THE Y CHROMOSOME IN CARDIOVASCULAR DISEASE

Thesis submitted for the degree of

Doctor of Philosophy

at the University of Leicester

by

Lisa Danielle Susan Bloomer, BSc MSc

Department of Cardiovascular Sciences

University of Leicester

January 2013

THE Y CHROMOSOME IN CARDIOVASCULAR DISEASE

Lisa Danielle Susan Bloomer

Men develop coronary artery disease (CAD) approximately 10 years earlier and die of CAD more frequently (2:1) than age-matched women. This sexual dimorphism is even more striking in abdominal aortic aneurysms (AAA) – the ratio of affected men to women is as high as 6:1. A major biological difference between both sexes is the exclusive presence of the Y chromosome in men.

Haplogroup I of the Y chromosome increases the risk of CAD by ~50% compared to other Y chromosome lineages. The analysis of ~1940 men from 3 cohorts recruited from the general European population revealed that the association between haplogroup I and CAD is not driven by conventional cardiovascular risk factors BMI, blood pressure, total cholesterol, HDL-C, triglycerides, LDL-C, glucose, aggression (total aggression, physical aggression, verbal aggression, hostility, anger) or sex steroids (testosterone, androstenedione, DHEA-S, estrone, oestradiol). Transcriptomic analysis revealed that none of "Kyoto Encyclopaedia of Genes and Genomes" (KEGG) pathways were differentially regulated in monocytes between men from haplogroup I and all other lineages. In contrast, 30 KEGG pathways showed differential expression between haplogroup I and other lineages of the Y chromosome in transcriptome-wide analysis of macrophages. Of these, 19 pathways were interconnected by common genes of adaptive immunity and the inflammatory response. In gene-centred analysis of the Y chromosome, carriers of haplogroup I also showed significant down-regulation of two X-degenerate male specific Y chromosome genes (*PRKY* and *UTY*) in macrophages (P=0.001 and P=0.0001) but not in monocytes (P=0.181, and P=0.611). Meta-analysis of 3 cohorts of British and Irish men revealed that the presence of haplogroup I was not associated with predisposition to AAA.

Neither conventional CAD risk factors nor male-associated/sex-specific phenotypes appear to drive the increased risk of CAD in carriers of haplogroup I. Furthermore, these results show that immune system and response to inflammation may mediate the association between Y chromosome and CAD. The absence of association between the Y chromosome and AAA suggest differences in the pathogenesis of sexual dimorphism of AAA and CAD.

Acknowledgements

I would like to dedicate this work to my loving grandparents. You instilled in me the traits I needed to get where I am today and I hope I make you both proud. Secondly, I would like to thank my parents for the time, money and effort they have put into me over the years. Without their unwavering support and encouragement I would never have believed I could come this far. I would also like to dedicate this work to my wonderful partner, Nick. You have been there throughout all the last three years has thrown at me; both supporting me when I was low and celebrating my achievements. You made me believe when I felt all was lost; I could not have completed my PhD without you. You are an inspiration to me and I hope I make you proud. Thank you all for believing in me.

I would also like to express my sincere gratitude to my two primary supervisors, Dr Maciej Tomaszewski and Mr Matt Bown. Their patience, knowledge and time have moulded me into the scientist I am today.

I would also like to thank Miss Evi Christofidou, Dr Radek Debiec and Mr Matthew Denniff for their friendship and support throughout the last three years. I hope we remain firm friends.

Finally, I would like to thank the Department of Cardiovascular Sciences, University of Leicester and NHS Charitable Funds for funding this work.

