09x10^{-6}$	rs6482466
SLC4A11	20	3208062	3219887	182	521.198	0.0008	$5.30x10^{-5}$	rs6084294
CLEC1A	12	10223079	10251605	162	442.051	0.0009	0.0012	rs3901532
DDRGK1	20	3171011	3185295	218	590.32	0.0009	$5.30x10^{-5}$	rs6084294
RWDD4	4	184560788	184580331	309	623.09	0.0009	$1.13x10^{-5}$	rs77219531
ITPA	20	3189513	3204516	212	587.68	0.0009	$5.30x10^{-5}$	rs6084294
CLEC7A	12	10269375	10282868	137	390.87	0.0011	0.0012	rs3901532
WT1	11	32409321	32457081	211	690.32	0.0011	$5.60x10^{-6}$	rs72907578
SLC22A12	11	64358281	64369825	84	rs11231813	rs502567	225.307	0.00112743

(see next page for legend)

Tables 14.22 and 14.23 legend: Top 15 protein-coding genes associated with ascending aortic distensibility. Chr: chromosome; Gene start/end: left and right-side boundaries of the gene (GRCh37); No.SNPs: number of SNPs occurring within gene contributing to the analysis; Chi-sq: sum of chi-squared test statistics of all SNPs within the gene region; P value: segment-based test p value; P value most sig SNP: smallest single-SNP p value in the segment; AGTR1: Angiotensin II Receptor Type I; C1ORF140: Chromosome 1 Open Reading Frame 140; PRTFDC1: Phosphoribosyl Transferase Domain Containing 1; NDST1: N-Deacetylase And N-Sulfotransferase 1; DNMT3B: DNA Methyltransferase 3 Beta; IMPACT: Impact RWD Domain Protein; HRH4: Histamine Receptor H4; DMRTC2: DMRT Like Family C2; LYPD4: LY6/PLAUR Domain Containing 4; ZNF404: Zinc Finger Protein 404; HEXIM1: Hexamethylene Bisacetamide Inducible 1; TUBB6: Tubulin Beta 6 Class V; COPG1: Coatomer Protein Complex Subunit Gamma 1; ACTC1: Actin, Alpha, Cardiac Muscle 1; HEXIM2: Hexamethylene Bisacetamide Inducible 2; GJD2: Gap Junction Protein Delta 2; GTF2H2C: General Transcription Factor IIH Subunit 2 Family Member C; PYROXD2: Pyridine Nucleotide-Disulphide Oxidoreductase Domain 2; THNSL2: Threonine Synthase Like 2; SLC4A11: Solute Carrier Family 4 Member 11; CLEC1A: C-Type Lectin Domain Family 1 Member A; DDRGK1: DDRGK Domain Containing 1; RWDD4: RWD Domain Containing 4; ITPA: Inosine Triphosphatase; CLEC7A: : C-Type Lectin Domain Family 7 Member A; WT1: Wilms Tumor 1; SLC22A12: Solute Carrier Family 22 Member 12.

Loci or genes which appear significant in both models are likely to be more robust biological signals than those which are significant solely in the log-transformed model. Whilst none of these gene-based signals remain significant after Bonferroni correction for multiple testing, the 3 genes which appear in the top 15 most highly associated genes with each model, represent biologically plausible and very interesting candidates for mediation of aortic distensibility. The fact that their association is robust to the model specifications makes them perhaps more likely to represent genuine biological associations.

The first of these genes is *AGTR1 – angiotensin II receptor type 1*. This is the most significant gene in both models. Angiotensin II is a peptide hormone which is a very potent regulator of blood pressure. It has pressor activity, causing vasoconstriction and volume retention. Its major cardiovascular effects are mediated through the type I receptor, encoded by this gene. Common anti-hypertensive medications (angiotensin II receptor inhibitors such as losartan) target this receptor. This current analysis suggests that it may additionally play a role in aortic elastic function. To what extent this effect is mediated through its impact on blood pressure is uncertain – it may be that this is its primary effect. However, previous research has also demonstrated that angiotensin II is a key component of TGF-ß signalling pathways – pathways which play a crucial role in aortic homeostasis and function. It is the TGF-ß-signalling pathway which is disrupted in severe aortopathies such as Loeys-Dietz syndrome and Marfan syndrome, and non-canonical TGF-ß signalling, possibly stimulated in part by angiotensin II, is a crucial component of their pathogenesis.

If we conclude from this study that *AGTR1* common variants have an impact on aortic distensibility, then this would support investigating pharmacological manipulation of this signalling pathway in aortic aneurysm. This is not a novel idea: the impact of angiotensin II on the TGF-ß signalling pathway stimulated interest in the use of angiotensin II receptor blockers (AIIRBs) such as losartan in Marfan syndrome. Unfortunately, these clinical trials have been less positive than hoped[150-152, 432]. It would be interesting to investigate the role of common variation in the *AGTR1* gene on aortic prognosis in Marfan, Loeys-Dietz and related syndromes, and thereafter to consider targeting AIIRBs at patients with a specific genetic profile.

The second gene which appears associated with ascending aortic distensibility in both models is *ACTC1 – cardiac muscle alpha actin*. Rare variants in this gene are recognised causes of hypertrophic cardiomyopathy, characterised particularly by high levels of cardiac fibrosis[430, 433, 434]. Variants are also implicated in familial atrial septal defect and other congenital heart disease[435]. Due to the common developmental origin of the heart and proximal portions of the thoracic aorta, it is again tempting to speculate that subtle variations in expression of this gene might lead to different mechanical properties of the proximal aorta and regulate the elastic function. Again, there are therapeutic implications here – ß-blockers such as bisoprolol have been shown to improve cardiac prognosis in patients with *ACTC1*-related dilated cardiomyopathy; these common medications might also serve to improve aortic stiffness in patients carrying SNPs affecting *ACTC1* expression or function.

14.5.4.5 Ascending aortic distensibility: Pathways analysis

Table 14.24 below shows the top 5 associated KEGG and GO biological processes derived from the genome-wide association results.

TABLE 14.24: Top 5 KEGG pathways and top 5 Gene Ontology biological processes associated with ascending aortic distensibility.

KEGG Pathway	FDR-corrected p value
KEGG: FOCAL ADHESION	0.0062
KEGG: PROXIMAL TUBULE BICARBONATE RECLAMATION	0.0070
KEGG: ECM RECEPTOR INTERACTION	0.0116
KEGG: NUCLEOTIDE EXCISION REPAIR	0.0122
KEGG: PYRIMIDINE METABOLISM	0.0130

GO pathway: Biological Process	FDR-corrected p value
GO: POSITIVE REGULATION OF NUCLEOBASENUCLEOSIDENUCLEOTIDE AND NUCLEIC ACID METABOLIC PROCESS	0.0030
GO: RESPONSE TO HYPOXIA	0.0030
GO: POSITIVE REGULATION OF TRANSCRIPTION	0.0035
GO: REGULATION OF DNA BINDING	0.0043
GO: POSITIVE REGULATION OF TRANSCRIPTION DNA DEPENDENT	0.0046

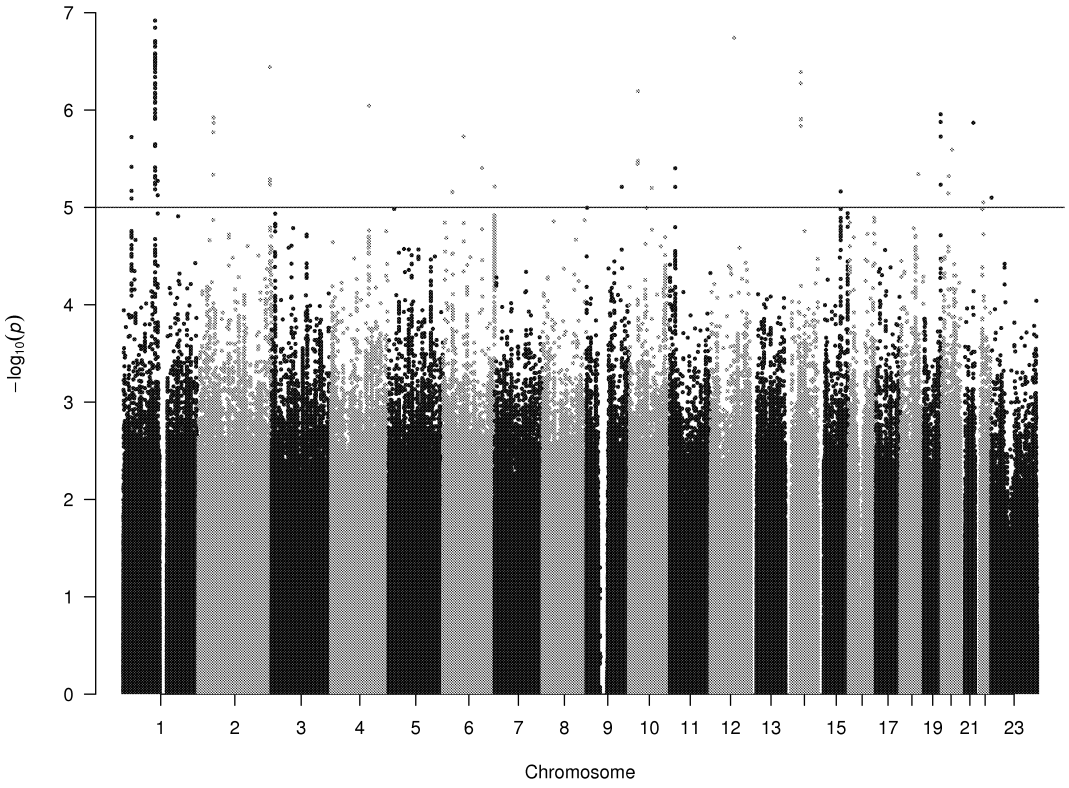
Significantly associated KEGG pathways again suggest a key role for adhesion– with both focal adhesion and extracellular matrix receptor interaction terms being identified. The term proximal tubule bicarbonate reclamation recurs here – this includes genes such as aquaporin 1 and carbonic anhydrase. It also includes many glutamate – signalling related genes. Glutamate metabolism is highly dependent on bicarbonate homeostasis, so this pathway association might also support the role of glutamatergic signalling in aortic homeostasis.

14.5.4.6 Descending aortic distensibility: single SNP analysis

This phenotype has been less extensively studied than ascending aortic distensibility, in part because the descending aorta is more difficult to image using techniques such as echocardiography. The heritability analysis presented in section 14.5.4.2 suggests that a limited proportion of variance is explained by genetic factors. However, it is important to note firstly that this estimate of heritability was non-significant, and secondly, the overall variance of the trait is lower than for ascending aortic distensibility so one would expect common variants to have a smaller absolute effect size. This would require greater power than this study provides.

As expected, no single SNPs reached the genome-wide significance threshold ($p<5x10^{-8}$). Nevertheless, there were some robust-looking association signals at suggestive levels of significance ($p<1x10^{-5}$). These results are presented in the Manhattan plots and tables below.

FIGURE 14.16: Manhattan plot of association with descending aortic distensibility



The y axis displays -log_{10} p values of association – i.e. the "higher" the point on the graph, the more significant the association. The x axis is the position on each chromosome. Each dot represents a single SNP; however, these often overlie one another. The horizontal threshold line represents a widely-accepted yet arbitrary "suggestive significance" level threshold of p = $1x10^{-5}$.

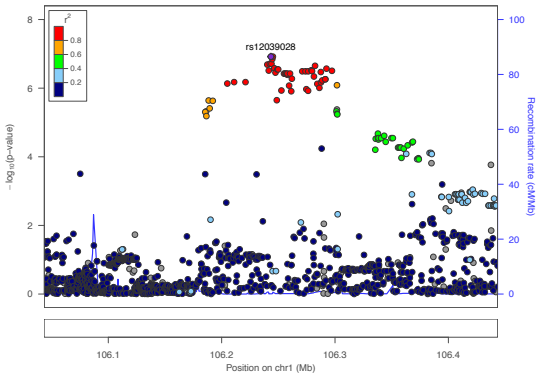
TABLE 14.25: Single SNP associations with descending aortic distensibility

dbSNP	Chr	Position	Ref	Alt	MAF	Beta	p-value	Closest protein-coding gene	Other nearby candidate genes	Comments
rs12039028	1	106243448	T	G	0.078	-0.312	1.21×10^{-7}	None within 500kb	*LINC01676* *LINC01677*	Unknown
rs397862470	1	28023222	GT	G	0.025	0.493	3.83×10^{-6}	*IFI6*	*FAM76A* *FGR*	No obvious links to cardiovascular
rs11894280	2	238509759	A	G	0.076	-0.312	3.61×10^{-7}	*LRRFIP*	*RAB17*	*LRRFIP* Involved in PDGFA signalling and others; may control smooth muscle proliferation.
rs140284601	2	50554099	A	G	0.013	0.698	1.19×10^{-6}	*NRXN1*		Primarily synaptic. Cardiovascular-neuronal cross-over?
rs2072917	6	170891856	A	C	0.491	0.147	6.10×10^{-6}	*PDCD2*	*PSMB1* *TBP*	*PSMB*1 expression is increased in pulmonary artery after exposure to chronic hypoxia.
rs9461633	6	30761168	G	A	0.198	-0.188	6.89×10^{-6}	*IER3* (top SNP eQTL)	*FLOT1*	*IER3* k/o mice have hypertension, less vasodilatation and LV hypertrophy
rs74413603	10	29102966	T	C	0.037	-0.437	6.40×10^{-7}	*C10orf126*	*LINC01517* *BAMBI*	*BAMBI* key regulator of TGF-ß
rs771664005	10	75493812	A	AC	0.011	-0.756	6.32×10^{-6}	*VCL*		Cell adhesion protein
rs2526551	11	17736848	A	C	0.332	-0.160	3.96×10^{-6}	*MYOD1*		*MYOD1* regulates fibroblast differentiation and controls *ACTC1* expression.
rs2039792	14	51335897	A	C	0.410	-0.167	4.06×10^{-7}	*ABHD12B*	*NIN, PYGL*	*NIN* required for angiogenesis and endothelial tube formation
rs28572423	15	76057173	C	G	0.166	0.211	6.85×10^{-6}	*UBE2Q2*	*CSPG4*	*CSPG4* encodes proteoglycan - role in microvascular development.
rs34837414	19	56052008	C	T	0.457	0.166	1.10×10^{-6}	*SBK2*	*SBK3,* *ZNF579*	*SBK2* highly expressed in heart; little known

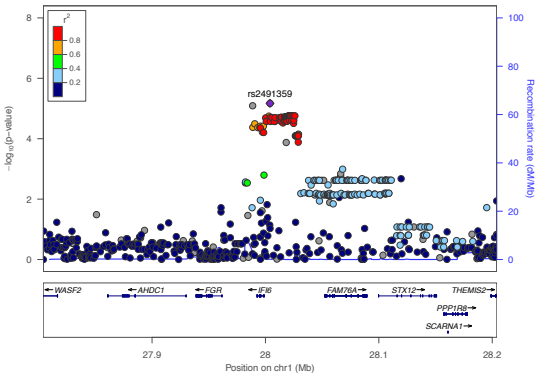
Chr: chromosome; Ref: Reference allele; Alt: effect allele; MAF minor allele frequency; Beta: effect size (change in diameter (mm) per copy of effect allele); Comments: annotations from multiple sources including GWAS catalog[36], GeneCards[369], Mouse Genome Informatics database[370], GTEx[265] and PubMed searches for gene function information; LINC01676: Long non-coding RNA 1676; IFI6: Interferon Alpha Inducible Protein 6; FAM76A: Family With Sequence Similarity 76 Member A; FGR: FGR Proto-Oncogene, Src Family Tyrosine Kinase; LRRFIP1: LRR Binding FLII Interacting Protein 1; RAB17: RAB17, Member RAS Oncogene Family; NRXN1: Neurexin1; PDCD2: Programmed Cell Death 2; PSMB1: Proteasome Subunit Beta 1; TBP: TATA-Box Binding Protein; IER3: Immediate Early Response 3; FLOT1: Flotillin 1; C10orf126: Chromosome 10 open reading frame 126; BAMBI: BMP And Activin Membrane Bound Inhibitor; VCL: Vinculin; MYOD1: Myogenic Differentiation 1; ABHD12B: Abhydrolase Domain Containing 12B; NIN: Ninein; UBE2Q2: Ubiquitin Conjugating Enzyme E2 Q2; CSPG4: Chondroitin Sulfate Proteoglycan 4; SBK2: SH3 Domain Binding Kinase Family Member 2; SBK3: SH3 Domain Binding Kinase Family Member 3; ZNF579: Zinc Finger Protein 579

FIGURE 14.17: LocusZoom plots showing loci associated with descending aortic distensibility (rank-normalised phenotype)

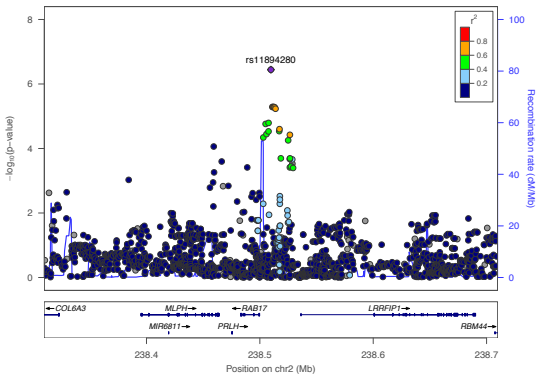
a. Chromosome 1: Unknown



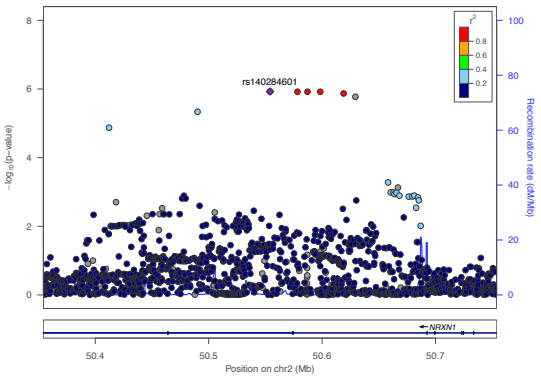
b. Chromosome 1: ?IFI6



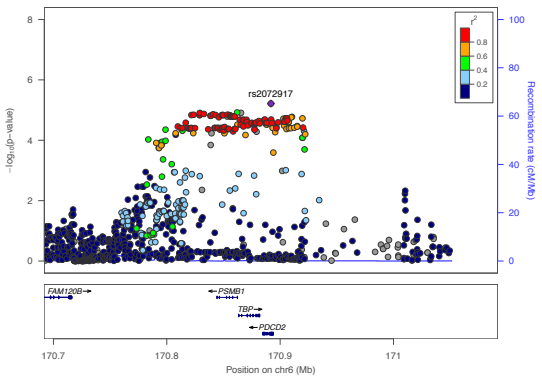
c. Chromosome 2: LRRFIP1



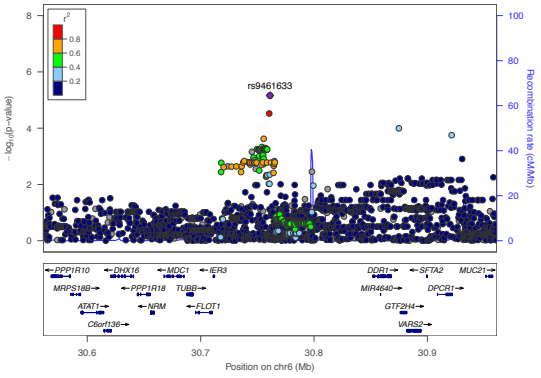
d. Chromosome 2: NRXN1



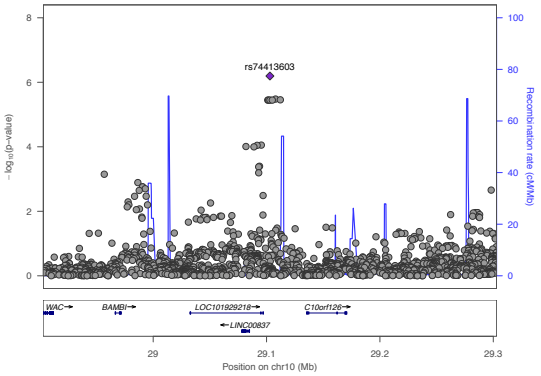
e. Chromosome 6: PSMB1?



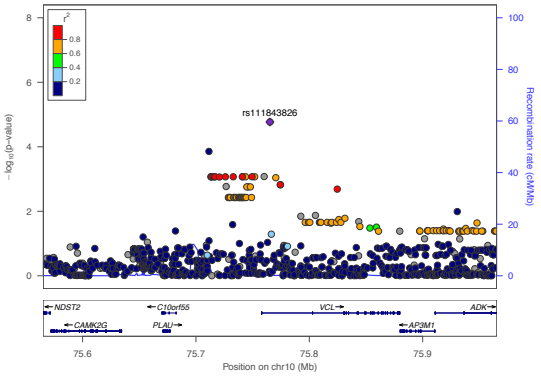
f. Chromosome 6: IER3



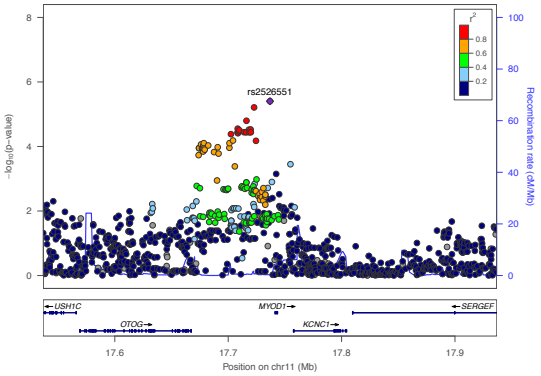
g. Chromosome 10: BAMBI?



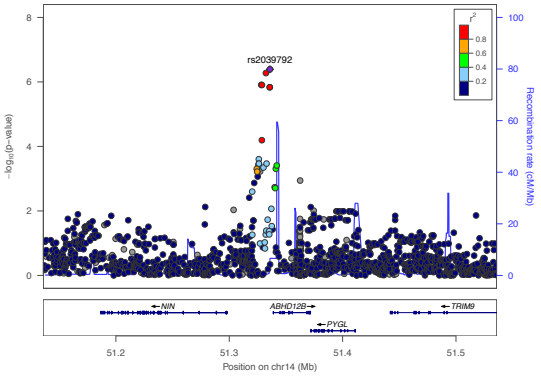
h. Chromosome 10:VCL



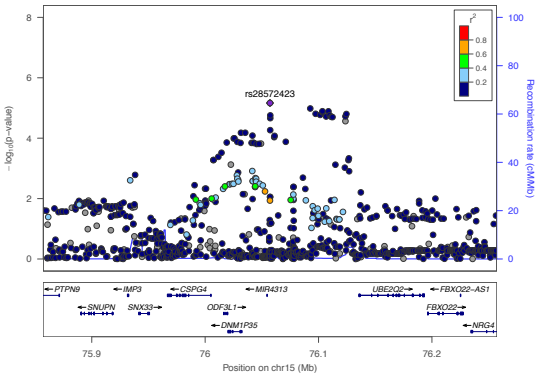
i. Chromosome11: MYOD1



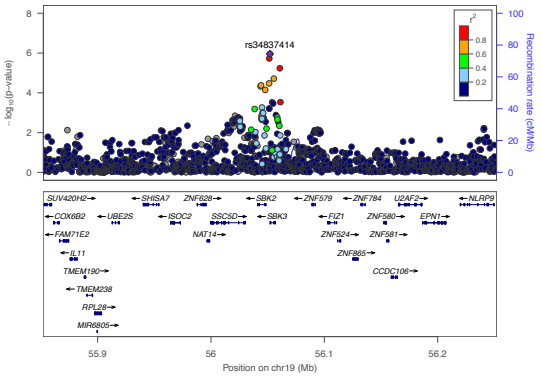
j. Chromosome 14: NIN



k. Chromosome 15:CSPG4



l. Chromosome 19: SBK2



Suggestive loci are shown as examples above in LocusZoom plots. These plots are essentially a "zoomed in" version of a Manhattan plot, but also demonstrate the linkage disequilibrium structure of the association peak. They demonstrate the spatial relation of the associated SNPs with annotated genes, and with nearby SNPs in linkage disequilibrium (LD) with the lead variant. The plots selected here demonstrate the expected structure of a "true positive" association, with SNPs in LD with the lead variant also associated with phenotype, albeit at lower significance. The labels a-l denote the chromosomal location of the loci, and the top possibilities for candidate genes at the locus, based on biological plausibility and functional evidence, detailed in Table 14.26 and below.

Perhaps the most interesting associations here are those at chromosomes 10 and 11. The first association peak at chromosome 10 tags *VCL (Vinculin 1),* a key mechanosensor which induces cytoskeletal remodelling in response to external stimuli. The top SNP is <5kb 5' of the transcription start site, and this region has enhancer histone marks in aorta. *VCL* is expressed both at cell-cell and cell-extracellular matrix interfaces ubiquitously, and particularly in large vessels such as the aorta. A recent study even suggested that targeting these interfaces, including the talin-vinculin interaction, could reduce ageing-induced stiffness[436]. This makes *VCL* a key gene of interest in the current analysis – both for its mechanistic interest and also as a potential therapeutic target.

Close to the other chromosome 10 peak, lies *BAMBI (BMP And Activin Membrane Bound Inhibitor)*. This gene encodes a TGF-ß receptor 1 antagonist which regulates TGFß1 signalling in vascular endothelial cells. *BAMBI*-knockout mice exhibit increased angiogenesis and increased TGFß1 signalling, without overt phenotype[437]. This is of potential therapeutic interest, and is clearly closely connected to the TGF-ß signalling pathway which is a known key driver of aortic phenotypes.

At chromosome 11,the top SNP tags *MYOD1 (Myogenic Differentiation 1)*. This encodes a regulator of fibroblast differentiation into myofibroblasts, and of *cardiac alpha actin* expression[438] (*ACTC1*; found above to be associated with ascending aortic distensibility at suggestive significance).

There is a very prominent peak at chromosome 1 associated with descending aortic distensibility. This is in a relative "gene desert", with the nearest protein-coding gene, *PRMT6*, is a whopping 1.5Mb away. There are 2 long non-coding RNAs in close proximity to the peak (*LINC01676* and *LINC01677*), but these have no known cardiovascular associations. There are some promoter and enhancer marks at this locus in cultured cells derived from embryonic stem cells[439], but no clear evidence to support regulatory function of this locus in a setting which might influence aortic distensibility.

The other chromosome 1 locus conversely is within a cluster of genes with diverse biological roles. None however stand out as possible mediators of an effect on aortic phenotypes.

A chromosome 2 association peak is intronic in *NRXN1, (Neurexin 1).* This gene encodes a synaptic protein which is also expressed in endothelial cells. There has been some interest in

the neurexin family as mediators of a neuronal-vascular "crossover" phenotype, affecting both systems. Neurexins are expressed at excitatory glutamatergic synapses and inhibitory GABA-ergic synapses, and differently spliced isoforms are also present in vascular cells – both endothelial and smooth muscle. At synapses, they may be involved in differentiation of the synapse, and in the vascular system, they have been shown to mediate angiogenesis[440]. The GABA and glutamate systems were identified above as potential contributors to aortic dimensions via central control of blood pressure and haemodynamic reflexes; neurexins could help to mediate the balance between these two opposing systems.

At chromosome 6, the top SNP is an eQTL for *IER3 (Immediate Early Response 3)* in whole blood and at sub-genome-wide significance in aorta. Knockout of this gene in mice leads to hypertension, reduced vasodilatation and LV hypertrophy, with decreased cardiac muscle contractility, indicating a potential role in aortic traits.

At the chromosome 12 locus, *NIN (ninein)* is another possible candidate. The product of this gene, ninein, is a centrosomal microtubule-anchoring protein which is highly expressed in vascular endothelium and is critical for vascular tube formation during angiogenesis[441].

At chromosome 19, there is a cluster of genes near the association peak. The top SNP is a strong eQTL for *SBK2* in adipose tissue; this gene is highly expressed in the heart, but there is little published about its specific role.

Other loci do not tag obvious candidates for mediation of phenotypic effect.

14.5.4.7 Descending aortic distensibility: gene-based analysis

TABLE 14.26: Gene-based associations with ascending aortic distensibility

Gene	Chr	Start	End	No. SNPs	Chisq (obs)	P value	Top SNP p value	Top SNP
PDCD2	6	170884659	170893780	63	354.6	$1.45x10^{-5}$	$6.10x10^{-6}$	rs2072917
TBP	6	170863420	170881958	70	367.7	$2.11x10^{-5}$	$6.10x10^{-6}$	rs2072917
PSMB1	6	170844203	170862417	79	375.1	$4.30x10^{-5}$	$6.10x10^{-6}$	rs2072917
MYOD1	11	17741109	17743678	151	498.3	$6.15x10^{-5}$	$3.96x10^{-6}$	rs2526551
PCDHB1	5	140430960	140433547	68	285.7	0.0001	$3.43x10^{-5}$	rs28276
PCDHB2	5	140474190	140476964	78	315.4	0.0001	$3.43x10^{-5}$	rs28276
IFI6	1	27992571	27998724	103	275.5	0.0002	$1.90x10^{-6}$	rs796849172
PCDHB3	5	140479829	140483406	84	307.6	0.0002	$3.43x10^{-5}$	rs28276
SBK3	19	56052022	56056909	320	806.5	0.0003	$1.10x10^{-6}$	rs34837414
SBK2	19	56041099	56048435	338	835.7	0.0003	$1.10x10^{-6}$	rs34837414

Top 15 protein-coding genes associated with descending aortic distensibility. Chr: chromosome; Gene start/end: left and right-side boundaries of the gene (GRCh37); No.SNPs: number of SNPs occurring within gene contributing to the analysis; Chi-sq: sum of chi-squared test statistics of all SNPs within the gene region; P value: segment-based test p value; P value most sig SNP: smallest single-SNP p value in the segment; PDCD2: Programmed Cell Death 2; TBP: TATA-Box Binding Protein; PSMB1: Proteasome Subunit Beta 1; MYOD1: Myogenic Differentiation 1; PCDHB1:Protocadherin beta 1; PCDHB2:Protocadherin beta 2; IFI6: Interferon Alpha Inducible Protein 6; PCDHB3:Protocadherin beta 3; SBK3: SH3 Domain Binding Kinase Family Member 3; SBK2: : SH3 Domain Binding Kinase Family Member 2

In the case of descending aortic distensibility, few new loci were identified by gene-based analysis. The beta-cadherin cluster on chromosome 5 was the only additional association. This regulates specific cell-cell adhesion, and may play a role in neural development, but does not have a well-defined vascular role. None of the additional loci had compelling links with cardiovascular phenotype. None remained significantly associated with descending aortic distensibility after multiple testing correction.

14.5.4.8 Descending aortic distensibility: pathways analysis

Table 14.27 below shows the top 5 associated KEGG and GO biological processes derived from the genome-wide association results.

TABLE 14.27: Top 5 KEGG pathways and top 5 Gene Ontology biological processes associated with descending aortic distensibility.

KEGG Pathway	FDR-corrected p value
KEGG: ANTIGEN PROCESSING AND PRESENTATION	0.0083
KEGG: DORSO VENTRAL AXIS FORMATION	0.0123
KEGG: FOCAL ADHESION	0.0150
KEGG: LINOLEIC ACID METABOLISM	0.0154
KEGG: INTESTINAL IMMUNE NETWORK FOR IGA PRODUCTION	0.0193

GO pathway: Biological Process	FDR-corrected p value
GO: NEGATIVE REGULATION OF TRANSCRIPTION FROM RNA POLYMERASE II PROMOTER	0.0030
GO: NEGATIVE REGULATION OF APOPTOSIS	0.0074
GO: NEGATIVE REGULATION OF RNA METABOLIC PROCESS	0.0077
GO: INFLAMMATORY RESPONSE	0.0080
GO: SYNAPSE ORGANIZATION AND BIOGENESIS	0.0084

These pathways associations are interesting for descending aortic distensibility – they perhaps suggest that immune / inflammatory mechanisms might play a more important role here than in the proximal aorta, with antigen processing and presentation identified as the top pathway. The relevance of association with linoleic acid metabolism is not immediately clear.

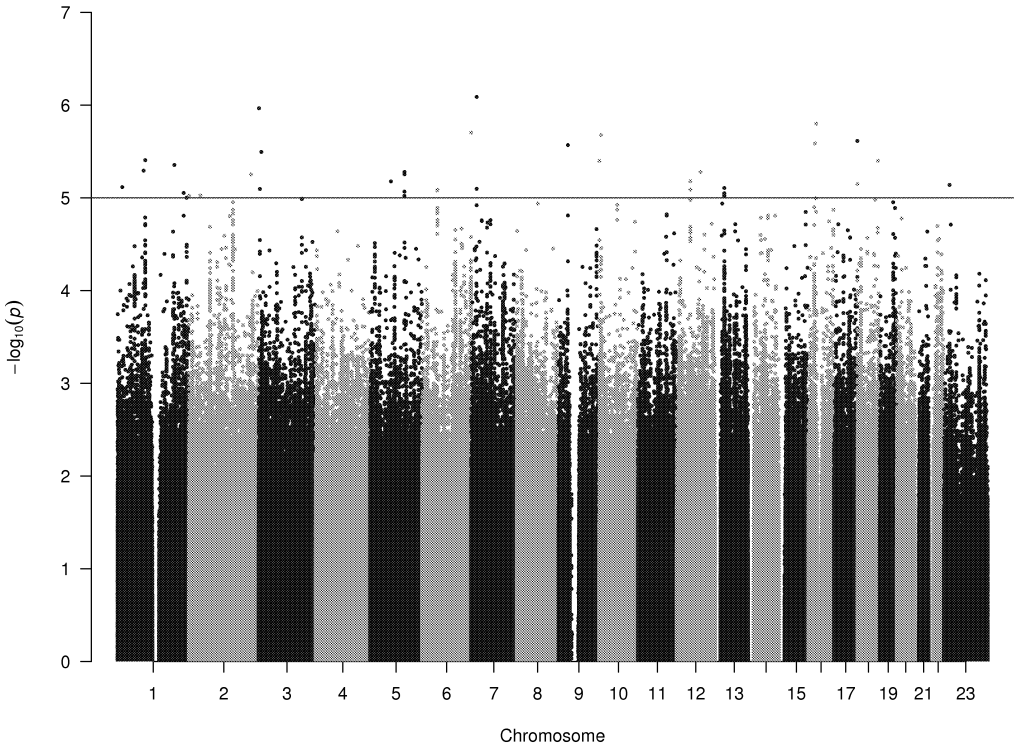
Again, however, we see the importance of cell adhesion and developmental pathways in determining aortic traits. The final GO pathway – synapse organization and biogenesis – also reminds us of the key interaction between the vascular and nervous system.

14.5.4.9 Aortic arch pulse wave velocity: single-SNP association analysis

The final aortic function phenotype was aortic arch pulse wave velocity (PWV). As with the other phenotypes, no single-SNP associations reached genome-wide significance. Again, there were multiple suggestive signals of association, detailed below. The association data for PWV was, in general, "noisier" than the other associations. This is to be expected – it is a measure with many components – measuring flows, transit time and path length, each of which will introduce some error into the final result. mPWV also integrates properties over the whole aortic arch and so, whilst it is much more site-specific than cfPWV, it still will be influenced by many other factors such as blood pressure, perhaps blood viscosity, heart rate and so on. Whilst we can correct for this in some ways with our regression analysis, the addition of multiple additional covariates significantly reduces statistical power, as each measured covariate adds further error to the model. The final consideration is that, once more, we needed to rank normalise the phenotype to construct a valid regression model, again losing the power to discriminate between fine and large differences in PWV. These factors all meant that we had a low and non-significant estimate of the heritability of this measure.

However, it would be wrong to dismiss the results entirely – amongst all the noise are some interesting association signals – presented in figure 14.18 and 14.19 and table 14.28 below and, if validated in independent cohorts, these could yield interesting insights into aortic pulse wave velocity. I have presented all the suggestive associations in the tables below, and picked out some potentially interesting candidates for discussion.

FIGURE 14.18: Manhattan plot of association with pulse wave velocity



The y axis displays -log$_{10}$ p values of association – i.e. the "higher" the point on the graph, the more significant the association. The x axis is the position on each chromosome. Each dot represents a single SNP; however, these often overlie one another. The horizontal threshold line represents a widely-accepted yet arbitrary "suggestive significance" level threshold of p = 1x10^{-5}.

TABLE 14.28: Single SNP associations with aortic pulse wave velocity

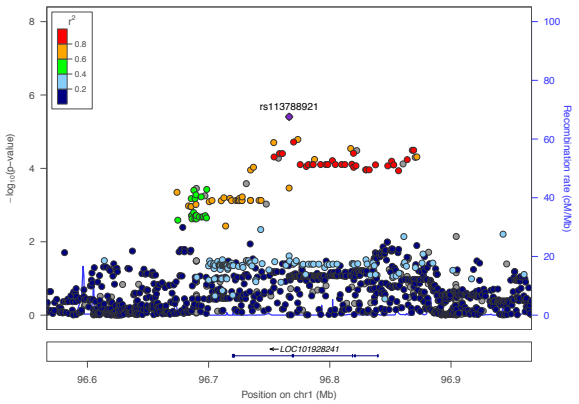
dbSNP	Chr	Position	Ref	Alt	MAF	Beta	p-value	Closest protein-coding gene	Other nearby candidate genes	Comments
rs113788921	1	96766484	C	T	0.137	-0.179	3.92×10^{-6}		*PTBP2*	No clear cardiovascular links
rs61829680	1	198567389	T	C	0.029	0.367	4.42×10^{-6}	*PTPRC*	*ATP6V1G3*	No clear cardiovascular links
rs112131264	1	241270025	C	A	0.120	0.198	9.97×10^{-6}	*RGS7*		Regulates G-protein coupled cascades
rs75404946	6	170730946	G	A	0.229	-0.164	1.98×10^{-6}	*PSMB1*	*TBP, FAM120B, DLL1*	Close to locus implicated in DA distensibility, above. *PSMB1* expression increased in pulmonary artery after exposure to chronic hypoxia. *DLL1* component of Notch1 signalling pathway. Down-regulates *MYOD1* – a candidate for descending aortic distensibility, above.
rs2478350	6	53532224	G	C	0.115	0.186	8.20×10^{-6}	*KLHL31*	*LRRC1, GCLC*	Glutathione synthesis regulated by *GCLC* is key mediator of vascular reactivity. SNPs in *GCLC* associated with ischaemic heart disease. Top SNP eQTL for *KLHL31* in adrenal and *GCLC* in fibroblasts
rs74398418	7	19489807	A	G	0.015	-0.554	8.15×10^{-7}	*TWISTNB (distant)*		No clear cardiovascular link
rs2762597	10	7251289	T	C	0.200	-0.160	2.09×10^{-6}	*SFMBT2*	-	No clear cardiovascular link
rs12573112	10	63678880	T	G	0.469	0.118	1.19×10^{-5}	*ARID5B*	-	Differential methylation & expression of *ARID5B* assoc. with atherosclerosis. Regulates adipogenesis; assoc with obesity. Top SNP spliceQTL for *ARID5B* in skin.
rs141880228	12	47703667	T	C	0.033	0.339	6.59×10^{-6}	*PCED1B*	*LINC02416, AMIGO2*	No clear cardiovascular link
rs372150510	13	32289984	C	CCT AA	0.041	-0.308	7.83×10^{-6}	*RXFP2*	*FRY*	Receptor for relaxin – a natural suppressor of age-related fibrosis in many tissues inc. heart Promotes extracellular remodelling (pregnancy) Serelaxin undergoing trials in acute heart failure; shown to reduce fibrosis in mouse hearts and kidneys
rs5816434	16	27378764	A	AT	0.187	-0.224	1.59×10^{-6}	*IL4R*	*IL21R*	No clear cardiovascular link
rs78732611	17	81126122	T	C	0.174	-0.335	2.43×10^{-6}	*METRNL*	*B3GNTL1*	No clear cardiovascular link

rs61388731	18	349362	G	A	0.149	0.173	$7.06x10^{-6}$	*COLEC12*	*THOC1, ROCK1P1*	*ROCK1P1* interesting pseudogene – highly expressed in heart, shares at least 4 functional exons with ROCK – key regulator of cardiac contractility.
rs141801749	19	46919826	C	CT	0.068	0.242	$1.12x10^{-5}$	*CCDC8*	*PPP5C*	*CCDC8* variants cause 3M syndrome (dwarfism etc.); may play role in growth by interaction with obscurin1. Top SNP eQTL for *CCDC8* in fibroblasts *PPP5C – Protein Phosphatase 5 Catalytic Subunit* – many roles inc. adipogenesis, TGF-beta signalling

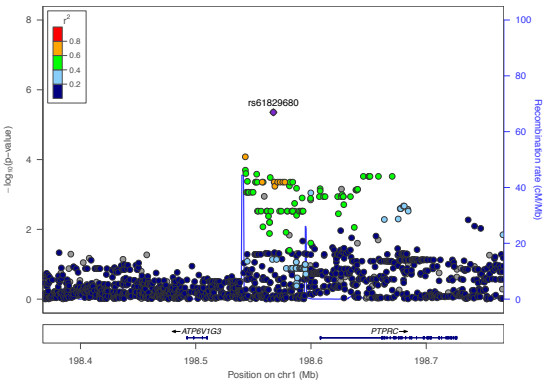
Chr: chromosome; Ref: Reference allele; Alt: effect allele; MAF minor allele frequency; Beta: effect size (change in diameter (mm) per copy of effect allele); Comments: annotations from multiple sources including GWAS catalog[36], GeneCards[369], Mouse Genome Informatics database[370], GTEx[265] and PubMed searches for gene function information; PTBP2: Polypyrimidine Tract Binding Protein 2; PTPRC: Protein Tyrosine Phosphatase, Receptor Type C; ATP6V1G3: ATPase H+ Transporting V1 Subunit G3; RGS7: Regulator Of G Protein Signaling 7; PSMB1: Proteasome Subunit Beta 1; TBP: TATA-Box Binding Protein; FAM120B: Family With Sequence Similarity 120B; DLL: Distal-Less Homeobox 2; KLHL31: Kelch Like Family Member 31; LRRC1: Leucine Rich Repeat Containing 1; GCLC: Glutamate-Cysteine Ligase Catalytic Subunit; TWISTNB: TWIST Neighbor; SFMBT2: Scm Like With Four Mbt Domains 2; ARID5B: AT-Rich Interaction Domain 5B; PCED1B: PC-Esterase Domain Containing 1B; AMIGO2: Adhesion Molecule With Ig Like Domain 2; RXFP2: Relaxin Family Peptide Receptor 2; IL4R: Interleukin 4 Receptor; IL21R: Interleukin 21 Receptor; METRNL: Meteorin Like, Glial Cell Differentiation Regulator; B3GNTL1: UDP-GlcNAc:BetaGal Beta-1,3-N-Acetylglucosaminyltransferase Like 1; COLEC12: Collectin Subfamily Member 12; THOC1: THO Complex 1; ROCK1P1: Rho Associated Coiled-Coil Containing Protein Kinase 1 Pseudogene 1; CCDC8: Coiled-Coil Domain Containing 8; PPP5C: Protein Phosphatase 5 Catalytic Subunit

FIGURE 14.19: LocusZoom plots showing loci associated with aortic pulse wave velocity (rank-normalised phenotype)

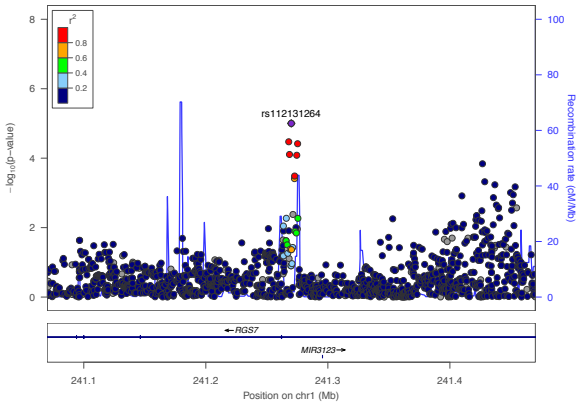
a. Chromosome 1: unknown (PTBP2)



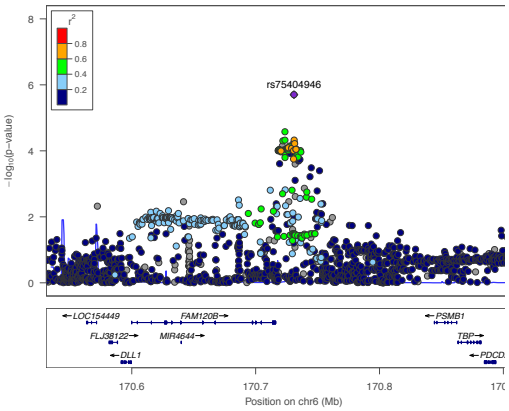
b. PTPRC



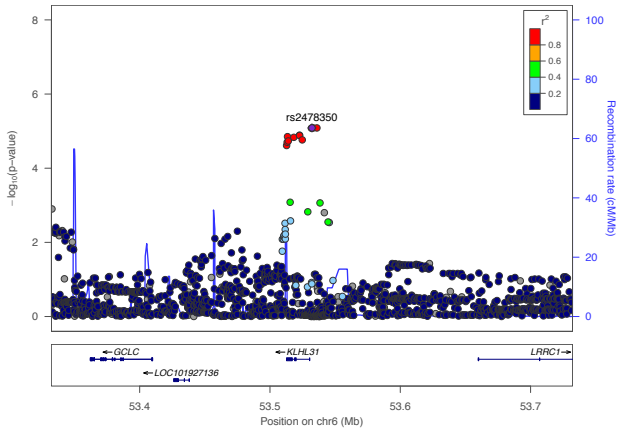
c. Chromosome 1: RGS7



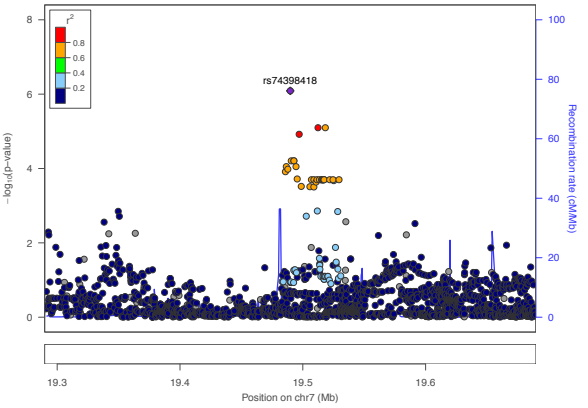
d. Chromosome 6: ?DLL1/PSMB1



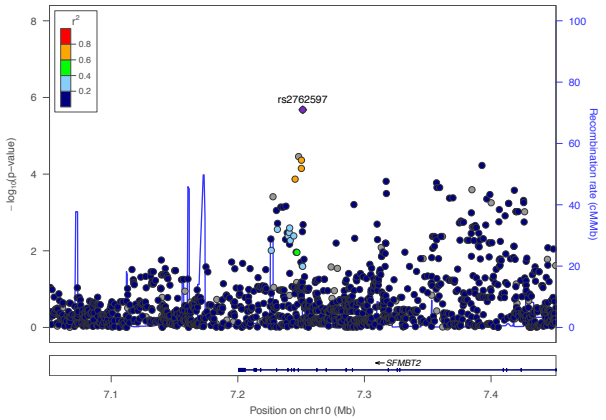
e. Chromosome 6: GCLC



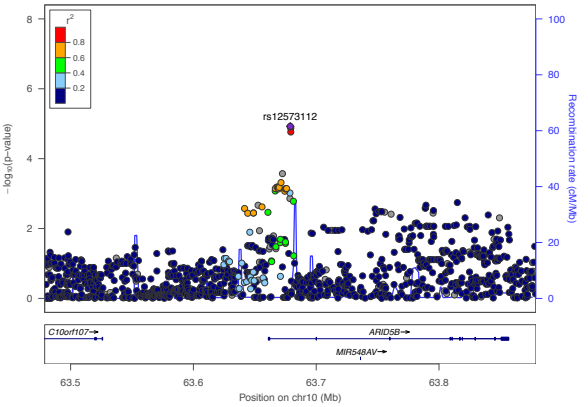
f. Chromosome 7 ? TWISTNB



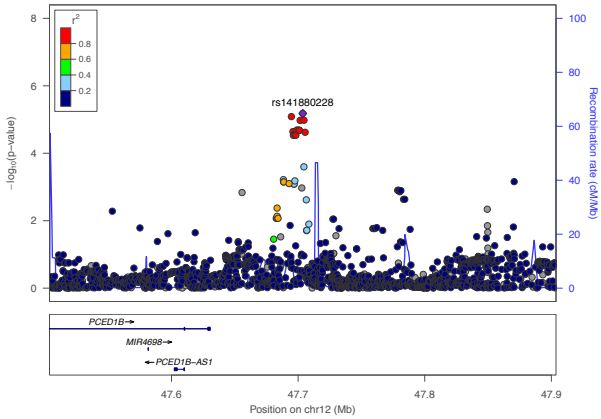
g. Chromosome 10 SFMBT2



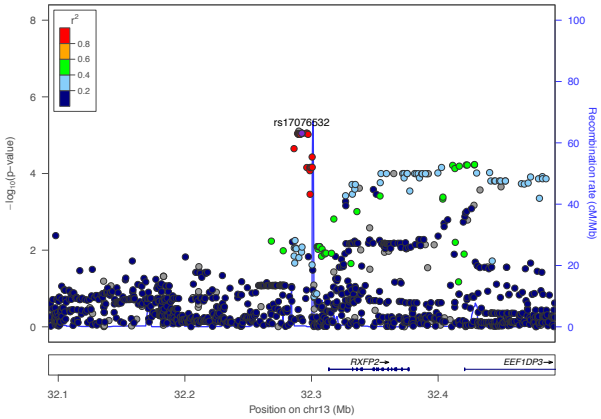
h. Chromosome 10: ARID5B



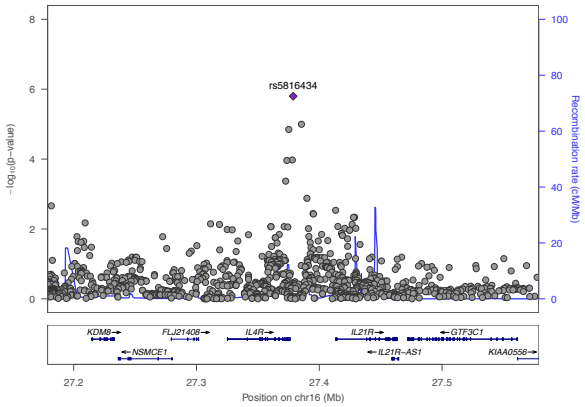
i. Chromosome 12: PCED1B



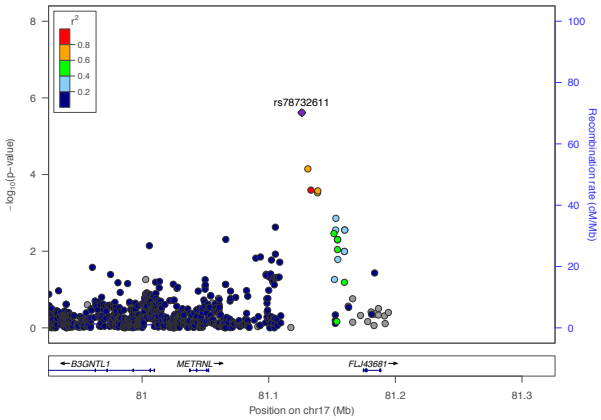
j. Chromosome 13: RXFP2



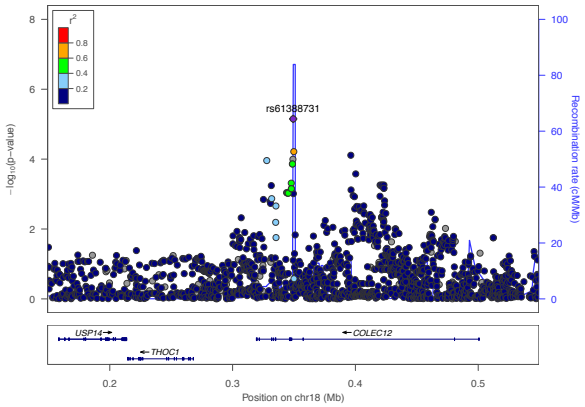
k. Chromosome 16: IL4R/IL21R



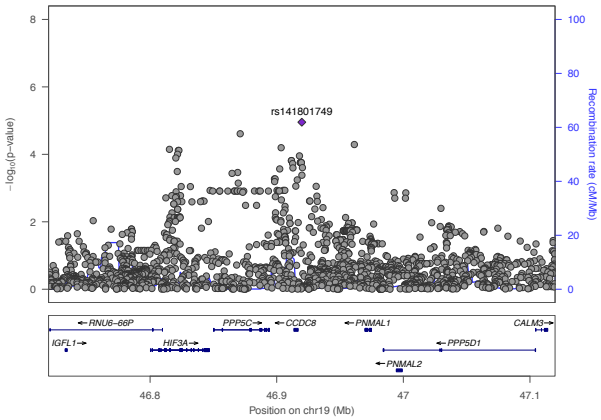
l. Chromosome 17: FLJ43681 / B3GNTL1



*m. Chromosome 18: ?ROCK1P1**



n. Chromosome 19: CCDC8



Suggestive loci are shown as examples above in LocusZoom plots. These plots are essentially a "zoomed in" version of a Manhattan plot, but also demonstrate the linkage disequilibrium structure of the association peak. They demonstrate the spatial relation of the associated SNPs with annotated genes, and with nearby SNPs in linkage disequilibrium (LD) with the lead variant. The plots selected here demonstrate the expected structure of a "true positive" association, with SNPs in LD with the lead variant also associated with phenotype, albeit at lower significance. The labels a-n denote the chromosomal location of the loci, and the top possibilities for candidate genes at the locus, based on biological plausibility and functional evidence, detailed in Table 14.29 and below.

GCLC (Glutamate-Cysteine Ligase Catalytic Subunit) is a critical component of the glutathione pathway, and is essential for cellular redox homeostasis as well as anti-fibrotic state. Mice with selective endothelial haplo-insufficiency demonstrated reduced endothelium-dependent vasodilatation and increased renal fibrosis[442]. Polymorphisms in *GCLC* have also been associated with impaired coronary endothelial vasomotor function and myocardial infarction[443]. The top SNP at this locus on chromosome 6 is an eQTL for *GCLC* in fibroblasts. *GCLC* has an important anti-fibrotic function, and the effect direction in our study is consistent with this. The C allele of rs2478350 is associated with reduced *GCLC* expression[29], and with an increase in aortic PWV. It is tempting to speculate that a reduction in the anti-fibrotic action of *GCLC* from reduced expression could mediate increased aortic fibrosis and therefore increased stiffness , measured as an increase in aortic PWV.

DLL1 (Distal-Less Homeobox 1) at the chromosome 6 locus is a strong candidate. It is essential for post-natal arteriogenesis, and interacts with a known key driver of aortic phenotype, Notch1. It also acts to suppress the activity of a gene identified in the current study as a potential modifier of descending aortic distensibility – *MYOD1.* However, *PSMB1 (Proteasome Subunit Beta 1)* is slightly closer to the main association peak here, and is differentially regulated in pulmonary artery hypoxia, playing a role in hypoxic vascular remodelling[444]. There are no eQTLs in strong linkage disequilibrium with the top SNP to steer us towards one or other candidate gene.

RXFP2 (Relaxin Family Peptide Receptor 2) is also a very interesting candidate gene. It acts as a receptor for both relaxin and insulin-like 3 (its cognate ligand). Most vascular relaxin receptors are encoded by *RXFP1*, and this has therefore been the main focus of research. However, the literature also shows that *RXFP2* is also expressed in vascular smooth muscle cells[445]. Relaxin promotes extracellular remodelling (pregnancy) and reduces fibrosis. Serelaxin, a recombinant relaxin analogue, is undergoing trials in acute heart failure and has been shown to reduce fibrosis in mouse hearts and kidneys. Variants in this gene have also very recently been shown to associate with HbA1c levels in Type I diabetics[446].

ARID5B is a transcriptome co-regulator which forms a chromatin derepressor complex with a histone demethylase. Mediation analysis has shown that *ARID5B* expression mediates effects of several cardiovascular risk factors on atherosclerosis, probably by switching

immunometabolism towards a more chronic inflammatory phenotype[447]. This inflammation-prone environment could certainly be associated with aortic stiffness. It has also been strongly associated with obesity[286], being one of the mediators of association at the mis-named "*FTO* locus" described in section 10.9.5.1. It is functionally linked to *IRX3*, identified here as a gene with putative association with AA area.

The chromosome 18 locus is also potentially of interest. Here, a pseudogene, *ROCK1P1 (Rho Associated Coiled-Coil Containing Protein Kinase 1 Pseudogene 1)* resides, which is, unusually for a true pseudogene, highly expressed in the heart and shares at least 4 functional exons with the ROCK family. Whether a functional protein product is transcribed or not is unknown. ROCK is known to be a key regulator of myosin contractility in heart, and inhibitors of ROCK have undergone trials in pulmonary hypertension[448]. If *ROCK1P1* retains some biological activity, this would be a key pathway to investigate further.

14.5.4.10 Aortic arch pulse wave velocity: gene-based association analysis

As with previous analyses, no gene remained significantly associated with phenotype after correction for multiple testing. However, amongst the top 15 genes were some interesting new candidates such as *SMYD1* not tagged by individual SNP associations. They are detailed in Table 14.29 below and in the following discussion.

TABLE 14.29: Gene-based associations with aortic pulse wave velocity

Gene	Chr	Start	End	No. SNPs	Chisq (obs)	P value	Top SNP p value	Top SNP
PPP5C	19	46850250	46894232	285	785.8	5.05×10^{-5}	1.12×10^{-5}	rs141801749
AFG3L2	18	12328942	12377275	204	710.3	7.36×10^{-5}	3.57×10^{-5}	rs478088
SLMO1	18	12407894	12432236	183	588.5	0.0002	3.57×10^{-5}	rs478088
SFMBT2	10	7200585	7453448	558	1171.1	0.0002	2.09×10^{-6}	rs2762597
FAM155A	13	107820878	108519460	1303	2073.9	0.0002	3.57×10^{-5}	rs61967187
SPRYD7	13	50486841	50510625	149	456.3	0.0003	0.0002	rs9568352
SFXN5	2	73169164	73298965	161	454.3	0.0003	2.05×10^{-5}	rs573555727
ARL6IP6	2	153574406	153617767	119	304.4	0.0004	1.11×10^{-5}	rs6713864
TUBB6	18	12308256	12326568	159	462.9	0.0005	9.87×10^{-5}	rs2509510
SBK3	19	56052022	56056909	320	765.9	0.0005	4.00×10^{-5}	rs7256324
EEF1DP3	13	32420919	32533721	227	535.3	0.0006	2.42×10^{-5}	rs3038105
SBK2	19	56041099	56048435	338	786.4	0.0007	4.00×10^{-5}	rs7256324
SMYD1	2	88367298	88412902	167	414.181	0.0007	0.0002	rs2919870
DCTN1	2	74588280	74619214	63	232.67	0.0009	0.0003	rs7588927
CCDC8	19	46913585	46916919	242	613.147	0.0010	1.12×10^{-5}	rs141801749

Top 15 protein-coding genes associated with aortic pulse wave velocity. Chr: chromosome; Gene start/end: left and right-side boundaries of the gene (GRCh37); No.SNPs: number of SNPs occurring within gene contributing to the analysis; Chi-sq: sum of chi-squared test statistics of all SNPs within the gene region; P value: segment-based test p value; P value most sig SNP: smallest single-SNP p value in the segment; ; PPP5C: Protein Phosphatase 5 Catalytic Subunit; AFG3L2: AFG3 Like Matrix AAA Peptidase Subunit 2; SLMO1 (aka PRELID3A): PRELI Domain Containing 3A; SFMBT2: : Scm Like With Four Mbt Domains 2; FAM155A: Family With Sequence Similarity 155A; SPRYD7: SPRY Domain Containing 7; SFXN5: Sideroflexin 5; ARL6IP6: ADP Ribosylation Factor Like GTPase 6 Interacting Protein 6; TUBB6: Tubulin Beta 6 Class V; SBK3: SH3 Domain Binding Kinase Family Member 3; SMYD1: SET And MYND Domain Containing 1; DCTN1: Dynactin Subunit 1; CCDC8: Coiled-Coil Domain Containing 8;

There is some overlap with the top genes identified in association with both ascending and descending aortic distensibility. *SBK2* and *SBK3* are implicated again, with *SBK2* highly expressed in cardiac tissue, but without a clearly defined role. *AFG3L2, SLMO1* and *TUBB6* seem to tag the same locus of association. There is not an obvious candidate here.

ARL6IP6 is within a known susceptibility locus for ischaemic stroke[449], and recessive variants have been associated with cerebral vascular malformation.

SMYD1 (SET And MYND Domain Containing 1) is a crucial gene for cardiac development and sarcomere organisation[450] and rare variants have been implicated in HCM[451]. The possible association of this gene with aortic pulse wave velocity is a further example of the overlap between aortic and cardiac developmental genes. It is possible too that variants which

directly affect cardiac contractility could have an impact on measured PWV – it is known that ejection time might influence measures of PWV[452] and makes intuitive sense that a more vigorous LV contraction would correlate with a higher PWV within a particular range of aortic stiffness.

DCTN1 (Dynactin Subunit 1) is also an interesting candidate gene. It is involved in neural development, regulating cytoskeletal formation. Rare variants are causes of Perry syndrome, a rare form of Parkinsonism, in which autonomic features such as orthostatic hypotension dominate. This implicates *DCTN1* as a key gene in the development of the autonomic nervous system and therefore it is possible that common variants could lead to a degree of autonomic dysregulation.

14.5.4.11 Aortic arch pulse wave velocity: pathways analysis

Table 14.30 below shows the top 5 associated KEGG and GO biological processes derived from the genome-wide association results. There were no pathways with statistically significant associations after FDR correction.

TABLE 14.30: Top 5 KEGG pathways and top 5 Gene Ontology biological processes associated with pulse wave velocity.

KEGG Pathway	FDR-corrected p value
KEGG: ASCORBATE AND ALDARATE METABOLISM	0.053
KEGG: RIG I LIKE RECEPTOR SIGNALING PATHWAY	0.075
KEGG: BUTANOATE METABOLISM	0.086
KEGG: PENTOSE AND GLUCURONATE INTERCONVERSIONS	0.104
KEGG: STARCH AND SUCROSE METABOLISM	0.120

GO pathway: Biological Process	FDR-corrected p value
GO: RHO PROTEIN SIGNAL TRANSDUCTION	0.073
GO: CELLULAR CATION HOMEOSTASIS	0.096
GO: CATION HOMEOSTASIS	0.101
GO: ION HOMEOSTASIS	0.102
GO: NEUROTRANSMITTER BINDING	0.105

Unlike most other traits, none of the pathways were significantly associated with PWV after multiple-testing correction. In general, however, the top pathways identified more metabolic

traits – with ascorbate, pentarose, starch and sucrose metabolism being highlighted – all sub-classifications of carbohydrate metabolic pathways. A key gene in all these pathways which recurs is *ALDH2*, a gene which is itself associated at nominal significance (p=0.01) in the gene-based analysis of PWV, and which was identified above as a potential driver of aortic dimensions. It has an impressive pedigree of data linking it to various cardiovascular traits and to the process of remodelling.

The top pathways also support the hypothesis that the autonomic nervous system is vitally important in regulating aortic traits: butanoate metabolism produces GABA – a major inhibitory neurotransmitter affecting cardiovascular homeostasis, and also in the GO pathways analysis, neurotransmitter binding appears among the top 5.

Joint trait analysis was not performed for aortic elastic traits due to the low and unstable estimates of heritability for PWV and DA distensibility.

14.5.5: SUMMARY OF ASSOCIATION RESULTS

Table 14.31 below shows a summary of selected results of interest from the single-SNP, gene-based and pathways analysis. This is not an exhaustive list of associations – but highlights some key interesting candidates for follow-up.

We must note that none of these associations, with the exception of several pathways associations, reach genome-wide significance thresholds. These are, therefore, speculative associations which require independent replication or validation in other cohorts.

TABLE 14.31: Summary of suggestive associations with each phenotype – candidate genes for follow-up

Phenotype	Heritability	SNP associations (p<1x10-5)	Additional gene-based associations (top 10-15)	Pathways (top 5-6)
Aortic valve annulus	0.21(NS)	*FLRT2* *KIT/PDGFRA* *PRRX1* *LDB3/BMPR1A*	*FGFRL1*	-
SoV	0.51	*PCDH7* *GABRB2 / GABA receptor cluster* *GABRG3* *EFCAB6* *SON* *MATK*	*IGF1R,* *IGF2R* *ESR* *PENK* *SYNDIG1* *ERBB2IP*	ERBB signalling Insulin signalling Proximal tubule bicarbonate reclamation Cell adhesion Cardiovascular development
STJ	0.65	*TMEM56/CNN3* *GRIA4* *SPIB/POLD1/* *MYBPC2* *PARP11*	*GABRA1* *SCUBE1* *XIRP1*	Linoleic acid metabolism VEGF signalling VSMC contraction
Joint trait aortic root	-	*TENM4* *PTN*	*NPY4R* *PDGFD* *NOMO3*	TGF-beta signalling Neurotrophin signalling Wnt signalling Focal adhesion
AA area	0.46 (NS)	*CYBRD1* *PDE1C* *ESM1* *EPHA5* *IRX3*	*HMCN1* *NUDT19* *CYP26B1*	Histidine metabolism Fatty acid metabolism Proximal tubule bicarbonate reclamation Adipocytokine signalling Aldosterone-regulated sodium reabsorption
DA area	0.85	*PTPN11/ALDH2* *LRRN1* *GYG1* *GRAMD3* *PCDH9* *LATS2* *SPLTC3* *ISM1*	*POMT2* *NGB*	Cell adhesion molecules Hedgehog signalling TGF beta signalling
AA distensibility	0.61	*RYR2* *GC* *NDST1* *PTPRN2* *FAAH*	*AGTR1* *ACTC1* *PRTFDC1*	Focal adhesion Prox. tubule bicarbonate Nucleotide excision repair Pyrimidine metabolism

DA distensibility	0.09 (NS)	*VCL* *BAMBI* *NRXN1* *IER3* *MYOD1* *NIN* *SBK2*	*PCDHB1-3*	Antigen processing Dorso-ventral axis formation Focal adhesion Linoleic acid metabolism
PWV	0.13 (NS)	*ROCK1P1* *RXFP2* *ARID5B* *GCLC* *DLL1/PSMB1*	*SMYD1* *SBK2 &3* *DCTN1*	-

14.6 DISCUSSION

The data presented here represent common genetic variants, genes and pathways which might influence key aortic traits in a healthy population.

14.6.1 Heritability

It is perhaps surprising that the heritability estimates for theoretically and measurably closely correlated traits are so different from one another. This is probably primarily a reflection of the limited power of our study. However, it also speaks to the complexity of the traits under investigation: pulse wave velocity, for example, will be affected by multiple different environmental influences – from exercise to diet, from heart rate to blood pressure and many other factors. Therefore, individual genetic variants are likely to have only very small individual effect on phenotype, particularly in a young age group where the effects of ageing are not yet apparent and there is a very confined variation in the phenotype.

14.6.2 Common themes

This work has identified a bewildering variety of genes and pathways which might be key determinants of aortic biology. Without replication or experimental validation, these putative associations remain hypothetical. Nevertheless, there are some interesting trends in the

associations which might be of great relevance to aortic biology, aortic disease and cardiovascular risk. Many of the variants, genes and pathways identified fall into one of four categories: 1) genes which drive cardiovascular development, 2) genes involved in the autonomic nervous system and 3) genes involved in fibrosis and 4) genes regulating cardiovascular risk traits.

14.6.2.1 The importance of developmental genes and pathways

Many of the genes likely to be driving associations at different loci are important in cardiovascular development, as reflected by the appearance of multiple development terms in the pathways associations. Developmental genes are, perhaps unsurprisingly, more prominent in the associations found with aortic dimensions – for example, *PCDH7* and *SON* and pathways terms mesoderm development and ERBB signalling associated with SoV diameter, and *PTPN11* with descending aortic area. This probably reflects simply the fact that small variations in expression or function of genes during the developmental process could have an impact upon the eventual adult dimensions of the aorta. However, there is good evidence that genes involved in cardiovascular development are also important in the response to injury or stress in adult life[453], and therefore a recapitulation of some developmental mechanisms might be important for remodelling and, potentially for responding to local environmental factors such as shear stress patterns (see Chapter 16).

14.6.2.2 The importance of the autonomic nervous system

This is the first large genomic dataset to suggest explicitly a key role for the autonomic nervous system in determining structural and functional aortic characteristics, particularly in the aortic root. Both GABA and glutamate signalling are implicated in determining aortic structure and function, with recurrent suggestive associations found across multiple traits. Examples include GABA receptor genes associated with diameter of the sinuses of Valsalva, *GRIA4* associated at the sino-tubular junction, *NPY4R* associated with joint aortic root diameter analysis and *NRXN1* associated with descending aortic distensibility.

14.6.2.3 The role of different aortic cells & signalling pathways

This study has identified genes and pathways which influence aortic phenotype via a variety of different signalling pathways. The genes identified clearly act via both vascular smooth muscle cells (e.g. *PSMB1, VCL*), endothelial cells (e.g. *GCLC, BAMBI*), as well as fibroblasts and adipose tissue. This highlights the multi-factorial nature of the aortic phenotype, and the importance of looking widely across cell types for mediators of aortic traits. Cell adhesion is clearly a key process in aortic phenotype regulation; many of the genes identified at all aortic levels regulate cellular interaction with extracellular matrix or focal adhesion.

Key pathways identified include TGF-ß, discussed further below. Wnt, VEGF, Hedgehog and neurotrophin signalling are also highlighted, and IGF signalling particularly in the aortic root. Intracellular calcium signalling may also play a role in ascending aortic distensibility.

14.6.2.4 Fibrosis

Particularly prominent in ascending and descending aortic dimensions, as well as in aortic elastic function, are associations with genes involved with fibrosis. *HMCN1* and *PDE1C* associated with ascending aortic area regulate fibrotic pathways, and *AGTR1* and *ACTC1* associated with ascending aortic distensibility both act as regulators of TGF-ß mediated fibrosis. Similarly in descending aorta, Hedgehog and TGF-ß signalling pathways are associated with area and *BAMBI* and *MYOD* also regulate fibrosis and are associated with distensibility. Associations with PWV also involve genes which regulate fibroblast activity such as *GCLC, DLL* and *RXFP2.* The *PDGFD* association with aortic root diameters also implicates fibrosis as a key regulator in the aortic root. Most of these pathways converge on TGF-ß signalling as a mediator of fibrosis, placing this at the centre of aortic biology, as suggested by its roles in aortic disease.

14.6.2.4 Associations with genes acting via cardiovascular risk factors

Many of the associated genes have known associations with cardiovascular risk factors.
In particular, genes associated with obesity and adiposity have been identified. It is particularly interesting to note that both of the genes eventually identified as mediators of the "*FTO* locus" association with obesity, *ARID5B* and *IRX3*, have been associated

independently with different aortic traits: PWV and ascending aortic area respectively. This supports the observation of body fat mass being a key determinant of aortic phenotype.

14.6.2.5 Overlap with dilated cardiomyopathy genes

Many of the phenotypes were significantly associated with the pathways term “dilated cardiomyopathy”. This reflects the interdependence of the cardiovascular system – both during development and during adaptive remodelling throughout life.

14.6.3 Relevance to cardiovascular risk

The phenotypes most closely related to cardiovascular risk, ascending aortic distensibility and aortic pulse wave velocity, have yielded several genes and pathways meriting further attention and research. Particularly interesting are the associations with genes or pathways which might be amenable to intervention to reduce risk. Examples of these are the associations with *AGTR1* – the angiotensin II receptor, and *ROCK1P1* and relaxin receptors. If these associations are replicated in UK Biobank, they could yield new or more precise therapeutic targets. Equally, the general observations of the importance of genes involved in the autonomic nervous system, gives a focus for further research efforts.

14.6.4. Relevance to aortopathy

Identification of genes which are associated with quantitative aortic traits in a healthy population might help to narrow down the search for genes causing aortic disease. Pathways and processes known to be important in aortopathy have been identified within these associations – for example, the TGF-ß signalling pathway and smooth muscle cell contraction. We will therefore be paying particular attention to the genes and pathways identified in this data when we analyse, for example, segregation data from families with aortopathy. Equally, variants and genes identified here might be important in determining disease severity, risk of aortic complications, or even penetrance in families with aortopathies.

14.7 LIMITATIONS

14.7.1 Size, power

The major limitation of this GWAS is the under-powering due to small sample size, and the lack of any replication to date. A sample size of 1218 is very small for a quantitative, polygenic trait. We can see this in the under-inflated Q-Q plots and the lack of statistically-significant loci. This is particularly a problem with skewed phenotypes such as the aortic elastic phenotypes, where phenotype transformations are necessary to make the models valid, but also reduce power significantly. Although I defined a genome-wide significance level of $p<5x10^{-8}$, widely accepted as a reasonable cut-off, our genomic dataset is in fact even larger than the original GWASs in which this threshold was defined, which contained up to 1 million SNPs. Our dataset has closer to 10 million SNPs, and therefore one could argue that the significance threshold here should be $5x10^{-9}$, or in fact even lower to account for the multiple phenotypes being tested. However, here, the argument is somewhat academic: none of our hits come close to those significance levels. The young age and healthy nature of our population is a double-edged sword: it perhaps means that any inter-individual differences are more likely to be genomic in origin; however, it also reduces the variance in our traits, so effect sizes will be smaller. As ageing is a key determinant of some of the aortic traits in question, it would perhaps be more fruitful to gather data from a broader and older age group. Whilst these factors are problematic in drawing definitive conclusions from our data, they also give us a unique dataset: a truly healthy, young population in whom a range of healthy trait variation can be defined and explored.

The lack of power means that any conclusions drawn about the loci associated at suggestive significance are largely conjecture. It is certain that there will be some false positive associations amongst them; equally, many could represent real associations with our traits of interest. There is certainly some merit in inspecting the "quality" of the associations in terms of the structure of the association peaks, the biological plausibility of the tagged genes, and the likelihood of functional role for the top associated SNPs. However, it is easy to be beguiled by interesting stories or possibilities for biological associations; we must remember that these are as yet unproven, however strong or plausible the potential link with aortic biology. Nevertheless, it is possible to identify some candidate genes and loci of interest for future

follow-up and start to consider the pathways involved. The initial analysis undertaken here can therefore be used to inform further research efforts into particular genes and pathways, and to rapidly generate replication data for specific hypotheses or associations detected in larger datasets. The best use for this dataset in the future will be for replication of associations detected in larger cohorts.

14.7.2 Replication

It has generally become accepted that the threat of false positive associations requires even those loci reaching genome-wide significance to be replicated in an independent dataset for confirmation. I had hoped to replicate any hits in the UK Biobank imaging dataset, but unfortunately, the data was not available in time to include in this thesis. Additionally, the UK Biobank cardiac MRI protocol is limited in the aortic images and phenotypes which are available. Nevertheless, I am closely involved in the analysis and interpretation of what will hopefully be the first large-scale GWAS (approximately 26,000 subjects) of MRI-derived cardiac and aortic parameters. In fact, because of the relative sizes of the datasets, the UK Biobank dataset will be used for "discovery" and we hope to replicate hits in the Digital Heart Project dataset presented in this thesis.

14.7.3 Limitations of statistical models

We chose to use a simple statistical model here which controls for anthropometric factors affecting phenotypes and for blood pressure, which might otherwise dominate the association analysis. By including blood pressure as a covariate however, we almost certainly reduce the association signal for our phenotypes, as with many, the influence on / of blood pressure will be bi-directional. Similarly, the use in SNPTEST of mixed regression models requiring transformation of the main phenotype means that we lose some power. The gene and pathways analysis is rudimentary: the gene-based analysis in GCTA operates simply on the position of the associated SNPs rather than incorporating any functional data. The pathways analysis does incorporate some functional data, but this is obviously blind to any specificity of phenotype or tissues of interest.

14.7.4 Need for further annotation and biological validation

The inherent limitations of GWAS as a research technique have been discussed above. Even if all the suggestive hits identified in this study were to be confirmed, it would still leave us with a huge problem of annotation and identification of mechanism of effect for each locus. This is where many large-scale GWAS fall down – they focus on one or two interesting hits, leaving many unexplored and unexplained. With this, as with any hypothesis-generating study, the key is going to be the follow-up: ensuring that hits replicate and undertaking careful locus dissection to identify causal variants and genes. Whilst some of this problem can be addressed through using the annotation resources available *in silico*, such as VEP, GTEx, MGI, Haploreg, and so on, these large datasets remain incomplete, underpowered and non-specific when one looks at individual traits. For example, GTEx contains data from just 267 aortic samples with associated genotype. Therefore any SNPs which are not extremely common and/or which have smaller effect sizes on expression will be missed. Equally, genes such as transcription factors where very small changes in expression levels could have a profound effect on phenotype will be underrepresented by eQTL analysis. There is a growing and significant literature on the limitations of eQTLs in assessing functional effects of SNPs and in prioritising genes at GWAS loci. There are regular publications of more relevant datasets – so it is possible to "update" the interpretation of association at each locus in light of new information.

14.8 CONCLUSIONS

The common variant associations presented in this chapter represent putative drivers of aortic phenotype which might be relevant both to cardiovascular risk and to phenotype in aortopathies. I found no genome-wide significant associations of individual SNPs or genes with aortic traits. Nevertheless, I have identified some key recurring themes and candidate genes which might mediate aortic phenotype. This data has suggested key roles in the aorta for genes involved in cardiovascular development, the autonomic nervous system, fibrosis and obesity, as well as specific molecules and pathways such as the TGF-ß pathway. The UK Biobank dataset will provide a key opportunity to replicate and validate the associations discovered here.

14.9 ACKNOWLEDGEMENTS

The Digital Heart Project was conceived and co-ordinated by Professor Stuart Cook and Dr Declan O'Regan. Dr Hannah Meyer generously shared her experience and her code, which I adapted for these association studies. Dr Inga Prokapenko kindly advised on joint trait analysis. Dr James Ware and Dr Nicky Whiffin contributed to fruitful discussions.

15: GENOTYPE-PHENOTYPE CORRELATION IN PATIENTS WITH BICUSPID AORTIC VALVE

"The gene that enables birds to learn songs can become cancer-causing. There is no normal physiological process that can't be bastardized by… disease." – Siddhartha Mukherjee

15.1 INTRODUCTION

Whilst GWAS and healthy population studies rely on the power of large cohorts to discover the impact of common variants on phenotype, smaller scale family or case-control studies using sequencing are needed to examine disease-causing rare variants.

The simplest paradigm for studying rare variants in disease is a fully penetrant Mendelian autosomal dominant disease: every person carrying a particular allele will express the disease phenotype; every person without it will be free of disease phenotype. Large families may be used for linkage analysis; cohorts of unrelated cases may be used for association studies by burden testing or genome-wide. Recently, sequencing whole exomes or whole genomes has become not only technically and bioinformatically possible, but also financially feasible, with studies such as the UK100,000 genomes project[454] and DDD[455] sequencing large numbers of probands and families. Until this point, panel sequencing of targeted genes has been the mainstay of both clinical diagnosis and research into Mendelian or oligogenic disease.

However, heritable diseases are rarely so simple. Reduced penetrance, variable expressivity and subclinical forms of phenotype can all muddy the waters. Bicuspid aortic valve is a case in point; this chapter will focus on a panel sequencing study to define whether in a mixed, real-life cohort of both sporadic and familial BAV, we can gain any insights into the genetic architecture of the condition, or particular genes which may be responsible for pathogenesis or phenotypic variability.

15.1.1 The clinical setting - Bicuspid Aortic Valve

BAV is the most common congenital cardiac malformation (see Chapter 10: Introduction), comprised not only of misshapen aortic valves, but also of fundamental abnormalities of haemodynamics and aortic function. Estimates of its prevalence in the general population

vary widely according to the techniques used - but range from around 0.5-2%[178, 179]. 36.7 per cent of these cases appear to be familial and inherited in Mendelian fashion[186] and yet, even in these apparently "straightforward" cases, we have been largely unable to characterise the genetic and molecular basis of this disorder[185].

This is not only an academic problem; the phenotypic heterogeneity of BAV is huge, making risk assessment and determination of appropriate treatment extremely difficult. Major sequelae of BAV include significant valve disease, risk of infective endocarditis, and aortopathy with concomitant risk of aortic dissection. Up to 1 in 3 patients with BAV will require valve replacement in their lifetime; all will require lifelong follow-up for risk of complications, and all will require family screening; requirements which have significant resource implications for any healthcare service. The art of managing these patients is knowing who will require closer follow-up, predicting the incidence of valve and aortic complications, and identifying family members who are at risk. This is made all the more challenging by the incomplete penetrance, variable expressivity and genetic complexity of BAV[185].

15.1.2 Genetic basis of BAV

Syndromic associations with BAV, such as Marfan and Turner's syndromes have provided clues as to the underlying genetic / molecular basis of the disorder (see Introduction Section 10.8.3). Usually, patients will be diagnosed due to the other systemic features of these syndromes and therefore receive screening for aortic features. However, many of these syndromes may be incompletely penetrant - or with subtle forms of systemic features, meaning that many patients may reach adulthood without a clear diagnosis. The presence of BAV should provide a motivation for careful examination for any other systemic features associated with these syndromes.

In addition to these syndromes, there are other single-gene disorders known to be associated with BAV, such as variants in *NOTCH1*[208], *GATA5*[203], *MATR3*[212] and *ROBO4*[456] (the latter 2 usually associated with aortic pathology). Despite identification of these "culprit" variants however, there remains a large amount of "missing" heritability. First degree relatives of patients with BAV have a 10-fold higher risk of BAV than the general population. Where other

cardiovascular features are associated, such as coarctation of the aorta, patent ductus arteriosus or aortopathy, heritability is estimated as high as 0.89[140].

15.1.3 Observational studies

The work presented here is an observational study of BAV, conceived to identify potentially pathogenic variants in candidate genes. This study design certainly has its flaws: it is difficult to assign pathogenicity to variants, and small cohorts can create spurious signals. Nevertheless, when approached with a healthy scepticism, this study design can be used to assess the utility of genetic testing in a cohort, and to generate pilot data or hypotheses to test at larger scale or using different methods.

15.2 AIMS

- To assess the diagnostic yield from panel testing of a range of known aortopathy genes and BAV genes in patients presenting with BAV
- To identify putative causative variants in candidate genes with known mechanistic associations with aortic or aortic valve disease and examine evidence for any phenotype:genotype correlations

15.3 METHODS

15.3.1 Recruitment

174 patients with Bicuspid Aortic Valve were recruited from two large tertiary referral centres, between 2009 and 2012 (Royal Brompton Hospital, London and St George's Hospital, London).

Exclusion criteria included: patients with clinical features of developmental syndromes, multiple major developmental abnormalities or major cytogenetic abnormalities; patients fulfilling standard diagnostic criteria for Turner syndrome or connective tissue disorders, such as Marfan and Ehlers-Danlos syndromes; patients with inflammatory aortic disease; patients with associated congenital cardiac defects other than ventricular septal defect, coarctation of the aorta and PDA; and patients unwilling or unable to provide written informed consent.

15.3.2 Phenotyping

Phenotyping was historic, with a combination of clinical cardiac MRI and echocardiography data used to identify presence of BAV, significant aortopathy (defined very simply as TAA with diameter >4cm at any level) or valve disease. Clinical notes were screened to identify patients with a family history of aortic valve or aortic disease.

15.3.3 DNA extraction

From each patient, 20 ml of venous blood were collected in EDTA tubes and stored at -80˚ C for future DNA extraction. In patients refusing blood sampling, Oragene DNA Self-Collection kits were used for storage of 2 ml of saliva at room temperature.

An automated DNA purification system (BioRobot EZ1) available at the Cardiovascular Biomedical Research Unit (BRU), Royal Brompton Hospital, was used for DNA extraction. The EZ1 utilizes silica-magnetic-particle technology for high-throughput DNA extraction. The EZ1 DNA Blood kit (Qiagen) was used for DNA purification as per the manufacturer's standardized protocol. The EZ1 DNA Tissue kit (Qiagen) was used for DNA purification from the collected saliva samples.

15.3.4 Next Generation Sequencing – panel design

For the design of the sequence capture array, current literature was reviewed, and 63 genes of interest were selected (See Table 15.1; extended version in Appendix 2 with rationale for choice of each gene - thanks to Dr Matina Prapa). In addition to genes known to be associated with syndromic and sporadic BAV disease, a number of genes related to inherited forms of aortopathy, aneurysm formation, AS, and valve formation were also included in the library (see Appendix 2). Gene sequences were retrieved from a reference database (http://www.ensembl.org) and RNA baits were designed for all exons of Ensembl transcripts of the selected genes, using Agilent's eArray platform (https://earray.chem.agilent.com/earray). The final microarray design included UTRs and flanking exon/intron boundaries (+/-100 bp) with a total number of 16,216 unique 120 mer RNA baits covering a target region of 479,673bp. Applied standard eArray parameters to generate RNA baits included tiling frequency = 5x, bait length = 120, standard repeats = off, avoid overlap = 20, and layout strategy = centred.

TABLE 15.1: Genes sequenced in BAV panel

ACE (angiotensin I converting enzyme)	*JAG1* (JAGGED1)
ACTA2 (alpha smooth muscle actin)	*KCNJ2 (Potassium Voltage-Gated Channel Subfamily J Member 2)*
ACTN1 (Alpha-actinin-1) ‡	*KLF15* (Kruppel-like factor 15) *
ACVRL1 (activin receptor-like kinase-1)	*KLF2* (Kruppel-like factor 2)
APOB (apolipoprotein B)	*LOX* (lysyl oxidase)
APOE (apolipoprotein E)	*MED12* (mediator complex subunit 12) ‡
AXIN1 (Axin-1)	*MMP9* (matrix metallopeptidase 9)
CCR5 (chemokine receptor 5)	*MTHFR* (methylene-tetrahydrofolate reductase)
COL11A1 (Collagen type XI)	*MYH11* (smooth muscle myosin, heavy chain 11)
COL1A1 (Collagen Type I)	*NF1* (Neurofibromin-1)
COL1A2 (Collagen type II)	*NFATc1* (Nuclear factor of activated T-cells calcineurin-dependant1)
COL3A1 (Collagen alpha-1 III)	*NKX2.5* (NK2 transcription factor related)
COL4A1 (Collagen alpha-1IV)	*NOS3/eNOS (endothelial nitric oxide synthase)*
COL4A5 (Collagen alpha-5 IV)	*NOTCH1 (Notch 1)*
CTGF (connective tissue growth factor)	*PAI1/SERPINE1* (serpin peptidase inhibitor)
DCHS1 (Dachsous 1)	*PDIA2 (Protein Disuphide Isomerase Family A Member 2)*
DDAH1 (Dimethylarginine Dimethylaminohydrolase 1)	*PGF* (placental growth factor)
DDAH2 (Dimethylarginine Dimethylaminohydrolase 2)	*PLOD1* (lysyl hydroxylase 1)
EFEMP2 (Fibulin-4)	*PLOD3* (lysyl hydroxylase 3)
EGFR (epidermal growth factor receptor)	*PTPN11* (Protein-tyrosine phosphatase 2C)
EIF2S1 (Eukaryotic Translation Initiation Factor 2 Subunit Alpha)	*S100A12 (S100 calcium binding protein A12)*
ELN (Elastin)	*SLC2A10* (glucose transporter type 10)
ENG (endoglin)	*SMAD3* (SMAD family member 3)
ESR1 (Estrogen receptor alpha)	*SOX9* (SRY-box 9)
FBN1 (Fibrillin-1)	*TGFB1* (TGF-beta1)
FLNA (Filamin-A)	*TGFBR1* (TGF-β receptor type 1)
FN1 (Fibronectin-1)	*TGFBR2* (TGF-β receptor type 1)
GAA (Glucosidase Alpha, Acid)	*TSC2* (tuberin)
GATA5 (GATA binding protein 5)	*UFD1L* (ubiquitin fusion degradation 1 like)
GJA1 (connexin-43)	*VCL (vinculin)*
HIF1A (hypoxia inducible factor 1, alpha subunit)	*VDR* (Vitamin D receptor)
IL10 (Interleukin 10)	

15.3.5 Library preparation and sequencing

This work was completed before the upgrade of our laboratory sequencing resources. The SOLiD 5500 platform was used for sequencing. Targeted exons and adjacent introns were enriched and barcoded, followed by next-generation sequencing to screen for sequence variants. First, 3µg of genomic DNA was sheared using the Covaris S2 system and libraries constructed using the SureSelectXT Target Enrichment System for SOLiD 4. Libraries were multiplexed and 32 samples pooled per lane for sequencing on the SOLiD 5500 platform, to generate paired end reads (75bp + 35bp).

15.3.6 Data analysis: genotype calling

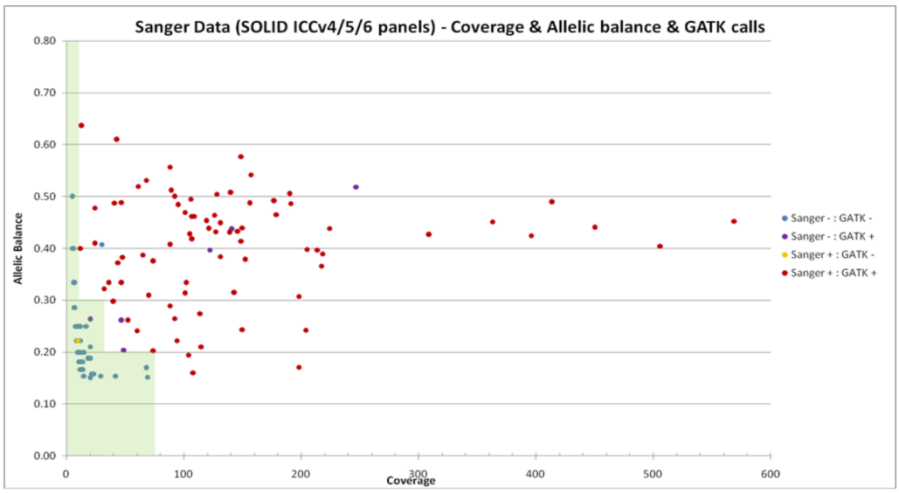
SOLiD 5500xl reads were demultiplexed and aligned in colour space using LifeScope v2.5.1 "Targeted re-sequencing" pipeline (http://www.lifetechnologies.com/us/en/home/technical-resources/software-downloads/lifescope-genomic-analysis-software.html). SOLiD Accuracy Enhancement Tool (SAET) was used to improve colour call accuracy prior to mapping. SAET Duplicate reads were marked by LifeScope and created a subset file (ontarget) based on reads mapping quality > 8.

LifeScope DiBayes and GATK Unified Genotyper algorithms were used for SNP calling and the "smallindel" algorithm to call small insertion/deletions (indels). Local realignment around indels, and base quality score recalibration processes were done in The Genome Analysis Toolkit (GATK) v1.5-20. Alignment summary metrics, callability and coverage report were calculated using Picard v1.65 (http://picard.sourceforge.net), BedTools v2.11.2 and in house perl scripts. The same "ontarget" file was used in LifeScope, GATK and Samtools 131 v0.1.18 programs to make consistent variant calls. Only SNPs that had at least one copy of the non-reference allele, a sequencing depth of >4x, mapping and base quality score >30 were considered for analysis for each test sample. Variants were functionally annotated using the Ensembl API v70_37[457] and HGMD Professional version 2012.2[458]. Variants were called by either DiBayes or GATK Unified Genotyper algorithms.

Variants were validated and quality controlled *in silico* using the Integrated Genome Viewer. Coverage filters were based on our own laboratory's standardised cut-offs, which have been shown to align closely with Sanger validation of variants for this platform (Figure 15.1) Variants were excluded with the following allelic balance and coverage cut-offs: coverage

<10x, or coverage <20x and allelic balance <30, or coverage <75 and allelic balance <20. This left 9184 instances of 502 variants.

FIGURE 15.1 Sanger validation of variant calls correlates with coverage & allelic balance from Integrated Genome Viewer & GATK calls



Each dot represents a variant which was called by GATK and validated by IGV. Each variant also underwent Sanger validation. The colours of the dots represent the outcomes of Sanger validation and IGV filtering, with blue and purple dots representing variants which did not validate by Sanger, and yellow and red dots representing variants which did validate. The green shaded are represents the variants which would be discarded by our allelic balance and coverage filtering – this removes all but 4 of the variants which did not validate with Sanger.

15.3.7 In silico analysis

In addition to the quality control steps outlined above, variants were also filtered on the basis of minor allele frequency (MAF) in both this cohort and population cohorts. Novel variants with a frequency >2% in this cohort were excluded, as these are likely to represent sequencing errors. Variants with an allele frequency of >5% in this cohort were also excluded. Finally, variants were excluded by population frequency cut-offs – if MAF in ExAC, 1000 Genomes or ESP populations was >1%. Any SNPs with frequencies greater than this are unlikely to contribute significantly to pathogenesis, although this may remove from analysis some polymorphisms which may contribute to expression of complex phenotypes (see discussion). A higher-than-normal cut-off frequency of 1% was used, due to the variable penetrance and relatively high prevalence of 0.5-2% of BAV in the general population[178].

Non-synonymous rare variants in coding regions were then annotated using a variety of algorithms to predict their likely functional impact. These included both SIFT ("SortIng tolerant From inTolerant")[459] [http://sift.jcvi.org/]and Polyphen 2[460] [http://genetics.bwh.harvard.edu/pph2/index.shtml] algorithms where appropriate. Both are algorithms which use sequence homology to predict whether a particular SNP is likely to affect protein function, based on the hypothesis that functionally important sequences will be conserved through evolution of functionally related proteins. Thus SIFT predicts whether a particular SNP will be "tolerated" or "damaging" to protein function. Polyphen 2 predicts whether a particular SNP will be "benign" or "possibly damaging" or "probably damaging," based on sequence, phylogenetic and structural features. Other prediction tools included Likelihood Ratio Testing (LRT)[461]. This assumes that some variants we detect are causal and some are not and computes the likelihood of the data under every possible combination of causal statuses. This allows LRT to compute likelihoods of null and alternative models where the null model is one that asserts no causal variants in a group while the alternative model asserts at least one causal variant. We also used MutationTaster[462], which combines assessment of evolutionary conservation, splice-site changes, loss of protein features and changes that might affect the amount of mRNA. Mutation Assessor[463] creates a multiple sequence alignment, partitioned into functionally specific domains, and generates conservation scores for each partition to represent the functional impact of a missense variant. The final in silico tool I used was FATHMM - Functional Analysis Through Hidden Markov Models[464]. This tool aligns homologous sequences and conserved protein domains, with "pathogenicity weights", representing the overall tolerance of the protein/domain to mutations.

We combined these outputs into a very simple, additive overall prediction score "PS," with a maximum score of 2 from each algorithm, or 12 overall, higher scores being more likely to be pathogenic, and lower scores tending towards the benign.

MaxEntScan[465] was used to make predictions of the functional impact of splice site variants; this tool uses the maximum entropy distribution to model short sequence motifs; this information can then be used to assess the difference between alternate and reference

alleles, with a commonly-used cut-off for significant splice effect of a 15% difference between alternate and reference alleles.

Variants were further annotated using CADD (Combined Annotation Dependent Depletion) scores[466]; a method which uses machine learning to predict pathogenicity from a combination of other different annotation tools (including SIFT and Polyphen as mentioned above). This score integrates information from the separate predictive tools in a "meta-annotation". The scaled "PHRED" CADD score scores missense variants according to the rank of their deleteriousness, allowing one to define cut-offs on this basis. To define a variant as of interest in pathogenicity, we used a stringent PHRED score of 23, which equates to a variant being in the most deleterious 0.005% of variation.

Variants were also annotated with ClinVar[467] classifications.

15.3.8 Statistics

Tests for independence were conducted using the R statistical package. Fisher's Exact Tests with FDR adjustment were used for variant burden testing for each gene of our 63 gene panel and overall.

15.4 RESULTS

15.4.1 Patient characteristics

174 patients were recruited in total. 5 were excluded: 2 due to relatedness to other patients in the cohort and 3 due to diagnoses of syndromes accounting for BAV (2x Marfan syndrome; 1 x Turner syndrome). The remaining 169 were taken forward for analysis. Patient characteristics are presented in Table 15.2 below.

TABLE 15.2: Patient characteristics

BAV population (n)	**169**
Gender (male n, %)	122 (72)
Age in years (mean, range)	47 (9-85)
Moderate- severe aortic valve disease (n, %)	129 (74)
Aortic regurgitation	67 (40)
Aortic stenosis	92 (54)
Concomitant lesions (n, %)	
Significant thoracic aortic dilatation/aneurysm (defined as diameter>4cm)	78 (46)
Coarctation of the aorta	28 (16)
Ventricular septal defect	14 (8)
Patent ductus arteriosus	5 (3)
Family history of BAV / aortopathy	
First degree relative with BAV / aortopathy	17 (10)
Broader family history of BAV / other congenital cardiac lesion	33 (20)

These figures demonstrate the male predominance of this condition. Many estimates put the male:female ratio of BAV at 2:1[468], and the current results are consistent with this balance.

This cohort is a relatively young one, but with a very high proportion of complications - indicating the phenotypic severity of these cases. The cohort described here has a very high prevalence of moderate-to-severe aortic valve disease, as well as a high prevalence of significant thoracic aortic aneurysm and other congenital abnormalities. This is likely to reflect the populations from which these patients are drawn; Royal Brompton and St George's are large cardiothoracic surgical centres, the former with a large congenital heart disease unit; and therefore the patients attending these centres are likely to be rather more complex than the standard population with BAV.

15.4.2 Sequencing metrics

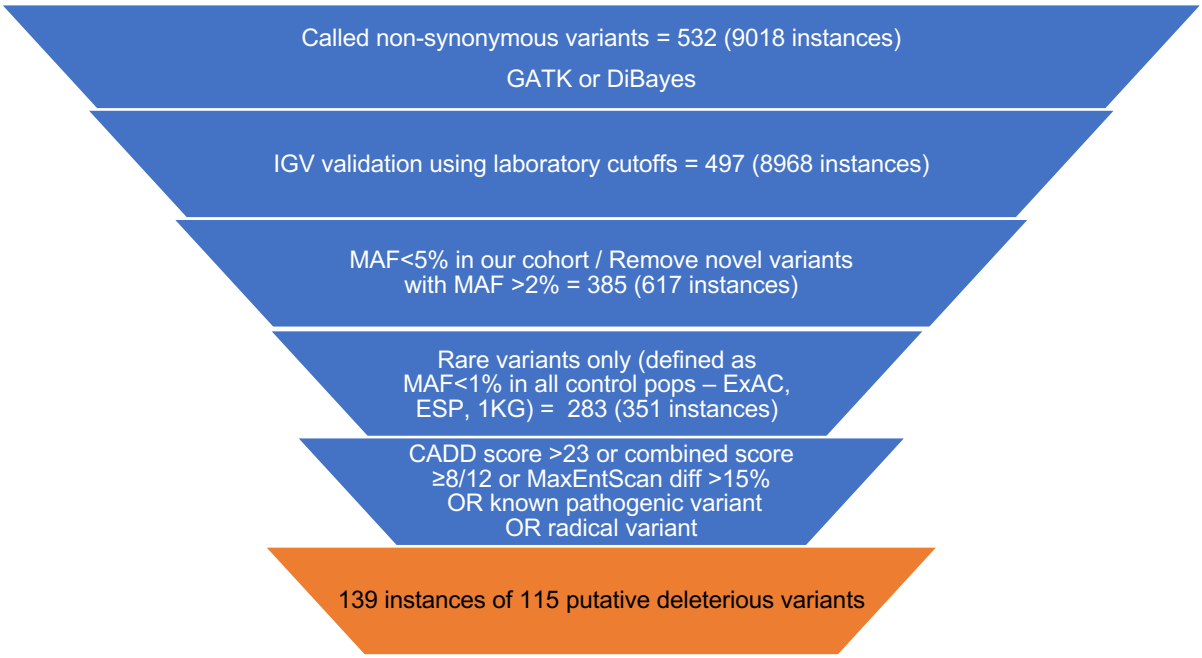
For the genetic analysis, 2 samples were excluded due to a sample mix-up. One further sample was excluded due to a diagnosis of previous Takayasu's arteritis which complicated phenotype. A further 12 samples did not pass quality control measures due to low coverage or low callable percentage. This left 155 patients for further analysis. The mean coverage was 135 (range per patient 18-224). The percentage of bases with coverage of >10 was 94.5%, and percentage of reads on target was 71.7%. Callable percentage was 96.3%, with variable results per gene.

Some particular genes such as GATA5 had low coverage overall, due to difficulties mapping reads to significantly repetitive regions.

15.4.3 Variant numbers

Variants were filtered according to the cut-offs outlined in the Methods section above. Figure 15.2 demonstrates each filtering step, with the number of variants remaining at each point.

FIGURE 15.2: Scheme of analysis showing numbers of variants at each stage



15.4.4 Pathogenic Mendelian disease-causing variants

No patients in this cohort had known Mendelian disease-causing variants defined as pathogenic or likely pathogenic for BAV in ClinVar.

However a variant was identified in one patient which was classified as pathogenic for mitral valve prolapse in ClinVar (in *DCHS1*; rs201457110 (c.7538G>A; p.Arg2513His). He had a normal mitral valve on review of echocardiogram. For further discussion of this patient, his phenotype and ongoing family studies, please see Discussion below.