Contents

List of figures

List of tables

Abbreviations list

A	Adenine
AAA	Abdominal aortic aneurysm
ACE	Angiotensin converting enzyme
ADCK5	aarF domain containing kinase 5
AIDS	Acquired immunodeficiency syndrome
AK1	Adenylate kinase
ALKBH7	AlkB, alkylation repair homolog 7
AMELX	Amelogenin X linked
AMELY	Amelogenin, Y-linked
ANGII	Angiotensin 2
ANRIL	Cyclin-dependent kinase inhibitor 2B antisense RNA
APOA1	Apolipoprotein A1
APOB	Apolipoprotein B
ATP	Adenosine triphosphate
AZF-C	Azoospermia factor C
B2M	Beta-microglobulin
BAB1	Bric a brac 1
BHF-FHS	British Heart Foundation Family Heart Study
BMI	Body mass index
BMP	Bone morphogenetic protein
BP	Blood pressure
BPY2	Basic charge, Y-linked 2
Byn3a1	butyrophilin subfamily 3A1
C	Cytosine
C12ORF45	Open reading frame chromosome 12
C22ORF36	Open reading frame chromosome 22
CAD	Coronary artery disease
CAD	Carbamoyl-phosphate synthetase 2
CAM	Cell adhesion molecules
CAR	Coxsackie-adenoviral receptor
CDKN2A	Cyclin-dependent kinase inhibitor 2A
CDKN2B	Cyclin-dependent kinase inhibitor 2B
cDNA	Complementary deoxyribonucleic acid
CDY	Chromodomain protein, Y-linked
CETP	Cholesterol ester transfer protein
CEU	Utah residents with Northern and Western European ancestry
CI	Confidence interval
CKAP5	Cytoskeleton associated protein 5
CNTNAP1	Contactin associated protein 1
COL2α1	Collagen type II alpha 1
COL3α1	Collagen type III alpha 1
COPD	Chronic obstructive pulmonary disease
CRP	C- reactive protein
CT	Cycle threshold
CTS	Cardiogenics transcrptomic study
CVB3	Coxsackievirus B3
CVD	Cardiovascular disease
CXCL12	Chemokine (C-X-C motif) ligand 12

CXCR2	Chemokine (C-X-C motif) receptor 2
CXorf15	Chromosome X open reading frame 15
CYorf15A	Chromosome Y open reading frame 15A
CYorf15B	Chromosome Y open reading frame 15B
CYP27A1	Cytochrome P450, family 27, subfamily A and polypeptide 1
DAB2	Disabled homolog 2
DAB2	Disabled homolog 2 interacting protein
DAB2IP	Disabled homolog 2 interacting protein
DAF	Decay-accelerating factor
DAZ	Deleted in azoospermia
DBP	Diastolic blood pressure
^{o}C	Degrees Celsius
DDX3X	DEAD (Asp-Glu-Ala-Asp) box polypeptide 3 X linked
dCT	Delta cycle threshold
DHEA-S	Dehydroepiandrosterone sulphate
DCM	Dilated cardiomyopathy
DNA	Deoxyribonucleic acid
dNTPs	Deoxynucleotide Triphosphates
DTT	Dithiothreitol
E. coli	Escherichia coli
ECM	Extra cellular matrix
EDNRA	Endothelin receptor type A
E.G.	For example
eGFR	Estimated Glomerular filtration rate
EIF2S1	Eukaryotic translation initiation factor 2 and subunit 1 alpha
EIF1AX	Eukaryotic translation initiation factor 1A X linked
EIF1AY	Eukaryotic translation initiation factor 1A, Y-linked
eIF4G	Eukaryote initiation factor 4G
eNOS	Endothelial nitric oxide synthase
FAM134A	Family with sequence similarity 134, member A
FAM116B	Family with sequence similarity 116, member B
FBN1	Fibrillin 1
FBXO10	F-box protein 10
FDR	False discovery rate
FSH	Follicle stimulating hormone
G	Guanine
g	G force
GFM1	G elongation factor, mitochondrial 1
GRAPHIC	Genetic Regulation of Arterial Pressure in Humans in the Community
GSEA	Gene set enrichment analysis
GTP	Guanosine triphosphate,
GWASs	Genome-wide association study
HDL-C	High density lipoprotein cholesterol
HIV	Human immunodeficiency virus
HLA	Human leukocyte antigen
HLA-A	Major histocompatability complex class 1A
HLA-DR	Human leukocyte antigen-D related
Hrs	Hours
HSFY	Heat shock transcription factor, Y-linked 1
HSP72	Heat shock protein 72