Two patients were heterozygous for a variant classified as pathogenic for the autosomal recessive form of type 3 hyperlipoproteinaemia. One male patient was hemizygous for a variant described as “pathogenic” in ClinVar for Alport syndrome; an X linked semi-dominant condition which causes progressive basement membrane dysfunction, leading to renal impairment, sensorineural hearing loss, and variable eye manifestations, and which has been linked to development of TAA in the past. However, on further inspection, there was just one published study underlying the “pathogenic” classification for this variant, and the female proband described had double-hit *COL4A5* variants which included this one. This would not fulfil criteria for a pathogenic variant under current ACMG guidelines. The patient in this cohort is not known to have any renal impairment or hearing loss. He has no other candidate variant.

One patient was heterozygous for a variant in *MTHFR* classified in ClinVar as pathogenic for homocysteinaemia and a VUS for neural tube defects. Its pathogenic classification however, does not stand up to a scrutiny of the functional paper in which it was reported, where there is no evidence of this variant altering *MTHFR* activity levels.

In line with previous studies, no significant (rare and predicted deleterious) variants in *TGFBR1* or *TGFBR2* were found, suggesting that these genes are not major causes of non-syndromic BAV.

15.4.5 Variant burden analysis

There was no overrepresentation of rare variants in genes we tested in our cohort compared with the control populations from the 1000 genomes project (1KG) and exome variant server (EVS). This study was, however, underpowered for this type of analysis.

15.4.6 Candidate variant identification

There were 115 possible candidate variants identified through in silico screening in our cohort, in 91 patients.

All variants which were classified as likely deleterious in any candidate gene are presented in Table 15.3, below.

TABLE 15.3: Candidate variants identified by in silico filtering in our BAV cohort

Variant dbSNP identifier (rsID)	Protein change *(or cDNA for splice variants)*	Consequence	MAF in ExAC	ClinVar status	Combined score	CADD score (or MaxEnt Scan score for splice variants)	Phenotypes in our cohort
ACE							
rs148943954	p.Pro505Ala	missense	0.0004	VUS (reported with renal dysplasia)	8	25.4	AR+TAA
rs149412997	p.Gly267Arg	missense	0.0003	-	10	31	AR+TAA
rs56394458 (x2)	p.Gly354Arg	missense	0.0078	-	7	27.4	AS/AS+AR
ACTN1							
rs762425067	p.Tyr466Cys	missense	$8x10^{-6}$	-	9	29.8	AS+TAA
APOB							
rs141763789	*(c.3507T>C)*	Splice region	0.0002	Likely benign	-	-19 (splice)	AS++
rs12713559	p.Arg3558Cys	Missense	0.0003	VUS	7	33	CoA
rs146341569	p.Gly1617Glu	missense	$5x10^{-5}$	-	7	27.6	None
rs147223101	p.Arg297Leu	missense	$8x10^{-5}$	-	5	34	AS++
2:21250792G>A	p.Leu659Phe	missense	Novel	-	7	23.1	AS
rs41288783	p.Pro994Leu	missense	0.0004	VUS, likely benign	7	29.9	AS
rs151009667 (x2)	p.Arg1689His	missense	0.0013	VUS	7	27.8	None / AS calcium+ TAA
APOE							
rs769452	p.Leu46Pro	missense	0.0024	VUS, Pathogenic (disease not specified)	4	11.43	AS
rs267606661	p.Arg269Gly	missense	0.0004	Pathogenic (Familial type 3 hyperlipoproteinemia; usually autosomal recessive)	6	25.6	AR
AXIN1							
rs546827136	p.Arg797Cys	missense	$1x10^{-5}$	-	9	27	AR+ TAA
CCR5							
rs56198941	p.Ala73Val	missense	0.0008	-	6	31	AS x calcium
3:46415254G>T	p.Met287Ile	missense	novel	-	3	23.3	AR+TAA
COL1A1							
rs66489345	p.Gly629Ser	missense	0	-	8	27	AS+TAA
rs1800211	p.Arg564His	missense	0.0002	VUS	6.5	29.5	AS

17:48273697C>T	p.Arg312His	missense	novel	-	6.5	26	AS,AR,TAA
COL3A1							
rs112371422	p.Arg1109Gly	missense	$6x10^{-5}$	VUS	9	32	AS
rs144614075	p.Lys1273Arg	missense	$8x10^{-5}$	VUS, likely benign	3	24.8	AR+AS
2:189860884G>A	p.Asp548Asn	missense	novel	-	6.5	24.7	AS, TAA
rs111840783	p.Lys1313Arg	missense	0.0008	VUS, likely benign	4.5	23.4	AR, TAA
rs1801183	p.Pro668Thr	missense	0.0017	benign, likely benign	7	23.5	TAA
rs35795890	p.Pro602Thr	missense	0.0044	benign	6.5	23.2	AR, TAA/TAA
COL4A1							
13:110864786A>C	p.Ile122Ser	missense	novel	-	6	24.2	AS
rs34004222	p.Pro54Leu	missense	0.0028	Likely benign	7	26.5	None
rs41275090	p.Val883Ile	missense	0.0045	Likely benign	4	23	AR,AS,TAA
COL4A5							
rs144282156	p.Arg1422Cys	missense	0.0003	Pathogenic (Alport syndrome, X-linked recessive)*	5	31	Sev AR+ AS+TAA (male)
X:107938662A>G	p.Met1657Val	missense	novel	-	9	27.8	Severe AR
COL11A1							
rs55821405	p.Asp1472Glu	missense	0.0029	Likely benign	6	23	AR,TAA
CTGF							
rs759434407	p.Tyr182His	missense, splice variant	$3x10^{-5}$	-	6	23.2	AS,AR
DCHS1							
rs368211314	p.Arg2768Cys	missense	0.0003	-	7	28.3	AR
rs141901540	p.Arg2403Trp	missense	0.0005	-	7	28.7	AR, AS, TAA
rs145132459	p.Arg387His	missense	$7x10^{-5}$	-	8.5	32	AR, TAA
rs367695682	p.Arg2170Cys	missense	$4x10^{-5}$	-	9	32	AR, AS, TAA
rs145132459	p.Arg387His	missense	$7x10^{-5}$	-	6.5	32	AR, TAA
rs145725245	p.Arg1991His	missense	0.0004	-	7	26	AR,AS,TAA
rs367695682	p.Arg2170Cys	missense	$4x10^{-5}$	-	7	32	AR, AS, TAA
11:6648134G>T		missense	novel	-	7	27.3	AR, TAA
rs201457110*	**p.Arg2513His**	**missense**	**0.0005**	**Pathogenic (MVP, autosomal dominant)**	**6**	**24.3**	**AS, TAA**
11:6643792C>G	p.Ala3039Pro	missense	novel	-	5	28.8	AR,TAA
rs369059057	c.5036-5C>G	Splice variant	0.0009	-	0	35.9 (splice)	AR, TAA
rs138340204	p.Trp1053Ser	missense	0.0034	VUS	6.5	24.4	AS, AR (note pt had Ross aged 34)
DDAH1							
rs139393958 (x2)	p.Thr87Met	missense	0.0013	-	7	34	AR/TAA
EGFR							
rs587778252	p.Arg165Trp	missense	$2x10^{-5}$	-	1.5	26.1	AS+TAA
rs143884981	p.Ala840Thr	missense	$8x10^{-6}$	-	6.5	34	AS
FBN1							
rs368650399	p.Pro1453Leu	missense	$1x10^{-5}$	-	8	25.2	AS, AR, borderline TAA
rs145105768	p.Gly2691Ser	missense	$6x10^{-5}$	VUS	8	27.7	AS
rs761886457	p.Thr1844Lys	missense	novel	-	8.5	23.9	AS,

FLNA							
rs371689052	p.Asp2323Asn	missense	$5x10^{-5}$	-	8	28.6	AS (male pt)
rs187029309	p.Ser1983Leu	missense	0.0026	benign	5	25.8	AS, TAA
rs200130356 (x2)	p.Gln1484Arg	missense	0.0012	-	6	23.1	CoA,VSD,PDA (F)/ AS, subAo stenosis
rs201656372	p.Val949Ile	missense	0.0004	VUS, benign	6	23.8	CoA, VSD (M) AR, CoA, floppy MV (M)
FN1							
rs369076813	pThr124Asn	missense	$7x10^{-5}$	-	7	28.4	AR, AS, TAA
2:216248144A>G	p.Arg592His	missense	novel	-	10	33	AR, AS, TAA
rs147831535	p.Arg653His	missense	0.0037	-	6.5	29.5	AR, AS
rs139078629 (x4)	p.Arg1496Trp	missense	0.0049	-	5	28.6	AR, TAA/ AS/AS/AS,TAA
rs140926439 (x4)	p.Gly357Glu	missense	0.0023	-	9	33	TAA, CoA/none/none/AS,TAA
rs766594997	p.Arg653His	missense	$2x10^{-5}$	-	2.5	23	AS, TAA
rs748612972	p.Met1856Arg	missense	novel	-	2	23.9	None
GAA							
rs367632754	p.Ala693Thr	missense	0.0001	VUS	9	34	AS+TAA
GATA5							
rs782759156	p.Trp391Cys	missense	novel	-	9	33	AS
rs116164480	p.Leu233Pro/ splice	missense, splice variant	0.0020	-	7	23.3	AS
HIF1A							
rs370608348	p.Gln717His	missense	$8x10^{-6}$	-	4	24.1	AS
rs149348765	p.Asp422Tyr	missense	0.0042	-	2.5	24.5	TAA
JAG1							
rs140330283	p.Thr767Met	missense	$4x10^{-5}$	VUS	8.5	27.8	None
20:10627751A>T	p.Val574Glu	missense, splice variant	novel	-	8.5	33	AR
rs145895196	p.Arg937Gln	missense	0.0022	Likely benign	6	23.1	None
rs773039210	p.Thr408Met	missense	$2x10^{-5}$	-	9	33	AR, TAA
KCNJ2							
rs776976697	p.Gln285Ter	stop gained	novel	-	-	38	AS
MMP9							
rs768365704	p.Thr246ProfsTer92	frameshift	novel	-	-	33	AS
rs749803246	p.His411Arg	missense	$8x10^{-6}$	-	8	25.3	AR, TAA
rs41529445 (x2)	p.Thr258Ile	missense	0.0009	VUS	7	29	AS,TAA/ dissection(no TAA)
rs144023823	p.Arg24Cys	missense	0.0039	-	1.5	23.2	AR,TAA
rs144098289	p.Gly296Ser	missense	0.0017	VUS	5	24	AR,TAA
MTHFR							

rs200100285	p.Ile225Val	missense	0.0001	VUS, pathogenic (homocysteinaemia)	4.5	4.957	AR,AS
1:11850861T>A	p.Glu470Val	missense	NA	-	6.5	23.2	AS
rs142617551	p.Glu470Val	missense	0.0014	VUS	7	23.3	AS, TAA
MYH11							
rs372247345	p.Arg439His	missense	$7x10^{-5}$	-	9	32	AS,TAA
rs775809843	p.Lys1449Gln	missense	$7x10^{-5}$	VUS	8	26.8	AR,TAA
rs762308378	p.Lys889Gln	missense	0.0003	VUS	6	23.2	AR,AS,TAA
rs147447269	p.Lys73Gln	missense	0.0003	VUS	9	26.1	AR,AS
rs111854563	p.Thr1558Met	missense	0.0007	VUS, likely benign	8	29.1	TAA
rs150759461 (x2)	p.Arg247Cys	missense	0.0019	VUS, likely benign	7	34	AR/AS
NF1							
rs763082717	p.Asn2331Ser	missense	8x10-6	VUS	4.5	24.4	AR,TAA
rs587781502	p.Lys2593Arg	missense	8x10-6	VUS	6.5	24	AR,TAA
NKX2-5							
rs759518211	p.Ala255Thr	missense	$1x10^{-5}$	-	6	28.5	AR, CoA, VSD
rs104893904	p.Glu21Gln	missense	0.0008	VUS, likely benign	7	26.4	AR, CoA, VSD, PDA
rs553883993	p.Tyr259Phe	missense	$9x10^{-5}$	-	6	23.4	AR,AS
NOS3							
rs566042414	p.Gly1135Ser	missense	0.0006	-	7	25.7	AS,TAA
rs149539813	p.Asp287Asn	missense	0.0020	-	8	33	AR, AS, TAA
rs139184126	p.Gly284Ser	missense	0.0002	-	4.5	24.2	AR, **AS**, TAA
NOTCH1							
rs372739350	*c.3510+3G>A*	splice variant	0.0001	-	-	-49 (Splice)	**AS**, AR
rs774680812	p.Pro2551Leu	missense	$1x10^{-5}$	-	9	29	AS
9:139410453-/T	p.Tyr550Ter	Stop gained, frameshift	novel	-	-	-	AS
9:139412717C/-	p.Cys376SerfsTer 255	frameshift	novel	-	-	-	AR, TAA
rs201620358 (x2)	p.Arg912Trp	missense	0.0019	VUS, likely benign	5	31	AS / AR+TAA
rs750318685	p.Arg1761Gln	missense	novel	-	8	31	AR, AS+ (calc++), TAA
rs182330532(x2)	p.Arg1279Cys	missense	0.0003	VUS	5.5	27.6	AS/AS
PDIA2							
rs182349041 (x6)	p.Thr115Met	missense	0.0047	-	1.5	23.5	AR,TAA/AR, TAA/AS/ none/AR,TA A/TAA
PLOD1							
rs148510973	p.Arg136His	missense	0	-	7	27.8	AS
rs138490756	p.Arg512Cys	missense	0.0035	VUS, benign ,likely benign	6	24.2	AR,AS,TAA
PLOD3							
7:100853723G>A	*c.1503C>T*	Splice variant	novel	-	2	+54(splice)	AS
rs373230090	p.Arg317Gln	missense	$2x10^{-5}$	-	6	24.1	AS
SERPINE1							
rs142959808	p.Leu175Pro	missense	$4x10^{-5}$	-	9	24.2	AS,TAA

SLC2A10							
rs746480018	p.Arg534His	missense	$8x10^{-6}$	-	4.5	24.7	AS
rs141310869	p.Ala58Val	missense	0	-	6	24.4	AR,AS,TAA
TGFB1							
rs200164212	p.Arg296Gln	missense	0.0004	-	8.5	34	No valve, CoA
19:41854253G>C (x2)	p.Leu155Val	missense	novel	-	9	23.3	AS/AS,TAA
rs547881966	p.Met261Leu	missense	$2x10^{-5}$	-	5	24.2	AR
TSC2							
rs1800729 (x2)	p.Ala583Thr	missense	0.0017	benign, likely benign	9	26.2	AS,TAA/AS, TAA
rs766451267	p.Arg639Trp	missense	$4.x10^{-5}$	VUS	9	27.3	AS,TAA
rs45484298	p.Gly440Ser	missense	0.0008	VUS, likely benign	4.5	27.2	AR,**AS**,TAA
16:2138262C>T	p.Pro1732Leu	missense	novel	-	8.5	27.6	AR,**AS**
VCL							
rs189242810	p.Thr197Ile	missense	0.0002	VUS	7	25.8	AS
10:75863667T>A	p.Asp704Glu	missense	novel	-	6.5	23.1	AS

Table showing potentially pathogenic variants identified in the BAV cohort. AS: aortic stenosis; AR: aortic regurgitation; TAA: thoracic aortic aneurysm. VUS: variant of uncertain significance

15.4.7. Observation of variant clusters and possible genotype:phenotype correlations

15.4.7.1 APOB variants

The 8 patients with variants in APOB seem to form a cluster of stenotic phenotypes (see Table 15.4, below), with 6 out of the 8 patients having a stenotic aortic valve, but all 8 having some form of obstructive / hypoplastic lesion, ranging from “simple” BAV stenosis to a complex array of subaortic stenosis and coarctation. This appears to be quite a marked clustering of phenotype with variants in this particular gene, but the numbers are small (see Discussion).

TABLE 15.4: Phenotype: genotype correlation: *APOB* variants

Study ID	dbSNP	Phenotype	Genes with other sig variants	Family history
20NA02076	2:21250792G>A	33y F. Severe AS with small aorta	*MTHFR*	No
20MP01994	rs151009667	49 y/o F. CoA, subAo stenosis, VSD, PDA, no valve issues	*FLNA, FN1*	No
20MP02000	rs146341569	37 y/o F. CoA, subAo stenosis, VSD, PDA, ALCAPA	-	No
20MP02044	rs41288783	29 y/o M. CoA, VSD, subAo stenosis, AS	*FLNA*	Yes – son has same
20NA02123	rs151009667	61 y/o M. AS, sig calcification	-	No
20NA02143	rs147223101	77 y/o M Severe AS.	-	Yes – daughter has BAV
20MP01992	rs12713559	26y/o M. CoA repair as baby. No significant valve disease.	-	No
20NA02090	rs141763789	75 y/o M Severe AS, TAA	-	No

Table reporting phenotypes of patients with rare protein-altering variants in APOB. F: female; M: male; AS; Aortic Stenosis; AR: Aortic Regurgitation; CoA: Coarctation of the Aorta; subAo: sub-aortic; VSD: Ventricular Septal Defect; PDA: Patent Ductus Arteriosus; ALCAPA: Abnormal origin of left coronary artery from the pulmonary artery; sig: significant; TAA; Thoracic Aortic Aneurysm

15.4.7.2 *DCHS1* variants

Of the 9 unrelated patients with candidate variants in *DCHS1*, 7 had thoracic aortic aneurysms (TAAs), and 1 further patient had had a Ross procedure at a young age and may therefore be protected from aortic root dilatation (see Table 15.5). There is significant difference in the prevalence of TAAs in this subgroup to the prevalence of TAAs in the rest of our tertiary centre BAV population (7/8 *DCHS1*+ve [with 1 excluded for Ross]; 71/146 *DCHS1*-negative), with chi-squared of 5.2 and p 0.02. However, it must be noted that the numbers are small and this is a post-hoc subgroup analysis. Variants in *DCHS1* have previously been identified as causal for mitral valve prolapse (MVP), and indeed this MVP phenotype has been linked with BAV, and coincides with it in conditions such as Marfan syndrome. On review of original echocardiography, none of the patients in the current study with *DCHS1* variants had an

abnormal mitral valve. Of note, a father and son (excluded from the main study) with BAV, share a rare variant in *DCHS1*. For further description of this family, see Discussion.

TABLE 15.5: Phenotype:genotype correlation: *DCHS1* variants

Study ID	dbSNP	Phenotype	Genes with other sig variants	Family history
20MP02004	rs367695682	31 y/o M. AR**, AS,** borderline **TAA.** Root & valve replacement for IE. Normal MV	*PLOD1*	No
20MP02034	rs145132459	53 y/o M. AR, **TAA;** AVR at age 22. Known rheumatic fever in childhood. Normal MV	-	No
20MP02060	rs138340204	68 y/o M. Severe **AS**, AR**. Ross** at age of 34, multiple redos. Normal MV	-	Father – ?SCD
20MP02068	rs369059057	28 y/o M. AR**, TAA,** CoA, requiring ARR aged 18. Normal MV	*MMP9*	
20NA02072	rs368211314	51 y/o M Severe AR. Normal MV	*COL4A5*	No
20NA02077*	*rs201457110*	*56 y/o M. Severe **AS, TAA** Normal MV*	*FN1, MYH11*	*Yes – son and brother with BAV*
20NA02096	rs145725245	77 y/o M. Moderate AR, severe **AS, large TAA** 6cm Normal MV	-	
20NA02124	11:6643792C>G 11:6648134G>T	36 y/o M. AR, **TAA.** Normal MV	*2xDCHS1, NOTCH1*	No
20NA02126	rs141901540	Moderate AS, moderate AR, **TAA** 5.5cm. Normal MV	*FN1*	No
*Excluded from main study as related**	*rs201457110*	*33 y/o M. Learning difficulties. AS, possibly monocusp morphology*	*GAA, MED12*	*Yes – father and uncle with BAV*

Table reporting phenotypes of patients with rare protein-altering variants in DCHS1. F: female; M: male; AS; Aortic Stenosis; AR: Aortic Regurgitation; AVR: Aortic Valve Replacement; MV: Mitral Valve; CoA: Coarctation of the Aorta; ARR: Aortic Root Replacement; SCD: Sudden Cardiac Death; subAo: sub-aortic; VSD: Ventricular Septal Defect; sig: significant; TAA; Thoracic Aortic Aneurysm

15.4.7.3 *MMP9* variants

5/6 patients with *MMP9* candidate variants had significant TAA or dissection. As differential expression of *MMP9* has been consistently found in dilated aortas, and the finely-controlled balance of matrix renewal and degradation is known to be of vital importance for aortic wall remodelling, this could be consistent with a phenotype-modifying role for some *MMP9* variants.

15.4.7.4 *NOTCH1* variants

Two frameshift variants were observed in *NOTCH1* in this cohort – whilst not pathogenic by ACMG criteria, there is strong suspicion of a causal role for these.

7/10 patients with candidate *NOTCH1* variants had significant AS, with several described in operation notes as "heavily calcified" (see Table 15.6). This fits well with previous descriptions of *NOTCH1* variants causing valve stenosis and calcification, although it is not statistically significantly different from the prevalence of AS in our overall population. The patients with *NOTCH1* variants who did not exhibit AS however had other compelling reasons for variable phenotype – one with a history of Takayasu's arteritis as well as additional variants in *FLNA* and *FN1*; one with double *DCHS1* variants which appear to be significant, as above, and the other with a *COL11A1* variant and more complex phenotype including coarctation of the aorta and VSD.

TABLE 15.6: Phenotype:genotype correlation: *NOTCH1* variants

Study ID	dbSNP and protein annotation	Phenotype	Genes with other sig variants	Family history
20MP01993	9:139410453-/T ***p.Tyr550Ter***	42y/o M. Heavily calcified AS	*HIF1A*	No
20MP02025	*rs370684825 p.Asn197Lys*	*46y/o female; AR, TAA but known Hx Takayasu's arteritis*	*FLNA, FN1*	*No*
20MP02026	rs750318685 ***p.Arg1761Gln***	59 y/o M. Heavily calcified AS, AR; known rheumatic fever in childhood	*TSC2, COL4A1*	No
20MP02055	rs774680812 ***p.Arg1279Cys***	38 y/o M. AS, TAA		Maternal grandmother "valve disease"
20MP02059	9:139412717C/- ***p.Cys376SerfsTer255***	43 y/o M. AR, TAA, CoA, VSD	*COL11A1*	No
20NA02117	rs201620358 ***p.Arg912Trp***	49 y/o M. AS		No
20NA02124	rs201620358 ***p.Arg912Trp***	36 y/o M AR, TAA	*DCHS1 x2*	No
20NA02132	rs182330532 ***p.Arg1279Cys***	64 y/o F. Severe AS	*SLC2A10*	No
20NA02135	rs182330532 ***p.Arg1279Cys***	50 y/o F. Mod-severe AS	*KCNJ2, PLOD1, VCL*	No
20NA02139	rs372739350 ***(c.3510+3G>A)***	85 y/o M. Severe AS	*FN1*	No

Table reporting phenotypes of patients with rare protein-altering variants in NOTCH1. F: female; M: male; AS; Aortic Stenosis; AR: Aortic Regurgitation; Hx: History; AVR: Aortic Valve Replacement; MV: Mitral Valve; CoA: Coarctation of the Aorta; SCD: Sudden Cardiac Death; subAo: sub-aortic; VSD: Ventricular Septal Defect; PDA: Patent Ductus Arteriosus; ALCAPA: Abnormal origin of left coronary artery from the pulmonary artery; sig: significant; TAA; Thoracic Aortic Aneurysm

15.4.7.5 Known phenotype modifiers: R247C in MYH11

This variant was found in 2 patients in this cohort. It has been shown previously to increase aortic vulnerability to additional risk factors for aortic aneurysm and dissection. Of note, neither patient carrying this variant had dilated thoracic aortas (1 male aged 42 with severe AR and no additional candidate variants identified; 1 female aged 70 with calcific AS and

additional candidate variants in *COL1A1* and *GATA5*). Overall, the rate of TAAs in patients with candidate variants in *MYH11* was no higher than in the rest of the cohort.

15.4.7.6 Patients with known family history of aortic disease

Interestingly, there was no difference in identification of candidate variants in patients with or without a family history of BAV or congenital disease, with 17/33 patients with family history having candidate variants, versus 74/122 patients with no known family history. The diagnostic yield was zero in both groups.

15.5 DISCUSSION

15.5.1 Clinical relevance

Genetic screening of genes known to be associated with BAV or aortic phenotypes has not identified any immediately clinically actionable variants in a tertiary cohort of 155 cases. Nevertheless, there are strong candidate variants identified such as two frameshift variants in NOTCH1 and a protein-truncating variant in KCNJ2.

Using in silico predictive tools, it is possible to identify candidate variants in 91 patients which are potentially worthy of further investigation. However, even if all of these proved to be causative mutations (which is highly unlikely), this approach would still only have identified a genetic cause in less than 60% of cases. Even in patients where BAV appears to be inherited in Mendelian fashion the clinically actionable yield from panel testing is zero, and candidate variants are identified only in 52%.

This gene panel was developed prior to more recent publications identifying further BAV-related genes such as *MATR3, SMAD6* and *ROBO4*. These more recent genes alone are unlikely to explain the missing heritability in the cohort. As with many other complex traits, despite all the advances in recent years with more easily accessible gene sequencing and the use of large cohorts to try to identify novel disease-causing genes, it is still not possible to explain inheritance even in our familial cases. There are many explanations for this; novel genes awaiting discovery (perhaps including *DCHS1*), inaccurate assessment of variant pathogenicity, regulatory region variation occurring outside the protein-coding exome, copy

number variation and the strong possibility that an oligogenic model of inheritance might be at play, rather than straightforward single-variant Mendelian inheritance. Additional haemodynamic considerations, perhaps most importantly during development, but also through later life, interact to define an individual's exact phenotype.

15.5.2 A note about family history and genetic screening

Clinical questioning about family history seems to be almost certainly underestimating the prevalence of familial BAV in this cohort - the number of familial cases in this cohort is low compared to previous publications. Family screening in this condition (as recommended by ESC and AHA guidelines[469, 470]) is patchy, and even documentation of family history is suboptimal. Both hospitals involved in this study have now established dedicated aortic clinics which can offer a more comprehensive screening service for such patients, including access to Clinical Genetics where indicated.

The findings of this study demonstrate the limited current role for routine gene sequencing in BAV; much more work in large cohorts is needed to identify risk alleles, new genes and possibly oligogenic models of inheritance before this would be fruitful. The eventual goal is to be able to carry out genetic screening of family members – and in particular, to discharge from follow-up those who do not carry causal or risk variants. However, there remains much uncertainty about the pathogenic role of most identified variants, in particular where interactions with other phenotypic modifiers are suspected, such as in the case of the *MYH11* (R247C) variant. Whilst the presence of this *MYH11* variant has been argued to merit closer screening and, more controversially, earlier intervention[471, 472], there was no evidence of a phenotype modifying effect in the two patients in this study who carry it.

15.5.3 The role of *in silico* predictions

In silico modelling is an interesting, relatively novel tool for assessment of genetic variants. Great care has to be taken with the interpretation of results. Several publications misleadingly classify variants as "likely pathogenic" on the basis of being found once or twice in a cohort and in silico predictions alone. This is an inappropriate claim for such variants; ACMG guidelines clearly delineate the very stringent classification rules which are needed for a variant to be clinically actionable "likely pathogenic" variants[273]; evidence less than this is

insufficient to make claims about the causal role of a variant in disease. In silico tools may, quite rightly, only be used as “suggestive” evidence.

However, with the huge expansion of gene sequencing and the numbers of new variants identified with each cohort, these tools can be a powerful adjunct to conventional *in vitro* functional gene assessment, which may be both costly, time-consuming and imperfect itself, due to the difficulty of accurately mimicking *in vivo* conditions. Combinations of concordant predictions from different in silico tools may be of use in predicting pathogenicity. At least for missense variants and where the predictors are concordant (as indicated by our combined score), these are fairly sensitive tests[274], although far from specific. It must also be noted that the stringent ACMG guidelines have been developed for assessing the contribution of highly-penetrant disease-causing variants to pathogenicity; they therefore are over-stringent for the assessment and analysis of the possible effects of variably penetrant risk alleles or phenotype modifiers. By using *in silico* scores to define variants which are likely to have an effect on gene function, rather than stringent pathogenicity criteria, it is possible to define a panel of variants which might well contribute to phenotype in our cohort, without being truly “causative” in a Mendelian sense. These scores are certainly imperfect assessments of the likely impacts of variants; for example, one of the "known" pathogenic variants in *APOE* models as "unlikely pathogenic" in all but one of the algorithms used in the current study. However, they can be used to filter down the variants and focus further research efforts on those most likely to have an effect on phenotype.

15.5.4 A novel candidate gene: *DCHS1* – further research ongoing

It was noted during this study that the one *DCHS1* variant classified as “pathogenic” in ClinVar (and indeed identified by segregation in the original scientific paper describing *DCHS1* variants in MVP) was present in both a father (proband who remained in our main BAV cohort) and his son who was recruited separately to the study, but who had been excluded from the subsequent general analysis. This prompted us to investigate this family further. There is one further family member with confirmed bicuspid aortic valve, and a very large family available for segregation studies. Recruitment of the wider family is ongoing, and we are poised to undertake Sanger confirmation of the variant (or absence of it) in as many family members as

possible. None of the other patients with *DCHS1* variants appear to have tractable family structures for segregation analysis.

Evidence from the original MVP data was examined to find any suggestions that *DCHS1* variants might cause abnormalities of the aortic valve as well. Having contacted several international groups, it appears that there is some emerging evidence from multiple angles for a potential role for *DCHS1* in BAV, and that the phenotype might be extremely variable. This fits well with the story in our family, where the father has "classic" BAV and the son has a more severely abnormal monocusp valve. A collaborative research effort with other groups is underway to define further the spectrum of cardiac abnormalities in patients with *DCHS1* variants, and the role of this gene in BAV pathogenesis.

15.5.5 Genotype:phenotype correlations

It is tempting to draw several conclusions from apparent phenotype clustering in the study population. Firstly, *APOB*, as a gene associated with atherosclerosis risk, appears to be associated with stenotic phenotypes in this cohort. This appears to be logically and scientifically consistent with *APOB*'s role as a mediator of atherosclerosis risk. Multiple lines of evidence support the role of apolipoprotein B in development of degenerative, calcific aortic stenosis. However, several of the "stenotic" phenotypes in the BAV cohort are rather more severe than one might expect from the influence of an atherosclerotic risk factor. Whilst, conceptually, one can argue that atherosclerotic risk factors can lead to endothelial dysfunction and valve calcification, it is rather harder to argue a mechanism for development of a major aortic coarctation or subaortic stenosis. This observation may therefore not reflect a major causal role for *APOB* in this phenotype. It is likely that the presence of *APOB* variants which correspond with an increased non-HDL lipid load could contribute to the more rapid calcification and sclerosis of a valve already predisposed to such a lesion; whether the magnitude of this effect is of clinical importance and therefore may be amenable to treatment such as statin therapy remains to be shown in larger studies.

Similarly, there is an apparent clustering of TAA in patients with variants in particular genes. MMP9 levels are associated with the presence of thoracic[197] and abdominal aortic aneurysm[473], and inhibitors of MMPs (doxycycline) as well as targeted gene disruption of

MMP9 have been shown to slow the growth of abdominal AAs in animal models[474]. However, data is lacking for a causative role of *MMP9* variants in thoracic aortic disease.

Finally, *DCHS1* variants seem also to be characterised by TAA also; this genotype-phenotype correlation remains to be fully investigated (see above).

15.6 LIMITATIONS

This study is a purely descriptive study. The cohort is small, with significant phenotypic variability, and therefore there is insufficient statistical power to undertake subgroup analysis or to make valid inferences about variant frequency. In addition, no functional or follow-up confirmation of *in silico* findings has been undertaken, either i*n vitro* or by segregation analysis. Therefore, assertions about the phenotype-modifying role of individual variants can be treated only as hypotheses and not as confirmed findings (see discussion above of the relative merits and limitations of in silico prediction tools). We hope to extend this work by selecting specific variants or genes of interest to investigate further using functional testing.

Of note is the fact that this cohort represents a specialist tertiary centre cohort, with much higher than standard prevalence of coincident abnormalities such as coarctation of the aorta, VSD and subaortic stenosis. To what extent the genetic variation found in this cohort reflects that of a standard population cohort of BAV patients is uncertain.

The sequencing panel was limited; firstly, it was developed prior to the identification of some important BAV genes such as *MATR3* and *SMAD6*, and, by definition, could not incorporate all genes of interest in BAV and aortopathy. In addition, it provided suboptimal coverage for one of the main genes of interest - *GATA5*, with only 80% of bases callable in this gene. This means that there may be "missed" pathogenic or candidate variants in this gene. This is most probably due to the presence of multiple sequence repeats in this gene, and therefore difficulty with alignment of reads. This makes *GATA5* one of the possible targets for re-sequencing. We also did not call copy number variants in our gene panel; the read length / coverage is insufficient to call them with our own sequencing results. These may be of importance in familial TAA, particularly in a paediatric population[475].

15.7 CONCLUSIONS

The current diagnostic yield of panel testing in a tertiary, non-syndromic BAV cohort is zero by current ACMG guidelines. Protein-truncating variants in NOTCH1 and KCNJ2 are strong candidates for pathogenicity. Potentially pathogenic variants have been identified in a number of candidate genes, with some suggestion of phenotype clustering with specific genes. *In silico* tools are useful, but insufficient, for prediction of variant pathogenicity. Variants in *DCHS1* are promising candidates for pathogenesis of varied valve disorders.

15.8 ACKNOWLEDGEMENTS

Dr Matina Prapa and Dr Nada Abdulkareem recruited the BAV cohort, under the supervision of Professor John Pepper and Professor Marjan Jahangiri. Sequencing and variant calling was carried out by Rachel Buchan and team at the Department of Cardiovascular Genetics and Genomics at Royal Brompton Hospital. Thanks to Dr Roddy Walsh for advice about quality control and pathogenicity interpretation.

16: AORTIC ENDOTHELIAL CELL TRANSCRIPTOMIC RESPONSE TO SHEAR STRESS

"Any living cell carries with it the experience of a billion years of experimentation by its ancestors" – Max Delbrück

16.1 INTRODUCTION

16.1.1 Shear stress

A paradigm shift in thinking about aortic biology has transformed thinking from considering the aorta as an inert, elastic tube, responsible only for a "Windkessel" effect, to its characterisation as a complete cellular "ecosystem," exquisitely responsive to external stimuli – in particular, the haemodynamic environment.

Key to this idea is mechanotransduction by different cells which make up the aortic wall; in particular, the endothelial cells which line the vessel lumen, and the vascular smooth muscle cells which form the contractile apparatus, alongside extracellular matrix components. Endothelial cells and vascular smooth muscle cells respond and adapt to different physical stimuli resulting from patterns of blood flow – predominantly shear stress and stretch or cyclic mechanical loading.

There has been particular interest in investigating the shear response of endothelial cells in the context of atherosclerosis[476]. Physiological changes in shear stress cause an acute vasomotor adaptation, with alterations in vessel diameter. Sustained shear stress variation induces gradual arterial wall remodelling, which may contribute both to arterial stiffness and to the process of plaque generation in atherosclerosis. It has long been recognised that atherosclerosis is particularly prone to develop at sites in the vascular tree exposed to different flow patterns[105, 477]. Specifically, exposure of endothelial cells to laminar, pulsatile shear stress appears to confer an atheroprotective phenotype, and exposure to oscillatory patterns of shear confers athero-prone properties on the vessel wall[478, 479]. This principle is illustrated well in the aorta, where the straight, tubular portions of the aorta are relatively protected from atherosclerosis, and branch points, where flow may be turbulent or non-laminar, are more athero-prone[105]. Physiological and anatomical variations in shear stress contribute to the spatial heterogeneity of endothelial cell gene expression[79]. These flow-

responsive, spatially-distinct mechanisms may not only be of importance in atherosclerosis in later life, but also during development, when the biomechanical milieu can influence arteriovenous differentiation and vascular bed formation[480]. There is also evidence that patterns of stretch and shear might influence development of other vascular pathology, such as the aortopathy associated with bicuspid aortic valve, in which haemodynamic factors play a major role[481].

16.1.2 Limitations of previous shear stress studies

The mechanisms of this polarisation of endothelial cell phenotype have been the subject of much investigation. Until recently, studies have been limited by the difficulties of maintaining endothelial cells in culture for prolonged periods of time, and also by restriction of investigation to a few genes or molecules of interest.

Whilst there is much published on endothelial cell response to shear, and the resultant polarisation of endothelial cell phenotype, research efforts are hampered by several factors. *In vivo* experiments make quantification and characterisation of shear stress difficult, and it is particularly hard to control for other mechanical influences on different cell types within the vessel wall, and for hormonally-mediated effects. On the other hand, *in vitro* experiments have been rather inconsistent in results[482] – probably in part due to differing characteristics of shear response in endothelial cells derived from different sources. Not all endothelial cells are the same – a fact often glossed over by researchers who favour easily-obtainable cellular models such as human umbilical vein endothelial cells (HUVECs). However, these have differing baseline transcriptomes from arterially-derived cells, and respond differently to shear stress[483, 484]. It therefore becomes difficult to draw conclusions from comparisons across cell types. Additionally, models for generating shear stress have differed – from cone-plate systems to tubular flow models, with differences in the magnitude and pattern of flow generating the shear stress (see section 16.1.5 below). Very few studies have examined chronic exposure to shear stress *in vitro*– almost all the experimental data examines shear stress up to 24 hours of cell culture. This is primarily due to the difficulties of maintaining endothelial cells in culture for longer periods. Finally, most research has been limited to a few genes or molecules of interest, making it difficult to replicate, compare or generalise results to create an overall picture of the endothelial transcriptomic response.

16.1.3 Molecular Biology of the shear stress response in endothelial cells

Initial studies of shear stress applied to endothelial cells *in vitro* demonstrated rapid release of prostacyclin[485, 486]and nitric oxide (NO)[487, 488] on exposure to laminar shear stress; two substances key for vasodilatation and endothelial homeostasis. In addition to this acute vasoregulatory release of factors, sustained changes in blood flow can lead to arterial wall remodelling via alterations in endothelial gene expression[482, 489, 490]. This blood flow-sensing response controls multiple functions; the regulation of blood vessel tone and blood pressure; the regulation of inflammation and local immune responses; regulation of cellular proliferation and also, during development, cardiovascular system differentiation[79].

NO generation by endothelial NO synthase has long been recognised as a key prerequisite for maintenance of endothelial cell function[491], and regulation of eNOS transcription and activity is central to endothelial shear stress response[492, 493]. The regulation of basal eNOS activity is complex, and involves both control of expression levels and also post-translational modification by phosphorylation[494]. Multiple protein kinase signalling cascades appear to contribute to this. Jin et al[493] found that laminar shear stress, both *in vitro* and *in vivo*, activated vascular endothelial growth factor receptor 2 (VEGF2), which in turn initiated signalling via phosphoinositide 3-kinase and activation of the serine/threonine kinase Akt to activate and induce eNOS via phosphorylation. The relevance of these pathways in vivo is demonstrated by the fact that *eNOS* knockout mice are hypertensive[495], and rings of aortic tissue from these mice fail to relax in response to acetylcholine, demonstrating a role for basal eNOS activity in maintenance of vascular tone. Additionally, *eNOS* transcription and expression has been found to be reduced in regions of the mouse aorta which are atherosclerosis-prone[496].

A key transcription factor identified as mediating the shear stress response in endothelial cells is Kruppel-like factor 2 (KLF2). This positively regulates *eNOS* expression. The expression of *KLF2* itself is increased by laminar shear stress via a signalling cascade involving extracellular-signal-related-kinase 5 (ERK5), myocyte enhancer factor 2, AMP-activated protein kinases (MEK5/ERK5/MEF2 pathway) and *miR-92a*[497-501]. These influences act on a shear stress response element in the promoter of *KLF-2*. KLF-2 not only induces *NOS3 (eNOS),* but also up-

regulates *thrombomodulin* and *haem-oxygenase 1 (HO-1)*[251, 502]; both cytoprotective mechanisms.

The anti-inflammatory properties of the laminar shear-exposed endothelium have also been investigated. Inflammatory activation was found to be suppressed by another key transcription factor, Nrf-2 (NF-E2-Related Factor 2), which regulates mitogen activated protein kinase (MAPK) signalling[503]. JNK and p38, both members of the MAPK family, induce pro-inflammatory cytokines[504] and adhesion molecules[505], and enhance cellular apoptosis[506]. Nrf-2 down-regulates this MAPK signalling both *in vivo* and *in vitro*[507], and induces a set of antioxidant genes including *HO-1, NADPH:quinine oxidoreductase-1, glutathione reductase, ferritin heavy chain* and others[508].

Induction of *MKP1* (also known as *DUSP1*) has also been shown to result from exposure of endothelial cells to laminar shear, and again down-regulates MAPK signalling pathways which are pro-apoptotic[507]. Laminar shear also promotes endothelial cell senescence, down-regulating proliferation pathways such as that mediated by mTOR.

Another key inflammatory signalling pathway; NFκB, is down-regulated by shear stress. When active, the NFκB pathways induce expression of the pro-inflammatory cytokines IL6 and IL8, as well as adhesion molecules including E-selectin, VCAM1 and the macrophage chemoattractant protein MCP-1[509-511]. Increased expression of *NFκB* has also been demonstrated *in vivo* at atheroprone sites in the mouse aorta, demonstrating the importance of this pathway in determining endothelial phenotype[512].

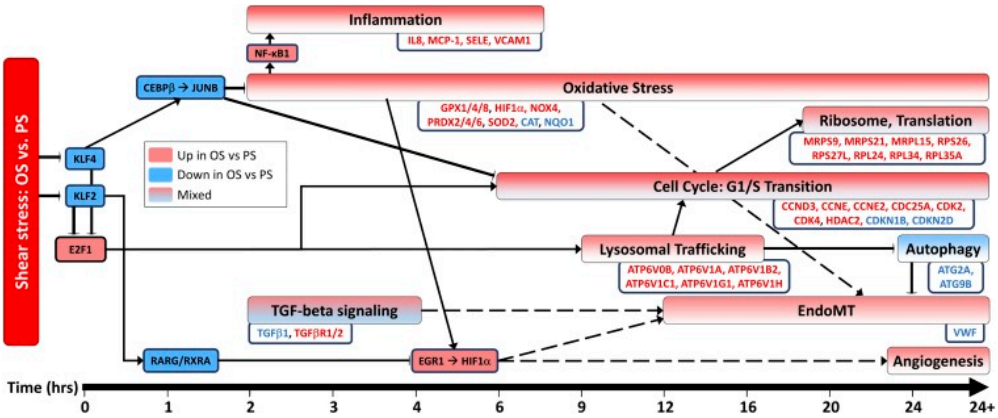
The complexity of the shear stress response is great; indeed it seems that there are multiple levels of cross-talk between all the above pathways[79, 512, 513]. A more global view of the pathways involved in endothelial cell response to shear stress is provided by the relatively recent techniques of transcriptomics. These have been used in a variety of in vitro and in vivo models to assess the pathways and functional characteristics of differential gene expression.

Perhaps the most comprehensive assessments of shear response to date are those published by Qiao et al in 2016[514] and Ajami et al in 2017[513]. They exposed primary human coronary artery endothelial cells (HCAECs) and HUVECs respectively to either oscillatory or laminar shear stress (OS / LS). Both groups used RNA sequencing to obtain data for global transcriptomic analysis. Qiao et al used a cone-plate system for shear stress and Ajami et al

used a parallel plate system, whereby one can fine-tune the magnitude and pattern of shear to which the cells are exposed. Qiao et al reported multiple genes regulated by shear stress patterns. The focus of their paper was on the similarity between cells cultured under static conditions and oscillatory shear stress. However, there was a comparison of OS with LS at 24 hours of shear. This confirmed altered expression of many known shear-sensitive genes such as *Kruppel-like factor 2 (KLF2), nitric oxide synthase 3 (NOS3), angiopoietin-2* (*ANGPT2*)[515], and *vascular cell adhesion molecule 1 (VCAM1)*[516]. They additionally identified many novel shear-sensitive genes and long non-coding RNAs, such as *KLF11, TEK tyrosine kinase (TEK), serpin peptidase inhibitor member 2 (SERPINE2), chromosome 10 open reading frame 10 (C10orf10), Ephrin-A1 (EFNA1), chemokine (C-C motif) ligand 14* (*CCL14*), and *lncRNA metastasis-associated lung adenocarcinoma transcript 1* (*MALAT1*). Ajami et al[513] examined in more detail the pathways regulated by shear pattern using temporal changes in gene expression up to 24 hours. They identified multiple transcription factor networks which were differentially regulated. They confirmed key roles for differential regulation of oxidative stress pathways, with genes encoding anti-oxidative mediators being downregulated by oscillatory shear (e.g. *catalase*, *G6PD, NQO1*)and those mediating superoxide generation being upregulated (e.g. NOX4, SOD2). However, the story is not quite this simple: some anti-oxidant genes: glutathione peroxidases *GPX1, GPX4*, and *GPX8*, peroxiredoxins *PRDX2, PRDX4*, and *PRDX6*, and metallothioneins such as *MT1X,* are upregulated by oscillatory shear. HIF1A seems to be an important transcription factor in mediating the response to oxidative stress, and is again upregulated by hour 4 in response to oscillatory shear.

The group found similarly conflicting patterns when looking at TGF-ß signalling pathway genes. This pathway is one of the key drivers of endothelial-to-mesenchymal transition, thought to be a key process in atherosclerosis generation in response to shear stress.

FIGURE 16.1: Genes and pathways showing differential expression up to 24 hours between laminar and oscillatory shear stress (reproduced from Ajami et al)[513]



The x axis here shows the time course of the shear experiment. Red colours indicate upregulation of genes or pathways by oscillatory shear stress; blue colours denote down-regulation.

The time-course of these expression changes is also of interest – genes have distinct temporal lags before activation or repression. Of particular note is that even by 24 hours, the response to shear is not "stable" – in other words, there are still significant differences between gene expression and pathway activation or repression at each timepoint, as illustrated in Figure 16.1 above. It is therefore difficult to draw conclusions about chronic response to shear from these i*n vitro* experiments. Additionally, the Ajami paper used HUVECs as their cellular model – cells which have different basal transcriptomes and different mechanotransductive responses to arterially-derived cells.

There remain therefore significant gaps in our understanding of shear response. None of the *in vitro* shear models have compared laminar versus turbulent shear effects beyond 48 hours. The literature on the role of shear stress in the specific context of the aorta is confounded by the use of multiple different lineages of endothelial cells in shear experiments; most commonly HUVECs. In addition, characterisation of the pathways and gene ontologies which differentiate the two types of shear stress (oscillatory versus laminar) has been limited by very few global expression studies. This study attempts to address these gaps by assessing the differential expression of genes in response to pulsatile laminar vs oscillatory shear up to 6 days of culture.

16.1.4 Shear stress and aortic disease

The relevance of these shear experiments to aortic disease comes from several directions. Firstly, haemodynamic influences are thought to be of vital importance in cardiovascular system development. Indeed, it has been shown that shear-responsive genes *KLF2, eNOS* and *ET-1* play a key role in the development of the cardiovascular system[246, 247, 517]. Therefore, it is conceivable that pathological changes in the genes responsible for shear stress response could cause structural cardiovascular abnormalities, such as that seen in bicuspid aortic valve.

Secondly, aortic structural and functional changes seen with ageing represent a complex remodelling process which is known to be driven, at least in part, by mechanotransduction and sensing by vascular cells of different haemodynamic conditions[79].

Thirdly, the aortopathy seen in BAV is known to represent a complex interplay between haemodynamic and genetic factors[182]; an interplay which is likely to be mediated by the vascular endothelial cells which are the primary mechanosensors of shear stress[182, 255, 256]Secondly, evidence shows changes in gene expression in the BAV aorta of several genes known to be regulated by flow, such as *eNOS*[252], MMPs[233, 237, 253], *PKD-2*[254] and genes in the TGF-β signalling pathway. *eNOS* knockout mice have a high incidence of bicuspid aortic valve[199]. Also, endothelial function has been shown to be impaired in patients with BAV[80]. These pieces of evidence together indicate that endothelial response to altered patterns of shear stress in BAV may be at least in part responsible for the aortopathy seen in that condition. Therefore, understanding the normal endothelial response to shear is vital for understanding how this process may malfunction in the diseased aorta.

16.1.5 Study models to assess effects of shear stress

A range of different experimental techniques have been used to identify the response of the endothelium to different patterns of shear stress. *In vivo* systems involve either en face imaging of vascular endothelial cells, or isolation of vascular endothelial cells for analysis; a process which is fraught with difficulty and in which results may be confounded by the presence of additional vascular cell types. Ni et al[518] examined the response of mouse carotid endothelium to disturbed flow by ligating one of the carotid arteries, and examining gene expression by microarray at 12 and 48 hours after ligation. They identified many novel mechanosensitive genes. Passerini et al[519] examined the transcriptomic profile of different

regions of porcine aorta, from athero-prone versus athero-protected regions, again finding multiple putative shear-regulated genes. However, these *in vivo* studies suffer from the problems of other cell types contributing to the results, and confounding from possible additional haemodynamic factors[482].

Multiple different model systems have been developed to expose cells in vitro to carefully controlled levels and patterns of shear stress. These include a cone-and-plate flow system, in which a teflon cone is rotated in the centre of a culture plate, thus causing the medium to flow over the surface of cells in a way which exposes the cells to uniform shear conditions[520]. An alternative model is to use a parallel-plate flow chamber-based system[521]. The advantage of these systems is that the magnitude of the shear stress is easily controlled. However, there are several limitations of this approach. Firstly, the shear stress is sometimes not pulsatile as *in vivo*. Secondly, different "batches" of cells have to be grown under different conditions – so any small differences in confluence or conditions could affect the results. Thirdly, there are temporal limitations to the duration of these experiments as the cells cannot easily be maintained in these culture conditions, and cultures using this technique have thus generally been limited to 48 hours of culture. Fourthly, cells are cultured in a monolayer and therefore the conditions do not mimic the *in vivo* environment, in which there is interaction with layers of vascular smooth muscle cells and a complex extracellular matrix.

A third *in vitro* method was developed which addresses the first three of these limitations; the orbital shaker, in which cells are cultured in standard plates on a device which moves in a single plane, "swilling" the culture medium over the top of the cells. This has the benefit of creating a range of shear conditions in one plate; the cells in the centre of the wells are exposed to oscillatory, non-directional shear, and those at the edge to laminar, pulsatile shear. Previous research has shown that the magnitude of the laminar shear stress at the edge of the well is analagous to that seen in the arterial system (approximately 11dynes/cm2)[522]. In the centre of the well, the shear stress was much lower and multi-directional (oscillatory). In addition, endothelial cells at the centre of the well appear to have the phenotypic characteristics of those exposed to oscillatory shear *in vivo*; they have increased proliferation, increased apoptosis and reduced Akt phosphorylation, whereas those at the periphery align with flow and demonstrate an anti-inflammatory phenotype[522].

16.1.6 Transcriptomics: microarray and RNA-sequencing

Defining cellular responses to stimuli has been hugely enriched by the development of transcriptomic methods. These, much like genome-wide association studies, allow hypothesis-free testing, and assimilation of results across pathways. Different methods are available for gene expression analysis: primarily RNA-sequencing and microarray assays. Large-scale transcriptomic datasets such as GTEx[28] make use of both techniques for data generation. More recently, single-cell transcriptomic methods have enabled even finer resolution of cellular differences in gene expression. Like GWAS, transcriptomics suffers from the multiple-testing problem: large effect sizes are required to reach genome-wide significance. At the time of writing, RNA sequencing has somewhat overtaken microarray transcriptomics as the method of choice, for several reasons[523]. Firstly, it provides a much more comprehensive view of the cellular transcriptome, including non-protein-coding transcripts such as lncRNAs and transcribed pseudogenes. Secondly, RNA-seq is unbiased by detection platforms. RNA-seq also demonstrates a broader dynamic range than microarray, allowing superior detection of low abundance transcripts, differentiating isoforms and allowing identification of more differentially expressed genes. There is also an advantage of RNA-sequencing in avoiding some technical issues which arise from the method, including cross- or non-specific hybridization, limited detection range and the need to annotate and pre-define probes. However, RNA-seq remains expensive and requires significant bioinformatics expertise and infrastructure for alignment, interpretation and data storage. Microarrays still form an accessible and broad overview of gene expression, being particularly designed to cover the protein-coding exome. At the time of inception of the experiments described in this thesis, funding and bioinformatics experience within the group was limited, and therefore microarrays were chosen for this project, with a view to progressing to RNA sequencing for future work arising from this initial data.

16.2 HYPOTHESIS

Exposure to different patterns of shear stress affects gene expression in aortic endothelial cells in ways which could contribute to pathology.

16.3 AIMS

1. To develop a protocol for culture and isolation of DNA from human aortic endothelial cells exposed to different patterns of shear stress
2. To define the impact of prolonged shear stress pattern on gene expression and identify key genes which might drive the phenotypic response to haemodynamics and genetic variation

16.4 METHODS

16.4.1 Cell culture

Human aortic endothelial cells (HAECs) isolated from deceased donors were purchased from Lonza (New Jersey, USA) as cryovials containing >500,000 cells at passage 3. These came from three separate donors, details of whom are given in Table 16.1.

TABLE 16.1: Details of HAEC donors

Donor code	Age of donor	Gender of donor	Smoking status of donor	Ethnic origin of donor
0000227764	35	M	Non-smoker	Caucasian
0000239151	18	F	Non-smoker	Caucasian
0000316663	49	M	Non-smoker	Caucasian

16.4.2 Cell maintenance and passage

Tests performed by the supplier indicate that these HAECs are positive for the endothelial cell marker von Willebrand Factor VIII (vWF) and acetylated low density lipoprotein uptake, and negative for α-smooth muscle actin. We also confirmed that these HAECs displayed typical morphological phenotypes of endothelial cells, adopting a “cobblestone” morphology when grown under static conditions, and an “aligned” morphology when grown under shear stress.

The HAECs in cryovials were plated according to manufacturer’s instructions. Four 25 cm^2 culture flasks were pre-prepared with 5ml of warmed EGM-2 complete medium (Lonza). The cryovial containing the cells was removed from liquid nitrogen and thawed briefly in a waterbath at 37^oC. Contents were immediately distributed between each of the four pre-

prepared flasks, and these were then placed in an incubator at 37°C and 5% CO2. After 24 hours, medium was changed to remove the cryopreservant; and thereafter, medium changes were carried out every 48 hours.

Cell passage was carried out when cells reached 80% confluence. The medium from the confluent flask was removed and the cells were rinsed twice with Hepes- buffered saline (HBS) at room temperature. The cells were treated with 2ml of Clonetics trypsin and ethylenediamenetetraacetic acid (EDTA) preparation for 2 minutes at room temperature. After cells became rounded and non-adherent, the trypsin preparation was neutralised with the Clonetics trypsin neutralising agent at 4ml per flask. If cells remained adherent, they were dislodged by gently tapping the flask. This solution was aspirated and transferred into a sterile falcon tube. This was spun at room temperature and at 200g for 5 minutes to produce a pellet of cells. The supernatant was discarded and the pellet re-suspended in 1ml medium. Cells were counted using a haemocytometer and re-plated as appropriate.

Cells from passage 4-6 were plated onto 6 well plates for the remainder of the experiment.

16.4.3 Shear stress

Cells were subjected to different patterns of shear stress using an orbital shaker model as described above (16.1.5 and below).

16.4.4 Protocol development

16.4.4.1 Defining centre vs edge of well

The orbital shaker model of shear stress (see Figures 16.2, 16.3 for more details) mimics the full range of shear stress experienced by aortic endothelium. Cells growing around the edge of the well are exposed to high levels of directional, or laminar shear stress, and cells in the centre are exposed to non-directional, or oscillatory shear. Modelling from our group has previously characterised the shear stress exerted on the endothelial cell monolayer at different locations in the wells of the 6 well plate (see Figure 16.2). This showed that the peak of laminar shear stress around the edge of the plate had diminished within a distance of 14mm from the centre, and cells in the very centre of the plate experience oscillatory, disordered shear.

FIGURE 16.2: Computational fluid dynamics modelling of the instantaneous shear stress experienced by endothelial cells in a standard 6-well plate on the orbital shaker at 200rpm.

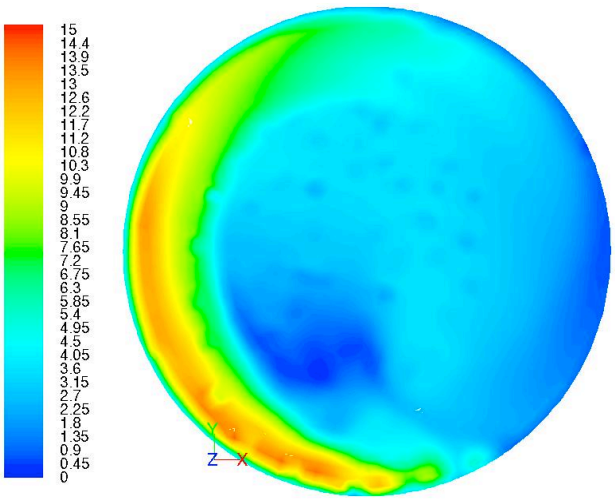
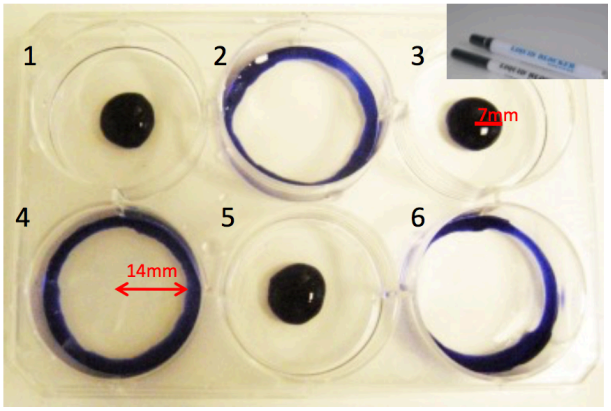


Figure from Claire Potter PhD Thesis: Role of Chronic Shear Stress in Endothelial Form and Function: Imperial College, 2012, with thanks. The figure shows the circular well in which the cells are cultured. The red and orange colouring demonstrate the higher, pulsatile and laminar shear stress exerted on the endothelial monolayer by the "wave" of medium swilling around the plate. In the centre, shear stress is low and non-laminar, or oscillatory (represented by the blue colours)

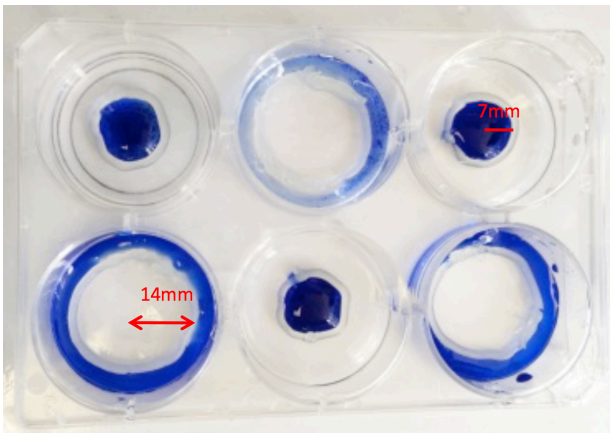
FIGURE 16.3: Photograph of orbital shaker equipment

In order to investigate the differing transcriptomic responses to these two types of shear stress, it was important to develop a way to isolate the RNA from only the centre or the edge of the well, without disturbing the fluid flow or shear patterns the cells experienced. The best plan seemed to be to subject the whole well to shear stress, and then to create a physical barrier to restrict the RNA lysis buffer to one area of the plate. We attempted to achieve this with different substances (See Figure 16.4), including a PAP pen, but this resulted in spillage of the lysis buffer into the non-target areas of the plate (visualised with ink added to lysis buffer). We eventually achieved target zone isolation using vaseline which had been autoclaved, syringed from a 20ml syringe into a circle onto the surface of the cell monolayer. The circle was "drawn" in this way, following a laminated template which was placed underneath the well.

FIGURE 16.4 attempts at isolating RNA from different zones of the 6-well plate



a. *First attempt at defining centre vs edge of wells in a standard 6-well plate. A PAP pen (pictured in insert) was used to draw a hydrophilic boundary after shear exposure at diameters defined by fluid dynamics modelling (7mm or 14mm from centre). 350µl RNA lysis buffer with a few drops of ink was carefully added to the centre or edge of the wells. No spillage occurred initially, but on scraping the well to mimic extraction of RNA, spillage did occur.*



b. *Second attempt at defining centre vs edge of wells. A 20ml syringe was used to syringe circles of petroleum jelly (Vaseline) onto the wells, using a laminated template placed under each well as an outline. 350µl RNA lysis buffer with a few drops of ink was carefully added to the centre or edge of the wells as before. No spillage occurred, and it was possible to scrape the wells without spillage too. Initial outlines were rough, but these improved with practice.*

I confirmed the phenotype of the cells from each region visually, looking for alignment with light microscopy and fluorescence microscopy, using staining with DAPI and VE-cadherin as per lab protocols.

16.4.5 RNA extraction

Cells from each donor were plated and cultured under either static conditions or subjected to shear stress using the orbital shaker model as described above. Cells were cultured and DNA extracted at different time points; 24 hours and 6 days for the differential shear expression analysis. RNA lysis buffer was applied to the plates as per manufacturer's instructions, and according to the zones defined above, and RNA extracted with Qiagen RNEasy mini-kit. Samples were scraped using either a pipette tip or cell scraper. 10 µl β-mercaptoethanol was added to 1ml Buffer RLT, and 350 µl of this solution was added to the sample, which was then homogenised by passage through a needle and syringe multiple times. RNA extraction was then performed as per manufacturer's instructions, using binding columns and centrifugation. DNAse digestion was also performed on-column. RNA purity was checked with a Nanodrop machine (see Table 16.2 for representative sample of RNA obtained from donor1, below).

TABLE 16.2 : RNA concentration and purity obtained from each sample (donor 1)

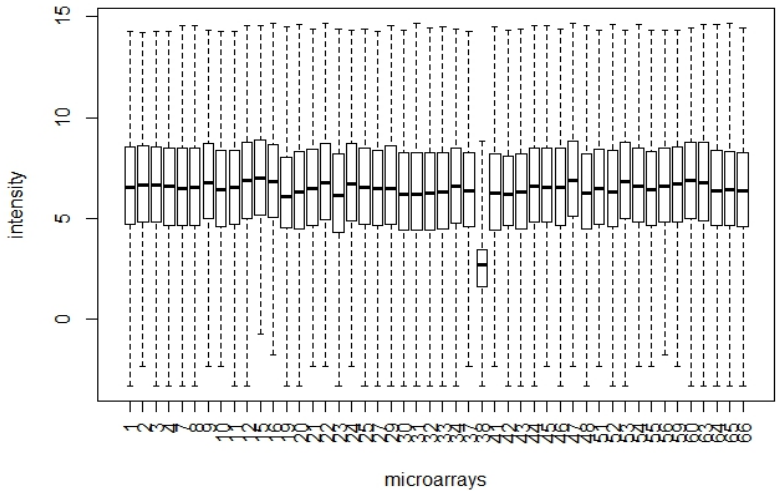
Sample number	Sample type	RNA conc. (ng/µl)	Purity (260:280)
1	1 hr shear (whole well)	111.8	2.06
2	1 hr static (whole well)	124.9	2.00
3	24 hr shear (whole well)	198.3	2.08
4	24 hr static (whole well)	193.8	2.06
5	24 hr shear edge	77.9	2.05
6	24 hr shear centre	49.6	2.03
7	24 hr static edge	72.9	2.07
8	24 hr static centre	28.8	2.01
9	72 hr shear (whole well)	271.9	2.08
10	72 hr static (whole well)	435.0	2.08
11	144 hr shear (whole well)	162.6	2.08
12	144 hr static (whole well)	480.4	2.09
13	144 hr shear edge	98.2	2.05
14	144 hr shear centre	62.3	2.03
15	144 hr static edge	105.4	2.07
16	144 hr static centre	58.6	1.94

RNA samples were frozen at -80C and sent to Source Bioscience for analysis, using the Illumina Human HT-12 V4 microarray. We received back expression data files, which had been background-corrected but not normalised.

16.4.6 Transcriptomic analysis in R

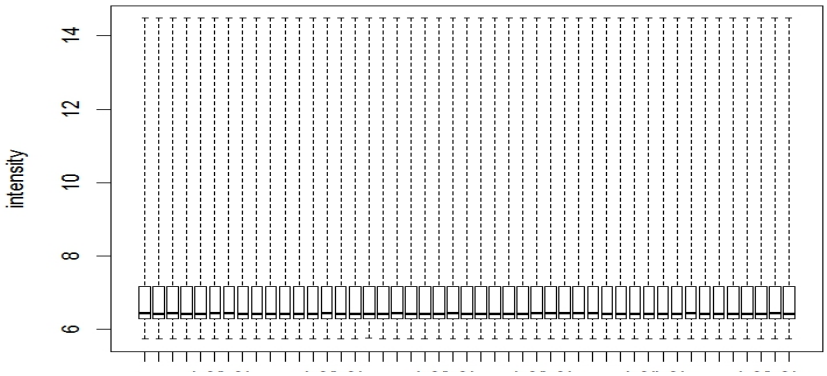
I performed analysis in R, using the lumi package for annotation and initial processing of the data. One sample was removed as the expression levels across the chip were consistently lower, and this was deemed to be a processing problem with the chip. See figures below for pre-normalisation box-plot including the outlying sample (number 38; donor 2, 144 hour shear edge) and post-normalisation box-plots.

FIGURE 16.5 Boxplot showing failure of array number 38 - expression data not globally in line with other samples and therefore this was removed from subsequent analysis.



Variance-stabilising transformation was applied across all samples, using the lumiT function in R. A robust spline normalisation algorithm was then applied to normalise the data, to enhance the power to detect significant differential expression.

FIGURE 16.6: Boxplot of array intensity post-VST transformation and RSN normalisation



I annotated probes from the lumiHumanAll database. Statistical model design was performed using linear regression models in the limma package. The donor was set up as a dependent variable in the regression model, with other variables being time (24 hours vs 6 days) and type of shear (laminar vs turbulent). Comparisons were made between turbulent (centre) and laminar (edge) shear at 24 hours and 6 days.

Multiple testing correction was applied using the Benjamini-Hochberg method, and genes were deemed to be significantly differentially expressed if the adjusted p value was <0.05. We opted not to specify fold-change cut-offs for our differential expression analysis, as many of our genes of interest represent transcription factors or members of signalling cascades, whereby small fold-change values may be of great biological significance. We also recognise that the cells are cultured under very similar conditions; both groups in the analysis are subjected to mechanical stimuli, and therefore the degree of change between the two conditions may be small.

eNOS (*NOS3*) was pre-selected as a known shear-induced gene to check our model and validate our results.

Gene ontology analysis (GO) was performed with the topGO package in R. Network analysis was undertaken in STRING 10.

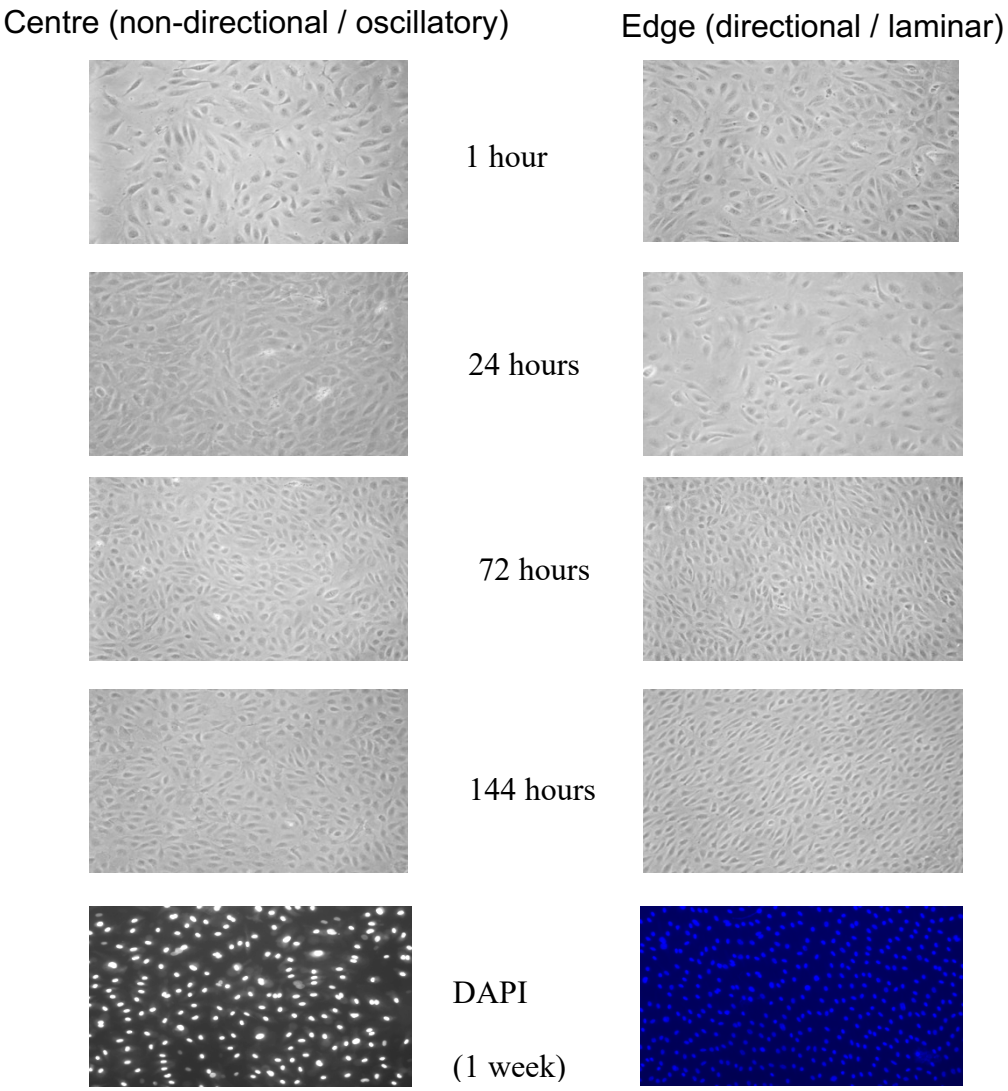
16.5 RESULTS

16.5.1 Cell morphology

The cells grown under conditions of shear demonstrated expected morphological changes with laminar shear (at the edge of the wells), becoming elongated and aligned with direction of flow. Cells at the centre of the wells exposed to oscillatory shear retained a cobblestone appearance which was not distinct from the cells cultured under static conditions (Figures 16.7 and 16.8). Interestingly, the morphological changes were not fully apparent until 144 hours of shear (see Figure 16.7).

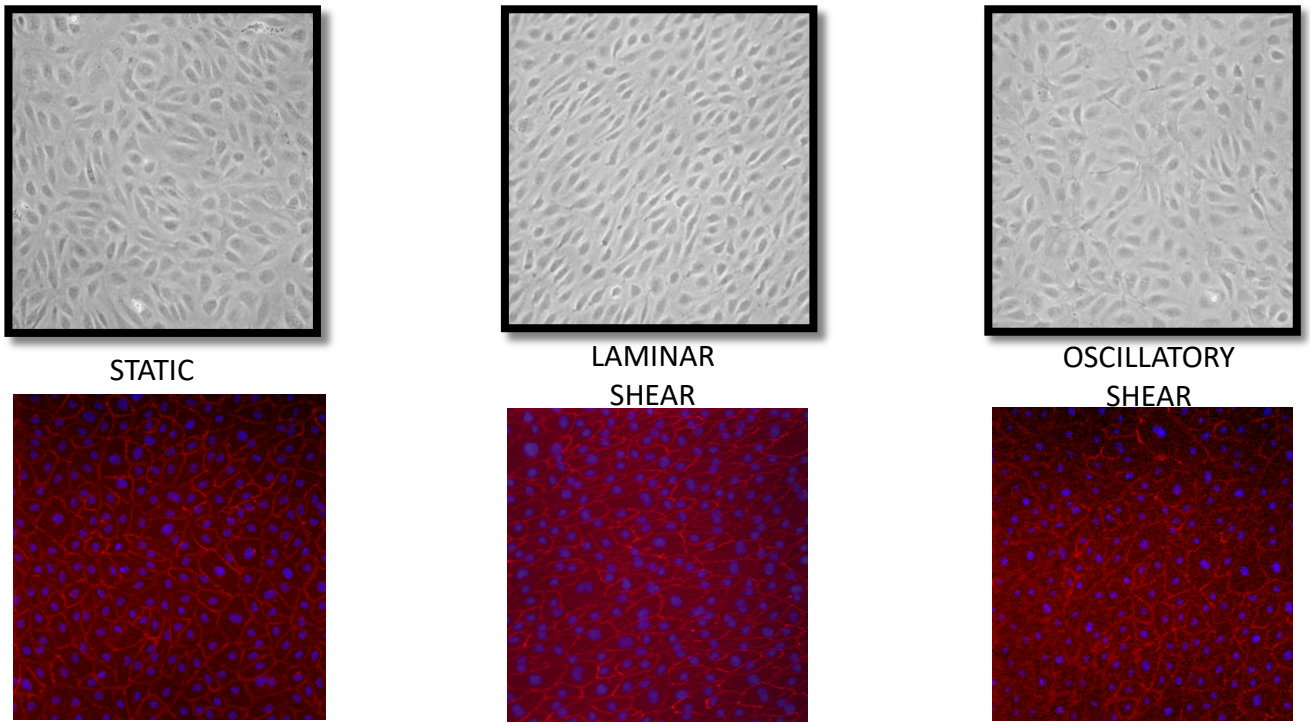
We confirmed these morphological changes and endothelial cell phenotype by staining with VE-cadherin and DAPI (Figure 16.8), which confirmed that at 144 hours, the cells cultured in the centre of the wells were indistinguishable morphologically from those cultured under static conditions, whereas those from the edge of the wells were elongated and aligned.

FIGURE 16.7 HAECs undergo morphological change in response to laminar shear after 72-144 hours



Human Aortic Endothelial Cells (HAECs) exposed to different patterns of shear stress in our model display expected morphological changes, with those exposed to laminar shear elongating and aligning with flow, and those exposed to oscillatory shear remaining rounded and "cobblestone" in appearance. The bottom panel is taken using a Cellomics fluorescence microscope, with DAPI staining; the remainder of the images are light microscopy . Images taken at different time points from either edge (within 10mm of edge of well) or centre (within 7mm of centre of well) ; corresponding to exposure to non-directional / oscillatory shear stress (centre) or directional / laminar shear stress (edge).

FIGURE 16.8: Morphological changes at 144 hours of culture: comparison of static with laminar and oscillatory shear



Top panels: light microscopy images (10x LEICA) : cells under static or oscillatory shear conditions retain a cobblestone appearance with no uniformity of cell polarity. When grown under laminar shear, the cells elongate and align.

Bottom panels: Cellomics fluorescence microscope images with staining for VE cadherin (red) and DAPI (blue), showing the same morphological changes.

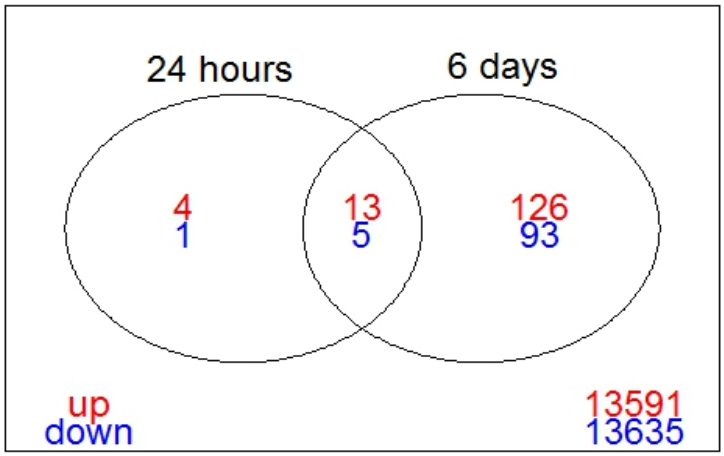
16.5.2 Transcriptomic Analysis

We firstly validated our model, showing that *eNOS* was, as expected, up-regulated in laminar shear as opposed to static conditions (laminar shear versus static at 144 hours: $\log_2$ fold change = 1.5, adjusted p=0.0025).

We also examined whether there was any difference in gene expression of cells cultured under static conditions, between the edge and centre of the well. There were no genes which were significantly altered here at either timepoint, showing that any transcriptomic differences between the edge and centre of the sheared cells are due solely to the mechanical forces, and not to the culture conditions or any difference in harvesting of the RNA.

We examined differential gene expression between cells exposed to laminar and oscillatory shear. At 24 hours, there were just 23 genes which displayed significant differential expression after multiple testing correction. By 144 hours (6 days), there were 237 differentially expressed genes

FIGURE 16.9: Venn diagram illustrating numbers of genes which were significantly up-regulated in oscillatory shear compared with laminar shear (red) or down-regulated (blue) at 24 hours and 6 days



16.5.3 Differentially expressed genes: 24 hour timepoint

We identified only 23 genes which were significantly differentially expressed at this timepoint (see Table 16.3). This has many potential explanations; the most likely of which is that we lack statistical power to identify more genes, with just 3 biological replicates per condition. It is also true that in our model, there is a gradient of shear. Whilst this is more representative of the gradients of shear stress found *in vivo*, it might also eliminate some of the differences between the two shear conditions. Given previous results from endothelial cells exposed to shear, we would expect a more significant transcriptomic shear response by 24 hours, although many previous publications use a much more lenient cut-off for significance (raw $p<0.05$[513]).

However, the 10-fold increase in number of significantly shear-regulated genes seen by 6 days (to 237) does suggest that the magnitude and scope of the shear response increases up to this timepoint.

TABLE 16.3: Differentially regulated genes at 24 hours; fold changes represent change from laminar to oscillatory shear (i.e. a positive fold-change represents genes which are up-regulated by oscillatory shear conditions)

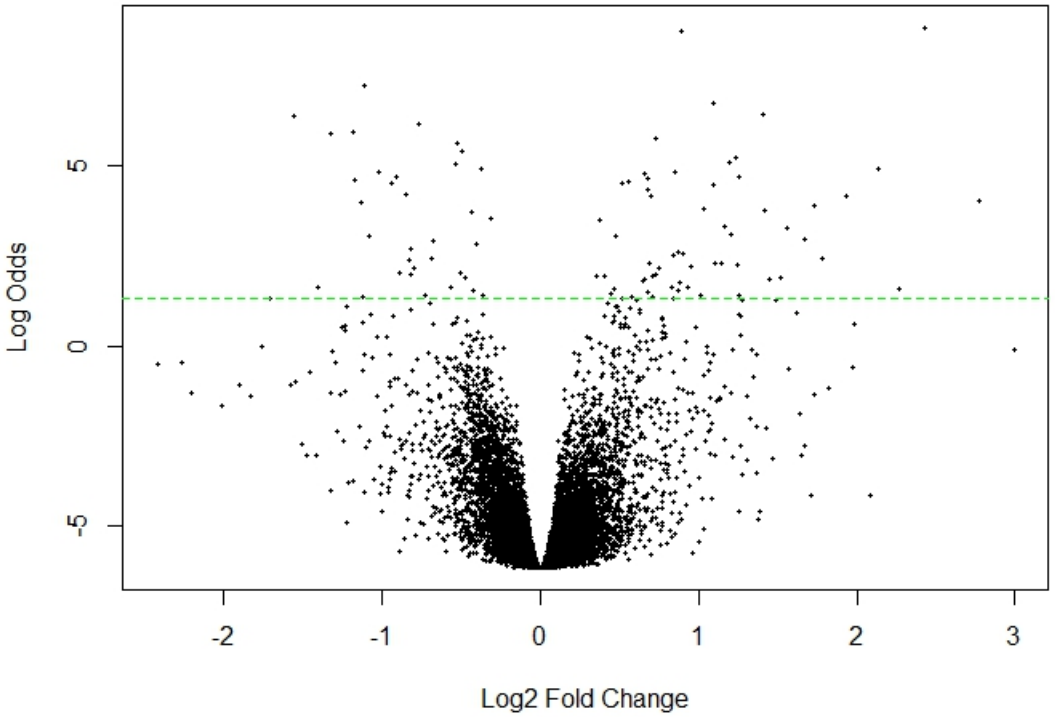
Gene symbol	Gene Name	log2 Fold Change	Average expression	t statistic	Adjusted p value
LPCAT4	*lysophosphatidylcholine acyltransferase 4*	-0.54	7.09	-6.47	0.003
C10orf10	*chromosome 10 open reading frame 10*	3.06	9.98	6.08	0.005
TCF4	*transcription factor 4*	1.17	9.56	5.66	0.011
GLUD1	*glutamate dehydrogenase 1*	-0.36	10.49	-5.48	0.015
EXTL3	*exostosin-like glycosyltransferase 3*	0.58	8.21	5.26	0.022
ID1	*inhibitor of DNA binding 1, dominant negative helix-loop-helix protein*	-1.60	10.27	-5.18	0.022
ADD3	*adducin 3 (gamma)*	1.16	9.27	5.15	0.022
TNFRSF4	*tumor necrosis factor receptor superfamily, member 4*	1.33	7.16	5.00	0.030
WASF2	*WAS protein family, member 2*	0.76	8.43	4.89	0.034
LRIG1	*leucine-rich repeats and immunoglobulin-like domains 1*	0.96	9.17	4.87	0.034
ADM	*adrenomedullin*	1.43	10.54	4.85	0.034
SFR1	*SWI5-dependent recombination repair 1*	-0.60	7.89	-4.80	0.035
GOSR2	*golgi SNAP receptor complex member 2*	-0.74	8.82	-4.77	0.035
AIM1	*absent in melanoma 1*	0.47	6.98	4.76	0.035
HLX	*H2.0-like homeobox*	1.20	7.33	4.71	0.035
CLDN5	*claudin 5*	1.37	11.19	4.70	0.035
CHST1	*carbohydrate (keratan sulfate Gal-6) sulfotransferase 1*	1.23	7.85	4.69	0.035
LOX	*lysyl oxidase*	1.55	9.23	4.67	0.036
ZFAND5	*zinc finger, AN1-type domain 5*	0.57	11.76	4.63	0.039
FAM89A	*family with sequence similarity 89, member A*	1.62	9.25	4.56	0.045
LXN	*latexin*	1.30	11.32	4.54	0.045
ABI3	*ABI family, member 3*	-0.82	8.57	-4.53	0.045
ARHGEF12	*Rho guanine nucleotide exchange factor (GEF) 12*	0.37	7.52	4.51	0.045

Due to the low number of differentially expressed genes at 24 hours, gene ontology analysis was not performed for this comparison. Many of these genes are known shear-response genes, reported in previous analyses – such as *C10orf10, LOX1, CLDN5* and *LXN*.

16.5.4 Differential gene expression at 6 days of shear

By 6 days of shear, 237 genes were found to be significantly differentially expressed. The volcano plot below shows the distribution of statistical significance and log fold change.

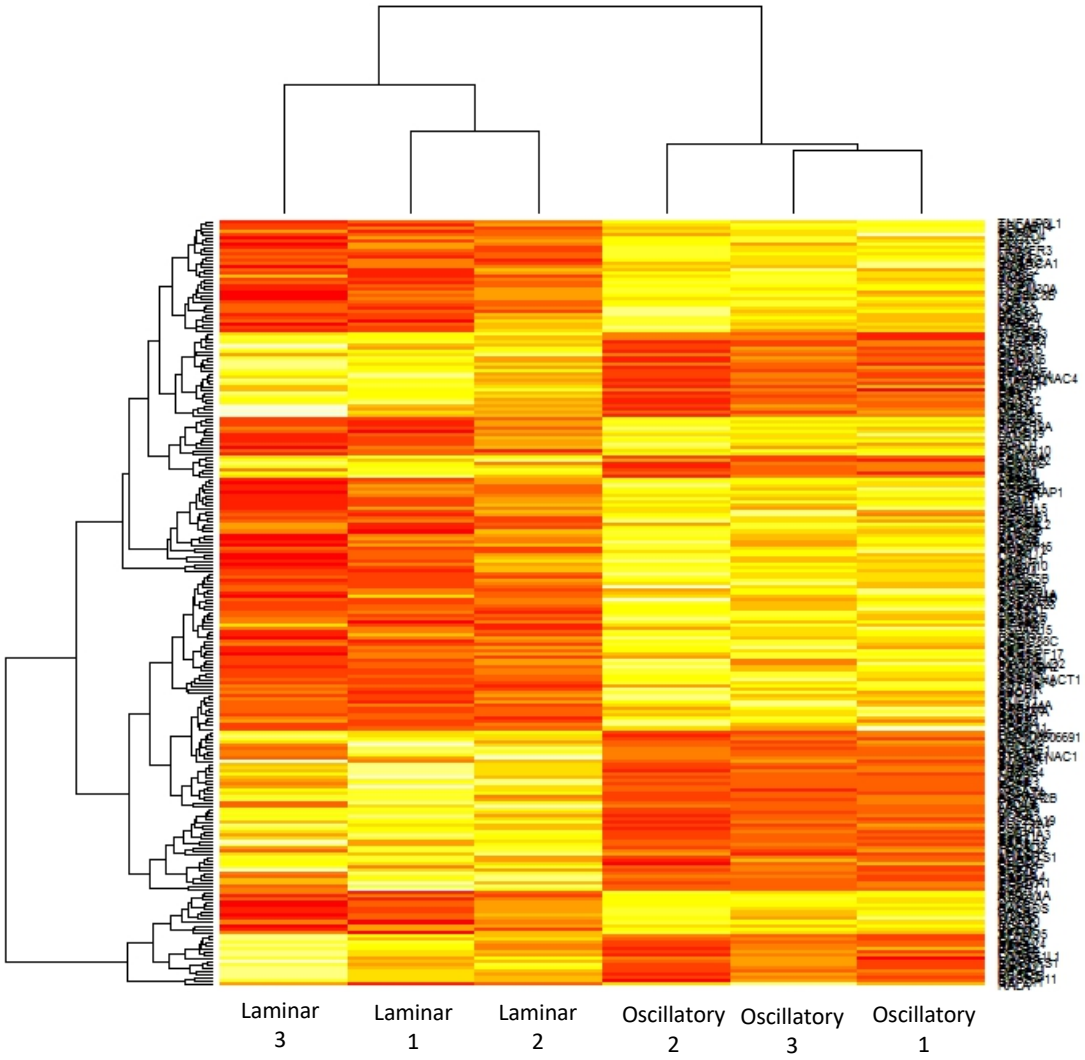
FIGURE 16.10: Volcano plot of comparison between gene expression under oscillatory and laminar shear conditions at 6 days (144 hours)



This volcano plot shows log_2 fold change of gene expression in oscillatory versus laminar conditions, plotted against log odds for each gene, ($-log_{10}$ adjusted p value). The green dashed line shows the adjusted p=0.05 significance level.

The heatmap in Figure 16.11, below, shows that there are discrete patterns of differential expression between laminar and oscillatory shear conditions at 144 hours of exposure. This heatmap is designed to give an overview of transcriptomic changes at this timepoint and therefore the gene names are not shown.

FIGURE 16.11: Heatmap showing significant differences in gene expression between laminar and oscillatory shear conditions at 6 days



The heatmap demonstrates clustering of differentially regulated genes into the laminar vs oscillatory conditions, with distinct patterns of differential expression between laminar and oscillatory shear. Laminar 3,2,1 etc., correspond to cells derived from different individuals numbered 1, 2 and 3.

The heatmap above uses hierarchical clustering of genes with similar expression levels and responses to shear, demonstrating that there are distinct patterns of differential expression, with replicability across samples. The table below (table 16.4) identifies the top 20 differentially regulated genes at 6 days (144 hours) by fold-change values.

TABLE 16.4: Top differentially regulated genes at 144 hours; fold changes represent change from laminar to oscillatory shear (i.e. a positive fold-change represents genes which are up-regulated by oscillatory shear conditions)

Gene symbol	Gene name	log2 Fold Change	Average Expression	t statistic	Adjusted p value
FABP4	*fatty acid binding protein 4, adipocyte*	2.99	9.86	3.88	0.037
ANGPT2	*angiopoietin 2*	2.77	9.00	5.40	0.002
FILIP1	*filamin A interacting protein 1*	2.43	7.99	7.19	0.000
ACKR1	*atypical chemokine receptor 1 (Duffy blood group)*	-2.42	7.09	-3.72	0.046
C10orf10 (DEPP1)	*chromosome 10 open reading frame 10*	2.27	9.98	4.51	0.013
PI16	*peptidase inhibitor 16*	-2.27	7.06	-3.74	0.044
APOD	*apolipoprotein D*	2.14	7.64	5.73	0.002
RGCC	*regulator of cell cycle*	1.99	11.26	4.15	0.023
A2M	*alpha-2-macroglobulin*	1.98	7.80	3.67	0.049
FAM89A	*family with sequence similarity 89, member A*	1.93	9.25	5.45	0.002
PGF	*placental growth factor*	1.78	10.00	4.82	0.008
ADAMTS1	*ADAM metallopeptidase with thrombospondin type 1 motif, 1*	-1.76	8.52	-3.91	0.036
KIT	*v-kit Hardy-Zuckerman 4 feline sarcoma viral oncogene homolog*	1.73	7.23	5.36	0.002
LFNG	*LFNG O-fucosylpeptide 3-beta-N-acetylglucosaminyltransferase*	-1.71	8.26	-4.41	0.015
LOX	*lysyl oxidase*	1.67	9.23	5.02	0.005
CTSK	*cathepsin K*	1.63	8.29	4.26	0.020
SPRY1	*sprouty homolog 1, antagonist of FGF signalling (Drosophila)*	1.57	9.71	3.66	0.050
DACH1	*dachshund family transcription factor 1*	1.56	7.69	5.13	0.004
MT1E	*metallothionein 1E*	-1.56	12.14	-6.26	0.001
CXADR	*coxsackie virus and adenovirus receptor*	1.52	7.83	4.62	0.011

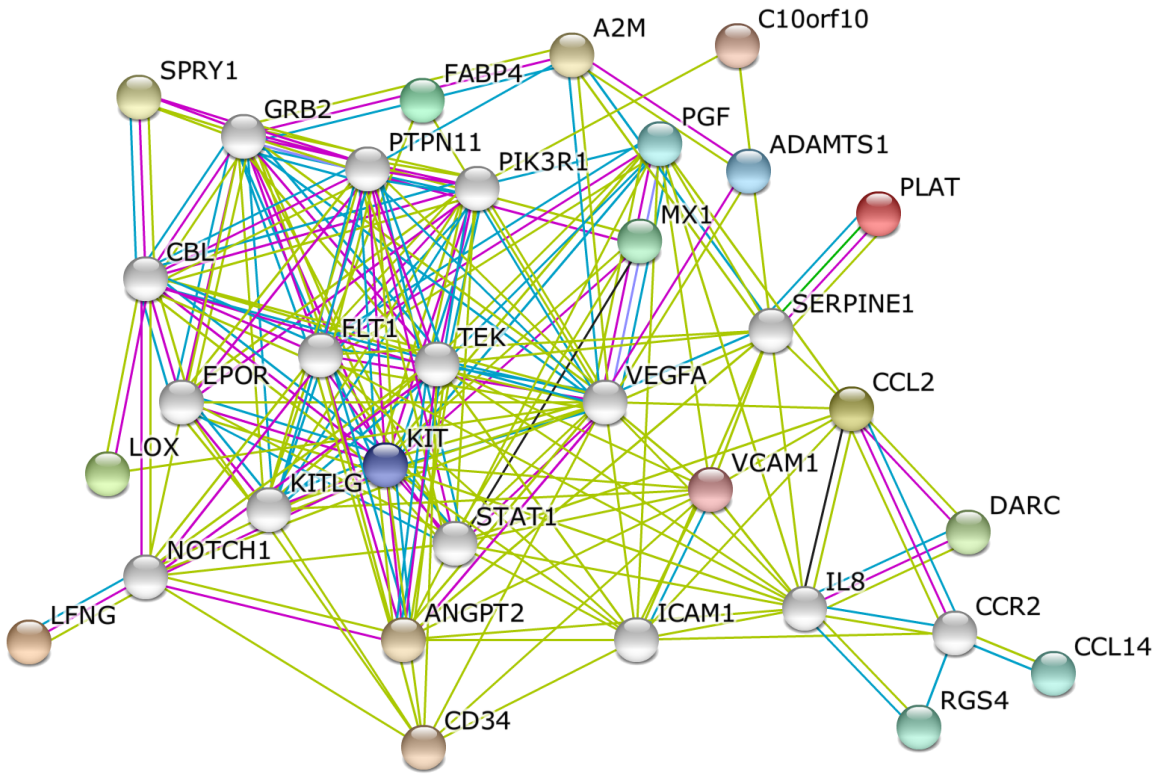
The genes in table 16.4 all demonstrate significant differential regulation by shear stress pattern. The complete table of differentially expressed genes may be found at Appendix 3. To emphasise the relevance of these findings, the top 5 genes are listed in Table 16.5 below with their mechanistic links to atherosclerosis or aortic disease. Many of the genes have proven impact upon atherogenesis – for example, infusion of recombinant angiopoetin 2 can ameliorate both aortic aneurysm formation and atherosclerosis in *apoE*-/- mice[524]

TABLE 16.5: Top 5 differentially expressed genes have mechanistic links to atherosclerosis and/or aortic disease

Gene	Gene name	log_2 fold change	Adjusted p value	Molecular function and clinical links
FABP4	*fatty acid binding protein 4, adipocyte*	2.99	0.04	Lipid binding; linked with coronary restenosis
ANGPT2	*angiopoietin 2*	2.77	<0.01	Disrupts vascular remodelling; induces endothelial apoptosis; rANGPT2 can attenuate aortic aneurysm and atherosclerosis in angiotensin II infused mice[524]
FILIP1	*filamin A interacting protein 1*	2.43	<0.01	Binds FLNA – mutations in FLNA cause aortopathy
ACKR1	*atypical chemokine receptor 1 (Duffy blood group)*	-2.42	0.05	Leukocyte recruitment and pro-inflammatory; major QTL for MCP-1 levels known to be associated with atheroprone endothelial phenotype
C10orf10	*chromosome 10 open reading frame 10*	2.27	0.01	Regulator of oxidative stress response

To enhance the overview of networks and pathways differentially regulated by shear stress pattern, we used STRING10 to identify networks, seeded with the top 25 most significantly regulated genes. This network analysis identified several "node" genes which are key regulators of the transcriptomic response to stress and also cardiovascular development, including *ICAM1, VCAM1, NOTCH1, VEGFA, FLT1*, and *TEK*. This network analysis also demonstrates the complexity of overlapping mechanosensitive pathways.

FIGURE 16.12: STRING network analysis seeded with top 25 differentially expressed genes (oscillatory vs laminar shear stress)



Coloured nodes indicate those significantly differentially regulated by shear stress in the current analysis.

The network analysis identifies a large network of genes known to be key to several signalling pathways mediating the shear stress response, and several with links with atherosclerosis.

Gene ontology analysis was performed to determine which biological pathways were significantly overrepresented in these results.

The significantly regulated pathways are presented below in the directed acyclic graph (Figure 16.13), and listed beneath in table 16.6.

FIGURE 16.13: Directed Acyclic Graph showing Gene Ontology hierarchy (biological processes) of genes differentially expressed in laminar shear compared with turbulent. The GO Terms are identified in Table 16.6 below

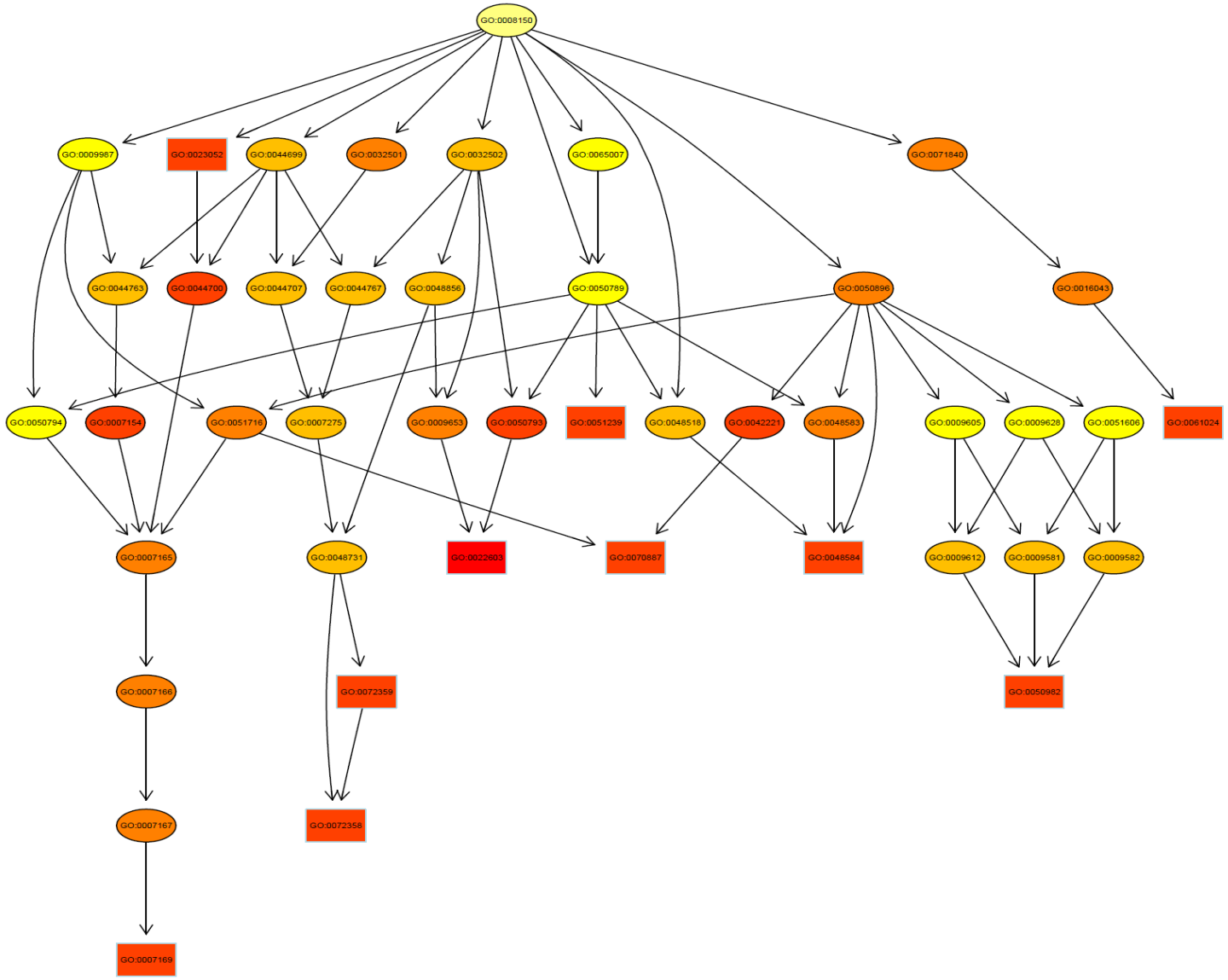


TABLE 16.6: Top 25 Gene Ontology terms demonstrating differential expression between turbulent and laminar shear at 144 hours

GO ID	Term	Annotated	Significant	Expected	p value (classic Fisher)
GO:0022603	regulation of anatomical structure morphogenesis	518	25	9.63	1.10E-05
GO:0072358	cardiovascular system development	632	27	11.75	4.30E-05
GO:0072359	circulatory system development	632	27	11.75	4.30E-05
GO:0007169	transmembrane receptor protein tyrosine kinase signalling pathway	532	24	9.89	5.10E-05
GO:0061024	membrane organization	677	27	12.59	0.00014
GO:0070887	cellular response to chemical stimulus	1625	50	30.22	0.00016
GO:0051239	regulation of multicellular organismal process	1453	46	27.02	0.00017
GO:0048584	positive regulation of response to stimulus	1078	37	20.05	0.00017
GO:0050982	detection of mechanical stimulus	16	4	0.3	0.00018
GO:0023052	signalling	3741	94	69.57	0.00024
GO:0044700	single organism signalling	3741	94	69.57	0.00024
GO:0007154	cell communication	3792	95	70.51	0.00024
GO:0050793	regulation of developmental process	1240	40	23.06	0.00033
GO:0042221	response to chemical	2360	65	43.89	0.00035
GO:2000026	regulation of multicellular organismal development	913	32	16.98	0.00035
GO:0010646	regulation of cell communication	1911	55	35.54	0.0004
GO:0050954	sensory perception of mechanical stimulus	75	7	1.39	0.00047
GO:0023051	regulation of signalling	1908	54	35.48	0.0007
GO:0001934	positive regulation of protein phosphorylation	488	20	9.07	0.00076
GO:0051716	cellular response to stimulus	4223	101	78.53	0.00083
GO:0051094	positive regulation of developmental process	569	22	10.58	0.0009
GO:0030814	regulation of cAMP metabolic process	61	6	1.13	0.00091
GO:0071840	cellular component organization or biogenesis	3926	95	73.01	0.00093
GO:0016043	cellular component organization	3825	93	71.13	0.00094
GO:0022604	regulation of cell morphogenesis	288	14	5.36	0.00098

It is interesting to note that the top 3 GO biological processes which are significantly overrepresented in our differential expression analysis include processes related to

anatomical and cardiovascular system development. We examined this in more detail by extracting information about the specific genes which were found to display significant differential expression in this gene ontology group (cardiovascular system development; GO:0072358).

TABLE 16.7: The 28 significantly differentially expressed genes within GO Term "Cardiovascular system development" (oscillatory vs laminar shear at 144 hours)

Gene ID	Gene name	Log Fold Change	Adjusted p value
APOD	*apolipoprotein D*	2.138	0.002
PRKD2	*protein kinase D2*	0.682	0.002
ANGPT2	*angiopoietin 2*	2.770	0.002
GAB1	*GRB2-associated binding protein 1*	0.471	0.005
LOX	*lysyl oxidase*	1.668	0.005
PGF	*placental growth factor*	1.779	0.008
JUP	*junction plakoglobin*	1.146	0.008
CXADR	*coxsackie virus and adenovirus receptor*	1.519	0.011
MYH10	*myosin, heavy chain 10, non-muscle*	1.443	0.011
KDR	*kinase insert domain receptor (a type III receptor tyrosine kinase)*	0.649	0.011
RHOJ	*ras homolog family member J*	0.869	0.013
PKD2	*polycystic kidney disease 2 (autosomal dominant)*	0.472	0.017
ADM	*adrenomedullin*	1.256	0.020
EPHB1	*EPH receptor B1*	0.767	0.022
RGCC	*regulator of cell cycle*	1.988	0.023
CRIP1	*cysteine-rich protein 1 (intestinal)*	-1.120	0.023
DCHS1	*dachsous cadherin-related 1*	0.780	0.024
AXIN2	*axin 2*	0.524	0.025
ZFAND5	*zinc finger, AN1-type domain 5*	0.504	0.025
ADAMTS1	*ADAM metallopeptidase with thrombospondin type 1 motif, 1*	-1.757	0.036
NPY5R	*neuropeptide Y receptor Y5*	-0.279	0.036
WARS	*tryptophanyl-tRNA synthetase*	1.096	0.039
VASH1	*vasohibin 1*	1.372	0.039
APLNR	*apelin receptor*	-1.063	0.041
CUL7	*cullin 7*	0.242	0.043
TGFBR3	*transforming growth factor, beta receptor III*	1.074	0.044
PI16	*peptidase inhibitor 16*	-2.267	0.044

There are many genes identified as differentially expressed within the "cardiovascular development" ontology with important effects on vascular risk and aortopathy risk (see Discussion).

16.6 DISCUSSION

The results of our comparison between oscillatory and laminar shear identify a set of genes which are differentially regulated by the pattern of shear stress.

16.6.1 Timing of transcriptomic changes

A very significant finding is that the transcriptomic - and indeed morphological- changes in these human aortic endothelial cells were not fully apparent until 144 hours of exposure to shear. This has significant implications for the interpretation and extrapolation of results from previous gene expression studies of cultured endothelial cells, which have tended to examine cells at a maximum of 48 hours of culture (usually 24 hours). In order for biological relevance to be preserved, further efforts must be made to assess the response to a chronic level of shear stress.

16.6.2 Genes differentially regulated by shear stress are important in atherosclerosis

Many of the differentially expressed genes (DEGs) are important drivers in atherosclerosis, and have evidence backing up their role. *FABP4*, our top DEG, plays a major role in atherosclerosis. Circulating FABP4 acts as an adipokine, stimulating atherogenesis, and the development of insulin resistance[525, 526]. *FABP4* expression in vascular endothelial cells worsened endothelial dysfunction and promoted proliferation of nearby vascular smooth muscle cells[527]; features which greatly contribute to atherosclerosis and restenosis after stenting. Of great interest is the fact that some of these local effects can be suppressed by anti-FABP4 antibody, making this a possible therapeutic target for cardiovascular disease[527, 528]. Many of the other top associated DEGs have similar links with coronary artery disease, including *ANGPT2*. This raises the possibility that other identified genes which are not so well linked with coronary disease might also be important mechanistically and for identification of therapeutic targets for atherosclerosis. Less well-characterised genes include *FILIP1* and *C10orf10*, a gene known to be important in the oxidative stress pathways and which is a

known shear-response gene, but which has not been intensively investigated in the context of atherosclerosis. *PI16* again has previously been shown to be upregulated by laminar shear stress, and acts in a protective manner to inhibit MMP2 and local endothelial cell migration. This protective effect is lost during inflammation[529]. *APOD* (the top DEG in the cardiovascular development ontology term) knockout in mice results in increased infarct size with induced ischaemia[530]. None of these latter 4 genes have been extensively investigated in the context of atherosclerosis, but could be interesting targets for future research.

16.6.2 Genes differentially regulated by shear stress are important in aortic disease

The results have been particularly interesting from the perspective of aortic disease and aortic traits. One of the top differentially expressed genes here, *KIT*, was also a candidate GWAS hit for aortic valve diameter in Chapter 14 above. This would be consistent with this gene's known role in cardiovascular development and its regulatory role for multiple signalling pathways including VEGFA. *KIT* here is demonstrated to be a key node in the network analysis of shear response traits – and is therefore an interesting candidate for future research on aortic valve development and disease and the contribution of shear stress to phenotype in related conditions such as BAV.

Gene ontology terms which are particularly overrepresented in the differential expression analysis include those pertaining to cardiovascular system development. Within this category are several genes of interest for aortic disease - such as *DCHS1* (identified in Chapter 15 as a likely BAV-related gene), *TGFBR3* and *LOX*. Pathways known to be important for aortic disease such as the TGF-B pathway were also significantly differentially regulated (data not presented as it was not amongst the top 20). This highlights the complex interplay of haemodynamic and genetic factors which may underlie aortic disease, and hints at complex patterns of pathogenesis, whereby subtle perturbations in exposure to shear stress during development, possibly coupled with aberrant genetics, could lead to profound changes in morphology. Similarly, rare variants in genes involved in shear stress response could lead to abnormal aortic remodelling in disease. Genes identified here are therefore candidates to "look out for" in segregation or population studies of aortic diseases.

16.7 LIMITATIONS

Our model still has its limitations. Firstly, despite our compartmentalisation of the well corresponding to polarised regions of laminar versus oscillatory shear, there is inevitably some degree of overlap between the cells exposed to laminar and to turbulent shear. This is because the shear experienced is not uniform even across the different "zones" of the well. However, this may be rather more representative of conditions in the in vivo state, where cells will be exposed to gradients of shear stress, even in regions where the pattern is predominantly oscillatory (e.g. at bifurcations) or laminar (in tubular sections). Secondly, our cells are cultured in isolation, in a monolayer, so the model does not include any degree of interaction with other cells or structures which form the vessel wall; neither does it allow consideration of response to external chemical mediators of response. 6 days is probably the limit of chronicity which can be assessed by this model; after this time, the cells become over-confluent, overlapping and displaying alterations in morphology, and therefore the model would not be suitable for assessing more chronic time points.

We also lack power as n=3 provides insufficient replicates to overcome the multiple testing hurdles. We are limited also by the inherent limitations of microarray testing, as discussed in section 16.1. Of course, transcriptomic analysis is an incomplete evaluation of functional response to stimuli – protein expression, post-translational modifications and regulatory RNAs not captured by microarray will all play key roles in determining endothelial cell phenotype in response to shear stress.

These limitations aside, this model has the potential to tease apart the complex pathways involved in shear response and atheroprotection.

16.8 CONCLUSIONS

Many genes involved in cardiovascular development, and many key in pathogenesis of atherosclerosis or aortic disease, are differentially regulated by shear stress patterns. This differential regulation continues to evolve until 144 hours at least. In some cases, there is overlap of transcriptomic and gene sequencing or genome-wide association findings in related phenotypes. This is the case with *DCHS1,* where rare variants may be found in BAV,

and with *KIT*, where common variants influence aortic valve annulus diameter. These genes are also differentially regulated by shear stress. These complementary studies may therefore provide extra evidence both to support functional roles for these genes in determining aortic phenotype, and to provide some insight into the mechanisms of that effect. The overlap also highlights the complexities of the interaction of genetic and haemodynamic effects in determining aortic phenotype.

16.9 ACKNOWLEDGEMENTS

This project was developed with the assistance of Professor Jane Mitchell, building on the previous work of Dr Claire Potter who refined modelling of the shear stress culture system. Dr Mark Paul-Clark, Hime Gashaw and Dr Daniel Reed provided invaluable assistance with cell culture and transcriptomic analysis. Dr Maxim Freydin provided help and code for evaluation of transcriptomic responses.

17: SUPPLEMENTARY, ONGOING AND FUTURE WORK

"Science is fun. Science is curiosity." – Sally Ride

The work presented in this thesis has opened the doors to many exciting and ongoing collaborative research opportunities. The following projects in which I am playing a key role will build upon and complement the work described in this thesis.

17.1 Common variant analysis in the general population: UK Biobank

UK Biobank has recruited >26,000 subjects who have undergone cardiovascular magnetic resonance imaging (CMR) and genotyping. Like the Digital Heart Project, the CMR is designed to gather information on a number of phenotypes, so aortic sequences are limited. Nevertheless, it allows measurement of ascending and descending aortic areas and distensibility. These phenotypes have been measured using automated segmentation of CMR images, which have been manually QCd. At the time of writing, an internal meta-analysis of GWAS for aortic phenotypes has just been completed, and analysis of results is underway. The DHP GWAS presented in this thesis will be used for independent replication of effect direction. Early indications are that there is some overlap with DHP loci, but UK Biobank has also identified additional loci of great interest in aortic biology.

17.2 Rare variants in the general population: UK Biobank

Many of the UK Biobank participants have also undergone whole exome sequencing, which has just been released. For candidate genes derived from results of the UK Biobank and Digital Heart Project GWAS, rare variant associations with extremes of phenotype or with cardiovascular, cerebrovascular and aortic disease will be assessed.

A pilot project in DHP also suggests that rare, protein-altering variants in known aortopathy disease genes might result in a sub-clinical phenotype of altered aortic elastic function. This hypothesis will be evaluated in the UK Biobank population.

17.3 Rare variants in disease: 100,000 genomes project

The 100,000 Genomes Project was an ambitious strategy to boost the implementation of Genomic Medicine within the National Health Service (NHS) and the UK as a whole. More than 100,000 probands and family members of patients with rare disease or cancer have undergone whole genome sequencing, and limited collection of clinical data.

There are, to date, 536 probands with thoracic aortic aneurysm or dissection, with whole genome sequencing available. Of these, 68 have been recruited as trios and there are 10 larger families. Systematic family segregation analysis is underway, looking for coding variants in new disease genes, non-coding variants and structural variants which might be associated with aortic disease. The DHP and UK Biobank GWAS is already informing this work, helping to prioritise genes identified. A case:control analysis is also planned which, whilst it perhaps lacks power, will at the very least form an ideal dataset for validation of hypotheses derived from other, more comprehensively phenotyped and / or larger cohorts.

18: PHENOTYPE, GENOMICS, GENETICS AND HAEMODYNAMICS: PUTTING IT ALL TOGETHER

"Information is not knowledge" – Albert Einstein

The work described in this thesis has generated the largest UK dataset to date for healthy aortic phenotyping by cardiovascular MRI. This incorporates aortic dimensions, morphology and function, measured using standard, quick cardiovascular MR acquisitions which are easy to replicate in clinical practice. These healthy reference values will provide useful reference data for research and clinical use.

Through regression modelling, the relationships between aortic traits and biometric variables were explored. The age-related changes in thoracic aortic structure, shape and function have been defined. With increasing age, the thoracic aorta dilates and stiffens, and the aortic arch lengthens, widens and tilts posteriorly. These changes have been quantified for the first time in a healthy cohort.

Whether the structural and morphological changes in the aorta are an adaptation to deteriorating elastic function of the aortic wall, or whether a single process causes functional and morphological remodelling is unclear. Classical cardiovascular risk factors or risk markers such as LDL, triglycerides, increasing heart rate and body fat percentage are, in general, associated with worse aortic elastic function but a decrease in the structural and morphological adaptations seen in "normal" ageing. There is a particularly important influence of body fat percentage on many of the traits measured. It is tempting to speculate that cardiovascular risk factors such as body fat might restrict the aorta's adaptive remodelling, thus exposing the aortic wall and left ventricle to more pathological haemodynamics. The major exception to this is blood pressure, which has a more complex relationship with aortic stiffness and structure.

Alongside the phenotyping, a framework for quality control and imputation of complex genotype data was developed and implemented. This generated a set of carefully quality-controlled, densely-imputed genotypes at > 9 million single nucleotide variants for 1218 healthy Caucasian volunteers. This will be a valuable healthy reference dataset, and is already being used in multiple studies.

The genotype and phenotype data were combined in a genome-wide association study of those aortic traits which were robust and which will be possible to replicate in larger studies such as UK Biobank. Whilst no individual loci or gene associations reaching genome-wide significance were identified due to under-powering, several loci reached suggestive significance thresholds. Candidate genes at several of these loci were identified which merit follow-up if replicated in independent cohorts.

Analysis of the functional roles of SNPs and genes at these loci suggest the importance of genes involved in cardiovascular development in regulating aortic traits in later life. Examples of this include *KIT*, *PDGFRA* and *PDGFD* in the aortic root. The suggestive associations also hint at the importance of the autonomic nervous system in regulation of aortic traits, with genes involved in both sympathetic and parasympathetic activity implicated at association peaks with different aortic traits. Examples include $GABA_A$ receptors associated with SoV diameter, and glutamate receptor subunit (encoded by *GRIA4*) associated with STJ diameter. A key role for fibrosis is also suggested by gene and pathways associations which are known to regulate this process, particularly via TGF-ß signalling regulation.

Genes influencing cardiovascular risk factors, and in particular obesity traits, such as ARID5B and IRX3, are also associated with aortic traits. These associations complement and reinforce the findings of the phenotype regression modelling undertaken, which identified body fat percentage as a strong contributor to aortic phenotype.

In ascending and descending aorta, genes involved in remodelling are also implicated, along with adhesion molecules such as vinculin, and pathways such as TGF-ß, Wnt, Hedgehog and angiotensin II signalling.

These results have provided possible insights into the genes and networks which regulate aortic size and function throughout life. Whilst replication and functional characterisation of loci will be mandatory, the genes discussed in this thesis represent key candidates for follow-up. They form a group of genes and pathways which might be key determinants of cardiovascular and local aortic risk, and which might be amenable to intervention. *PDE1C, AGTR1, PENK, VCL, RXFP2* and *ROCK1P1* are of particular interest in this regard, as their products (or putative product in the case of the pseudogene) are already targeted by known drugs.

The power of combining traits in joint trait analysis was also demonstrated, with an aortic root dimension joint trait analysis identifying two further loci of interest in vascular biology – *TENM4* and *PTN*. Again, these will be interesting candidates to follow up in larger studies.

UK Biobank will provide a valuable resource for replication and further investigation of these loci.

Rare variants affecting aortic traits and aortic diseases are less easily tractable using current resources. The work in this thesis included a candidate gene observational study of patients with bicuspid aortic valve. Whilst some might argue that this form of study design should be consigned to the history books (and indeed the study reported here was originally conceived >10 years ago!), it nevertheless can provide exploratory data to identify potential patterns and key associations for follow-up. The study demonstrated the very limited utility of panel sequencing in BAV – the clinically actionable diagnostic yield was 0. There were many potentially pathogenic variants identified in different genes, but insufficient evidence for any to be classed as causal.

There were some suggestions of phenotype clustering with rare variants in particular genes – for example, obstructive phenotypes appeared to cluster with APOB variants. These data, whilst far from conclusive, suggest that rare genetic variants might regulate aortic phenotype in pathology. Common variants in just one gene (*VCL*) under study in the bicuspid valve cohort were also identified as associated with aortic traits in the healthy population GWAS. Other pathways identified by the healthy population GWAS overlap with BAV candidate genes – such as the TGF-ß pathways. There were no significant phenotype clusters associated with rare variants in genes in these pathways in the BAV cohort.

The value, however, of this data, is perhaps in keeping an eye open for the interesting tangential associations – here, a family was identified in whom rare variants in DCHS1 look likely to contribute to the BAV phenotype. This is a potential novel gene:disease association. If confirmed through segregation analysis and additional supportive information, this also will form a clear example of variable expressivity, whereby a family with the same causative variant had a mitral valve prolapse phenotype[531], whilst the family reported here have normal mitral valves, but abnormal aortic valves. Collaborative studies are underway to examine

more closely the role of this gene in BAV, and further rare variant analysis is currently underway in UK Biobank and the 100,000 Genomes Project.

Finally, mechanisms of genetic influence on aortic phenotype were examined by studying aortic endothelial cell response to a key haemodynamic driver: shear stress pattern. A technique was developed to expose cells to different patterns of shear stress within the same well, and RNA expression levels were assessed by microarray. This revealed that the evolution of expression patterns in response to differing shear stress exposures takes some time – certainly more than the usual 24 hours used in most shear stress experiments. The genes which were differentially expressed in response to shear stress pattern included some which overlapped with the GWAS hits (e.g. *KIT*) and with genes implicated in bicuspid aortic valve, aortopathy, or of interest in these conditions (e.g. *DCHS1, LOX, TGFBR3*). This study also reaffirms the importance of genes involved in cardiovascular development in determining dynamic aortic responses – a fact which may be relevant not only during development itself, but also in remodelling and aortic adaptation throughout life.

There are, of course, challenges ahead: the generation of these phenotypes, genotypes, associations and observations is far from the end of the story. Instead, the work in this thesis forms a contribution to a framework on which to build a greater understanding of aortic biology. It has also allowed an exploration of the power, pitfalls, advantages and limitations of different genetic and genomic tools, resources and study designs. This work has highlighted gaping holes in our understanding of both the genome and the aorta – gaps which create exciting opportunities for future research, learning from the limitations of the current studies.

By combining the associations found here with data from UK Biobank, it will be possible to create more robust common variant associations with aortic phenotype, and explore in more detail the interactions between them. Rare variants in genes of interest will also be investigated in the 100,000 Genomes Project, and cellular and molecular work to understand individual associations will also be necessary and is under way.

CONCLUSIONS

In summary, this thesis has examined thoracic aortic phenotype, and the anthropometrics, cardiovascular risk factors, common genetic variation, rare variation and haemodynamics which might affect aortic form and function at population, disease, and cellular levels.

The aorta is a dynamic, responsive and vital component of the cardiovascular system, which remodels and adapts throughout life. This work has defined normal aortic phenotypes and demonstrated that common variants in genes involved in aortic development, aortic homeostasis, autonomic cardiovascular response, and multiple regulatory pathways including the TGF-ß network, shape and refine the aorta and its function in individuals. Rare variants in developmental genes as well as genes regulating smooth muscle contraction and cellular signalling pathways such as the TGF-ß network, can cause aortic disease, and might modify phenotype where the aorta is weakened by pathology. Finally, expression of genes within the same networks in cells lining the aorta is greatly influenced by local shear stress patterns. The genes identified here show promise for further understanding the genetics and genomics of aortic form and function.

19. REFERENCES

1. Ohyama Y, Redheuil A, Kachenoura N, Ambale Venkatesh B, Lima JAC. Imaging Insights on the Aorta in Aging. Circulation Cardiovascular imaging. 2018;11(4):e005617.
2. Ben-Shlomo Y, Spears M, Boustred C, May M, Anderson SG, Benjamin EJ, et al. Aortic pulse wave velocity improves cardiovascular event prediction: an individual participant meta-analysis of prospective observational data from 17,635 subjects. Journal of the American College of Cardiology. 2014;63(7):636-46.
3. Lim HE, Park CG, Shin SH, Ahn JC, Seo HS, Oh DJ. Aortic pulse wave velocity as an independent marker of coronary artery disease. Blood pressure. 2004;13(6):369-75.
4. Mattace-Raso FU, van der Cammen TJ, Hofman A, van Popele NM, Bos ML, Schalekamp MA, et al. Arterial stiffness and risk of coronary heart disease and stroke: the Rotterdam Study. Circulation. 2006;113(5):657-63.
5. Redheuil A, Wu CO, Kachenoura N, Ohyama Y, Yan RT, Bertoni AG, et al. Proximal Aortic Distensibility Is an Independent Predictor of All-Cause Mortality and Incident CV Events: The MESA Study. Journal of the American College of Cardiology. 2014;64(24):2619-29.
6. Ohyama Y, Ambale-Venkatesh B, Noda C, Kim JY, Tanami Y, Teixido-Tura G, et al. Aortic Arch Pulse Wave Velocity Assessed by Magnetic Resonance Imaging as a Predictor of Incident Cardiovascular Events: The MESA (Multi-Ethnic Study of Atherosclerosis). Hypertension. 2017;70(3):524-30.
7. Bechlioulis A, Vakalis K, Naka KK, Bourantas CV, Papamichael ND, Kotsia A, et al. Increased aortic pulse wave velocity is associated with the presence of angiographic coronary artery disease in overweight and obese patients. American journal of hypertension. 2013;26(2):265-70.
8. Mitchell GF. Effects of central arterial aging on the structure and function of the peripheral vasculature: implications for end-organ damage. Journal of applied physiology. 2008;105(5):1652-60.
9. Coutinho T, Turner ST, Kullo IJ. Aortic pulse wave velocity is associated with measures of subclinical target organ damage. JACC Cardiovascular imaging. 2011;4(7):754-61.
10. Poels MM, Zaccai K, Verwoert GC, Vernooij MW, Hofman A, van der Lugt A, et al. Arterial stiffness and cerebral small vessel disease: the Rotterdam Scan Study. Stroke; a journal of cerebral circulation. 2012;43(10):2637-42.
11. Sedaghat S, Dawkins Arce FG, Verwoert GC, Hofman A, Ikram MA, Franco OH, et al. Association of renal function with vascular stiffness in older adults: the Rotterdam study. Age and ageing. 2014;43(6):827-33.
12. Groenink M, de Roos A, Mulder BJ, Spaan JA, van der Wall EE. Changes in aortic distensibility and pulse wave velocity assessed with magnetic resonance imaging following beta-blocker therapy in the Marfan syndrome. The American journal of cardiology. 1998;82(2):203-8.

13. Teixido-Tura G, Redheuil A, Rodriguez-Palomares J, Gutierrez L, Sanchez V, Forteza A, et al. Aortic biomechanics by magnetic resonance: early markers of aortic disease in Marfan syndrome regardless of aortic dilatation? International journal of cardiology. 2014;171(1):56-61.

14. Cuspidi C, Facchetti R, Bombelli M, Re A, Cairoa M, Sala C, et al. Aortic root diameter and risk of cardiovascular events in a general population: data from the PAMELA study. Journal of hypertension. 2014;32(9):1879-87.

15. Kamimura D, Suzuki T, Musani SK, Hall ME, Samdarshi TE, Correa A, et al. Increased Proximal Aortic Diameter is Associated With Risk of Cardiovascular Events and All-Cause Mortality in Blacks The Jackson Heart Study. Journal of the American Heart Association. 2017;6(6).

16. Lam CS, Gona P, Larson MG, Aragam J, Lee DS, Mitchell GF, et al. Aortic root remodeling and risk of heart failure in the Framingham Heart study. JACC Heart failure. 2013;1(1):79-83.

17. Bensalah MZ, Bollache E, Kachenoura N, Giron A, De Cesare A, Macron L, et al. Geometry is a major determinant of flow reversal in proximal aorta. American journal of physiology Heart and circulatory physiology. 2014;306(10):H1408-16.

18. Redheuil A, Yu WC, Mousseaux E, Harouni AA, Kachenoura N, Wu CO, et al. Age-related changes in aortic arch geometry: relationship with proximal aortic function and left ventricular mass and remodeling. Journal of the American College of Cardiology. 2011;58(12):1262-70.

19. Totaro S, Rabbia F, Milan A, Urbina EM, Veglio F. Aortic root dilatation in the children and young adults: prevalence, determinants, and association with target organ damage. Journal of the American Society of Hypertension : JASH. 2016;10(10):782-9.

20. Heuts S, Adriaans BP, Gerretsen S, Natour E, Vos R, Cheriex EC, et al. Aortic elongation part II: the risk of acute type A aortic dissection. Heart. 2018;104(21):1778-82.

21. Humphrey JD, Milewicz DM. Aging, Smooth Muscle Vitality, and Aortic Integrity. Circulation research. 2017;120(12):1849-51.

22. Humphrey JD, Schwartz MA, Tellides G, Milewicz DM. Role of mechanotransduction in vascular biology: focus on thoracic aortic aneurysms and dissections. Circulation research. 2015;116(8):1448-61.

23. Mazurek R, Dave JM, Chandran RR, Misra A, Sheikh AQ, Greif DM. Vascular Cells in Blood Vessel Wall Development and Disease. Adv Pharmacol. 2017;78:323-50.

24. Chen KD, Li YS, Kim M, Li S, Yuan S, Chien S, et al. Mechanotransduction in response to shear stress. Roles of receptor tyrosine kinases, integrins, and Shc. The Journal of biological chemistry. 1999;274(26):18393-400.

25. Smigielski EM, Sirotkin K, Ward M, Sherry ST. dbSNP: a database of single nucleotide polymorphisms. Nucleic acids research. 2000;28(1):352-5.

26. The Genome Aggregation Database: https://gnomad.broadinstitute.org [

27. Lek M, Karczewski KJ, Minikel EV, Samocha KE, Banks E, Fennell T, et al. Analysis of protein-coding genetic variation in 60,706 humans. Nature. 2016;536(7616):285-91.

28. Carithers LJ, Moore HM. The Genotype-Tissue Expression (GTEx) Project. Biopreserv Biobank. 2015;13(5):307-8.

29. Consortium G. GTEx v7 2019: https://gtexportal.org/home/ [

30. Rentzsch P, Witten D, Cooper GM, Shendure J, Kircher M. CADD: predicting the deleteriousness of variants throughout the human genome. Nucleic acids research. 2019;47(D1):D886-D94.

31. Consortium EP, Birney E, Stamatoyannopoulos JA, Dutta A, Guigo R, Gingeras TR, et al. Identification and analysis of functional elements in 1% of the human genome by the ENCODE pilot project. Nature. 2007;447(7146):799-816.

32. Boyle AP, Hong EL, Hariharan M, Cheng Y, Schaub MA, Kasowski M, et al. Annotation of functional variation in personal genomes using RegulomeDB. Genome research. 2012;22(9):1790-7.

33. Ward LD, Kellis M. HaploReg v4: systematic mining of putative causal variants, cell types, regulators and target genes for human complex traits and disease. Nucleic acids research. 2016;44(D1):D877-81.

34. Ward LD, Kellis M. HaploReg: a resource for exploring chromatin states, conservation, and regulatory motif alterations within sets of genetically linked variants. Nucleic acids research. 2012;40(Database issue):D930-4.

35. Leslie R, O'Donnell CJ, Johnson AD. GRASP: analysis of genotype-phenotype results from 1390 genome-wide association studies and corresponding open access database. Bioinformatics. 2014;30(12):i185-94.

36. Buniello A, MacArthur JAL, Cerezo M, Harris LW, Hayhurst J, Malangone C, et al. The NHGRI-EBI GWAS Catalog of published genome-wide association studies, targeted arrays and summary statistics 2019. Nucleic acids research. 2019;47(D1):D1005-D12.

37. Welter D, MacArthur J, Morales J, Burdett T, Hall P, Junkins H, et al. The NHGRI GWAS Catalog, a curated resource of SNP-trait associations. Nucleic acids research. 2014;42(Database issue):D1001-6.

38. Eppig JT. Mouse Genome Informatics (MGI) Resource: Genetic, Genomic, and Biological Knowledgebase for the Laboratory Mouse. ILAR J. 2017;58(1):17-41.

39. Shaw D. Searching the Mouse Genome Informatics (MGI) resources for information on mouse biology from genotype to phenotype. Current protocols in bioinformatics / editoral board, Andreas D Baxevanis [et al]. 2004;Chapter 1:Unit 1 7.

40. CP B. The Papyrus Ebers: The Garden City Press LTD; 1930.

41. Steckerl F. The Fragments of Praxagoras of Cos and His School: Brill Archive; 1958.

42. Westerhof N, Lankhaar JW, Westerhof BE. The arterial Windkessel. Medical & biological engineering & computing. 2009;47(2):131-41.

43. Greenwald SE. Ageing of the conduit arteries. The Journal of pathology. 2007;211(2):157-72.

44. Cocciolone AJ, Hawes JZ, Staiculescu MC, Johnson EO, Murshed M, Wagenseil JE. Elastin, arterial mechanics, and cardiovascular disease. American journal of physiology Heart and circulatory physiology. 2018;315(2):H189-H205.

45. Dagenais F. Anatomy of the thoracic aorta and of its branches. Thorac Surg Clin. 2011;21(2):219-27, viii.

46. Ho SY. Structure and anatomy of the aortic root. Eur J Echocardiogr. 2009;10(1):i3-10.

47. Kaess BM, Rong J, Larson MG, Hamburg NM, Vita JA, Cheng S, et al. Relations of Central Hemodynamics and Aortic Stiffness with Left Ventricular Structure and Function: The Framingham Heart Study. Journal of the American Heart Association. 2016;5(3):e002693.

48. Schillaci G, Mannarino MR, Pucci G, Pirro M, Helou J, Savarese G, et al. Age-specific relationship of aortic pulse wave velocity with left ventricular geometry and function in hypertension. Hypertension. 2007;49(2):317-21.

49. Tsamis A, Krawiec JT, Vorp DA. Elastin and collagen fibre microstructure of the human aorta in ageing and disease: a review. Journal of the Royal Society, Interface / the Royal Society. 2013;10(83):20121004.

50. El-Hamamsy I, Yacoub MH. Cellular and molecular mechanisms of thoracic aortic aneurysms. Nature reviews Cardiology. 2009;6(12):771-86.

51. Duca L, Blaise S, Romier B, Laffargue M, Gayral S, El Btaouri H, et al. Matrix ageing and vascular impacts: focus on elastin fragmentation. Cardiovascular research. 2016;110(3):298-308.

52. O'Rourke MF, Safar ME, Dzau V. The Cardiovascular Continuum extended: aging effects on the aorta and microvasculature. Vasc Med. 2010;15(6):461-8.

53. Erdheim. J. Medionecrosis aortae idiopathica (cystica). . Archiv für pathologische Anatomie und Physiologie und für klinische Medizin. 1929;273:454-79.

54. Maleszewski JJ, Miller DV, Lu J, Dietz HC, Halushka MK. Histopathologic findings in ascending aortas from individuals with Loeys-Dietz syndrome (LDS). The American journal of surgical pathology. 2009;33(2):194-201.

55. Wahart A, Hocine T, Albrecht C, Henry A, Sarazin T, Martiny L, et al. Role of elastin peptides and elastin receptor complex in metabolic and cardiovascular diseases. The FEBS journal. 2019.

56. Phinikaridou A, Lacerda S, Lavin B, Andia ME, Smith A, Saha P, et al. Tropoelastin: A novel marker for plaque progression and instability. Circulation Cardiovascular imaging. 2018;11(8).

57. Curran ME, Atkinson DL, Ewart AK, Morris CA, Leppert MF, Keating MT. The elastin gene is disrupted by a translocation associated with supravalvular aortic stenosis. Cell. 1993;73(1):159-68.

58. Ewart AK, Morris CA, Atkinson D, Jin W, Sternes K, Spallone P, et al. Hemizygosity at the elastin locus in a developmental disorder, Williams syndrome. Nature genetics. 1993;5(1):11-6.

59. Keating M. Elastin and vascular disease. Trends in cardiovascular medicine. 1994;4(4):165-9.

60. Karnik SK, Brooke BS, Bayes-Genis A, Sorensen L, Wythe JD, Schwartz RS, et al. A critical role for elastin signaling in vascular morphogenesis and disease. Development. 2003;130(2):411-23.

61. Li DY, Brooke B, Davis EC, Mecham RP, Sorensen LK, Boak BB, et al. Elastin is an essential determinant of arterial morphogenesis. Nature. 1998;393(6682):276-80.

62. Schmelzer CEH, Heinz A, Troilo H, Lockhart-Cairns MP, Jowitt TA, Marchand MF, et al. Lysyl oxidase-like 2 (LOXL2)-mediated cross-linking of tropoelastin. FASEB journal : official publication of the Federation of American Societies for Experimental Biology. 2019;33(4):5468-81.

63. Guo DC, Regalado ES, Gong L, Duan X, Santos-Cortez RL, Arnaud P, et al. LOX Mutations Predispose to Thoracic Aortic Aneurysms and Dissections. Circulation research. 2016;118(6):928-34.

64. Karimi A, Milewicz DM. Structure of the Elastin-Contractile Units in the Thoracic Aorta and How Genes That Cause Thoracic Aortic Aneurysms and Dissections Disrupt This Structure. The Canadian journal of cardiology. 2016;32(1):26-34.

65. Dietz HC, Cutting GR, Pyeritz RE, Maslen CL, Sakai LY, Corson GM, et al. Marfan syndrome caused by a recurrent de novo missense mutation in the fibrillin gene. Nature. 1991;352(6333):337-9.

66. Dietz HC, Pyeritz RE. Mutations in the human gene for fibrillin-1 (FBN1) in the Marfan syndrome and related disorders. Human molecular genetics. 1995;4 Spec No:1799-809.

67. Kappanayil M, Nampoothiri S, Kannan R, Renard M, Coucke P, Malfait F, et al. Characterization of a distinct lethal arteriopathy syndrome in twenty-two infants associated with an identical, novel mutation in FBLN4 gene, confirms fibulin-4 as a critical determinant of human vascular elastogenesis. Orphanet J Rare Dis. 2012;7:61.

68. Renard M, Holm T, Veith R, Callewaert BL, Ades LC, Baspinar O, et al. Altered TGFbeta signaling and cardiovascular manifestations in patients with autosomal recessive cutis laxa type I caused by fibulin-4 deficiency. European journal of human genetics : EJHG. 2010;18(8):895-901.

69. Guo DC, Pannu H, Tran-Fadulu V, Papke CL, Yu RK, Avidan N, et al. Mutations in smooth muscle alpha-actin (ACTA2) lead to thoracic aortic aneurysms and dissections. Nature genetics. 2007;39(12):1488-93.

70. Bellini C, Wang S, Milewicz DM, Humphrey JD. Myh11(R247C/R247C) mutations increase thoracic aorta vulnerability to intramural damage despite a general biomechanical adaptivity. Journal of biomechanics. 2015;48(1):113-21.

71. Pannu H, Tran-Fadulu V, Papke CL, Scherer S, Liu Y, Presley C, et al. MYH11 mutations result in a distinct vascular pathology driven by insulin-like growth factor 1 and angiotensin II. Human molecular genetics. 2007;16(20):2453-62.

72. Zhu L, Vranckx R, Khau Van Kien P, Lalande A, Boisset N, Mathieu F, et al. Mutations in myosin heavy chain 11 cause a syndrome associating thoracic aortic aneurysm/aortic dissection and patent ductus arteriosus. Nature genetics. 2006;38(3):343-9.

73. Wang L, Guo DC, Cao J, Gong L, Kamm KE, Regalado E, et al. Mutations in myosin light chain kinase cause familial aortic dissections. American journal of human genetics. 2010;87(5):701-7.

74. Guo DC, Regalado E, Casteel DE, Santos-Cortez RL, Gong L, Kim JJ, et al. Recurrent gain-of-function mutation in PRKG1 causes thoracic aortic aneurysms and acute aortic dissections. American journal of human genetics. 2013;93(2):398-404.

75. Ando J, Tsuboi H, Korenaga R, Takahashi K, Kosaki K, Isshiki M, et al. Differential display and cloning of shear stress-responsive messenger RNAs in human endothelial cells. Biochemical and biophysical research communications. 1996;225(2):347-51.

76. Li M, Qian M, Kyler K, Xu J. Endothelial-Vascular Smooth Muscle Cells Interactions in Atherosclerosis. Front Cardiovasc Med. 2018;5:151.

77. Garcia-Cardena G, Comander J, Anderson KR, Blackman BR, Gimbrone MA, Jr. Biomechanical activation of vascular endothelium as a determinant of its functional phenotype. Proceedings of the National Academy of Sciences of the United States of America. 2001;98(8):4478-85.

78. Gimbrone MA, Jr., Garcia-Cardena G. Endothelial Cell Dysfunction and the Pathobiology of Atherosclerosis. Circulation research. 2016;118(4):620-36.

79. Davies PF. Hemodynamic shear stress and the endothelium in cardiovascular pathophysiology. Nature clinical practice Cardiovascular medicine. 2009;6(1):16-26.

80. Tzemos N, Lyseggen E, Silversides C, Jamorski M, Tong JH, Harvey P, et al. Endothelial function, carotid-femoral stiffness, and plasma matrix metalloproteinase-2 in men with bicuspid aortic valve and dilated aorta. Journal of the American College of Cardiology. 2010;55(7):660-8.

81. Gittenberger-de Groot AC, Bartelings MM, Deruiter MC, Poelmann RE. Basics of cardiac development for the understanding of congenital heart malformations. Pediatric research. 2005;57(2):169-76.

82. Laforest B, Nemer M. Genetic insights into bicuspid aortic valve formation. Cardiology research and practice. 2012;2012:180297.

83. Phillips HM, Mahendran P, Singh E, Anderson RH, Chaudhry B, Henderson DJ. Neural crest cells are required for correct positioning of the developing outflow cushions and pattern the arterial valve leaflets. Cardiovascular research. 2013;99(3):452-60.

84. Grewal N, DeRuiter MC, Jongbloed MR, Goumans MJ, Klautz RJ, Poelmann RE, et al. Normal and abnormal development of the aortic wall and valve: correlation with clinical entities. Netherlands heart journal : monthly journal of the Netherlands Society of Cardiology and the Netherlands Heart Foundation. 2014;22(9):363-9.

85. MacFarlane EG, Parker SJ, Shin JY, Kang BE, Ziegler SG, Creamer TJ, et al. Lineage-specific events underlie aortic root aneurysm pathogenesis in Loeys-Dietz syndrome. The Journal of clinical investigation. 2019;129(2):659-75.

86. Chang AC, Fu Y, Garside VC, Niessen K, Chang L, Fuller M, et al. Notch initiates the endothelial-to-mesenchymal transition in the atrioventricular canal through autocrine activation of soluble guanylyl cyclase. Developmental cell. 2011;21(2):288-300.

87. Fu Y, Chang A, Chang L, Niessen K, Eapen S, Setiadi A, et al. Differential regulation of transforming growth factor beta signaling pathways by Notch in human endothelial cells. The Journal of biological chemistry. 2009;284(29):19452-62.

88. Niessen K, Karsan A. Notch signaling in the developing cardiovascular system. American journal of physiology Cell physiology. 2007;293(1):C1-11.

89. Noseda M, McLean G, Niessen K, Chang L, Pollet I, Montpetit R, et al. Notch activation results in phenotypic and functional changes consistent with endothelial-to-mesenchymal transformation. Circulation research. 2004;94(7):910-7.

90. Timmerman LA, Grego-Bessa J, Raya A, Bertran E, Perez-Pomares JM, Diez J, et al. Notch promotes epithelial-mesenchymal transition during cardiac development and oncogenic transformation. Genes & development. 2004;18(1):99-115.

91. Chakraborty S, Combs MD, Yutzey KE. Transcriptional regulation of heart valve progenitor cells. Pediatric cardiology. 2010;31(3):414-21.

92. Combs MD, Yutzey KE. Heart valve development: regulatory networks in development and disease. Circulation research. 2009;105(5):408-21.

93. Gary C Schoenwolf SBB, Philip R Brauer, Philippa H Francis-West. Larsen's Human Embryology. 5th Edition ed: Elsevier; 2014.

94. Roger VL, Go AS, Lloyd-Jones DM, Adams RJ, Berry JD, Brown TM, et al. Heart disease and stroke statistics--2011 update: a report from the American Heart Association. Circulation. 2011;123(4):e18-e209.

95. Key R. https://radiologykey.com/embryology-and-physiology-of-the-fetal-heart/ [

96. Majesky MW. Developmental basis of vascular smooth muscle diversity. Arteriosclerosis, thrombosis, and vascular biology. 2007;27(6):1248-58.

97. Waldo KL, Hutson MR, Ward CC, Zdanowicz M, Stadt HA, Kumiski D, et al. Secondary heart field contributes myocardium and smooth muscle to the arterial pole of the developing heart. Developmental biology. 2005;281(1):78-90.

98. Majesky MW, Dong XR, Hoglund VJ. Parsing aortic aneurysms: more surprises. Circulation research. 2011;108(5):528-30.

99. Gabriela Espinosa M, Catalin Staiculescu M, Kim J, Marin E, Wagenseil JE. Elastic Fibers and Large Artery Mechanics in Animal Models of Development and Disease. Journal of biomechanical engineering. 2018;140(2).

100. Karakaya C, Goktas S, Celik M, Kowalski WJ, Keller BB, Pekkan K. Asymmetry in Mechanosensitive Gene Expression during Aortic Arch Morphogenesis. Scientific reports. 2018;8(1):16948.

101. Redheuil A, Yu WC, Wu CO, Mousseaux E, de Cesare A, Yan R, et al. Reduced ascending aortic strain and distensibility: earliest manifestations of vascular aging in humans. Hypertension. 2010;55(2):319-26.

102. Turkbey EB, Jain A, Johnson C, Redheuil A, Arai AE, Gomes AS, et al. Determinants and normal values of ascending aortic diameter by age, gender, and race/ethnicity in the Multi-Ethnic Study of Atherosclerosis (MESA). Journal of magnetic resonance imaging : JMRI. 2014;39(2):360-8.

103. Mitchell GF, Parise H, Benjamin EJ, Larson MG, Keyes MJ, Vita JA, et al. Changes in arterial stiffness and wave reflection with advancing age in healthy men and women: the Framingham Heart Study. Hypertension. 2004;43(6):1239-45.

104. Mitchell GF, Vita JA, Larson MG, Parise H, Keyes MJ, Warner E, et al. Cross-sectional relations of peripheral microvascular function, cardiovascular disease risk factors, and aortic stiffness: the Framingham Heart Study. Circulation. 2005;112(24):3722-8.

105. Caro CG, Fitz-Gerald JM, Schroter RC. Arterial wall shear and distribution of early atheroma in man. Nature. 1969;223(5211):1159-60.

106. Ohyama Y, Ambale-Venkatesh B, Noda C, Chugh AR, Teixido-Tura G, Kim JY, et al. Association of Aortic Stiffness With Left Ventricular Remodeling and Reduced Left Ventricular Function Measured by Magnetic Resonance Imaging: The Multi-Ethnic Study of Atherosclerosis. Circulation Cardiovascular imaging. 2016;9(7).

107. Cooper LL, Musani SK, Washington F, Moore J, Tripathi A, Tsao CW, et al. Relations of Microvascular Function, Cardiovascular Disease Risk Factors, and Aortic Stiffness in Blacks: The Jackson Heart Study. Journal of the American Heart Association. 2018;7(20):e009515.

108. Cooper LL, Palmisano JN, Benjamin EJ, Larson MG, Vasan RS, Mitchell GF, et al. Microvascular Function Contributes to the Relation Between Aortic Stiffness and Cardiovascular Events: The Framingham Heart Study. Circulation Cardiovascular imaging. 2016;9(12).

109. Woodard T, Sigurdsson S, Gotal JD, Torjesen AA, Inker LA, Aspelund T, et al. Mediation analysis of aortic stiffness and renal microvascular function. Journal of the American Society of Nephrology : JASN. 2015;26(5):1181-7.

110. Davis EC. Elastic lamina growth in the developing mouse aorta. The journal of histochemistry and cytochemistry : official journal of the Histochemistry Society. 1995;43(11):1115-23.

111. Tare M, Parkington HC, Bubb KJ, Wlodek ME. Uteroplacental insufficiency and lactational environment separately influence arterial stiffness and vascular function in adult male rats. Hypertension. 2012;60(2):378-86.

112. Fall CHD, Kumaran K. Metabolic programming in early life in humans. Philos Trans R Soc Lond B Biol Sci. 2019;374(1770):20180123.

113. Ferrucci L, Fabbri E. Inflammageing: chronic inflammation in ageing, cardiovascular disease, and frailty. Nature reviews Cardiology. 2018;15(9):505-22.

114. Abaci O, Kocas C, Kilickesmez KO, Uner S, Kucukoglu S. Matrix metalloproteinase-2 and -9 levels in patients with dilated ascending aorta and bicuspid aortic valve. Echocardiography. 2013;30(2):121-6.

115. Sugawara J, Hayashi K, Yokoi T, Tanaka H. Age-associated elongation of the ascending aorta in adults. JACC Cardiovascular imaging. 2008;1(6):739-48.

116. Lam CS, Xanthakis V, Sullivan LM, Lieb W, Aragam J, Redfield MM, et al. Aortic root remodeling over the adult life course: longitudinal data from the Framingham Heart Study. Circulation. 2010;122(9):884-90.

117. Chuang ML, Gona PN, Qazi S, Musgrave RM, Fox CS, Massaro JM, et al. Aortic Arch Width and Cardiovascular Disease in Men and Women in the Community. Journal of the American Heart Association. 2018;7(12).

118. Vlachopoulos C, Aznaouridis K, Stefanadis C. Aortic stiffness for cardiovascular risk prediction: just measure it, just do it! Journal of the American College of Cardiology. 2014;63(7):647-9.

119. Lalande A, Khau Van Kien P, Walker PM, Zhu L, Legrand L, Claustres M, et al. Compliance and pulse wave velocity assessed by MRI detect early aortic impairment in young patients with mutation of the smooth muscle myosin heavy chain. Journal of magnetic resonance imaging : JMRI. 2008;28(5):1180-7.

120. de Wit A, Vis K, Jeremy RW. Aortic stiffness in heritable aortopathies: relationship to aneurysm growth rate. Heart, lung & circulation. 2013;22(1):3-11.

121. Willum-Hansen T, Staessen JA, Torp-Pedersen C, Rasmussen S, Thijs L, Ibsen H, et al. Prognostic value of aortic pulse wave velocity as index of arterial stiffness in the general population. Circulation. 2006;113(5):664-70.

122. Mitchell GF, Hwang SJ, Vasan RS, Larson MG, Pencina MJ, Hamburg NM, et al. Arterial stiffness and cardiovascular events: the Framingham Heart Study. Circulation. 2010;121(4):505-11.

123. Meaume S, Benetos A, Henry OF, Rudnichi A, Safar ME. Aortic pulse wave velocity predicts cardiovascular mortality in subjects >70 years of age. Arteriosclerosis, thrombosis, and vascular biology. 2001;21(12):2046-50.

124. Safar ME, Henry O, Meaume S. Aortic pulse wave velocity: an independent marker of cardiovascular risk. The American journal of geriatric cardiology. 2002;11(5):295-8.

125. Townsend RR, Anderson AH, Chirinos JA, Feldman HI, Grunwald JE, Nessel L, et al. Association of Pulse Wave Velocity With Chronic Kidney Disease Progression and Mortality: Findings From the CRIC Study (Chronic Renal Insufficiency Cohort). Hypertension. 2018;71(6):1101-7.

126. Adenwalla SF, Graham-Brown MPM, Leone FMT, Burton JO, McCann GP. The importance of accurate measurement of aortic stiffness in patients with chronic kidney disease and end-stage renal disease. Clin Kidney J. 2017;10(4):503-15.

127. Blacher J, Guerin AP, Pannier B, Marchais SJ, Safar ME, London GM. Impact of aortic stiffness on survival in end-stage renal disease. Circulation. 1999;99(18):2434-9.

128. Mark PB, Doyle A, Blyth KG, Patel RK, Weir RA, Steedman T, et al. Vascular function assessed with cardiovascular magnetic resonance predicts survival in patients with advanced chronic kidney disease. Journal of cardiovascular magnetic resonance : official journal of the Society for Cardiovascular Magnetic Resonance. 2008;10:39.

129. Feistritzer HJ, Klug G, Reinstadler SJ, Reindl M, Niess L, Nalbach T, et al. Prognostic Value of Aortic Stiffness in Patients After ST-Elevation Myocardial Infarction. Journal of the American Heart Association. 2017;6(9).

130. Blacher J, Safar ME, Guerin AP, Pannier B, Marchais SJ, London GM. Aortic pulse wave velocity index and mortality in end-stage renal disease. Kidney international. 2003;63(5):1852-60.

131. Maroules CD, Khera A, Ayers C, Goel A, Peshock RM, Abbara S, et al. Cardiovascular outcome associations among cardiovascular magnetic resonance measures of arterial stiffness: the Dallas heart study. Journal of cardiovascular magnetic resonance : official journal of the Society for Cardiovascular Magnetic Resonance. 2014;16:33.

132. Alberta HB, Secor JL, Smits TC, Farber MA, Jordan WD, Matsumura JS. Differences in aortic arch radius of curvature, neck size, and taper in patients with traumatic and aortic disease. The Journal of surgical research. 2013;184(1):613-8.

133. Dumfarth J, Chou AS, Ziganshin BA, Bhandari R, Peterss S, Tranquilli M, et al. Atypical aortic arch branching variants: A novel marker for thoracic aortic disease. The Journal of thoracic and cardiovascular surgery. 2015;149(6):1586-92.

134. Nollen GJ, Groenink M, Tijssen JG, Van Der Wall EE, Mulder BJ. Aortic stiffness and diameter predict progressive aortic dilatation in patients with Marfan syndrome. European heart journal. 2004;25(13):1146-52.

135. Hannuksela M, Johansson B, Carlberg B. Aortic stiffness in families with inherited non-syndromic thoracic aortic disease. Scand Cardiovasc J. 2018;52(6):301-7.

136. Grotenhuis HB, Westenberg JJ, Steendijk P, van der Geest RJ, Ottenkamp J, Bax JJ, et al. Validation and reproducibility of aortic pulse wave velocity as assessed with velocity-encoded MRI. Journal of magnetic resonance imaging : JMRI. 2009;30(3):521-6.

137. Campo D, Khettab H, Yu R, Genain N, Edouard P, Buard N, et al. Measurement of Aortic Pulse Wave Velocity With a Connected Bathroom Scale. American journal of hypertension. 2017;30(9):876-83.

138. Poullis MP, Warwick R, Oo A, Poole RJ. Ascending aortic curvature as an independent risk factor for type A dissection, and ascending aortic aneurysm formation: a mathematical model. European journal of cardio-thoracic surgery : official journal of the European Association for Cardio-thoracic Surgery. 2008;33(6):995-1001.

139. Schnell S, Smith DA, Barker AJ, Entezari P, Honarmand AR, Carr ML, et al. Altered aortic shape in bicuspid aortic valve relatives influences blood flow patterns. European heart journal cardiovascular Imaging. 2016;17(11):1239-47.

140. Cripe L, Andelfinger G, Martin LJ, Shooner K, Benson DW. Bicuspid aortic valve is heritable. Journal of the American College of Cardiology. 2004;44(1):138-43.

141. Loscalzo ML, Goh DL, Loeys B, Kent KC, Spevak PJ, Dietz HC. Familial thoracic aortic dilation and bicommissural aortic valve: a prospective analysis of natural history and inheritance. American journal of medical genetics Part A. 2007;143A(17):1960-7.

142. Biner S, Rafique AM, Ray I, Cuk O, Siegel RJ, Tolstrup K. Aortopathy is prevalent in relatives of bicuspid aortic valve patients. Journal of the American College of Cardiology. 2009;53(24):2288-95.

143. Evangelista A, Isselbacher EM, Bossone E, Gleason TG, Eusanio MD, Sechtem U, et al. Insights From the International Registry of Acute Aortic Dissection: A 20-Year Experience of Collaborative Clinical Research. Circulation. 2018;137(17):1846-60.

144. Fleischer KJ, Nousari HC, Anhalt GJ, Stone CD, Laschinger JC. Immunohistochemical abnormalities of fibrillin in cardiovascular tissues in Marfan's syndrome. The Annals of thoracic surgery. 1997;63(4):1012-7.

145. Pereira L, Andrikopoulos K, Tian J, Lee SY, Keene DR, Ono R, et al. Targetting of the gene encoding fibrillin-1 recapitulates the vascular aspect of Marfan syndrome. Nature genetics. 1997;17(2):218-22.

146. Habashi JP, Judge DP, Holm TM, Cohn RD, Loeys BL, Cooper TK, et al. Losartan, an AT1 antagonist, prevents aortic aneurysm in a mouse model of Marfan syndrome. Science. 2006;312(5770):117-21.

147. Neptune ER, Frischmeyer PA, Arking DE, Myers L, Bunton TE, Gayraud B, et al. Dysregulation of TGF-beta activation contributes to pathogenesis in Marfan syndrome. Nature genetics. 2003;33(3):407-11.

148. Lindsay ME, Dietz HC. Lessons on the pathogenesis of aneurysm from heritable conditions. Nature. 2011;473(7347):308-16.

149. Carta L, Smaldone S, Zilberberg L, Loch D, Dietz HC, Rifkin DB, et al. p38 MAPK is an early determinant of promiscuous Smad2/3 signaling in the aortas of fibrillin-1 (Fbn1)-null mice. The Journal of biological chemistry. 2009;284(9):5630-6.

150. Milleron O, Arnoult F, Ropers J, Aegerter P, Detaint D, Delorme G, et al. Marfan Sartan: a randomized, double-blind, placebo-controlled trial. European heart journal. 2015;36(32):2160-6.

151. Chiu HH, Wu MH, Wang JK, Lu CW, Chiu SN, Chen CA, et al. Losartan added to beta-blockade therapy for aortic root dilation in Marfan syndrome: a randomized, open-label pilot study. Mayo Clinic proceedings. 2013;88(3):271-6.

152. Groenink M, den Hartog AW, Franken R, Radonic T, de Waard V, Timmermans J, et al. Losartan reduces aortic dilatation rate in adults with Marfan syndrome: a randomized controlled trial. European heart journal. 2013;34(45):3491-500.

153. Loeys BL, Chen J, Neptune ER, Judge DP, Podowski M, Holm T, et al. A syndrome of altered cardiovascular, craniofacial, neurocognitive and skeletal development caused by mutations in TGFBR1 or TGFBR2. Nature genetics. 2005;37(3):275-81.

154. Hiratzka LF, Bakris GL, Beckman JA, Bersin RM, Carr VF, Casey DE, Jr., et al. 2010 ACCF/AHA/AATS/ACR/ASA/SCA/SCAI/SIR/STS/SVM guidelines for the diagnosis and management of patients with Thoracic Aortic Disease: a report of the American College of Cardiology Foundation/American Heart Association Task Force on Practice Guidelines, American Association for Thoracic Surgery, American College of Radiology, American Stroke Association, Society of Cardiovascular Anesthesiologists, Society for

Cardiovascular Angiography and Interventions, Society of Interventional Radiology, Society of Thoracic Surgeons, and Society for Vascular Medicine. Circulation. 2010;121(13):e266-369.

155. Boileau C, Guo DC, Hanna N, Regalado ES, Detaint D, Gong L, et al. TGFB2 mutations cause familial thoracic aortic aneurysms and dissections associated with mild systemic features of Marfan syndrome. Nature genetics. 2012;44(8):916-21.

156. Lindsay ME, Schepers D, Bolar NA, Doyle JJ, Gallo E, Fert-Bober J, et al. Loss-of-function mutations in TGFB2 cause a syndromic presentation of thoracic aortic aneurysm. Nature genetics. 2012;44(8):922-7.

157. Bertoli-Avella AM, Gillis E, Morisaki H, Verhagen JMA, de Graaf BM, van de Beek G, et al. Mutations in a TGF-beta ligand, TGFB3, cause syndromic aortic aneurysms and dissections. Journal of the American College of Cardiology. 2015;65(13):1324-36.

158. Micha D, Guo DC, Hilhorst-Hofstee Y, van Kooten F, Atmaja D, Overwater E, et al. SMAD2 Mutations Are Associated with Arterial Aneurysms and Dissections. Human mutation. 2015;36(12):1145-9.

159. van de Laar IM, Oldenburg RA, Pals G, Roos-Hesselink JW, de Graaf BM, Verhagen JM, et al. Mutations in SMAD3 cause a syndromic form of aortic aneurysms and dissections with early-onset osteoarthritis. Nature genetics. 2011;43(2):121-6.

160. Doyle AJ, Doyle JJ, Bessling SL, Maragh S, Lindsay ME, Schepers D, et al. Mutations in the TGF-beta repressor SKI cause Shprintzen-Goldberg syndrome with aortic aneurysm. Nature genetics. 2012;44(11):1249-54.

161. Gallo EM, Loch DC, Habashi JP, Calderon JF, Chen Y, Bedja D, et al. Angiotensin II-dependent TGF-beta signaling contributes to Loeys-Dietz syndrome vascular pathogenesis. The Journal of clinical investigation. 2014;124(1):448-60.

162. Pope FM, Martin GR, Lichtenstein JR, Penttinen R, Gerson B, Rowe DW, et al. Patients with Ehlers-Danlos syndrome type IV lack type III collagen. Proceedings of the National Academy of Sciences of the United States of America. 1975;72(4):1314-6.

163. Pope FM, Nicholls AC, Narcisi P, Temple A, Chia Y, Fryer P, et al. Type III collagen mutations in Ehlers Danlos syndrome type IV and other related disorders. Clin Exp Dermatol. 1988;13(5):285-302.

164. Meester JA, Vandeweyer G, Pintelon I, Lammens M, Van Hoorick L, De Belder S, et al. Loss-of-function mutations in the X-linked biglycan gene cause a severe syndromic form of thoracic aortic aneurysms and dissections. Genetics in medicine : official journal of the American College of Medical Genetics. 2017;19(4):386-95.

165. Milewicz DM, Guo DC, Tran-Fadulu V, Lafont AL, Papke CL, Inamoto S, et al. Genetic basis of thoracic aortic aneurysms and dissections: focus on smooth muscle cell contractile dysfunction. Annu Rev Genomics Hum Genet. 2008;9:283-302.

166. Milewicz DM, Trybus KM, Guo DC, Sweeney HL, Regalado E, Kamm K, et al. Altered Smooth Muscle Cell Force Generation as a Driver of Thoracic Aortic Aneurysms and Dissections. Arteriosclerosis, thrombosis, and vascular biology. 2017;37(1):26-34.

167. Guo DC, Papke CL, Tran-Fadulu V, Regalado ES, Avidan N, Johnson RJ, et al. Mutations in smooth muscle alpha-actin (ACTA2) cause coronary artery disease, stroke, and Moyamoya disease, along with thoracic aortic disease. American journal of human genetics. 2009;84(5):617-27.

168. Shalata A, Mahroom M, Milewicz DM, Limin G, Kassum F, Badarna K, et al. Fatal thoracic aortic aneurysm and dissection in a large family with a novel MYLK gene mutation: delineation of the clinical phenotype. Orphanet J Rare Dis. 2018;13(1):41.

169. Wallace SE, Regalado ES, Gong L, Janda AL, Guo DC, Russo CF, et al. MYLK pathogenic variants aortic disease presentation, pregnancy risk, and characterization of pathogenic missense variants. Genetics in medicine : official journal of the American College of Medical Genetics. 2019;21(1):144-51.

170. Kuang SQ, Medina-Martinez O, Guo DC, Gong L, Regalado ES, Reynolds CL, et al. FOXE3 mutations predispose to thoracic aortic aneurysms and dissections. The Journal of clinical investigation. 2016;126(3):948-61.

171. Guo DC, Gong L, Regalado ES, Santos-Cortez RL, Zhao R, Cai B, et al. MAT2A mutations predispose individuals to thoracic aortic aneurysms. American journal of human genetics. 2015;96(1):170-7.

172. Tischfield MA, Bosley TM, Salih MA, Alorainy IA, Sener EC, Nester MJ, et al. Homozygous HOXA1 mutations disrupt human brainstem, inner ear, cardiovascular and cognitive development. Nature genetics. 2005;37(10):1035-7.

173. Elliott DA, Kirk EP, Yeoh T, Chandar S, McKenzie F, Taylor P, et al. Cardiac homeobox gene NKX2-5 mutations and congenital heart disease: associations with atrial septal defect and hypoplastic left heart syndrome. Journal of the American College of Cardiology. 2003;41(11):2072-6.

174. Kamath BM, Spinner NB, Emerick KM, Chudley AE, Booth C, Piccoli DA, et al. Vascular anomalies in Alagille syndrome: a significant cause of morbidity and mortality. Circulation. 2004;109(11):1354-8.

175. Pannu H, Fadulu VT, Chang J, Lafont A, Hasham SN, Sparks E, et al. Mutations in transforming growth factor-beta receptor type II cause familial thoracic aortic aneurysms and dissections. Circulation. 2005;112(4):513-20.

176. Renard M, Francis C, Ghosh R, Scott AF, Witmer PD, Ades LC, et al. Clinical Validity of Genes for Heritable Thoracic Aortic Aneurysm and Dissection. Journal of the American College of Cardiology. 2018;72(6):605-15.

177. Osler W. The Bicuspid Condition of the Aortic Valves. Transactions of the Association of American Physicians. 1886:185-92.

178. Braverman AC, Guven H, Beardslee MA, Makan M, Kates AM, Moon MR. The bicuspid aortic valve. Current problems in cardiology. 2005;30(9):470-522.

179. Basso C, Boschello M, Perrone C, Mecenero A, Cera A, Bicego D, et al. An echocardiographic survey of primary school children for bicuspid aortic valve. The American journal of cardiology. 2004;93(5):661-3.

180. Masri A, Svensson LG, Griffin BP, Desai MY. Contemporary natural history of bicuspid aortic valve disease: a systematic review. Heart. 2017;103(17):1323-30.

181. Siu SC, Silversides CK. Bicuspid aortic valve disease. Journal of the American College of Cardiology. 2010;55(25):2789-800.

182. Tadros TM, Klein MD, Shapira OM. Ascending aortic dilatation associated with bicuspid aortic valve: pathophysiology, molecular biology, and clinical implications. Circulation. 2009;119(6):880-90.

183. Ward C. Clinical significance of the bicuspid aortic valve. Heart. 2000;83(1):81-5.

184. Benjamin EJ, Muntner P, Alonso A, Bittencourt MS, Callaway CW, Carson AP, et al. Heart Disease and Stroke Statistics-2019 Update: A Report From the American Heart Association. Circulation. 2019;139(10):e56-e528.

185. Prakash SK, Bosse Y, Muehlschlegel JD, Michelena HI, Limongelli G, Della Corte A, et al. A roadmap to investigate the genetic basis of bicuspid aortic valve and its complications: insights from the International BAVCon (Bicuspid Aortic Valve Consortium). Journal of the American College of Cardiology. 2014;64(8):832-9.

186. Huntington K, Hunter AG, Chan KL. A prospective study to assess the frequency of familial clustering of congenital bicuspid aortic valve. Journal of the American College of Cardiology. 1997;30(7):1809-12.

187. Panayotova R, Macnab A, Waterworth PD. A pilot project of familial screening in patients with bicuspid aortic valve disease. The Journal of heart valve disease. 2013;22(2):150-5.

188. Duran AC, Frescura C, Sans-Coma V, Angelini A, Basso C, Thiene G. Bicuspid aortic valves in hearts with other congenital heart disease. The Journal of heart valve disease. 1995;4(6):581-90.

189. Hinton RB, Martin LJ, Rame-Gowda S, Tabangin ME, Cripe LH, Benson DW. Hypoplastic left heart syndrome links to chromosomes 10q and 6q and is genetically related to bicuspid aortic valve. Journal of the American College of Cardiology. 2009;53(12):1065-71.

190. McBride KL, Pignatelli R, Lewin M, Ho T, Fernbach S, Menesses A, et al. Inheritance analysis of congenital left ventricular outflow tract obstruction malformations: Segregation, multiplex relative risk, and heritability. American journal of medical genetics Part A. 2005;134A(2):180-6.

191. Hiratzka LF, Bakris GL, Beckman JA, Bersin RM, Carr VF, Casey DE, Jr., et al. 2010 ACCF/AHA/AATS/ACR/ASA/SCA/SCAI/SIR/STS/SVM Guidelines for the diagnosis and management of patients with thoracic aortic disease. A Report of the American College of Cardiology Foundation/American Heart Association Task Force on Practice Guidelines, American Association for Thoracic Surgery, American College of Radiology,American Stroke Association, Society of Cardiovascular Anesthesiologists, Society for Cardiovascular Angiography and Interventions, Society of Interventional Radiology, Society of Thoracic Surgeons,and Society for Vascular Medicine. Journal of the American College of Cardiology. 2010;55(14):e27-e129.

192. Mu TS, McAdams RM, Bush DM. A case of hypoplastic left heart syndrome and bicuspid aortic valve in monochorionic twins. Pediatric cardiology. 2005;26(6):884-5.

193. Abdulkareem N, Smelt J, Jahangiri M. Bicuspid aortic valve aortopathy: genetics, pathophysiology and medical therapy. Interactive cardiovascular and thoracic surgery. 2013;17(3):554-9.

194. Kim HK, Gottliebson W, Hor K, Backeljauw P, Gutmark-Little I, Salisbury SR, et al. Cardiovascular anomalies in Turner syndrome: spectrum, prevalence, and cardiac MRI findings in a pediatric and young adult population. AJR American journal of roentgenology. 2011;196(2):454-60.

195. Gomez D, Al Haj Zen A, Borges LF, Philippe M, Gutierrez PS, Jondeau G, et al. Syndromic and non-syndromic aneurysms of the human ascending aorta share activation of the Smad2 pathway. The Journal of pathology. 2009;218(1):131-42.

196. Fedak PW, de Sa MP, Verma S, Nili N, Kazemian P, Butany J, et al. Vascular matrix remodeling in patients with bicuspid aortic valve malformations: implications for aortic dilatation. The Journal of thoracic and cardiovascular surgery. 2003;126(3):797-806.

197. Rabkin SW. Differential expression of MMP-2, MMP-9 and TIMP proteins in thoracic aortic aneurysm - comparison with and without bicuspid aortic valve: a meta-analysis. VASA Zeitschrift fur Gefasskrankheiten. 2014;43(6):433-42.

198. Arrington CB, Sower CT, Chuckwuk N, Stevens J, Leppert MF, Yetman AT, et al. Absence of TGFBR1 and TGFBR2 mutations in patients with bicuspid aortic valve and aortic dilation. The American journal of cardiology. 2008;102(5):629-31.

199. Lee TC, Zhao YD, Courtman DW, Stewart DJ. Abnormal aortic valve development in mice lacking endothelial nitric oxide synthase. Circulation. 2000;101(20):2345-8.

200. Aicher D, Urbich C, Zeiher A, Dimmeler S, Schafers HJ. Endothelial nitric oxide synthase in bicuspid aortic valve disease. The Annals of thoracic surgery. 2007;83(4):1290-4.

201. Laforest B, Andelfinger G, Nemer M. Loss of Gata5 in mice leads to bicuspid aortic valve. The Journal of clinical investigation. 2011;121(7):2876-87.

202. Nemer G, Nemer M. Cooperative interaction between GATA5 and NF-ATc regulates endothelial-endocardial differentiation of cardiogenic cells. Development. 2002;129(17):4045-55.

203. Padang R, Bagnall RD, Richmond DR, Bannon PG, Semsarian C. Rare non-synonymous variations in the transcriptional activation domains of GATA5 in bicuspid aortic valve disease. Journal of molecular and cellular cardiology. 2012;53(2):277-81.

204. Foffa I, Ait Ali L, Panesi P, Mariani M, Festa P, Botto N, et al. Sequencing of NOTCH1, GATA5, TGFBR1 and TGFBR2 genes in familial cases of bicuspid aortic valve. BMC medical genetics. 2013;14:44.

205. Bonachea EM, Chang SW, Zender G, LaHaye S, Fitzgerald-Butt S, McBride KL, et al. Rare GATA5 sequence variants identified in individuals with bicuspid aortic valve. Pediatric research. 2014;76(2):211-6.

206. Qu XK, Qiu XB, Yuan F, Wang J, Zhao CM, Liu XY, et al. A novel NKX2.5 loss-of-function mutation associated with congenital bicuspid aortic valve. The American journal of cardiology. 2014;114(12):1891-5.

207. Niessen K, Karsan A. Notch signaling in cardiac development. Circulation research. 2008;102(10):1169-81.

208. Garg V, Muth AN, Ransom JF, Schluterman MK, Barnes R, King IN, et al. Mutations in NOTCH1 cause aortic valve disease. Nature. 2005;437(7056):270-4.

209. Jain D, Dietz HC, Oswald GL, Maleszewski JJ, Halushka MK. Causes and histopathology of ascending aortic disease in children and young adults. Cardiovascular pathology : the official journal of the Society for Cardiovascular Pathology. 2011;20(1):15-25.

210. Thomas PS, Sridurongrit S, Ruiz-Lozano P, Kaartinen V. Deficient signaling via Alk2 (Acvr1) leads to bicuspid aortic valve development. PloS one. 2012;7(4):e35539.

211. Biben C, Weber R, Kesteven S, Stanley E, McDonald L, Elliott DA, et al. Cardiac septal and valvular dysmorphogenesis in mice heterozygous for mutations in the homeobox gene Nkx2-5. Circulation research. 2000;87(10):888-95.

212. Quintero-Rivera F, Xi QJ, Keppler-Noreuil KM, Lee JH, Higgins AW, Anchan RM, et al. MATR3 Disruption in Human and Mouse Associated with Bicuspid Aortic Valve, Aortic Coarctation and Patent Ductus Arteriosus. Human molecular genetics. 2015.

213. Gillis E, Kumar AA, Luyckx I, Preuss C, Cannaerts E, van de Beek G, et al. Candidate Gene Resequencing in a Large Bicuspid Aortic Valve-Associated Thoracic Aortic Aneurysm Cohort: SMAD6 as an Important Contributor. Frontiers in physiology. 2017;8:400.

214. Luyckx I, MacCarrick G, Kempers M, Meester J, Geryl C, Rombouts O, et al. Confirmation of the role of pathogenic SMAD6 variants in bicuspid aortic valve-related aortopathy. European journal of human genetics : EJHG. 2019.

215. Mitchell GF, Verwoert GC, Tarasov KV, Isaacs A, Smith AV, Yasmin, et al. Common genetic variation in the 3'-BCL11B gene desert is associated with carotid-femoral pulse wave velocity and excess cardiovascular disease risk: the AortaGen Consortium. Circulation Cardiovascular genetics. 2012;5(1):81-90.

216. Charitos EI, Stierle U, Petersen M, Mohamed SA, Hanke T, Schmidtke C, et al. The fate of the bicuspid valve aortopathy after aortic valve replacement. European journal of cardio-thoracic surgery : official journal of the European Association for Cardio-thoracic Surgery. 2014;45(5):e128-35.

217. Forte A, Della Corte A, Grossi M, Bancone C, Provenzano R, Finicelli M, et al. Early cell changes and TGFbeta pathway alterations in the aortopathy associated with bicuspid aortic valve stenosis. Clinical science. 2013;124(2):97-108.

218. Kim YG, Sun BJ, Park GM, Han S, Kim DH, Song JM, et al. Aortopathy and bicuspid aortic valve: haemodynamic burden is main contributor to aortic dilatation. Heart. 2012;98(24):1822-7.

219. Michelena HI, Prakash SK, Della Corte A, Bissell MM, Anavekar N, Mathieu P, et al. Bicuspid aortic valve: identifying knowledge gaps and rising to the challenge from the

International Bicuspid Aortic Valve Consortium (BAVCon). Circulation. 2014;129(25):2691-704.

220. Milewicz DM, Regalado ES, Shendure J, Nickerson DA, Guo DC. Successes and challenges of using whole exome sequencing to identify novel genes underlying an inherited predisposition for thoracic aortic aneurysms and acute aortic dissections. Trends in cardiovascular medicine. 2014;24(2):53-60.

221. Robinson PN, Arteaga-Solis E, Baldock C, Collod-Beroud G, Booms P, De Paepe A, et al. The molecular genetics of Marfan syndrome and related disorders. Journal of medical genetics. 2006;43(10):769-87.

222. Hinton RB, Jr., Martin LJ, Tabangin ME, Mazwi ML, Cripe LH, Benson DW. Hypoplastic left heart syndrome is heritable. Journal of the American College of Cardiology. 2007;50(16):1590-5.

223. Acharya A, Hans CP, Koenig SN, Nichols HA, Galindo CL, Garner HR, et al. Inhibitory role of Notch1 in calcific aortic valve disease. PloS one. 2011;6(11):e27743.

224. Bilen E, Akcay M, Bayram NA, Kocak U, Kurt M, Tanboga IH, et al. Aortic elastic properties and left ventricular diastolic function in patients with isolated bicuspid aortic valve. The Journal of heart valve disease. 2012;21(2):189-94.

225. Gurses D, Ozyurek AR, Levent E, Ulger Z. Elastic properties of the abdominal aorta in the children with bicuspid aortic valve: an observational study. Anadolu kardiyoloji dergisi : AKD = the Anatolian journal of cardiology. 2012;12(5):413-9.

226. Nistri S, Sorbo MD, Basso C, Thiene G. Bicuspid aortic valve: abnormal aortic elastic properties. The Journal of heart valve disease. 2002;11(3):369-73; discussion 73-4.

227. Oulego-Erroz I, Alonso-Quintela P, Mora-Matilla M, Gautreaux Minaya S, Lapena-Lopez de Armentia S. Ascending aorta elasticity in children with isolated bicuspid aortic valve. International journal of cardiology. 2013;168(2):1143-6.

228. Pees C, Michel-Behnke I. Morphology of the bicuspid aortic valve and elasticity of the adjacent aorta in children. The American journal of cardiology. 2012;110(9):1354-60.

229. Pichamuthu JE, Phillippi JA, Cleary DA, Chew DW, Hempel J, Vorp DA, et al. Differential tensile strength and collagen composition in ascending aortic aneurysms by aortic valve phenotype. The Annals of thoracic surgery. 2013;96(6):2147-54.

230. Pepe G, Nistri S, Giusti B, Sticchi E, Attanasio M, Porciani C, et al. Identification of fibrillin 1 gene mutations in patients with bicuspid aortic valve (BAV) without Marfan syndrome. BMC medical genetics. 2014;15:23.

231. Ikonomidis JS, Ivey CR, Wheeler JB, Akerman AW, Rice A, Patel RK, et al. Plasma biomarkers for distinguishing etiologic subtypes of thoracic aortic aneurysm disease. The Journal of thoracic and cardiovascular surgery. 2013;145(5):1326-33.

232. Ikonomidis JS, Ruddy JM, Benton SM, Jr., Arroyo J, Brinsa TA, Stroud RE, et al. Aortic dilatation with bicuspid aortic valves: cusp fusion correlates to matrix metalloproteinases and inhibitors. The Annals of thoracic surgery. 2012;93(2):457-63.

233. LeMaire SA, Wang X, Wilks JA, Carter SA, Wen S, Won T, et al. Matrix metalloproteinases in ascending aortic aneurysms: bicuspid versus trileaflet aortic valves. The Journal of surgical research. 2005;123(1):40-8.

234. Mohamed SA, Noack F, Schoellermann K, Karluss A, Radtke A, Schult-Badusche D, et al. Elevation of matrix metalloproteinases in different areas of ascending aortic aneurysms in patients with bicuspid and tricuspid aortic valves. TheScientificWorldJournal. 2012;2012:806261.

235. Folkersen L, Wagsater D, Paloschi V, Jackson V, Petrini J, Kurtovic S, et al. Unraveling divergent gene expression profiles in bicuspid and tricuspid aortic valve patients with thoracic aortic dilatation: the ASAP study. Molecular medicine. 2011;17(11-12):1365-73.

236. Kuang SQ, Kwartler CS, Byanova KL, Pham J, Gong L, Prakash SK, et al. Rare, nonsynonymous variant in the smooth muscle-specific isoform of myosin heavy chain, MYH11, R247C, alters force generation in the aorta and phenotype of smooth muscle cells. Circulation research. 2012;110(11):1411-22.

237. Balistreri CR, Pisano C, Candore G, Maresi E, Codispoti M, Ruvolo G. Focus on the unique mechanisms involved in thoracic aortic aneurysm formation in bicuspid aortic valve versus tricuspid aortic valve patients: clinical implications of a pilot study. European journal of cardio-thoracic surgery : official journal of the European Association for Cardio-thoracic Surgery. 2013;43(6):e180-6.

238. Foffa I, Murzi M, Mariani M, Mazzone AM, Glauber M, Ait Ali L, et al. Angiotensin-converting enzyme insertion/deletion polymorphism is a risk factor for thoracic aortic aneurysm in patients with bicuspid or tricuspid aortic valves. The Journal of thoracic and cardiovascular surgery. 2012;144(2):390-5.

239. Martin M, Pichel IA, Florez Munoz JP, Naves-Diaz M, Palacin M, Cannata-Andia JB, et al. Low transcriptional activity haplotype of matrix metalloproteinase 1 is less frequent in bicuspid aortic valve patients. Gene. 2013;524(2):304-8.

240. Pisano C, Maresi E, Balistreri CR, Candore G, Merlo D, Fattouch K, et al. Histological and genetic studies in patients with bicuspid aortic valve and ascending aorta complications. Interactive cardiovascular and thoracic surgery. 2012;14(3):300-6.

241. Hope MD, Hope TA, Meadows AK, Ordovas KG, Urbania TH, Alley MT, et al. Bicuspid aortic valve: four-dimensional MR evaluation of ascending aortic systolic flow patterns. Radiology. 2010;255(1):53-61.

242. Mahadevia R, Barker AJ, Schnell S, Entezari P, Kansal P, Fedak PW, et al. Bicuspid aortic cusp fusion morphology alters aortic three-dimensional outflow patterns, wall shear stress, and expression of aortopathy. Circulation. 2014;129(6):673-82.

243. Merritt BA, Turin A, Markl M, Malaisrie SC, McCarthy PM, Carr JC. Association between leaflet fusion pattern and thoracic aorta morphology in patients with bicuspid aortic valve. Journal of magnetic resonance imaging : JMRI. 2014;40(2):294-300.

244. Schaefer BM, Lewin MB, Stout KK, Gill E, Prueitt A, Byers PH, et al. The bicuspid aortic valve: an integrated phenotypic classification of leaflet morphology and aortic root shape. Heart. 2008;94(12):1634-8.

245. Vergara C, Viscardi F, Antiga L, Luciani GB. Influence of bicuspid valve geometry on ascending aortic fluid dynamics: a parametric study. Artificial organs. 2012;36(4):368-78.

246. Groenendijk BC, Hierck BP, Gittenberger-De Groot AC, Poelmann RE. Development-related changes in the expression of shear stress responsive genes KLF-2, ET-1, and NOS-3 in the developing cardiovascular system of chicken embryos. Developmental dynamics : an official publication of the American Association of Anatomists. 2004;230(1):57-68.

247. Groenendijk BC, Van der Heiden K, Hierck BP, Poelmann RE. The role of shear stress on ET-1, KLF2, and NOS-3 expression in the developing cardiovascular system of chicken embryos in a venous ligation model. Physiology. 2007;22:380-9.

248. Laforest B, Nemer M. GATA5 interacts with GATA4 and GATA6 in outflow tract development. Developmental biology. 2011;358(2):368-78.

249. Liu A, Nickerson A, Troyer A, Yin X, Cary R, Thornburg K, et al. Quantifying blood flow and wall shear stresses in the outflow tract of chick embryonic hearts. Computers & structures. 2011;89(11-12):855-67.

250. Chiu JJ, Usami S, Chien S. Vascular endothelial responses to altered shear stress: pathologic implications for atherosclerosis. Annals of medicine. 2009;41(1):19-28.

251. Dekker RJ, van Soest S, Fontijn RD, Salamanca S, de Groot PG, VanBavel E, et al. Prolonged fluid shear stress induces a distinct set of endothelial cell genes, most specifically lung Kruppel-like factor (KLF2). Blood. 2002;100(5):1689-98.

252. Grewal N, Gittenberger-de Groot AC, DeRuiter MC, Klautz RJ, Poelmann RE, Duim S, et al. Bicuspid aortic valve: phosphorylation of c-Kit and downstream targets are prognostic for future aortopathy. European journal of cardio-thoracic surgery : official journal of the European Association for Cardio-thoracic Surgery. 2014;46(5):831-9.

253. Drapisz S, Goralczyk T, Jamka-Miszalski T, Olszowska M, Undas A. Nonstenotic bicuspid aortic valve is associated with elevated plasma asymmetric dimethylarginine. Journal of cardiovascular medicine. 2013;14(6):446-52.

254. Maleki S, Bjorck HM, Folkersen L, Nilsson R, Renner J, Caidahl K, et al. Identification of a novel flow-mediated gene expression signature in patients with bicuspid aortic valve. Journal of molecular medicine. 2013;91(1):129-39.

255. Barker AJ, Markl M, Burk J, Lorenz R, Bock J, Bauer S, et al. Bicuspid aortic valve is associated with altered wall shear stress in the ascending aorta. Circulation Cardiovascular imaging. 2012;5(4):457-66.

256. Bissell MM, Hess AT, Biasiolli L, Glaze SJ, Loudon M, Pitcher A, et al. Aortic dilation in bicuspid aortic valve disease: flow pattern is a major contributor and differs with valve fusion type. Circulation Cardiovascular imaging. 2013;6(4):499-507.

257. Borger MA, Fedak PWM, Stephens EH, Gleason TG, Girdauskas E, Ikonomidis JS, et al. The American Association for Thoracic Surgery consensus guidelines on bicuspid aortic valve-related aortopathy: Full online-only version. The Journal of thoracic and cardiovascular surgery. 2018;156(2):e41-e74.

258. Consortium EP. An integrated encyclopedia of DNA elements in the human genome. Nature. 2012;489(7414):57-74.

259. Grody WW. The transformation of medical genetics by clinical genomics: hubris meets humility. Genetics in medicine : official journal of the American College of Medical Genetics. 2019.

260. Consortium HGP. Building on the DNA revolution: Special Issue Science. 2003;300(5617).

261. Herper M. https://www.forbes.com/sites/matthewherper/2017/01/09/illumina-promises-to-sequence-human-genome-for-100-but-not-quite-yet/#5cc73126386d 2017 [

262. Government H. The Health Secretary announces 5 million genomes project. https://www.gov.uk/government/news/matt-hancock-announces-ambition-to-map-5-million-genomes. 2018.

263. Genomes Project C, Auton A, Brooks LD, Durbin RM, Garrison EP, Kang HM, et al. A global reference for human genetic variation. Nature. 2015;526(7571):68-74.

264. Consortium GT. Human genomics. The Genotype-Tissue Expression (GTEx) pilot analysis: multitissue gene regulation in humans. Science. 2015;348(6235):648-60.

265. Consortium GT. The Genotype-Tissue Expression (GTEx) project. Nature genetics. 2013;45(6):580-5.

266. Evangelou E, Warren HR, Mosen-Ansorena D, Mifsud B, Pazoki R, Gao H, et al. Genetic analysis of over 1 million people identifies 535 new loci associated with blood pressure traits. Nature genetics. 2018;50(10):1412-25.

267. Visscher PM, Yang J, Goddard ME. A commentary on 'common SNPs explain a large proportion of the heritability for human height' by Yang et al. (2010). Twin research and human genetics : the official journal of the International Society for Twin Studies. 2010;13(6):517-24.

268. Yang J, Benyamin B, McEvoy BP, Gordon S, Henders AK, Nyholt DR, et al. Common SNPs explain a large proportion of the heritability for human height. Nature genetics. 2010;42(7):565-9.

269. Speed D, Cai N, Consortium U, Johnson MR, Nejentsev S, Balding DJ. Reevaluation of SNP heritability in complex human traits. Nature genetics. 2017;49(7):986-92.

270. Pierrick Wainschtein DPJ, Loic Yengo, Zhili Zheng, TOPMed Anthropometry Working Group, Trans-Omics for Precision Medicine Consortium, L. Adrienne Cupples, Aladdin H. Shadyab, Barbara McKnight, Benjamin M. Shoemaker, Braxton D. Mitchell, Bruce M. Psaty, Charles Kooperberg, Dan Roden, Dawood Darbar, Donna K. Arnett, Elizabeth A. Regan, Eric Boerwinkle, Jerome I. Rotter, Matthew A. Allison, Merry-Lynn N. McDonald, Mina K Chung, Nicholas L. Smith, Patrick T. Ellinor, Ramachandran S. Vasan,

Rasika A. Mathias, Stephen S. Rich, Susan R. Heckbert, Susan Redline, Xiuqing Guo, Y.-D Ida Chen, Ching-Ti Liu, Mariza de Andrade, Lisa R. Yanek, Christine M. Albert, Ryan D. Hernandez, Stephen T. McGarvey, Kari E. North, Leslie A. Lange, Bruce S. Weir, Cathy C. Laurie, Jian Yang, Peter M. Visscher. Recovery of trait heritability from whole genome sequence data. Preprint at https://wwwbiorxivorg/content/101101/588020v1. 2019.

271. Tennessen JA, Bigham AW, O'Connor TD, Fu W, Kenny EE, Gravel S, et al. Evolution and functional impact of rare coding variation from deep sequencing of human exomes. Science. 2012;337(6090):64-9.

272. Li B, Leal SM. Methods for detecting associations with rare variants for common diseases: application to analysis of sequence data. American journal of human genetics. 2008;83(3):311-21.

273. Richards S, Aziz N, Bale S, Bick D, Das S, Gastier-Foster J, et al. Standards and guidelines for the interpretation of sequence variants: a joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology. Genetics in medicine : official journal of the American College of Medical Genetics. 2015;17(5):405-24.

274. Ghosh R, Oak N, Plon SE. Evaluation of in silico algorithms for use with ACMG/AMP clinical variant interpretation guidelines. Genome biology. 2017;18(1):225.

275. Schafer S, de Marvao A, Adami E, Fiedler LR, Ng B, Khin E, et al. Titin-truncating variants affect heart function in disease cohorts and the general population. Nature genetics. 2017;49(1):46-53.

276. Klein RJ, Zeiss C, Chew EY, Tsai JY, Sackler RS, Haynes C, et al. Complement factor H polymorphism in age-related macular degeneration. Science. 2005;308(5720):385-9.

277. Khera AV, Chaffin M, Aragam KG, Haas ME, Roselli C, Choi SH, et al. Genome-wide polygenic scores for common diseases identify individuals with risk equivalent to monogenic mutations. Nature genetics. 2018;50(9):1219-24.

278. Tam V, Patel N, Turcotte M, Bosse Y, Pare G, Meyre D. Benefits and limitations of genome-wide association studies. Nature reviews Genetics. 2019.

279. Brest P, Lapaquette P, Souidi M, Lebrigand K, Cesaro A, Vouret-Craviari V, et al. A synonymous variant in IRGM alters a binding site for miR-196 and causes deregulation of IRGM-dependent xenophagy in Crohn's disease. Nature genetics. 2011;43(3):242-5.

280. Murthy A, Li Y, Peng I, Reichelt M, Katakam AK, Noubade R, et al. A Crohn's disease variant in Atg16l1 enhances its degradation by caspase 3. Nature. 2014;506(7489):456-62.

281. Visscher PM, Wray NR, Zhang Q, Sklar P, McCarthy MI, Brown MA, et al. 10 Years of GWAS Discovery: Biology, Function, and Translation. American journal of human genetics. 2017;101(1):5-22.

282. Hirschhorn JN. Genomewide association studies--illuminating biologic pathways. The New England journal of medicine. 2009;360(17):1699-701.

283. Doche ME, Bochukova EG, Su HW, Pearce LR, Keogh JM, Henning E, et al. Human SH2B1 mutations are associated with maladaptive behaviors and obesity. The Journal of clinical investigation. 2012;122(12):4732-6.

284. Liu R, Zou Y, Hong J, Cao M, Cui B, Zhang H, et al. Rare Loss-of-Function Variants in NPC1 Predispose to Human Obesity. Diabetes. 2017;66(4):935-47.

285. Saeed S, Bonnefond A, Tamanini F, Mirza MU, Manzoor J, Janjua QM, et al. Loss-of-function mutations in ADCY3 cause monogenic severe obesity. Nature genetics. 2018;50(2):175-9.

286. Claussnitzer M, Dankel SN, Kim KH, Quon G, Meuleman W, Haugen C, et al. FTO Obesity Variant Circuitry and Adipocyte Browning in Humans. The New England journal of medicine. 2015;373(10):895-907.

287. Roadmap Epigenomics C, Kundaje A, Meuleman W, Ernst J, Bilenky M, Yen A, et al. Integrative analysis of 111 reference human epigenomes. Nature. 2015;518(7539):317-30.

288. Chau KH, Elefteriades JA. Natural history of thoracic aortic aneurysms: size matters, plus moving beyond size. Prog Cardiovasc Dis. 2013;56(1):74-80.

289. Davies RR, Goldstein LJ, Coady MA, Tittle SL, Rizzo JA, Kopf GS, et al. Yearly rupture or dissection rates for thoracic aortic aneurysms: simple prediction based on size. The Annals of thoracic surgery. 2002;73(1):17-27; discussion -8.

290. Pape LA, Tsai TT, Isselbacher EM, Oh JK, O'Gara P T, Evangelista A, et al. Aortic diameter >or = 5.5 cm is not a good predictor of type A aortic dissection: observations from the International Registry of Acute Aortic Dissection (IRAD). Circulation. 2007;116(10):1120-7.

291. Fatehi Hassanabad A, Barker AJ, Guzzardi D, Markl M, Malaisrie C, McCarthy PM, et al. Evolution of Precision Medicine and Surgical Strategies for Bicuspid Aortic Valve-Associated Aortopathy. Frontiers in physiology. 2017;8:475.

292. Masri A, Kalahasti V, Svensson LG, Alashi A, Schoenhagen P, Roselli EE, et al. Aortic Cross-Sectional Area/Height Ratio and Outcomes in Patients With Bicuspid Aortic Valve and a Dilated Ascending Aorta. Circulation Cardiovascular imaging. 2017;10(6):e006249.

293. Masri A, Kalahasti V, Svensson LG, Roselli EE, Johnston D, Hammer D, et al. Aortic Cross-Sectional Area/Height Ratio and Outcomes in Patients With a Trileaflet Aortic Valve and a Dilated Aorta. Circulation. 2016;134(22):1724-37.

294. Erbel R, Aboyans V, Boileau C, Bossone E, Bartolomeo RD, Eggebrecht H, et al. 2014 ESC Guidelines on the diagnosis and treatment of aortic diseases: Document covering acute and chronic aortic diseases of the thoracic and abdominal aorta of the adult. The Task Force for the Diagnosis and Treatment of Aortic Diseases of the European Society of Cardiology (ESC). European heart journal. 2014;35(41):2873-926.

295. Accf/Aha/Aats/Acr/Asa/Sca/Scai/Sir/Sts/Svm Guidelines For The D, Management Of Patients With Thoracic Aortic Disease Representative M, Hiratzka LF, Creager MA, Isselbacher EM, Svensson LG, et al. Surgery for Aortic Dilatation in Patients With

Bicuspid Aortic Valves: A Statement of Clarification From the American College of Cardiology/American Heart Association Task Force on Clinical Practice Guidelines. Circulation. 2016;133(7):680-6.

296. Davis AE, Lewandowski AJ, Holloway CJ, Ntusi NA, Banerjee R, Nethononda R, et al. Observational study of regional aortic size referenced to body size: production of a cardiovascular magnetic resonance nomogram. Journal of cardiovascular magnetic resonance : official journal of the Society for Cardiovascular Magnetic Resonance. 2014;16:9.

297. Choi HW, Luo T, Navia JA, Kassab GS. Role of Aortic Geometry on Stroke Propensity based on Simulations of Patient-Specific Models. Scientific reports. 2017;7(1):7065.

298. Peiffer V, Rowland EM, Cremers SG, Weinberg PD, Sherwin SJ. Effect of aortic taper on patterns of blood flow and wall shear stress in rabbits: association with age. Atherosclerosis. 2012;223(1):114-21.

299. Rylski B, Desjardins B, Moser W, Bavaria JE, Milewski RK. Gender-related changes in aortic geometry throughout life. European journal of cardio-thoracic surgery : official journal of the European Association for Cardio-thoracic Surgery. 2014;45(5):805-11.

300. Tomita H, Zhilicheva S, Kim S, Maeda N. Aortic arch curvature and atherosclerosis have overlapping quantitative trait loci in a cross between 129S6/SvEvTac and C57BL/6J apolipoprotein E-null mice. Circulation research. 2010;106(6):1052-60.

301. Craiem D, Alsac JM, Casciaro ME, El Batti S, Mousseaux E, Sirieix ME, et al. Association Between Thoracic Aorta Calcium and Thoracic Aorta Geometry in a Cohort of Asymptomatic Participants at Increased Cardiovascular Risk. Revista espanola de cardiologia. 2016;69(9):827-35.

302. Petersen SE, Matthews PM, Bamberg F, Bluemke DA, Francis JM, Friedrich MG, et al. Imaging in population science: cardiovascular magnetic resonance in 100,000 participants of UK Biobank - rationale, challenges and approaches. Journal of cardiovascular magnetic resonance : official journal of the Society for Cardiovascular Magnetic Resonance. 2013;15:46.

303. Armstrong AC, Gidding S, Gjesdal O, Wu C, Bluemke DA, Lima JA. LV mass assessed by echocardiography and CMR, cardiovascular outcomes, and medical practice. JACC Cardiovascular imaging. 2012;5(8):837-48.

304. Jensky-Squires NE, Dieli-Conwright CM, Rossuello A, Erceg DN, McCauley S, Schroeder ET. Validity and reliability of body composition analysers in children and adults. Br J Nutr. 2008;100(4):859-65.

305. Stergiou GS, Palatini P, Asmar R, Bilo G, de la Sierra A, Head G, et al. Blood pressure monitoring: theory and practice. European Society of Hypertension Working Group on Blood Pressure Monitoring and Cardiovascular Variability Teaching Course Proceedings. Blood pressure monitoring. 2018;23(1):1-8.

306. Woodbridge M, Fagiolo G, O'Regan DP. MRIdb: medical image management for biobank research. Journal of digital imaging. 2013;26(5):886-90.

307. Schulz-Menger J, Bluemke DA, Bremerich J, Flamm SD, Fogel MA, Friedrich MG, et al. Standardized image interpretation and post processing in cardiovascular magnetic resonance: Society for Cardiovascular Magnetic Resonance (SCMR) board of trustees task force on standardized post processing. Journal of cardiovascular magnetic resonance : official journal of the Society for Cardiovascular Magnetic Resonance. 2013;15:35.

308. Demertzis S, Hurni S, Stalder M, Gahl B, Herrmann G, Van den Berg J. Aortic arch morphometry in living humans. Journal of anatomy. 2010;217(5):588-96.

309. Burman ED, Keegan J, Kilner PJ. Aortic root measurement by cardiovascular magnetic resonance: specification of planes and lines of measurement and corresponding normal values. Circulation Cardiovascular imaging. 2008;1(2):104-13.

310. Vasan RS, Larson MG, Levy D. Determinants of echocardiographic aortic root size. The Framingham Heart Study. Circulation. 1995;91(3):734-40.

311. de Simone G, Roman MJ, De Marco M, Bella JN, Izzo R, Lee ET, et al. Hemodynamic Correlates of Abnormal Aortic Root Dimension in an Adult Population: The Strong Heart Study. Journal of the American Heart Association. 2015;4(10):e002309.

312. Teixido-Tura G, Almeida AL, Choi EY, Gjesdal O, Jacobs DR, Jr., Dietz HC, et al. Determinants of Aortic Root Dilatation and Reference Values Among Young Adults Over a 20-Year Period: Coronary Artery Risk Development in Young Adults Study. Hypertension. 2015;66(1):23-9.

313. Agmon Y, Khandheria BK, Meissner I, Schwartz GL, Sicks JD, Fought AJ, et al. Is aortic dilatation an atherosclerosis-related process? Clinical, laboratory, and transesophageal echocardiographic correlates of thoracic aortic dimensions in the population with implications for thoracic aortic aneurysm formation. Journal of the American College of Cardiology. 2003;42(6):1076-83.

314. Osler W. The Principles and Practice of Medicine. 1892.

315. Redheuil A, Wu CO, Kachenoura N, Ohyama Y, Yan RT, Bertoni AG, et al. Proximal aortic distensibility is an independent predictor of all-cause mortality and incident CV events: the MESA study. Journal of the American College of Cardiology. 2014;64(24):2619-29.

316. Abhayaratna WP, Srikusalanukul W, Budge MM. Aortic stiffness for the detection of preclinical left ventricular diastolic dysfunction: pulse wave velocity versus pulse pressure. Journal of hypertension. 2008;26(4):758-64.

317. Van Bortel LM, Laurent S, Boutouyrie P, Chowienczyk P, Cruickshank JK, De Backer T, et al. Expert consensus document on the measurement of aortic stiffness in daily practice using carotid-femoral pulse wave velocity. Journal of hypertension. 2012;30(3):445-8.

318. Wentland AL, Grist TM, Wieben O. Review of MRI-based measurements of pulse wave velocity: a biomarker of arterial stiffness. Cardiovascular diagnosis and therapy. 2014;4(2):193-206.

319. Dogui A, Redheuil A, Lefort M, DeCesare A, Kachenoura N, Herment A, et al. Measurement of aortic arch pulse wave velocity in cardiovascular MR: comparison of transit time estimators and description of a new approach. Journal of magnetic resonance imaging : JMRI. 2011;33(6):1321-9.

320. Alfakih K, Plein S, Thiele H, Jones T, Ridgway JP, Sivananthan MU. Normal human left and right ventricular dimensions for MRI as assessed by turbo gradient echo and steady-state free precession imaging sequences. Journal of magnetic resonance imaging : JMRI. 2003;17(3):323-9.

321. Nethononda RM, Lewandowski AJ, Stewart R, Kylinterias I, Whitworth P, Francis J, et al. Gender specific patterns of age-related decline in aortic stiffness: a cardiovascular magnetic resonance study including normal ranges. Journal of cardiovascular magnetic resonance : official journal of the Society for Cardiovascular Magnetic Resonance. 2015;17:20.

322. Voges I, Jerosch-Herold M, Hedderich J, Pardun E, Hart C, Gabbert DD, et al. Normal values of aortic dimensions, distensibility, and pulse wave velocity in children and young adults: a cross-sectional study. Journal of cardiovascular magnetic resonance : official journal of the Society for Cardiovascular Magnetic Resonance. 2012;14:77.

323. Ogola BO, Zimmerman MA, Clark GL, Abshire CM, Gentry KM, Miller KS, et al. New insights into arterial stiffening: does sex matter? American journal of physiology Heart and circulatory physiology. 2018;315(5):H1073-H87.

324. Subramanya V, Ambale-Venkatesh B, Ohyama Y, Zhao D, Nwabuo CC, Post WS, et al. Relation of Sex Hormone Levels With Prevalent and 10-Year Change in Aortic Distensibility Assessed by MRI: The Multi-Ethnic Study of Atherosclerosis. American journal of hypertension. 2018;31(7):774-83.

325. Yaron M, Greenman Y, Rosenfeld JB, Izkhakov E, Limor R, Osher E, et al. Effect of testosterone replacement therapy on arterial stiffness in older hypogonadal men. European journal of endocrinology / European Federation of Endocrine Societies. 2009;160(5):839-46.

326. Smith JC, Bennett S, Evans LM, Kynaston HG, Parmar M, Mason MD, et al. The effects of induced hypogonadism on arterial stiffness, body composition, and metabolic parameters in males with prostate cancer. J Clin Endocrinol Metab. 2001;86(9):4261-7.

327. Yue P, Chatterjee K, Beale C, Poole-Wilson PA, Collins P. Testosterone relaxes rabbit coronary arteries and aorta. Circulation. 1995;91(4):1154-60.

328. Chakrabarti S, Morton JS, Davidge ST. Mechanisms of estrogen effects on the endothelium: an overview. The Canadian journal of cardiology. 2014;30(7):705-12.

329. Gompel A, Boutouyrie P, Joannides R, Christin-Maitre S, Kearny-Schwartz A, Kunz K, et al. Association of menopause and hormone replacement therapy with large artery remodeling. Fertility and sterility. 2011;96(6):1445-50.

330. Manrique C, Lastra G, Ramirez-Perez FI, Haertling D, DeMarco VG, Aroor AR, et al. Endothelial Estrogen Receptor-alpha Does Not Protect Against Vascular Stiffness Induced by Western Diet in Female Mice. Endocrinology. 2016;157(4):1590-600.

331. Manrique-Acevedo C, Ramirez-Perez FI, Padilla J, Vieira-Potter VJ, Aroor AR, Barron BJ, et al. Absence of Endothelial ERalpha Results in Arterial Remodeling and Decreased Stiffness in Western Diet-Fed Male Mice. Endocrinology. 2017;158(6):1875-85.

332. Douglas G, Cruz MN, Poston L, Gustafsson JA, Kublickiene K. Functional characterization and sex differences in small mesenteric arteries of the estrogen receptor-beta knockout mouse. American journal of physiology Regulatory, integrative and comparative physiology. 2008;294(1):R112-20.

333. Liu L, Kashyap S, Murphy B, Hutson DD, Budish RA, Trimmer EH, et al. GPER activation ameliorates aortic remodeling induced by salt-sensitive hypertension. American journal of physiology Heart and circulatory physiology. 2016;310(8):H953-61.

334. Decano JL, Pasion KA, Black N, Giordano NJ, Herrera VL, Ruiz-Opazo N. Sex-specific genetic determinants for arterial stiffness in Dahl salt-sensitive hypertensive rats. BMC genetics. 2016;17:19.

335. Coutinho T, Bailey KR, Turner ST, Kullo IJ. Arterial stiffness is associated with increase in blood pressure over time in treated hypertensives. Journal of the American Society of Hypertension : JASH. 2014;8(6):414-21.

336. Kaess BM, Rong J, Larson MG, Hamburg NM, Vita JA, Levy D, et al. Aortic stiffness, blood pressure progression, and incident hypertension. Jama. 2012;308(9):875-81.

337. Chen W, Li S, Fernandez C, Sun D, Lai CC, Zhang T, et al. Temporal Relationship Between Elevated Blood Pressure and Arterial Stiffening Among Middle-Aged Black and White Adults: The Bogalusa Heart Study. American journal of epidemiology. 2016;183(7):599-608.

338. Anderson CA, Pettersson FH, Clarke GM, Cardon LR, Morris AP, Zondervan KT. Data quality control in genetic case-control association studies. Nature protocols. 2010;5(9):1564-73.

339. Johnson EO, Hancock DB, Levy JL, Gaddis NC, Saccone NL, Bierut LJ, et al. Imputation across genotyping arrays for genome-wide association studies: assessment of bias and a correction strategy. Human genetics. 2013;132(5):509-22.

340. Howie B, Fuchsberger C, Stephens M, Marchini J, Abecasis GR. Fast and accurate genotype imputation in genome-wide association studies through pre-phasing. Nature genetics. 2012;44(8):955-9.

341. Howie B, Marchini J, Stephens M. Genotype imputation with thousands of genomes. G3 (Bethesda). 2011;1(6):457-70.

342. Marchini J, Howie B. Genotype imputation for genome-wide association studies. Nature reviews Genetics. 2010;11(7):499-511.

343. McLaren W, Gil L, Hunt SE, Riat HS, Ritchie GR, Thormann A, et al. The Ensembl Variant Effect Predictor. Genome biology. 2016;17(1):122.

344. Sherry ST, Ward MH, Kholodov M, Baker J, Phan L, Smigielski EM, et al. dbSNP: the NCBI database of genetic variation. Nucleic acids research. 2001;29(1):308-11.

345. Yang J, Lee SH, Goddard ME, Visscher PM. GCTA: a tool for genome-wide complex trait analysis. American journal of human genetics. 2011;88(1):76-82.

346. Yang J, Lee SH, Wray NR, Goddard ME, Visscher PM. GCTA-GREML accounts for linkage disequilibrium when estimating genetic variance from genome-wide SNPs. Proceedings of the National Academy of Sciences of the United States of America. 2016;113(32):E4579-80.

347. Teo YY, Inouye M, Small KS, Gwilliam R, Deloukas P, Kwiatkowski DP, et al. A genotype calling algorithm for the Illumina BeadArray platform. Bioinformatics. 2007;23(20):2741-6.

348. Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MA, Bender D, et al. PLINK: a tool set for whole-genome association and population-based linkage analyses. American journal of human genetics. 2007;81(3):559-75.

349. Wellcome Trust Case Control C. Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls. Nature. 2007;447(7145):661-78.

350. Consortium UK, Walter K, Min JL, Huang J, Crooks L, Memari Y, et al. The UK10K project identifies rare variants in health and disease. Nature. 2015;526(7571):82-90.

351. Huang J, Howie B, McCarthy S, Memari Y, Walter K, Min JL, et al. Improved imputation of low-frequency and rare variants using the UK10K haplotype reference panel. Nature communications. 2015;6:8111.

352. Chou WC, Zheng HF, Cheng CH, Yan H, Wang L, Han F, et al. A combined reference panel from the 1000 Genomes and UK10K projects improved rare variant imputation in European and Chinese samples. Scientific reports. 2016;6:39313.

353. Sherry ST, Ward M, Sirotkin K. dbSNP-database for single nucleotide polymorphisms and other classes of minor genetic variation. Genome research. 1999;9(8):677-9.

354. Delaneau O, Zagury JF, Marchini J. Improved whole-chromosome phasing for disease and population genetic studies. Nature methods. 2013;10(1):5-6.

355. International HapMap C, Altshuler DM, Gibbs RA, Peltonen L, Altshuler DM, Gibbs RA, et al. Integrating common and rare genetic variation in diverse human populations. Nature. 2010;467(7311):52-8.

356. Marchini J, Howie B, Myers S, McVean G, Donnelly P. A new multipoint method for genome-wide association studies by imputation of genotypes. Nature genetics. 2007;39(7):906-13.

357. Howie BN, Donnelly P, Marchini J. A flexible and accurate genotype imputation method for the next generation of genome-wide association studies. PLoS genetics. 2009;5(6):e1000529.

358. Bakshi A, Zhu Z, Vinkhuyzen AA, Hill WD, McRae AF, Visscher PM, et al. Fast set-based association analysis using summary data from GWAS identifies novel gene loci for human complex traits. Scientific reports. 2016;6:32894.

359. Magi R, Suleimanov YV, Clarke GM, Kaakinen M, Fischer K, Prokopenko I, et al. SCOPA and META-SCOPA: software for the analysis and aggregation of genome-wide

association studies of multiple correlated phenotypes. BMC bioinformatics. 2017;18(1):25.

360. Ray D, Boehnke M. Methods for meta-analysis of multiple traits using GWAS summary statistics. Genetic epidemiology. 2018;42(2):134-45.

361. Myocardial Infarction Genetics C, Kathiresan S, Voight BF, Purcell S, Musunuru K, Ardissino D, et al. Genome-wide association of early-onset myocardial infarction with single nucleotide polymorphisms and copy number variants. Nature genetics. 2009;41(3):334-41.

362. Maskari RA, Hardege I, Cleary S, Figg N, Li Y, Siew K, et al. Functional characterization of common BCL11B gene desert variants suggests a lymphocyte-mediated association of BCL11B with aortic stiffness. European journal of human genetics : EJHG. 2018;26(11):1648-57.

363. Levy D, Larson MG, Benjamin EJ, Newton-Cheh C, Wang TJ, Hwang SJ, et al. Framingham Heart Study 100K Project: genome-wide associations for blood pressure and arterial stiffness. BMC medical genetics. 2007;8 Suppl 1:S3.

364. Tarasov KV, Sanna S, Scuteri A, Strait JB, Orru M, Parsa A, et al. COL4A1 is associated with arterial stiffness by genome-wide association scan. Circulation Cardiovascular genetics. 2009;2(2):151-8.

365. Yasmin, Maskari RA, McEniery CM, Cleary SE, Li Y, Siew K, et al. The matrix proteins aggrecan and fibulin-1 play a key role in determining aortic stiffness. Scientific reports. 2018;8(1):8550.

366. Mangino M, Cecelja M, Menni C, Tsai PC, Yuan W, Small K, et al. Integrated multiomics approach identifies calcium and integrin-binding protein-2 as a novel gene for pulse wave velocity. Journal of hypertension. 2016;34(1):79-87.

367. Wild PS, Felix JF, Schillert A, Teumer A, Chen MH, Leening MJG, et al. Large-scale genome-wide analysis identifies genetic variants associated with cardiac structure and function. The Journal of clinical investigation. 2017;127(5):1798-812.

368. Vasan RS, Glazer NL, Felix JF, Lieb W, Wild PS, Felix SB, et al. Genetic variants associated with cardiac structure and function: a meta-analysis and replication of genome-wide association data. Jama. 2009;302(2):168-78.

369. Rebhan M, Chalifa-Caspi V, Prilusky J, Lancet D. GeneCards: integrating information about genes, proteins and diseases. Trends Genet. 1997;13(4):163.

370. Law M, Shaw DR. Mouse Genome Informatics (MGI) Is the International Resource for Information on the Laboratory Mouse. Methods in molecular biology. 2018;1757:141-61.

371. Provencio I, Rodriguez IR, Jiang G, Hayes WP, Moreira EF, Rollag MD. A novel human opsin in the inner retina. The Journal of neuroscience : the official journal of the Society for Neuroscience. 2000;20(2):600-5.

372. Vatta M, Mohapatra B, Jimenez S, Sanchez X, Faulkner G, Perles Z, et al. Mutations in Cypher/ZASP in patients with dilated cardiomyopathy and left ventricular non-compaction. Journal of the American College of Cardiology. 2003;42(11):2014-27.

373. El-Bizri N, Wang L, Merklinger SL, Guignabert C, Desai T, Urashima T, et al. Smooth muscle protein 22alpha-mediated patchy deletion of Bmpr1a impairs cardiac contractility but protects against pulmonary vascular remodeling. Circulation research. 2008;102(3):380-8.

374. Calva-Cerqueira D, Chinnathambi S, Pechman B, Bair J, Larsen-Haidle J, Howe JR. The rate of germline mutations and large deletions of SMAD4 and BMPR1A in juvenile polyposis. Clinical genetics. 2009;75(1):79-85.

375. Zhang X, Szeto C, Gao E, Tang M, Jin J, Fu Q, et al. Cardiotoxic and cardioprotective features of chronic beta-adrenergic signaling. Circulation research. 2013;112(3):498-509.

376. Group TMGI. Mouse Genome Informatics, The Jackson Laboratory, Bar Harbor, Maine. World Wide Web (URL: http://www.informatics.jax.org) [

377. Aghajanian H, Cho YK, Rizer NW, Wang Q, Li L, Degenhardt K, et al. Pdgfralpha functions in endothelial-derived cells to regulate neural crest cells and the development of the great arteries. Dis Model Mech. 2017;10(9):1101-8.

378. Trueb B. Biology of FGFRL1, the fifth fibroblast growth factor receptor. Cell Mol Life Sci. 2011;68(6):951-64.

379. Grabowski K, Riemenschneider M, Schulte L, Witten A, Schulz A, Stoll M, et al. Fetal-adult cardiac transcriptome analysis in rats with contrasting left ventricular mass reveals new candidates for cardiac hypertrophy. PloS one. 2015;10(2):e0116807.

380. GeneCards: The human gene database. 2019; https://www.genecards.org.

381. Yeganeh F, Ranjbar A, Hatam M, Nasimi A. Mechanism of the cardiovascular effects of the GABAA receptors of the ventral tegmental area of the rat brain. Neuroscience letters. 2015;600:193-8.

382. Stevenson ER, Johns EM, Marques FZ, Jackson KL, Davern PJ, Evans RG, et al. Positive allosteric modulation of GABAA receptors attenuates high blood pressure in Schlager hypertensive mice. Journal of hypertension. 2017;35(3):546-57.

383. van der Harst P, Verweij N. Identification of 64 Novel Genetic Loci Provides an Expanded View on the Genetic Architecture of Coronary Artery Disease. Circulation research. 2018;122(3):433-43.

384. Barroso MC, Kramer F, Greene SJ, Scheyer D, Kohler T, Karoff M, et al. Serum insulin-like growth factor-1 and its binding protein-7: potential novel biomarkers for heart failure with preserved ejection fraction. BMC cardiovascular disorders. 2016;16(1):199.

385. Wei CD, Zheng HY, Wu W, Dai W, Tong YQ, Wang M, et al. Meta-analysis of the association of the rs2234693 and rs9340799 polymorphisms of estrogen receptor alpha gene with coronary heart disease risk in Chinese Han population. International journal of medical sciences. 2013;10(4):457-66.

386. Arbit B, Marston N, Shah K, Lee EL, Aramin H, Clopton P, et al. Prognostic Usefulness of Proenkephalin in Stable Ambulatory Patients With Heart Failure. The American journal of cardiology. 2016;117(8):1310-4.

387. Matsue Y, Ter Maaten JM, Struck J, Metra M, O'Connor CM, Ponikowski P, et al. Clinical Correlates and Prognostic Value of Proenkephalin in Acute and Chronic Heart Failure. J Card Fail. 2017;23(3):231-9.

388. Ng LL, Squire IB, Jones DJL, Cao TH, Chan DCS, Sandhu JK, et al. Proenkephalin, Renal Dysfunction, and Prognosis in Patients With Acute Heart Failure: A GREAT Network Study. Journal of the American College of Cardiology. 2017;69(1):56-69.

389. Siong Chan DC, Cao TH, Ng LL. Proenkephalin in Heart Failure. Heart failure clinics. 2018;14(1):1-11.

390. Kanagala P, Squire IB, Jones DJL, Cao TH, Chan DCS, McCann G, et al. Proenkephalin and prognosis in heart failure with preserved ejection fraction: a GREAT network study. Clinical research in cardiology : official journal of the German Cardiac Society. 2019.

391. Lovero KL, Blankenship SM, Shi Y, Nicoll RA. SynDIG1 promotes excitatory synaptogenesis independent of AMPA receptor trafficking and biophysical regulation. PloS one. 2013;8(6):e66171.

392. Sanchez-Soria P, Camenisch TD. ErbB signaling in cardiac development and disease. Semin Cell Dev Biol. 2010;21(9):929-35.

393. Wu JH, Lemaitre RN, Manichaikul A, Guan W, Tanaka T, Foy M, et al. Genome-wide association study identifies novel loci associated with concentrations of four plasma phospholipid fatty acids in the de novo lipogenesis pathway: results from the Cohorts for Heart and Aging Research in Genomic Epidemiology (CHARGE) consortium. Circulation Cardiovascular genetics. 2013;6(2):171-83.

394. Joehanes R, Ying S, Huan T, Johnson AD, Raghavachari N, Wang R, et al. Gene expression signatures of coronary heart disease. Arteriosclerosis, thrombosis, and vascular biology. 2013;33(6):1418-26.

395. Zhou X, Yang H, Song X, Wang J, Shen L, Wang J. Central blockade of the AT1 receptor attenuates pressor effects via reduction of glutamate release and downregulation of NMDA/AMPA receptors in the rostral ventrolateral medulla of rats with stress-induced hypertension. Hypertension research : official journal of the Japanese Society of Hypertension. 2019.

396. Ferreira-Junior NC, Lagatta DC, Resstel LBM. Glutamatergic, GABAergic, and endocannabinoid neurotransmissions within the dorsal hippocampus modulate the cardiac baroreflex function in rats. Pflugers Archiv : European journal of physiology. 2018;470(2):395-411.

397. Fujita S, Yokoyama U, Ishiwata R, Aoki R, Nagao K, Masukawa D, et al. Glutamate Promotes Contraction of the Rat Ductus Arteriosus. Circulation journal : official journal of the Japanese Circulation Society. 2016;80(11):2388-96.

398. Huan T, Zhang B, Wang Z, Joehanes R, Zhu J, Johnson AD, et al. A systems biology framework identifies molecular underpinnings of coronary heart disease. Arteriosclerosis, thrombosis, and vascular biology. 2013;33(6):1427-34.

399. https://www.proteomicsdb.org/.

400. Dai DF, Thajeb P, Tu CF, Chiang FT, Chen CH, Yang RB, et al. Plasma concentration of SCUBE1, a novel platelet protein, is elevated in patients with acute coronary syndrome and ischemic stroke. Journal of the American College of Cardiology. 2008;51(22):2173-80.

401. Ozkan G, Ulusoy S, Mentese A, Karahan SC, Cansiz M. New marker of platelet activation, SCUBE1, is elevated in hypertensive patients. American journal of hypertension. 2013;26(6):748-53.

402. Wang Q, Lin JL, Erives AJ, Lin CI, Lin JJ. New insights into the roles of Xin repeat-containing proteins in cardiac development, function, and disease. Int Rev Cell Mol Biol. 2014;310:89-128.

403. Nakamura H, Cook RN, Justice MJ. Mouse Tenm4 is required for mesoderm induction. BMC Dev Biol. 2013;13:9.

404. Besse S, Comte R, Frechault S, Courty J, Joel de L, Delbe J. Pleiotrophin promotes capillary-like sprouting from senescent aortic rings. Cytokine. 2013;62(1):44-7.

405. Chao YX, Lin Ng EY, Tio M, Kumar P, Tan L, Au WL, et al. Essential tremor linked TENM4 mutation found in healthy Chinese individuals. Parkinsonism Relat Disord. 2016;31:139-40.

406. Smith-White MA, Herzog H, Potter EK. Cardiac function in neuropeptide Y Y4 receptor-knockout mice. Regulatory peptides. 2002;110(1):47-54.

407. Zhang J, Zhang H, Chen Y, Fu J, Lei Y, Sun J, et al. Plateletderived growth factor D promotes the angiogenic capacity of endothelial progenitor cells. Molecular medicine reports. 2019;19(1):125-32.

408. Alehagen U, Olsen RS, Lanne T, Matussek A, Wagsater D. PDGF-D gene polymorphism is associated with increased cardiovascular mortality in elderly men. BMC medical genetics. 2016;17(1):62.

409. Zhang ZB, Ruan CC, Lin JR, Xu L, Chen XH, Du YN, et al. Perivascular Adipose Tissue-Derived PDGF-D Contributes to Aortic Aneurysm Formation During Obesity. Diabetes. 2018;67(8):1549-60.

410. De Cario R, Sticchi E, Lucarini L, Attanasio M, Nistri S, Marcucci R, et al. Role of TGFBR1 and TGFBR2 genetic variants in Marfan syndrome. Journal of vascular surgery. 2018;68(1):225-33 e5.

411. Hiltunen JO, Arumae U, Moshnyakov M, Saarma M. Expression of mRNAs for neurotrophins and their receptors in developing rat heart. Circulation research. 1996;79(5):930-9.

412. Malhotra R, Wunderer F, Barnes HJ, Bagchi A, Buswell MD, O'Rourke CD, et al. Hepcidin Deficiency Protects Against Atherosclerosis. Arteriosclerosis, thrombosis, and vascular biology. 2019;39(2):178-87.

413. Chen S, Knight WE, Yan C. Roles of PDE1 in Pathological Cardiac Remodeling and Dysfunction. J Cardiovasc Dev Dis. 2018;5(2).

414. Knight WE, Chen S, Zhang Y, Oikawa M, Wu M, Zhou Q, et al. PDE1C deficiency antagonizes pathological cardiac remodeling and dysfunction. Proceedings of the National Academy of Sciences of the United States of America. 2016;113(45):E7116-E25.

415. Cai Y, Nagel DJ, Zhou Q, Cygnar KD, Zhao H, Li F, et al. Role of cAMP-phosphodiesterase 1C signaling in regulating growth factor receptor stability, vascular smooth muscle cell growth, migration, and neointimal hyperplasia. Circulation research. 2015;116(7):1120-32.

416. Krivospitskaya O, Elmabsout AA, Sundman E, Soderstrom LA, Ovchinnikova O, Gidlof AC, et al. A CYP26B1 polymorphism enhances retinoic acid catabolism and may aggravate atherosclerosis. Molecular medicine. 2012;18:712-8.

417. Chowdhury A, Hasselbach L, Echtermeyer F, Jyotsana N, Theilmeier G, Herzog C. Fibulin-6 regulates pro-fibrotic TGF-beta responses in neonatal mouse ventricular cardiac fibroblasts. Scientific reports. 2017;7:42725.

418. Chowdhury A, Herzog C, Hasselbach L, Khouzani HL, Zhang J, Hammerschmidt M, et al. Expression of fibulin-6 in failing hearts and its role for cardiac fibroblast migration. Cardiovascular research. 2014;103(4):509-20.

419. Tartaglia M, Mehler EL, Goldberg R, Zampino G, Brunner HG, Kremer H, et al. Mutations in PTPN11, encoding the protein tyrosine phosphatase SHP-2, cause Noonan syndrome. Nature genetics. 2001;29(4):465-8.

420. Dassanayaka S, Zheng Y, Gibb AA, Cummins TD, McNally LA, Brittian KR, et al. Cardiac-specific overexpression of aldehyde dehydrogenase 2 exacerbates cardiac remodeling in response to pressure overload. Redox Biol. 2018;17:440-9.

421. Chu AY, Deng X, Fisher VA, Drong A, Zhang Y, Feitosa MF, et al. Multiethnic genome-wide meta-analysis of ectopic fat depots identifies loci associated with adipocyte development and differentiation. Nature genetics. 2017;49(1):125-30.

422. Bis JC, White CC, Franceschini N, Brody J, Zhang X, Muzny D, et al. Sequencing of 2 subclinical atherosclerosis candidate regions in 3669 individuals: Cohorts for Heart and Aging Research in Genomic Epidemiology (CHARGE) Consortium Targeted Sequencing Study. Circulation Cardiovascular genetics. 2014;7(3):359-64.

423. Xu S, Koroleva M, Yin M, Jin ZG. Atheroprotective laminar flow inhibits Hippo pathway effector YAP in endothelial cells. Transl Res. 2016;176:18-28 e2.

424. Matsui Y, Nakano N, Shao D, Gao S, Luo W, Hong C, et al. Lats2 is a negative regulator of myocyte size in the heart. Circulation research. 2008;103(11):1309-18.

425. Martinez HR, Craigen WJ, Ummat M, Adesina AM, Lotze TE, Jefferies JL. Novel cardiovascular findings in association with a POMT2 mutation: three siblings with alpha-dystroglycanopathy. European journal of human genetics : EJHG. 2014;22(4):486-91.

426. Khan AA, Wang Y, Sun Y, Mao XO, Xie L, Miles E, et al. Neuroglobin-overexpressing transgenic mice are resistant to cerebral and myocardial ischemia. Proceedings of the National Academy of Sciences of the United States of America. 2006;103(47):17944-8.

427. Doyle AJ, Redmond EM, Gillespie DL, Knight PA, Cullen JP, Cahill PA, et al. Differential expression of Hedgehog/Notch and transforming growth factor-beta in human abdominal aortic aneurysms. Journal of vascular surgery. 2015;62(2):464-70.

428. Salybekov AA, Salybekova AK, Pola R, Asahara T. Sonic Hedgehog Signaling Pathway in Endothelial Progenitor Cell Biology for Vascular Medicine. International journal of molecular sciences. 2018;19(10).

429. Wang Y, Lu P, Zhao D, Sheng J. Targeting the hedgehog signaling pathway for cardiac repair and regeneration. Herz. 2017;42(7):662-8.

430. Olson TM, Doan TP, Kishimoto NY, Whitby FG, Ackerman MJ, Fananapazir L. Inherited and de novo mutations in the cardiac actin gene cause hypertrophic cardiomyopathy. Journal of molecular and cellular cardiology. 2000;32(9):1687-94.

431. Adhikari N, Billaud M, Carlson M, Lake SP, Montaniel KR, Staggs R, et al. Vascular biomechanical properties in mice with smooth muscle specific deletion of Ndst1. Mol Cell Biochem. 2014;385(1-2):225-38.

432. Trindade PT. Losartan treatment in adult patients with Marfan syndrome: can we finally COMPARE? European heart journal. 2013;34(45):3469-71.

433. Frustaci A, De Luca A, Guida V, Biagini T, Mazza T, Gaudio C, et al. Novel alpha-Actin Gene Mutation p.(Ala21Val) Causing Familial Hypertrophic Cardiomyopathy, Myocardial Noncompaction, and Transmural Crypts. Clinical-Pathologic Correlation. Journal of the American Heart Association. 2018;7(4).

434. Ingles J, Goldstein J, Thaxton C, Caleshu C, Corty EW, Crowley SB, et al. Evaluating the Clinical Validity of Hypertrophic Cardiomyopathy Genes. Circ Genom Precis Med. 2019;12(2):e002460.

435. Matsson H, Eason J, Bookwalter CS, Klar J, Gustavsson P, Sunnegardh J, et al. Alpha-cardiac actin mutations produce atrial septal defects. Human molecular genetics. 2008;17(2):256-65.

436. Nicholson CJ, Singh K, Saphirstein RJ, Gao YZ, Li Q, Chiu JG, et al. Reversal of Aging-Induced Increases in Aortic Stiffness by Targeting Cytoskeletal Protein-Protein Interfaces. Journal of the American Heart Association. 2018;7(15).

437. Guillot N, Kollins D, Gilbert V, Xavier S, Chen J, Gentle M, et al. BAMBI regulates angiogenesis and endothelial homeostasis through modulation of alternative TGFbeta signaling. PloS one. 2012;7(6):e39406.

438. Sartorelli V, Webster KA, Kedes L. Muscle-specific expression of the cardiac alpha-actin gene requires MyoD1, CArG-box binding factor, and Sp1. Genes & development. 1990;4(10):1811-22.

439. Institute B. Haploreg v 4.1 queried in 2019: https://pubs.broadinstitute.org/mammals/haploreg/haploreg.php [

440. Bottos A, Destro E, Rissone A, Graziano S, Cordara G, Assenzio B, et al. The synaptic proteins neurexins and neuroligins are widely expressed in the vascular system and contribute to its functions. Proceedings of the National Academy of Sciences of the United States of America. 2009;106(49):20782-7.

441. Matsumoto T, Schiller P, Dieterich LC, Bahram F, Iribe Y, Hellman U, et al. Ninein is expressed in the cytoplasm of angiogenic tip-cells and regulates tubular morphogenesis of endothelial cells. Arteriosclerosis, thrombosis, and vascular biology. 2008;28(12):2123-30.

442. Espinosa-Diez C, Miguel V, Vallejo S, Sanchez FJ, Sandoval E, Blanco E, et al. Role of glutathione biosynthesis in endothelial dysfunction and fibrosis. Redox Biol. 2018;14:88-99.

443. Koide S, Kugiyama K, Sugiyama S, Nakamura S, Fukushima H, Honda O, et al. Association of polymorphism in glutamate-cysteine ligase catalytic subunit gene with coronary vasomotor dysfunction and myocardial infarction. Journal of the American College of Cardiology. 2003;41(4):539-45.

444. Wang J, Xu L, Yun X, Yang K, Liao D, Tian L, et al. Proteomic analysis reveals that proteasome subunit beta 6 is involved in hypoxia-induced pulmonary vascular remodeling in rats. PloS one. 2013;8(7):e67942.

445. Adams J, Schott S, Bern A, Renz M, Ikenberg K, Garbe C, et al. A novel role for relaxin-2 in the pathogenesis of primary varicosis. PloS one. 2012;7(6):e39021.

446. Syreeni A, Sandholm N, Cao J, Toppila I, Maahs DM, Rewers MJ, et al. Genetic Determinants of Glycated Hemoglobin in Type 1 Diabetes. Diabetes. 2019;68(4):858-67.

447. Liu Y, Reynolds LM, Ding J, Hou L, Lohman K, Young T, et al. Blood monocyte transcriptome and epigenome analyses reveal loci associated with human atherosclerosis. Nature communications. 2017;8(1):393.

448. Dai Y, Luo W, Chang J. Rho kinase signaling and cardiac physiology. Curr Opin Physiol. 2018;1:14-20.

449. Cheng YC, O'Connell JR, Cole JW, Stine OC, Dueker N, McArdle PF, et al. Genome-wide association analysis of ischemic stroke in young adults. G3 (Bethesda). 2011;1(6):505-14.

450. Just S, Meder B, Berger IM, Etard C, Trano N, Patzel E, et al. The myosin-interacting protein SMYD1 is essential for sarcomere organization. Journal of cell science. 2011;124(Pt 18):3127-36.

451. Fan LL, Ding DB, Huang H, Chen YQ, Jin JY, Xia K, et al. A de novo mutation of SMYD1 (p.F272L) is responsible for hypertrophic cardiomyopathy in a Chinese patient. Clin Chem Lab Med. 2019;57(4):532-9.

452. Nurnberger J, Opazo Saez A, Dammer S, Mitchell A, Wenzel RR, Philipp T, et al. Left ventricular ejection time: a potential determinant of pulse wave velocity in young, healthy males. Journal of hypertension. 2003;21(11):2125-32.

453. Begeman IJ, Kang J. Transcriptional Programs and Regeneration Enhancers Underlying Heart Regeneration. J Cardiovasc Dev Dis. 2018;6(1).

454. Aitman TJ, Cooper LD, Norsworthy PJ, Wahid FN, Gray JK, Curtis BR, et al. Malaria susceptibility and CD36 mutation. Nature. 2000;405(6790):1015-6.

455. Deciphering Developmental Disorders S. Large-scale discovery of novel genetic causes of developmental disorders. Nature. 2015;519(7542):223-8.

456. Gould RA, Aziz H, Woods CE, Seman-Senderos MA, Sparks E, Preuss C, et al. ROBO4 variants predispose individuals to bicuspid aortic valve and thoracic aortic aneurysm. Nature genetics. 2019;51(1):42-50.

457. McLaren W, Pritchard B, Rios D, Chen Y, Flicek P, Cunningham F. Deriving the consequences of genomic variants with the Ensembl API and SNP Effect Predictor. Bioinformatics. 2010;26(16):2069-70.

458. Stenson PD, Ball EV, Mort M, Phillips AD, Shaw K, Cooper DN. The Human Gene Mutation Database (HGMD) and its exploitation in the fields of personalized genomics and molecular evolution. Current protocols in bioinformatics / editoral board, Andreas D Baxevanis [et al]. 2012;Chapter 1:Unit1 13.

459. Sim NL, Kumar P, Hu J, Henikoff S, Schneider G, Ng PC. SIFT web server: predicting effects of amino acid substitutions on proteins. Nucleic acids research. 2012;40(Web Server issue):W452-7.

460. Adzhubei I, Jordan DM, Sunyaev SR. Predicting functional effect of human missense mutations using PolyPhen-2. Current protocols in human genetics / editorial board, Jonathan L Haines [et al]. 2013;Chapter 7:Unit7 20.

461. Chun S, Fay JC. Identification of deleterious mutations within three human genomes. Genome research. 2009;19(9):1553-61.

462. Schwarz JM, Rodelsperger C, Schuelke M, Seelow D. MutationTaster evaluates disease-causing potential of sequence alterations. Nature methods. 2010;7(8):575-6.

463. Reva B, Antipin Y, Sander C. Predicting the functional impact of protein mutations: application to cancer genomics. Nucleic acids research. 2011;39(17):e118.

464. Shihab HA, Gough J, Cooper DN, Stenson PD, Barker GL, Edwards KJ, et al. Predicting the functional, molecular, and phenotypic consequences of amino acid substitutions using hidden Markov models. Human mutation. 2013;34(1):57-65.

465. Shamsani J, Kazakoff SH, Armean IM, McLaren W, Parsons MT, Thompson BA, et al. A plugin for the Ensembl Variant Effect Predictor that uses MaxEntScan to predict variant spliceogenicity. Bioinformatics. 2018.

466. Kircher M, Witten DM, Jain P, O'Roak BJ, Cooper GM, Shendure J. A general framework for estimating the relative pathogenicity of human genetic variants. Nature genetics. 2014;46(3):310-5.

467. ClinVar. https://www.ncbi.nlm.nih.gov/clinvar/.

468. Holmes KW, Maslen CL, Kindem M, Kroner BL, Song HK, Ravekes W, et al. GenTAC registry report: gender differences among individuals with genetically triggered thoracic aortic aneurysm and dissection. American journal of medical genetics Part A. 2013;161A(4):779-86.

469. Hiratzka LF, Creager MA, Isselbacher EM, Svensson LG, Nishimura RA, Bonow RO, et al. Surgery for Aortic Dilatation in Patients With Bicuspid Aortic Valves: A Statement of

Clarification From the American College of Cardiology/American Heart Association Task Force on Clinical Practice Guidelines. Journal of the American College of Cardiology. 2016;67(6):724-31.

470. Nishimura RA, Otto CM, Bonow RO, Carabello BA, Erwin JP, 3rd, Guyton RA, et al. 2014 AHA/ACC guideline for the management of patients with valvular heart disease: a report of the American College of Cardiology/American Heart Association Task Force on Practice Guidelines. Journal of the American College of Cardiology. 2014;63(22):e57-185.

471. Brownstein AJ, Ziganshin BA, Kuivaniemi H, Body SC, Bale AE, Elefteriades JA. Genes Associated with Thoracic Aortic Aneurysm and Dissection: An Update and Clinical Implications. Aorta (Stamford). 2017;5(1):11-20.

472. Brownstein AJ, Kostiuk V, Ziganshin BA, Zafar MA, Kuivaniemi H, Body SC, et al. Genes Associated with Thoracic Aortic Aneurysm and Dissection: 2018 Update and Clinical Implications. Aorta (Stamford). 2018;6(1):13-20.

473. Takagi H, Manabe H, Kawai N, Goto SN, Umemoto T. Circulating matrix metalloproteinase-9 concentrations and abdominal aortic aneurysm presence: a meta-analysis. Interactive cardiovascular and thoracic surgery. 2009;9(3):437-40.

474. Pyo R, Lee JK, Shipley JM, Curci JA, Mao D, Ziporin SJ, et al. Targeted gene disruption of matrix metalloproteinase-9 (gelatinase B) suppresses development of experimental abdominal aortic aneurysms. The Journal of clinical investigation. 2000;105(11):1641-9.

475. Zarate YA, Sellars E, Lepard T, Tang X, Collins RT, 2nd. Aortic dilation, genetic testing, and associated diagnoses. Genetics in medicine : official journal of the American College of Medical Genetics. 2016;18(4):356-63.

476. Glagov S, Zarins C, Giddens DP, Ku DN. Hemodynamics and atherosclerosis. Insights and perspectives gained from studies of human arteries. Archives of pathology & laboratory medicine. 1988;112(10):1018-31.

477. Cornhill JF, Roach MR. A quantitative study of the localization of atherosclerotic lesions in the rabbit aorta. Atherosclerosis. 1976;23(3):489-501.

478. Caro CG. Discovery of the role of wall shear in atherosclerosis. Arteriosclerosis, thrombosis, and vascular biology. 2009;29(2):158-61.

479. Caro CG, Fitz-Gerald JM, Schroter RC. Atheroma and arterial wall shear. Observation, correlation and proposal of a shear dependent mass transfer mechanism for atherogenesis. Proceedings of the Royal Society of London Series B, Containing papers of a Biological character Royal Society. 1971;177(1046):109-59.

480. Lehoux S, Jones EA. Shear stress, arterial identity and atherosclerosis. Thrombosis and haemostasis. 2016;115(3):467-73.

481. Fatehi Hassanabad A, Garcia J, Verma S, White JA, Fedak PWM. Utilizing wall shear stress as a clinical biomarker for bicuspid valve-associated aortopathy. Current opinion in cardiology. 2019;34(2):124-31.

482. Chistiakov DA, Orekhov AN, Bobryshev YV. Effects of shear stress on endothelial cells: go with the flow. Acta physiologica. 2017;219(2):382-408.

483. Aird WC. Mechanisms of endothelial cell heterogeneity in health and disease. Circulation research. 2006;98(2):159-62.

484. Regan ER, Aird WC. Dynamical systems approach to endothelial heterogeneity. Circulation research. 2012;111(1):110-30.

485. Frangos JA, Eskin SG, McIntire LV, Ives CL. Flow effects on prostacyclin production by cultured human endothelial cells. Science. 1985;227(4693):1477-9.

486. Grabowski EF, Jaffe EA, Weksler BB. Prostacyclin production by cultured endothelial cell monolayers exposed to step increases in shear stress. The Journal of laboratory and clinical medicine. 1985;105(1):36-43.

487. Korenaga R, Ando J, Tsuboi H, Yang W, Sakuma I, Toyo-oka T, et al. Laminar flow stimulates ATP- and shear stress-dependent nitric oxide production in cultured bovine endothelial cells. Biochemical and biophysical research communications. 1994;198(1):213-9.

488. Kuchan MJ, Frangos JA. Role of calcium and calmodulin in flow-induced nitric oxide production in endothelial cells. The American journal of physiology. 1994;266(3 Pt 1):C628-36.

489. Langille BL, O'Donnell F. Reductions in arterial diameter produced by chronic decreases in blood flow are endothelium-dependent. Science. 1986;231(4736):405-7.

490. Zhang H, Sunnarborg SW, McNaughton KK, Johns TG, Lee DC, Faber JE. Heparin-binding epidermal growth factor-like growth factor signaling in flow-induced arterial remodeling. Circulation research. 2008;102(10):1275-85.

491. Forstermann U, Munzel T. Endothelial nitric oxide synthase in vascular disease: from marvel to menace. Circulation. 2006;113(13):1708-14.

492. Boo YC, Kim HJ, Song H, Fulton D, Sessa W, Jo H. Coordinated regulation of endothelial nitric oxide synthase activity by phosphorylation and subcellular localization. Free radical biology & medicine. 2006;41(1):144-53.

493. Jin ZG, Ueba H, Tanimoto T, Lungu AO, Frame MD, Berk BC. Ligand-independent activation of vascular endothelial growth factor receptor 2 by fluid shear stress regulates activation of endothelial nitric oxide synthase. Circulation research. 2003;93(4):354-63.

494. Corson MA, James NL, Latta SE, Nerem RM, Berk BC, Harrison DG. Phosphorylation of endothelial nitric oxide synthase in response to fluid shear stress. Circulation research. 1996;79(5):984-91.

495. Huang PL, Huang Z, Mashimo H, Bloch KD, Moskowitz MA, Bevan JA, et al. Hypertension in mice lacking the gene for endothelial nitric oxide synthase. Nature. 1995;377(6546):239-42.

496. Won D, Zhu SN, Chen M, Teichert AM, Fish JE, Matouk CC, et al. Relative reduction of endothelial nitric-oxide synthase expression and transcription in atherosclerosis-prone

regions of the mouse aorta and in an in vitro model of disturbed flow. The American journal of pathology. 2007;171(5):1691-704.

497. Young A, Wu W, Sun W, Benjamin Larman H, Wang N, Li YS, et al. Flow activation of AMP-activated protein kinase in vascular endothelium leads to Kruppel-like factor 2 expression. Arteriosclerosis, thrombosis, and vascular biology. 2009;29(11):1902-8.

498. Wang N, Miao H, Li YS, Zhang P, Haga JH, Hu Y, et al. Shear stress regulation of Kruppel-like factor 2 expression is flow pattern-specific. Biochemical and biophysical research communications. 2006;341(4):1244-51.

499. Wu W, Xiao H, Laguna-Fernandez A, Villarreal G, Jr., Wang KC, Geary GG, et al. Flow-Dependent Regulation of Kruppel-Like Factor 2 Is Mediated by MicroRNA-92a. Circulation. 2011;124(5):633-41.

500. Lin Z, Kumar A, SenBanerjee S, Staniszewski K, Parmar K, Vaughan DE, et al. Kruppel-like factor 2 (KLF2) regulates endothelial thrombotic function. Circulation research. 2005;96(5):e48-57.

501. Parmar KM, Larman HB, Dai G, Zhang Y, Wang ET, Moorthy SN, et al. Integration of flow-dependent endothelial phenotypes by Kruppel-like factor 2. The Journal of clinical investigation. 2006;116(1):49-58.

502. Ali F, Zakkar M, Karu K, Lidington EA, Hamdulay SS, Boyle JJ, et al. Induction of the cytoprotective enzyme heme oxygenase-1 by statins is enhanced in vascular endothelium exposed to laminar shear stress and impaired by disturbed flow. The Journal of biological chemistry. 2009;284(28):18882-92.

503. Zakkar M, Chaudhury H, Sandvik G, Enesa K, Luong le A, Cuhlmann S, et al. Increased endothelial mitogen-activated protein kinase phosphatase-1 expression suppresses proinflammatory activation at sites that are resistant to atherosclerosis. Circulation research. 2008;103(7):726-32.

504. Read MA, Whitley MZ, Gupta S, Pierce JW, Best J, Davis RJ, et al. Tumor necrosis factor alpha-induced E-selectin expression is activated by the nuclear factor-kappaB and c-JUN N-terminal kinase/p38 mitogen-activated protein kinase pathways. The Journal of biological chemistry. 1997;272(5):2753-61.

505. Tamura DY, Moore EE, Johnson JL, Zallen G, Aiboshi J, Silliman CC. p38 mitogen-activated protein kinase inhibition attenuates intercellular adhesion molecule-1 up-regulation on human pulmonary microvascular endothelial cells. Surgery. 1998;124(2):403-7; discussion 8.

506. Gratton JP, Morales-Ruiz M, Kureishi Y, Fulton D, Walsh K, Sessa WC. Akt down-regulation of p38 signaling provides a novel mechanism of vascular endothelial growth factor-mediated cytoprotection in endothelial cells. The Journal of biological chemistry. 2001;276(32):30359-65.

507. Zakkar M, Van der Heiden K, Luong le A, Chaudhury H, Cuhlmann S, Hamdulay SS, et al. Activation of Nrf2 in endothelial cells protects arteries from exhibiting a proinflammatory state. Arteriosclerosis, thrombosis, and vascular biology. 2009;29(11):1851-7.

508. Dai G, Vaughn S, Zhang Y, Wang ET, Garcia-Cardena G, Gimbrone MA, Jr. Biomechanical forces in atherosclerosis-resistant vascular regions regulate endothelial redox balance via phosphoinositol 3-kinase/Akt-dependent activation of Nrf2. Circulation research. 2007;101(7):723-33.

509. Tsao PS, Lewis NP, Alpert S, Cooke JP. Exposure to shear stress alters endothelial adhesiveness. Role of nitric oxide. Circulation. 1995;92(12):3513-9.

510. Tsao PS, Buitrago R, Chan JR, Cooke JP. Fluid flow inhibits endothelial adhesiveness. Nitric oxide and transcriptional regulation of VCAM-1. Circulation. 1996;94(7):1682-9.

511. Chiu JJ, Chen LJ, Chang SF, Lee PL, Lee CI, Tsai MC, et al. Shear stress inhibits smooth muscle cell-induced inflammatory gene expression in endothelial cells: role of NF-kappaB. Arteriosclerosis, thrombosis, and vascular biology. 2005;25(5):963-9.

512. Cuhlmann S, Van der Heiden K, Saliba D, Tremoleda JL, Khalil M, Zakkar M, et al. Disturbed blood flow induces RelA expression via c-Jun N-terminal kinase 1: a novel mode of NF-kappaB regulation that promotes arterial inflammation. Circulation research. 2011;108(8):950-9.

513. Ajami NE, Gupta S, Maurya MR, Nguyen P, Li JY, Shyy JY, et al. Systems biology analysis of longitudinal functional response of endothelial cells to shear stress. Proceedings of the National Academy of Sciences of the United States of America. 2017;114(41):10990-5.

514. Qiao C, Meng F, Jang I, Jo H, Chen YE, Zhang J. Deep transcriptomic profiling reveals the similarity between endothelial cells cultured under static and oscillatory shear stress conditions. Physiological genomics. 2016;48(9):660-6.

515. Li R, Beebe T, Jen N, Yu F, Takabe W, Harrison M, et al. Shear stress-activated Wnt-angiopoietin-2 signaling recapitulates vascular repair in zebrafish embryos. Arteriosclerosis, thrombosis, and vascular biology. 2014;34(10):2268-75.

516. Demolli S, Doebele C, Doddaballapur A, Lang V, Fisslthaler B, Chavakis E, et al. MicroRNA-30 mediates anti-inflammatory effects of shear stress and KLF2 via repression of angiopoietin 2. Journal of molecular and cellular cardiology. 2015;88:111-9.

517. Groenendijk BC, Hierck BP, Vrolijk J, Baiker M, Pourquie MJ, Gittenberger-de Groot AC, et al. Changes in shear stress-related gene expression after experimentally altered venous return in the chicken embryo. Circulation research. 2005;96(12):1291-8.

518. Ni CW, Qiu H, Rezvan A, Kwon K, Nam D, Son DJ, et al. Discovery of novel mechanosensitive genes in vivo using mouse carotid artery endothelium exposed to disturbed flow. Blood. 2010;116(15):e66-73.

519. Passerini AG, Polacek DC, Shi C, Francesco NM, Manduchi E, Grant GR, et al. Coexisting proinflammatory and antioxidative endothelial transcription profiles in a disturbed flow region of the adult porcine aorta. Proceedings of the National Academy of Sciences of the United States of America. 2004;101(8):2482-7.

520. Reinhart-King CA, Fujiwara K, Berk BC. Physiologic stress-mediated signaling in the endothelium. Methods in enzymology. 2008;443:25-44.

521. Koslow AR, Stromberg RR, Friedman LI, Lutz RJ, Hilbert SL, Schuster P. A flow system for the study of shear forces upon cultured endothelial cells. Journal of biomechanical engineering. 1986;108(4):338-41.

522. Dardik A, Chen L, Frattini J, Asada H, Aziz F, Kudo FA, et al. Differential effects of orbital and laminar shear stress on endothelial cells. Journal of vascular surgery. 2005;41(5):869-80.

523. Zhao S, Fung-Leung WP, Bittner A, Ngo K, Liu X. Comparison of RNA-Seq and microarray in transcriptome profiling of activated T cells. PloS one. 2014;9(1):e78644.

524. Yu H, Moran CS, Trollope AF, Woodward L, Kinobe R, Rush CM, et al. Angiopoietin-2 attenuates angiotensin II-induced aortic aneurysm and atherosclerosis in apolipoprotein E-deficient mice. Scientific reports. 2016;6:35190.

525. Cao H, Sekiya M, Ertunc ME, Burak MF, Mayers JR, White A, et al. Adipocyte lipid chaperone AP2 is a secreted adipokine regulating hepatic glucose production. Cell Metab. 2013;17(5):768-78.

526. Furuhashi M, Fuseya T, Murata M, Hoshina K, Ishimura S, Mita T, et al. Local Production of Fatty Acid-Binding Protein 4 in Epicardial/Perivascular Fat and Macrophages Is Linked to Coronary Atherosclerosis. Arteriosclerosis, thrombosis, and vascular biology. 2016;36(5):825-34.

527. Fuseya T, Furuhashi M, Matsumoto M, Watanabe Y, Hoshina K, Mita T, et al. Ectopic Fatty Acid-Binding Protein 4 Expression in the Vascular Endothelium is Involved in Neointima Formation After Vascular Injury. Journal of the American Heart Association. 2017;6(9).

528. Furuhashi M, Tuncman G, Gorgun CZ, Makowski L, Atsumi G, Vaillancourt E, et al. Treatment of diabetes and atherosclerosis by inhibiting fatty-acid-binding protein aP2. Nature. 2007;447(7147):959-65.

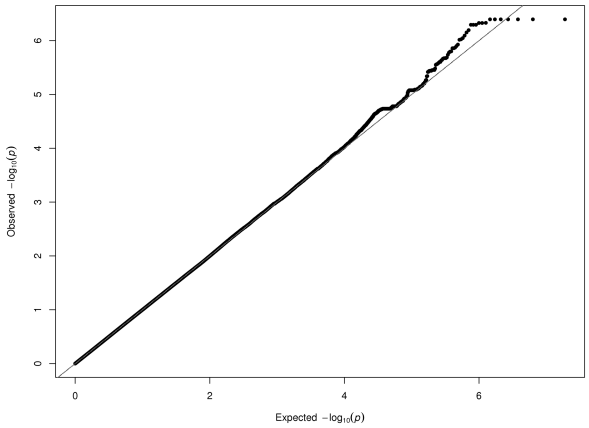
529. Hazell GG, Peachey AM, Teasdale JE, Sala-Newby GB, Angelini GD, Newby AC, et al. PI16 is a shear stress and inflammation-regulated inhibitor of MMP2. Scientific reports. 2016;6:39553.

530. Tsukamoto K, Mani DR, Shi J, Zhang S, Haagensen DE, Otsuka F, et al. Identification of apolipoprotein D as a cardioprotective gene using a mouse model of lethal atherosclerotic coronary artery disease. Proceedings of the National Academy of Sciences of the United States of America. 2013;110(42):17023-8.

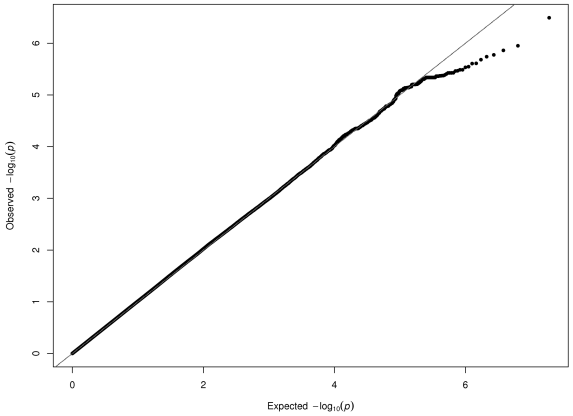
531. Durst R, Sauls K, Peal DS, deVlaming A, Toomer K, Leyne M, et al. Mutations in DCHS1 cause mitral valve prolapse. Nature. 2015;525(7567):109-13.

APPENDIX 1: Q-Q plots of aortic phenotype genome-wide association p values

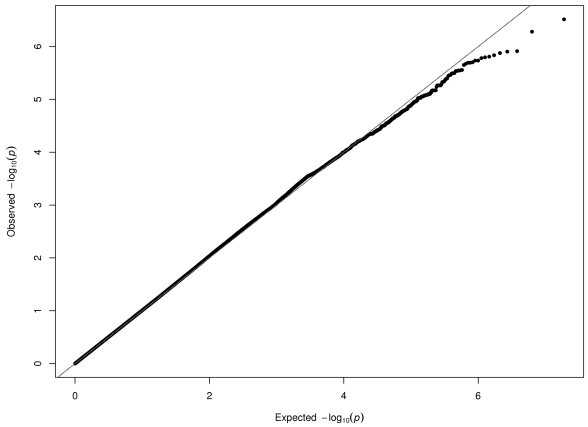
Aortic valve annulus diameter



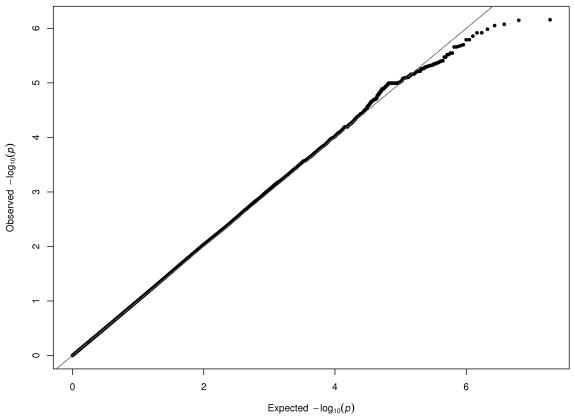
Sinuses of Valsalva diameter



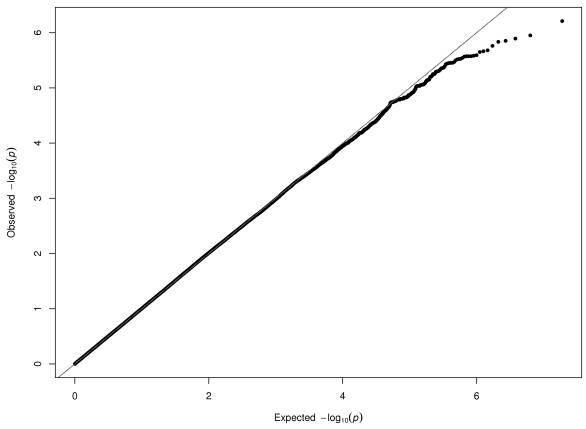
Sino-Tubular Junction diameter



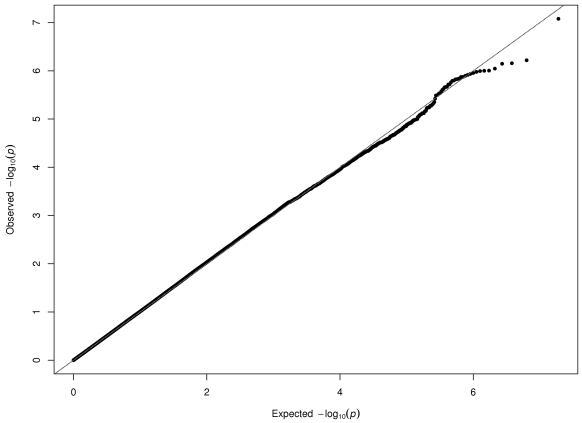
Joint Trait analysis: Aortic Root



Ascending aortic area (log-transformed)

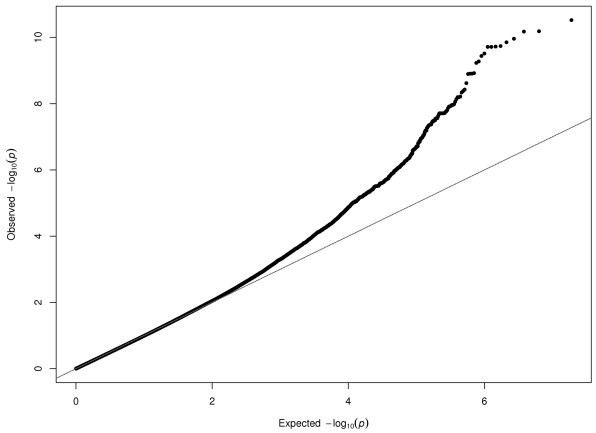


Descending aortic area (log-transformed)

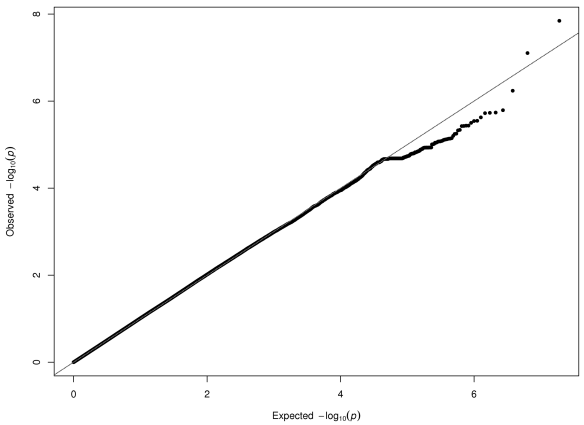


APPENDIX 1: Q-Q plots of aortic phenotype genome-wide association p values

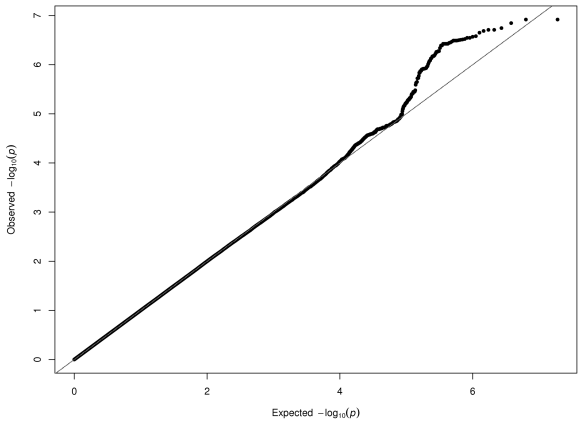
Ascending aortic distensibility (log-transformed)



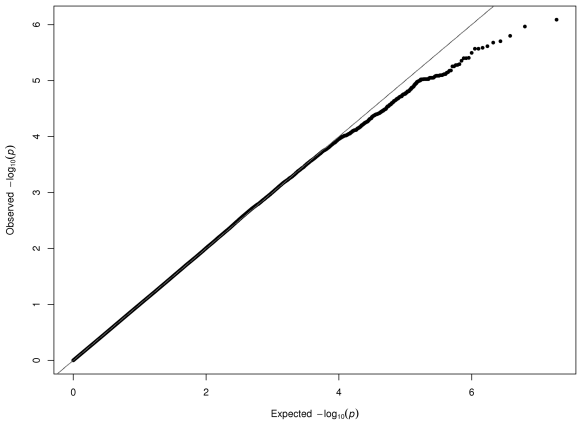
Ascending aortic distensibility (rank-normalised)



Descending aortic distensibility (rank-normalised)



Aortic arch pulse wave velocity (rank-normalised)



APPENDIX 2 : Genes selected from literature for sequencing in BAV cohort; thanks to Matina Prapa

Ensembl Gene ID	Gene	Chromosome	Function/signalling pathways	Human phenotype	Reference
			Extracellular matrix proteins		
ENSG00000166147	*FBN1 (Fibrillin-1)*	15	Tissue elasticity, TGFβ	MFS, BAV Ao- increased expression	(Dietz et al., 1991, Gomez et al., 2009)
ENSG00000115414	*FN1* (Fibronectin-1)	2	Marker of VSMC synthetic phenotype	Increased expression in convexity of Asc ao in stenotic BAV	(Della Corte et al., 2008)
ENSG00000049540	*ELN (Elastin)*	7	Tissue elasticity	Williams-Beuren syndrome; BAV with supravalvular AS	(Szabo et al., 2006, Li et al., 1998)
ENSG00000168542	*COL3A1* (Collagen alpha-1 III)	2	Collagen metabolism	EDS, type 4; frequent arterial dissection with infrequent aneurysm	(Liu et al., 1997, Superti-Furga et al., 1988)
ENSG00000188153	*COL4A5* (Collagen alpha-5 IV)	X	Collagen metabolism	X-linked Alport syndrome; Asc ao & abdominal aneurysm	(Kashtan et al., 2010)
ENSG00000187498	*COL4A1* (Collagen alpha-1IV)	13	Collagen metabolism	Hereditary angiopathy,nephropathy, aneurysms and muscle cramps	(Plaisier et al., 2007, Poschl et al., 2004)
ENSG00000083444	*PLOD1* (lysyl hydroxylase 1)	1	Collagen metabolism	EDS type 6, rare aneurysm	(Wenstrup et al., 1989, Takaluoma et al., 2007)
ENSG00000106397	*PLOD3* (lysyl hydroxylase 3)	7	Collagen metabolism	Bone fragility, arterial rupture and deafness	(Salo et al., 2008, Ruotsalainen et al., 2006)
ENSG00000113083	*LOX* (lysyl oxidase)	5	Collagen metabolism, TGFβ	Unknown, Ao aneurysm in KO mice	(Maki et al., 2002)
ENSG00000172638	*EFEMP2* (fibulin-4)	11	Elastic fiber formation, connective tissue development	Cutis laxa with Asc ao aneurysm and arterial tortuosity	(Dasouki et al., 2007, Huang et al., 2010)
			Transcription factors		
ENSG00000130700	*GATA5* (GATA binding protein 5)	20	Cardiac development (aortic valve)	BAV, BAV in 25% of KO mice	(Laforest et al., 2011)
ENSG00000125398	*SOX9* (SRY-box 9)	17	Chondrogenesis	Unknown, calcific valvular disease in KO mice	(Peacock et al., 2010)
ENSG00000131196	*NFATc1* (Nuclear factor of activated T-cells calcineurin-dependant1)	18	Cardiac development (endocardial cushion growth & remodelling)	Unknown	(de la Pompa et al., 1998)
ENSG00000183072	*NKX2.5* (NK2 transcription factor related)	5	Cardiac development (cardiac homeobox gene)	BAV-TAA, ASD, VSD, TOF, Ebstein's, DORV	(Biben et al., 2000, Schott et al., 1998, Majumdar et al., 2006, Wang et al., 2011)

APPENDIX 2 : Genes selected from literature for sequencing in BAV cohort; thanks to Matina Prapa

ENSG00000070010	*UFD1L* (ubiquitin fusion degradation 1 like)	22	Cardiac development (cardiac outflow tract)	BAV-TAA, DGS/VCFS (conotruncal cardiac defects)	(Mohamed et al., 2005)
Transmembrane proteins					
ENSG00000148400	*NOTCH1*	9	Cardiac development (cardiac outflow tract)	BAV, calcific aortic stenosis, VSD, TOF, mitral stenosis	(Garg et al., 2005, Niessen and Karsan, 2008)
ENSG00000106799	*TGFBR1* (TGFβ receptor type 1)	9	Connective tissue degradation, TGFβ	MFS-type 2, LDS, BAV Ao-increased expression	(Loeys et al., 2006, Gomez et al., 2009)
ENSG00000163513	*TGFBR2* (TGFβ receptor type 1)	3	Connective tissue degradation, TGFβ	MFS-type 2, LDS, BAV Ao-increased expression	(Loeys et al., 2006, Gomez et al., 2009)
ENSG00000106991	*ENG (endoglin)*	9	Cardiac development (aortic valve formation), TGFβ superfamily	BAV	(Wooten et al., 2010)
ENSG00000101384	*JAG1* (JAGGED1)	20	Cardiac development (aortic valve formation)	Allagile syndrome; BAV with characteristic facies, jaundice & skeletal abnormalities	(McElhinney et al., 2002)
ENSG00000123700	*KCNJ2*	17	Excitable myocardial tissue (inward-rectifying potassium current Kir2.1)	Andersen syndrome; BAV with periodic paralysis, ventricular arrhythmias & dysmorphic features	(Andelfinger et al., 2002)
ENSG00000152661	*GJA1* (connexin-43)	6	Connexin gap junction-development of normal cardiac architecture and ventricular conduction	Hypoplastic left heart syndrome	(Yu et al., 2004)
ENSG00000139567	*ACVRL1* (activin receptor-like kinase-1)	12	TGFβ superfamily	Hereditary haemorrahagic telangiectasia, arterial aneurysms	(Andersen et al., 2010, Oh et al., 2000)
ENSG00000197496	*SLC2A10* (glucose transporter type 10)	20	Glucose homeostasis	Arterial tortuosity syndrome	(Coucke et al., 2006)
Cytoplasmic proteins					
ENSG00000166949	*SMAD3* (SMAD family member 3)	15	Connective tissue degradation, TGFβ	LDS, aortic aneurysm with osteoarthritis	(van de Laar et al., 2011)
ENSG00000107796	*ACTA2* (alpha smooth muscle actin)	10	Vascular contractility, TGFβ	Familial aortic aneurysm; BAV with livedo reticularis	(Guo et al., 2007)
ENSG00000133392	*MYH11* (smooth muscle myosin, heavy chain 11) ‡	16	Vascular contractility, angiotensin II	Familial aortic aneurysm with patent ductus arteriosus	(Zhu et al., 2006, Pannu et al., 2007)
ENSG00000196924	*FLNA* (Filamin-A) ‡	X	Actin cytoskeleton, TGFβ	Periventricular nodular heterotopia with EDS, Asc ao aneurysm and valvular dystrophy	(Sheen et al., 2005, Feng et al., 2006)

APPENDIX 2 : Genes selected from literature for sequencing in BAV cohort; thanks to Matina Prapa

ENSG00000196712	*NF1* (Neurofibromin-1) ‡	17	Ras-MEK-ERK	Neurofibromatosis, arterial aneurysm and stenosis	(Friedman et al., 2002)
ENSG00000179295	*PTPN11* (Protein-tyrosine phosphatase 2C)	12	Ras-MEK-ERK	Noonan syndrome, coronary artery aneurysm and rare Asc ao aneurysm	(Purnell et al., 2005, Araki et al., 2004, Iwasaki et al., 2009)
ENSG00000103197	*TSC2* (tuberin) ‡	16	Tumour suppression, mammalian target of	Tuberous sclerosis, diffuse	(Cao et al., 2010)
ENSG00000163221	*S100A12 (S100 calcium binding protein A12)*	1	Interleukin-6, TGFβ	Increased expression in MYH11-mutation aneurysm	(Hofmann Bowman et al., 2010)
ENSG00000164867	*NOS3/eNOS (endothelial nitric oxide synthase) **	7	Cardiac development, stress-induced vascular remodelling	Abdominal aortic aneurysm, BAV in KO mice	(Atli et al., 2010, Fernandez et al., 2009)
ENSG00000103126	*AXIN1 (Axin-1)* ǂ	16	Cardiac development (cardiac valve, outflow tract), Wnt	BAV	(Wooten et al., 2010)
ENSG00000185615	*PDIA2* ǂ	16	Protein disulfide isomerase family A, member 2	BAV	(Wooten et al., 2010)
ENSG00000171298	*GAA* ǂ	17	Lysosomal alpha-glucosidase	Acid maltase deficiency, intracranial aneurysm, lysosomal accumulation in heart and aorta of KO mice	(Raben et al., 1998)
Nuclear protein					
ENSG00000127528	*KLF2* (Kruppel-like factor 2)	19	Unknown	Unknown, Ao aneurysm and dissection in KO mice	(Kuo et al., 1997)
ENSG00000163884	*KLF15* (Kruppel-like factor 15) *	3	Inhibition of VSMC proliferation and migration, thrombospondin-2, p53, TGFβ	Unknown, Ao aneurysm and cardiomyopathy in KO mice	(Haldar et al., 2010, Lu et al., 2010)
ENSG00000184634	*MED12* (mediator complex subunit 12) ‡	X	WNT-β-catenin, WNT-PCP	Lujan–Fryns syndrome with ao root dilation	(Schwartz et al., 2007, Rocha et al., 2010)
Vascular endothelial growth factor (VEGF) pathway members					
ENSG00000100644	*HIF1A* (hypoxia inducible factor 1, alpha subunit) ‡	14	Proangiogenic transcriptional factor, VEGF	Hypoplastic left heart syndrome with BAV	(Hinton et al., 2009)
ENSG00000119630	*PGF* (placental growth factor)	14	Embryogenesis (angiogenesis), VEGF	Hypoplastic left heart syndrome with BAV, atherosclerosis	(Hinton et al., 2009)

APPENDIX 2 : Genes selected from literature for sequencing in BAV cohort; thanks to Matina Prapa

ENSG00000072110	*ACTN1* (Alpha-actinin-1) ‡	14	Cardiac development (cardiac valve), VEGF	Hypoplastic left heart syndrome with BAV	(Hinton et al., 2009)
ENSG00000134001	*EIF2S1*	14	Eukaryotic translation initiation factor 2, subunit 1 alpha, VEGF	Hypoplastic left heart syndrome with BAV	(Hinton et al., 2009)
ENSG00000035403	*VCL (vinculin)* *	10	Cytoskeletal protein, VEGF	Hypoplastic left heart syndrome, dilated cardiomyopathy	(Hinton et al., 2009)
Enzymes regulating nitric oxide generation					
ENSG00000153904	DDAH1*	1	dimethylarginine dimethylaminohydrolase 1	Type II diabetes	(Abhary et al., 2010)
ENSG00000213722	*DDAH2*	6	dimethylarginine dimethylaminohydrolase 2	Type II diabetes	(Abhary et al., 2010)
Associated with abdominal aortic aneurysm (AAA)					
ENSG00000100985	*MMP9* (matrix metallopeptidase 9)	20	Connective tissue degradation, TGFβ	BAV-TAA with autoimmune disease (case report), AAA	(Foffa et al., 2009)
ENSG00000159640	*ACE* (angiotensin I converting enzyme)	17	Connective tissue degradation, angiotensin II	BAV-TAA, AAA, left ventricular hypertrophy	(Foffa et al., 2009, Foffa et al., 2012)
ENSG00000177000	*MTHFR* (methylene-tetrahydrofolate reductase	1	Connective tissue degradation	BAV-TAA with autoimmune disease (case report), AAA, coronary artery disease	(Foffa et al., 2009)
ENSG00000106366	*PAI1/SERPINE1* (serpin peptidase inhibitor)	7	Connective tissue degradation	BAV-TAA with autoimmune disease (case report), AAA, coronary artery disease	(Foffa et al., 2009)
Associated with aortic stenosis					
ENSG00000130203	*APOE* (apolipoprotein E)	19	Catabolism of triglyceride-rich lipoprotein constituents	Aortic stenosis-BAV (case report), atherosclerosis, Alzheimer's	(Saravanan and Kadir, 2009)
ENSG00000084674	*APOB* (apolipoprotein B)	2	Primary apolipoproteins of chylomicrons & low-density lipoproteins	Calcific aortic stenosis, atherosclerosis	(Anger et al., 2006)
ENSG00000111424	*VDR* (Vitamin D receptor)	12	Transcription factor, mineral metabolism	Calcific aortic stenosis, type II vitamin D-resistant rickets	(Ortlepp et al., 2001)
ENSG00000091831	*ESR1* (Estrogen receptor alpha)	6	Transcription factor	Calcific aortic stenosis, cancer, osteoporosis	(Anger et al., 2006)
ENSG00000136634	*IL10* (Interleukin 10)	1	Anti-inflammatory cytokine	Calcific aortic stenosis, HIV-1 infection, rheumatoid arthritis	(Ortlepp et al., 2004)

APPENDIX 2 : Genes selected from literature for sequencing in BAV cohort; thanks to Matina Prapa

ENSG00000160791	*CCR5* (chemokine receptor 5)	3	Transmembrane protein, anti-inflammatory properties	Calcific aortic stenosis, HIV-1 infection	(Ortlepp et al., 2004)
ENSG00000118523	*CTGF* (connective tissue growth factor)	6	Chondrocyte proliferation	Calcific aortic stenosis, systemic sclerosis	(Ortlepp et al., 2004)
ENSG00000146648	*EGFR* (epidermal growth factor receptor)	7	Transmembrane protein, cell proliferation	Calcific aortic stenosis, lung cancer	(Anger et al., 2006)
Associated with mitral valve prolapse (MVP)					
ENSG00000105329	*TGFB1* (TGF-beta1)	19	Connective tissue degradation, TGFβ	MVP, increased expression in LDS aorta	(Lindsay et al., 2012, Hagler et al., 2013)
ENSG00000108821	*COL1A1* (Collagen Type I)	17	Collagen metabolism	MVP, EDS, osteogenesis imperfecta	(Sykes et al., 1990, De Paepe, 1998, Malfait et al., 2007)
ENSG00000164692	*COL1A2* (Collagen type II)	7	Collagen metabolism	MVP, osteogenesis imperfecta	(Sykes et al., 1990, De Paepe, 1998)
ENSG00000060718	*COL11A1 (Collagen type XI)*	1	Collagen metabolism	MVP, EDS, Stickler syndrome, Marshall syndrome	(Griffith et al., 1998, Khalifa et al., 2012)
ENSG00000166341	*DCHS1 (Dachsous 1)*	11	Cadherin family member, expressed in fibroblasts	MVP, zebrafish model - no valve	(Freed et al., 2003); [456, 457]

Abbreviations: AAA, abdominal aortic aneurysm; Ao, aorta; AS, aortic stenosis; Asc Ao, ascending aortic; ASD, atrial septal defect; BAV, bicuspid aortic valve; DGS/VCFS, DiGeorge syndrome/Velo-cardio-facial syndrome; DORV, double outlet right ventricle; EDS, Ehlers Danlos syndrome; KO, knock-out; LDS, Loeys Dietz syndrome; MFS, Marfan syndrome; MVP, mitral valve prolapse; TAA, thoracic aortic aneurysm; TOF, Tetralogy of Fallot; VSD, ventricular septal defect.

APPENDIX 3: Differentially expressed genes at 144 hours of culture; comparison is turbulent versus laminar shear (i.e. positive fold-change is upregulation in oscillatory shear); ordered by p value

Log FC= log fold change

Gene Symbol	Gene Name	logFC	Average Express.	t statistic	P.Value	adj.P.Val
SLC22A23	solute carrier family 22, member 23	0.89	6.96	7.16	3.00E-08	0.0002
FILIP1	filamin A interacting protein 1	2.43	7.99	7.19	2.75E-08	0.0002
SMAGP	small cell adhesion glycoprotein	-1.11	9.47	-6.58	1.63E-07	0.0007
ADD3	adducin 3 (gamma)	1.41	9.27	6.29	3.81E-07	0.0010
CYP4X1	cytochrome P450, family 4, subfamily X, polypeptide 1	1.09	6.84	6.40	2.77E-07	0.0010
MT1E	metallothionein 1E	-1.56	12.14	-6.26	4.15E-07	0.0010
SFR1	SWI5-dependent recombination repair 1	-0.77	7.89	-6.19	5.11E-07	0.0010
COL17A1	collagen, type XVII, alpha 1	-1.33	7.35	-6.09	6.87E-07	0.0010
GJC2	gap junction protein, gamma 2, 47kDa	-1.19	7.48	-6.10	6.79E-07	0.0010
ACSS1	acyl-CoA synthetase short-chain family member 1	0.73	6.85	6.04	8.02E-07	0.0011
OAZ3	ornithine decarboxylase antizyme 3	-0.52	6.68	-5.99	9.39E-07	0.0012
MYOM3	myomesin 3	-0.50	6.50	-5.90	1.21E-06	0.0014
TMEM30A	transmembrane protein 30A	1.23	9.60	5.85	1.43E-06	0.0015
ABI3	ABI family, member 3	-1.03	8.57	-5.69	2.27E-06	0.0016
APOD	apolipoprotein D	2.14	7.64	5.73	2.00E-06	0.0016
ARHGEF17	Rho guanine nucleotide exchange factor (GEF) 17	0.66	7.35	5.69	2.31E-06	0.0016
GLUD1	glutamate dehydrogenase 1	-0.38	10.49	-5.73	2.03E-06	0.0016
HOMER3	homer homolog 3 (Drosophila)	0.86	8.60	5.70	2.20E-06	0.0016
SNCG	synuclein, gamma (breast cancer-specific protein 1)	-0.54	6.84	-5.77	1.77E-06	0.0016
TCF4	transcription factor 4	1.19	9.56	5.80	1.65E-06	0.0016
AIM1	absent in melanoma 1	0.55	6.98	5.60	2.96E-06	0.0016
FILIP1L	filamin A interacting protein 1-like	1.26	9.10	5.66	2.51E-06	0.0016
LCN6	lipocalin 6	-0.94	6.59	-5.58	3.12E-06	0.0016
MT1X	metallothionein 1X	-1.17	9.15	-5.61	2.85E-06	0.0016
PRKD2	protein kinase D2	0.68	7.85	5.63	2.73E-06	0.0016
PROCR	protein C receptor, endothelial	-0.91	11.30	-5.65	2.59E-06	0.0016
UVRAG	UV radiation resistance associated	0.52	8.78	5.59	3.11E-06	0.0016
LRIG1	leucine-rich repeats and immunoglobulin-like domains 1	1.10	9.17	5.57	3.25E-06	0.0016
ATP6V1A	ATPase, H+ transporting, lysosomal 70kDa, V1 subunit A	0.68	11.08	5.51	3.84E-06	0.0018
FAM89A	family with sequence similarity 89, member A	1.93	9.25	5.45	4.59E-06	0.0020
GOSR2	golgi SNAP receptor complex member 2	-0.85	8.82	-5.48	4.31E-06	0.0020
SORT1	sortilin 1	0.70	7.23	5.46	4.57E-06	0.0020

APPENDIX 3: Differentially expressed genes at 144 hours of culture; comparison is turbulent versus laminar shear (i.e. positive fold-change is upregulation in oscillatory shear); ordered by p value

ANGPT2	angiopoietin 2	2.77	9.00	5.40	5.40E-06	0.0022
SEMA3F	sema domain, immunoglobulin domain (Ig), short basic domain, secreted, (semaphorin) 3F	-1.13	8.81	-5.40	5.47E-06	0.0022
KIT	v-kit Hardy-Zuckerman 4 feline sarcoma viral oncogene homolog	1.73	7.23	5.36	6.04E-06	0.0024
SCHIP1	schwannomin interacting protein 1	1.04	9.03	5.33	6.63E-06	0.0025
TNFRSF4	tumor necrosis factor receptor superfamily, member 4	1.41	7.16	5.30	7.23E-06	0.0027
LPCAT4	lysophosphatidylcholine acyltransferase 4	-0.44	7.09	-5.29	7.51E-06	0.0027
PLEKHA6	pleckstrin homology domain containing, family A member 6	-0.31	6.57	-5.23	8.94E-06	0.0031
BCL2L11	BCL2-like 11 (apoptosis facilitator)	0.37	6.57	5.20	9.70E-06	0.0033
CSGALNACT1	chondroitin sulfate N-acetylgalactosaminyltransferase 1	1.16	7.17	5.15	1.13E-05	0.0038
DACH1	dachshund family transcription factor 1	1.56	7.69	5.13	1.20E-05	0.0039
ABCA6	ATP-binding cassette, sub-family A (ABC1), member 6	1.21	7.61	5.07	1.44E-05	0.0046
GAB1	GRB2-associated binding protein 1	0.47	6.66	5.05	1.54E-05	0.0047
RPS7	ribosomal protein S7	-1.08	9.14	-5.06	1.51E-05	0.0047
LOX	lysyl oxidase	1.67	9.23	5.02	1.68E-05	0.0050
TEAD4	TEA domain family member 4	-0.67	9.82	-5.00	1.75E-05	0.0051
DHH	desert hedgehog	-0.40	6.57	-4.96	1.97E-05	0.0056
EFHD2	EF-hand domain family, member D2	-0.82	10.90	-4.92	2.24E-05	0.0063
LHFPL2	lipoma HMGIC fusion partner-like 2	0.87	8.88	4.89	2.45E-05	0.0067
SGIP1	SH3-domain GRB2-like (endophilin) interacting protein 1	0.90	7.33	4.87	2.59E-05	0.0070
RGL1	ral guanine nucleotide dissociation stimulator-like 1	0.84	9.44	4.86	2.70E-05	0.0071
PGF	placental growth factor	1.78	10.00	4.82	2.98E-05	0.0077
SEP6	septin 6	-0.69	7.60	-4.81	3.08E-05	0.0078
CALML4	calmodulin-like 4	-0.83	7.22	-4.80	3.17E-05	0.0079
JUP	junction plakoglobin	1.15	9.79	4.77	3.55E-05	0.0084
MID1	midline 1	0.69	8.04	4.77	3.55E-05	0.0084
NTN4	netrin 4	1.10	10.65	4.77	3.46E-05	0.0084
FLRT2	fibronectin leucine rich transmembrane protein 2	1.25	9.20	4.75	3.68E-05	0.0086
C6orf141	chromosome 6 open reading frame 141	0.95	7.27	4.73	3.89E-05	0.0088
HPCAL1	hippocalcin-like 1	-0.80	11.70	-4.73	3.99E-05	0.0088
RALA	v-ral simian leukemia viral oncogene homolog A (ras related)	0.75	12.57	4.73	3.99E-05	0.0088

APPENDIX 3: Differentially expressed genes at 144 hours of culture; comparison is turbulent versus laminar shear (i.e. positive fold-change is upregulation in oscillatory shear); ordered by p value

RPS15	ribosomal protein S15	-0.89	8.36	-4.67	4.74E-05	0.0102
SEC11C	SEC11 homolog C (S. cerevisiae)	-0.50	10.12	-4.67	4.68E-05	0.0102
WASF2	WAS protein family, member 2	0.73	8.43	4.66	4.82E-05	0.0102
ADORA2B	adenosine A2b receptor	-0.82	7.46	-4.66	4.89E-05	0.0102
CORO7	coronin 7	0.40	7.38	4.64	5.19E-05	0.0104
FOXR1	forkhead box R1	0.36	6.57	4.63	5.23E-05	0.0104
STK32B	serine/threonine kinase 32B	0.70	8.14	4.63	5.25E-05	0.0104
CXADR	coxsackie virus and adenovirus receptor	1.52	7.83	4.62	5.42E-05	0.0106
NUP37	nucleoporin 37kDa	-0.48	10.15	-4.62	5.49E-05	0.0106
MYH10	myosin, heavy chain 10, non-muscle	1.44	10.67	4.60	5.75E-05	0.0109
SH3D19	SH3 domain containing 19	0.66	9.91	4.60	5.77E-05	0.0109
KDR	kinase insert domain receptor (a type III receptor tyrosine kinase)	0.65	7.72	4.59	5.96E-05	0.0111
YPEL2	yippee-like 2 (Drosophila)	0.88	8.49	4.58	6.16E-05	0.0113
GABBR2	gamma-aminobutyric acid (GABA) B receptor, 2	-1.40	9.31	-4.53	7.09E-05	0.0126
RALGDS	ral guanine nucleotide dissociation stimulator	0.94	10.44	4.53	7.19E-05	0.0126
SLC25A19	solute carrier family 25 (mitochondrial thiamine pyrophosphate carrier), member 19	-0.57	8.28	-4.53	7.04E-05	0.0126
TNFAIP8L1	tumor necrosis factor, alpha-induced protein 8-like 1	0.83	9.13	4.52	7.22E-05	0.0126
C10orf10	chromosome 10 open reading frame 10	2.27	9.98	4.51	7.42E-05	0.0127
RC3H2	ring finger and CCCH-type domains 2	0.47	8.17	4.51	7.53E-05	0.0128
PTGES2	prostaglandin E synthase 2	-0.43	9.32	-4.49	7.93E-05	0.0132
RHOJ	ras homolog family member J	0.87	10.14	4.49	7.99E-05	0.0132
SMARCA1	SWI/SNF related, matrix associated, actin dependent regulator of chromatin, subfamily a, member 1	0.68	8.81	4.48	8.16E-05	0.0133
EPHX4	epoxide hydrolase 4	1.25	7.81	4.45	9.06E-05	0.0142
C2orf27A	chromosome 2 open reading frame 27A	0.45	6.81	4.45	8.91E-05	0.0142
ST6GALNAC1	ST6 (alpha-N-acetyl-neuraminyl-2,3-beta-galactosyl-1,3)-N-acetylgalactosaminide alpha-2,6-sialyltransferase 1	-0.36	6.37	-4.45	9.07E-05	0.0142
WDR4	WD repeat domain 4	-0.73	9.50	-4.45	9.01E-05	0.0142
MID2	midline 2	1.01	8.65	4.44	9.25E-05	0.0143
PSAT1	phosphoserine aminotransferase 1	-1.13	8.92	-4.43	9.47E-05	0.0145
IQCK	IQ motif containing K	0.70	9.49	4.42	9.71E-05	0.0145
TSPAN15	tetraspanin 15	0.58	7.15	4.42	9.67E-05	0.0145
CCDC74A	coiled-coil domain containing 74A	0.52	6.83	4.41	0.0001	0.0146

APPENDIX 3: Differentially expressed genes at 144 hours of culture; comparison is turbulent versus laminar shear (i.e. positive fold-change is upregulation in oscillatory shear); ordered by p value

HSD17B2	hydroxysteroid (17-beta) dehydrogenase 2	0.84	6.83	4.41	0.0001	0.0146
LFNG	LFNG O-fucosylpeptide 3-beta-N-acetylglucosaminyltransferase	-1.71	8.26	-4.41	0.0001	0.0146
SLC18B1	solute carrier family 18, subfamily B, member 1	0.61	7.98	4.39	0.0001	0.0148
CLDN5	claudin 5	1.28	11.19	4.39	0.0001	0.0148
LIMCH1	LIM and calponin homology domains 1	1.49	8.16	4.40	0.0001	0.0148
RRAGA	Ras-related GTP binding A	0.43	10.93	4.37	0.0001	0.0159
CD320	CD320 molecule	-0.69	9.58	-4.36	0.0001	0.0162
CLIP2	CAP-GLY domain containing linker protein 2	0.48	7.38	4.33	0.0001	0.0173
TBC1D2	TBC1 domain family, member 2	-1.22	7.98	-4.33	0.0001	0.0173
PKD2	polycystic kidney disease 2 (autosomal dominant)	0.47	8.86	4.32	0.0001	0.0174
RAD54L2	RAD54-like 2 (S. cerevisiae)	0.57	7.91	4.32	0.0001	0.0174
TBC1D4	TBC1 domain family, member 4	0.63	8.97	4.30	0.0001	0.0182
GPER1	G protein-coupled estrogen receptor 1	-0.82	7.00	-4.29	0.0001	0.0184
EIF5A2	eukaryotic translation initiation factor 5A2	0.50	7.32	4.28	0.0001	0.0188
LAMB2	laminin, beta 2 (laminin S)	0.62	9.62	4.26	0.0002	0.0196
CTSK	cathepsin K	1.63	8.29	4.26	0.0002	0.0198
ADM	adrenomedullin	1.26	10.54	4.25	0.0002	0.0201
SULT1A3	sulfotransferase family, cytosolic, 1A, phenol-preferring, member 3	-1.07	7.78	-4.24	0.0002	0.0202
CCDC88C	coiled-coil domain containing 88C	0.47	7.25	4.23	0.0002	0.0203
GPRC5B	G protein-coupled receptor, class C, group 5, member B	0.41	6.89	4.23	0.0002	0.0203
TIMM9	translocase of inner mitochondrial membrane 9 homolog (yeast)	-0.36	9.34	-4.24	0.0002	0.0203
RPL29	ribosomal protein L29	-0.94	9.68	-4.23	0.0002	0.0204
WFS1	Wolfram syndrome 1 (wolframin)	1.27	9.65	4.22	0.0002	0.0204
PDE5A	phosphodiesterase 5A, cGMP-specific	0.53	6.93	4.21	0.0002	0.0212
QPCT	glutaminyl-peptide cyclotransferase	-0.53	9.51	-4.21	0.0002	0.0212
DYSF	dysferlin	0.77	10.69	4.20	0.0002	0.0213
EPHB1	EPH receptor B1	0.77	6.84	4.19	0.0002	0.0220
SULT1E1	sulfotransferase family 1E, estrogen-preferring, member 1	-0.47	6.58	-4.17	0.0002	0.0226
WAC	WW domain containing adaptor with coiled-coil	0.46	10.26	4.17	0.0002	0.0226
EXTL3	exostosin-like glycosyltransferase 3	0.46	8.21	4.17	0.0002	0.0227
RGCC	regulator of cell cycle	1.99	11.26	4.15	0.0002	0.0232
CRIP1	cysteine-rich protein 1 (intestinal)	-1.12	7.35	-4.15	0.0002	0.0232
ENDOG	endonuclease G	-0.54	7.71	-4.15	0.0002	0.0232

APPENDIX 3: Differentially expressed genes at 144 hours of culture; comparison is turbulent versus laminar shear (i.e. positive fold-change is upregulation in oscillatory shear); ordered by p value

NPM3	nucleophosmin/nucleoplasmin 3	-0.68	9.46	-4.15	0.0002	0.0232
GPR160	G protein-coupled receptor 160	0.54	6.61	4.14	0.0002	0.0235
BOP1	block of proliferation 1	-0.56	10.43	-4.13	0.0002	0.0237
NT5C2	5'-nucleotidase, cytosolic II	0.50	9.55	4.13	0.0002	0.0237
CMKLR1	chemokine-like receptor 1	-1.23	6.73	-4.13	0.0002	0.0239
DCHS1	dachsous cadherin-related 1	0.78	7.79	4.12	0.0002	0.0241
MT1IP	metallothionein 1I, pseudogene	-1.25	10.07	-4.11	0.0002	0.0245
AXIN2	axin 2	0.52	7.00	4.10	0.0002	0.0247
CDR2	cerebellar degeneration-related protein 2, 62kDa	0.56	9.42	4.11	0.0002	0.0247
SLC7A7	solute carrier family 7 (amino acid transporter light chain, y+L system), member 7	0.98	9.29	4.11	0.0002	0.0247
ZFAND5	zinc finger, AN1-type domain 5	0.50	11.76	4.10	0.0002	0.0247
CD36	CD36 molecule (thrombospondin receptor)	0.68	6.50	4.08	0.0003	0.0260
LTC4S	leukotriene C4 synthase	-1.23	6.60	-4.08	0.0003	0.0261
TAPBP	TAP binding protein (tapasin)	0.77	9.43	4.07	0.0003	0.0267
PRR7	proline rich 7 (synaptic)	-0.40	6.98	-4.05	0.0003	0.0276
PRRG1	proline rich Gla (G-carboxyglutamic acid) 1	0.79	9.30	4.05	0.0003	0.0279
SNCAIP	synuclein, alpha interacting protein	1.27	7.49	4.03	0.0003	0.0289
MRPL24	mitochondrial ribosomal protein L24	-0.45	10.55	-4.02	0.0003	0.0293
NAGLU	N-acetylglucosaminidase, alpha	0.60	8.71	4.02	0.0003	0.0296
MRPL9	mitochondrial ribosomal protein L9	-0.46	8.40	-4.02	0.0003	0.0296
CDCA7	cell division cycle associated 7	-0.97	8.73	-4.00	0.0003	0.0298
LOC100506691	uncharacterized LOC100506691	-0.32	6.58	-4.00	0.0003	0.0298
ITPKB	inositol-trisphosphate 3-kinase B	0.79	7.84	4.01	0.0003	0.0298
LMCD1	LIM and cysteine-rich domains 1	-1.03	7.74	-4.01	0.0003	0.0298
IL7	interleukin 7	0.29	6.70	4.00	0.0003	0.0299
C1orf174	chromosome 1 open reading frame 174	-0.36	8.93	-3.99	0.0003	0.0301
EFCAB14	EF-hand calcium binding domain 14	0.77	9.10	4.00	0.0003	0.0301
CLTB	clathrin, light chain B	-0.41	10.68	-3.98	0.0003	0.0308
UBXN2B	UBX domain protein 2B	0.32	7.12	3.98	0.0003	0.0308
EPB41L5	erythrocyte membrane protein band 4.1 like 5	0.71	7.88	3.98	0.0004	0.0309
ATP5I	ATP synthase, H+ transporting, mitochondrial Fo complex, subunit E	-0.42	9.50	-3.97	0.0004	0.0318
PAWR	PRKC, apoptosis, WT1, regulator	0.67	9.04	3.96	0.0004	0.0322
HERC1	HECT and RLD domain containing E3 ubiquitin protein ligase family member 1	0.62	7.90	3.96	0.0004	0.0323
BACE2	beta-site APP-cleaving enzyme 2	0.92	10.51	3.95	0.0004	0.0325

APPENDIX 3: Differentially expressed genes at 144 hours of culture; comparison is turbulent versus laminar shear (i.e. positive fold-change is upregulation in oscillatory shear); ordered by p value

BCS1L	BC1 (ubiquinol-cytochrome c reductase) synthesis-like	-0.50	8.33	-3.93	0.0004	0.0337
BSDC1	BSD domain containing 1	0.37	9.52	3.93	0.0004	0.0337
ARL4D	ADP-ribosylation factor-like 4D	-0.37	6.61	-3.92	0.0004	0.0345
ADAMTS1	ADAM metallopeptidase with thrombospondin type 1 motif, 1	-1.76	8.52	-3.91	0.0004	0.0360
FAM189A2	family with sequence similarity 189, member A2	0.61	7.20	3.90	0.0004	0.0360
CCDC74B	coiled-coil domain containing 74B	0.42	6.95	3.90	0.0004	0.0360
KLHDC8B	kelch domain containing 8B	1.05	8.92	3.90	0.0004	0.0360
NPY5R	neuropeptide Y receptor Y5	-0.28	6.84	-3.89	0.0004	0.0360
RCOR2	REST corepressor 2	-0.23	6.46	-3.89	0.0004	0.0360
TOMM22	translocase of outer mitochondrial membrane 22 homolog (yeast)	-0.45	10.20	-3.90	0.0004	0.0360
TRAF5	TNF receptor-associated factor 5	0.32	7.18	3.89	0.0004	0.0360
TSEN54	TSEN54 tRNA splicing endonuclease subunit	-0.38	6.93	-3.89	0.0005	0.0363
FMNL2	formin-like 2	0.50	7.97	3.88	0.0005	0.0364
FABP4	fatty acid binding protein 4, adipocyte	2.99	9.86	3.88	0.0005	0.0367
FNBP1L	formin binding protein 1-like	0.68	10.05	3.87	0.0005	0.0367
PAPLN	papilin, proteoglycan-like sulfated glycoprotein	1.22	7.55	3.87	0.0005	0.0367
PRKAB1	protein kinase, AMP-activated, beta 1 non-catalytic subunit	0.63	7.74	3.87	0.0005	0.0367
RAB15	RAB15, member RAS oncogene family	0.46	7.43	3.87	0.0005	0.0368
GALNT15	polypeptide N-acetylgalactosaminyltransferase 15	1.34	7.15	3.87	0.0005	0.0370
DDX60L	DEAD (Asp-Glu-Ala-Asp) box polypeptide 60-like	0.55	7.40	3.85	0.0005	0.0384
RPS6KA4	ribosomal protein S6 kinase, 90kDa, polypeptide 4	-0.43	8.95	-3.85	0.0005	0.0384
ZNF185	zinc finger protein 185 (LIM domain)	-1.31	9.93	-3.85	0.0005	0.0384
RAB3A	RAB3A, member RAS oncogene family	0.21	6.52	3.84	0.0005	0.0389
FGFR3	fibroblast growth factor receptor 3	-0.70	6.65	-3.83	0.0005	0.0390
MTMR10	myotubularin related protein 10	1.05	9.37	3.84	0.0005	0.0390
PPP2R2A	protein phosphatase 2, regulatory subunit B, alpha	0.46	10.31	3.83	0.0005	0.0390
STX3	syntaxin 3	0.58	8.63	3.82	0.0005	0.0390
SULT1A4	sulfotransferase family, cytosolic, 1A, phenol-preferring, member 4	-1.11	7.93	-3.82	0.0005	0.0390
UCK2	uridine-cytidine kinase 2	-0.47	8.40	-3.83	0.0005	0.0390
WASF3	WAS protein family, member 3	0.82	8.05	3.83	0.0005	0.0390
ZNF500	zinc finger protein 500	0.25	6.95	3.83	0.0005	0.0390

APPENDIX 3: Differentially expressed genes at 144 hours of culture; comparison is turbulent versus laminar shear (i.e. positive fold-change is upregulation in oscillatory shear); ordered by p value

DNASE1L1	deoxyribonuclease I-like 1	-0.95	10.47	-3.82	0.0006	0.0392
MEPCE	methylphosphate capping enzyme	0.49	8.12	3.81	0.0006	0.0392
RPS26P11	ribosomal protein S26 pseudogene 11	-0.29	13.25	-3.81	0.0006	0.0392
WARS	tryptophanyl-tRNA synthetase	1.10	10.43	3.82	0.0006	0.0392
VASH1	vasohibin 1	1.37	9.46	3.81	0.0006	0.0394
UQCRFS1	ubiquinol-cytochrome c reductase, Rieske iron-sulfur polypeptide 1	-0.33	11.79	-3.80	0.0006	0.0401
FYN	FYN proto-oncogene, Src family tyrosine kinase	0.47	8.88	3.80	0.0006	0.0405
GRK5	G protein-coupled receptor kinase 5	0.76	10.30	3.80	0.0006	0.0405
HAUS7	HAUS augmin-like complex, subunit 7	-0.46	8.73	-3.79	0.0006	0.0405
AGPAT9	1-acylglycerol-3-phosphate O-acyltransferase 9	-0.66	7.30	-3.79	0.0006	0.0405
APLNR	apelin receptor	-1.06	6.58	-3.79	0.0006	0.0405
POLD2	polymerase (DNA directed), delta 2, accessory subunit	-0.53	9.15	-3.79	0.0006	0.0405
RNF144A	ring finger protein 144A	0.38	6.67	3.79	0.0006	0.0407
ELOVL1	ELOVL fatty acid elongase 1	-0.39	8.64	-3.78	0.0006	0.0408
GEMIN6	gem (nuclear organelle) associated protein 6	-0.59	9.22	-3.78	0.0006	0.0412
TGFBRAP1	transforming growth factor, beta receptor associated protein 1	0.70	8.12	3.77	0.0006	0.0414
SDPR	serum deprivation response	0.91	11.26	3.77	0.0006	0.0415
LIMS2	LIM and senescent cell antigen-like domains 2	-0.78	8.57	-3.77	0.0006	0.0420
POLE4	polymerase (DNA-directed), epsilon 4, accessory subunit	-0.57	10.75	-3.76	0.0006	0.0420
MARVELD2	MARVEL domain containing 2	0.73	7.39	3.76	0.0006	0.0421
CUL7	cullin 7	0.24	6.75	3.75	0.0007	0.0428
PDGFD	platelet derived growth factor D	0.59	7.55	3.75	0.0007	0.0428
PI16	peptidase inhibitor 16	-2.27	7.06	-3.74	0.0007	0.0443
TGFBR3	transforming growth factor, beta receptor III	1.07	8.74	3.74	0.0007	0.0443
MALL	mal, T-cell differentiation protein-like	-1.30	10.77	-3.73	0.0007	0.0443
IPO9	importin 9	0.40	8.59	3.73	0.0007	0.0449
ADRB2	adrenoceptor beta 2, surface	-0.64	7.79	-3.72	0.0007	0.0452
ST6GALNAC4	ST6 (alpha-N-acetyl-neuraminyl-2,3-beta-galactosyl-1,3)-N-acetylgalactosaminide alpha-2,6-sialyltransferase 4	-0.51	8.81	-3.72	0.0007	0.0453
ACKR1	atypical chemokine receptor 1 (Duffy blood group)	-2.42	7.09	-3.72	0.0007	0.0457
RNMT	RNA (guanine-7-) methyltransferase	0.44	8.18	3.71	0.0007	0.0459

APPENDIX 3: Differentially expressed genes at 144 hours of culture; comparison is turbulent versus laminar shear (i.e. positive fold-change is upregulation in oscillatory shear); ordered by p value

ATL3	atlastin GTPase 3	-0.47	9.70	-3.69	0.0008	0.0485
C1orf131	chromosome 1 open reading frame 131	-0.39	9.10	-3.68	0.0008	0.0485
DPH5	diphthamide biosynthesis 5	-0.61	8.58	-3.69	0.0008	0.0485
DUSP4	dual specificity phosphatase 4	-0.39	6.51	-3.69	0.0008	0.0485
EIF2AK3	eukaryotic translation initiation factor 2-alpha kinase 3	0.46	7.21	3.69	0.0008	0.0485
GAPDH	glyceraldehyde-3-phosphate dehydrogenase	-0.44	12.65	-3.68	0.0008	0.0485
ZNF511	zinc finger protein 511	-0.52	9.13	-3.69	0.0008	0.0485
CHTF8	CTF8, chromosome transmission fidelity factor 8 homolog (S. cerevisiae)	0.54	7.23	3.68	0.0008	0.0491
RPS21	ribosomal protein S21	-0.55	10.63	-3.68	0.0008	0.0491
A2M	alpha-2-macroglobulin	1.98	7.80	3.67	0.0008	0.0492
RIN1	Ras and Rab interactor 1	-0.36	7.07	-3.67	0.0008	0.0495
CHEK1	checkpoint kinase 1	-0.49	7.99	-3.67	0.0008	0.0498
SPRY1	sprouty homolog 1, antagonist of FGF signaling (Drosophila)	1.57	9.71	3.66	0.0008	0.0498
P2RY2	purinergic receptor P2Y, G-protein coupled, 2	-0.60	7.12	-3.66	0.0009	0.0498
RWDD1	RWD domain containing 1	-0.33	11.67	-3.66	0.0009	0.0498
TMEFF1	transmembrane protein with EGF-like and two follistatin-like domains 1	0.75	7.08	3.66	0.0009	0.0499

21. PUBLICATIONS AND PRIZES ARISING FROM OR COMPLEMENTING THIS WORK

PUBLICATIONS

- Renard M, **Francis C**, Gosh R, Scott A, Witmer PD, Adès LC, Andelfinger G, Arnaud P, Boileau C, Callewaert B, Guo D, Hanna N, Lindsay ME, Morisaki H, Morisaki T, Pachter N, Robert L, Van Laer L, Dietz H, Loeys B, Milewicz D, De Backer J. Clinical Validity of Genes for Heritable Thoracic Aortic Aneurysm and Dissections. *J Am Coll Cardiol.* 2018 Aug; 72(6):605-615
- Whiffin N, Walsh R, Govind R, Edwards M, Ahmad M, Zhang X, Tayal U, Buchan R, Midwinter W, Wilk A, Najgebauer H, **Francis C,** Wilkinson S, Monk T,Brett L, O'Regan D, Prasad SK, Morris-Rosendahl DJ, Barton PJR, Edwards E, Ware JS, Cook SA. CardioClassifier: demonstrating the power of disease- and gene-specific computational decision support for clinical genome interpretation. *Genet Med.* 2018 Oct; 20(10):1246-1254
- Biffi C, de Marvao A, Attard MI, Dawes TJW, Whiffin N, Bai W, Shi W, **Francis C,** Meyer H, Buchan R, Cook SA, Rueckert D, O'Regan DP. Three-dimensional Cardiovascular Imaging-Genetics: A Mass Univariate Framework. *Bioinformatics* 2018 Jan 1;34(1):97-103
- **Francis CM**. Coarctation of the aorta: are genes relevant? *J Cardiovasc Surg (Torino)*. 2016 May 31
- **Francis K**, Pepper J. Genetic Influences on Phenotype in Bicuspid Aortic Valve: Screening, Diagnosis and Prognosis. *Chapter in: Bicuspid Aortic Valve: Diagnosis, Surgical Treatment and Complications*, Nova Publishers 2015
- Reed DM, Foldes G, Kirkby NS, Ahmetaj-Shala B, Mataragka S, Mohamed NA, **Francis C**, Gara E, Harding SE, Mitchell JA. Morphology and vasoactive hormone profiles from endothelial cells derived from stem cells of different sources. *Biochem Biophys Res Commun*. 2014 Dec 12;455(3-4):172-7

ABSTRACTS

- **Francis C,** Jamshidi Y, Gkoutos G, Fowler T, Keavney B, Clift P. Update on familial thoracic aortic aneurysm disease in the 100,000 Genomes Project. BCS, 2019
- **Francis C,** Jamshidi Y, Gkoutos G, Keavney B, Fowler T, Clift P. Familial Thoracic Aortic Aneurysm Disease in the 100,000 Genomes Project: much still to discover. EUROGUCH, 2019
- **Francis C**, Meyer H, Ware J, Dawes TJW, de Marvao AS, Walsh R, John S, Buchan R, Quinlan M, Barton PJ, O'Regan DP, Cook SA. Genome-wide association studies reveal shared genetic influences on measures of aortic form and function in healthy volunteers. NHLI Postgraduate Day 2018
- **Francis C**, Dawes TJW, de Marvao AS, Walsh R, John S, Buchan R, Gandhi A, Alenaini W, Quinlan M, Barton PJ, O'Regan DP, Cook SA. Aortopathy-causing variants increase aortic stiffness in healthy individuals. BCS 2015
- **Francis C**, Prapa M, Abdulkareem N, Walsh R, John S, Buchan R, Barton PJ, Jahangiri M, Gatzoulis MA, Pepper J, Cook SA. Identifying genetic causes of bicuspid aortic valve: a long way to go. Heart Valve Society 2015
- **Francis C,** Freidin MB, Gashaw H, Reed D, Potter CM, Paul-Clark M, Kirkby NS, Pepper JR, Mitchell JA. Shear stress pattern regulates aortic endothelial cell transcriptome: global gene expression profiles and implications for aortic disease. ESC 2015 – won Best Poster (endothelial function)
- **Francis K**, Prapa M, Abdulkareem N, Shibu J, Buchan R, Barton PJ, Jahangiri M, Gatzoulis MA, Pepper J, Cook SA. Identification of likely pathogenic variants in patients with bicuspid aortic valve: correlation of complex genotype with a more severe aortic phenotype. BCS 2014: Best Abstract (ACHD category)

PRIZES

- Highly commended oral presentation, NHLI Postgraduate Research Day, 2018
- Best Poster (Endothelial Function), European Society of Cardiology Congress, 2015
- Best abstract, Congenital Heart Disease Category, BCS Conference, 2014