HSPD1	Heat shock protein 1
IAAA	Inflammatory abdominal aortic aneurysm
ICAM-1	Intercellular adhesion molecule 1
IFNγ	Interferon gamma
IFI6	Interferon 6
Ig	Immunoglobulin
IL	Interleukin
ITGA6	Integrin, α6
ITGAL	Integrin α L
KEGG	Kyoto Encyclopaedia of Genes and Genomes
kg	Kilogram
JAP+CHB	Japanese in Tokyo, Japan and Han Chinese in Beijing, China
LCAT	Lecithin-cholesterol acyltransferase
LD	Linkage disequilibrium
LDL-C	Low density lipoprotein cholesterol
LDLR	Low density lipoprotein receptor
LH	Luteinising hormone
LDLRAP1	Low density lipoprotein receptor adapter protein 1
LINE-1	Long interspersed nuclear element 1
LIPC	Hepatic lipase
LOD	Likelihood of linkage
LPL	Lipoprotein lipase
LPS	Lipopolysaccharide
LRP1	Low density lipoprotein receptor related protein 1
LRP3	Low density lipoprotein receptor related protein 3
LRP6	Low density lipoprotein receptor related protein 6
LTR	Long tandem repeat
MAF	Minor allele frequency
MAPK	Mitogen activated protein kinase
MCIP1	Myocyte-enriched calcineurin-interacting protein
M-CSF	Monocyte colony stimulating factor
$MgCl_2$	Magnesium chloride
MGC71993	Ribonuclease, RNase K
MHC	Major histocompatability complex
MI	Myocardial infarction
µl	Microlitre
µg	Micrograms
µmil/L	Micrimoles per litre
Min	Minute
Ml	Millilitre
mmHg	Millimetres of mercury
mmol/L	Millimoles per litre
MMP	Matrix metalloproteinases
mRNA	messenger ribonucleic acid
MSY	Male specific region of the Y chromosome
MTAP	Methylthioadenosine phosphorylase
MTHFR	Methylenetetrahydrofolate reductase
n	Sample number
NADH	Nicotinamide adenine dinucleotide
NADPH	Nicotine adenine disphosphonucleotide, reduced

NDFIP2	Nedd4 family interacting protein 2
NFκB	Nuclear factor kappa light chain enhancer of activated B cells
ng	Nanograms
ng/µl	Nanograms per microlitre
NK cells	Natural killer cells
NLGN4X	Neuroligin 4 Y linked
NLGN4Y	Neuroligin 4, Y-linked
nm	Nanometres
OR	Odds ratio
P	P value
P2Rx4	Purinergic receptor P2X, ligand-gated ion channel, 4
PAD	Peripheral arterial disease
PAF-AH	Platelet activating factor acetylhydrolase
PAR1	Pseudoautosomal region 1
PAR2	Pseudoautosomal region 2
PCDH11X	Protocadherin 11 X linked
PCDH11Y	Protocadherin 11 Y-linked
PCSK9	Proprotein convertase subtilisin type 9
PCR	Polymerase chain reaction
PDCD5	Programmed cell death 5
PDGFRA	Platelet derived growth factor receptor alpha polypeptide
PDIA5	Protein disulfide isomerase family A, member 5
PEPD	Peptidase D
PGC-1α	Peroxisome proliferator-activated receptor-γ coactivator
PI3K-Act	Phosphatidylinositol-3 kinase Act
PPAR	Peroxisome proliferator-activated receptors
PRKX	Protein kinase, X linked
PRKY	Protein kinase, Y-linked
PRY	Protein tyrosine phosphatase, non-receptor type 13 (APO-1/CD95 (Fas)-associated phosphatase)-like protein, Y-linked
PSMA2	Proteasome (prosome, macropain) subunit, αtype, 2
PTP4A3	Protein tyrosine phosphatase type IVA and member 3
PVD	Peripheral vascular disease
RBMY	RNA binding motif protein, Y-linked
r	Pearson's correlation coefficient
RHR	Relative hazard ratio
RhoA	Ras homolog gene family, member A
rRNA	Ribosomal ribonucleic acid
RNA	Ribonucleic acid
RPS4X	Ribosomal protein S4, X-linked
RPS4Y1	Ribosomal protein S4, Y-linked 1
RPS4Y2	Ribosomal protein S4, Y-linked 2
RT-PCR	Reverse transcription polymerase chain reaction
RRAS2	Related RAS viral (r-ras) oncogene homolog 2
SBP	Systolic blood pressure
SCG5	Secretogranin V (7B2 protein)
SD	Standard deviation
SE	Standard error
s	Seconds
SHBG	Sex hormone binding globulin

SMCX	Lysine (K)-specific demethylase 5D X linked
KDM5D	Lysine (K)-specific demethylase 5D Y linked
SNP	Single nucleotide polymorphism
SPINK2	Serine peptidase inhibitor Kazal type 2
SRY	Sex determining region, Y
SRTB	Silesian Renal Tissue Bank
STR	Short tandem repeat
STS	Steroid sulfatase
T	Thymine
TBL1X	Transducin (beta)-like 1 X linked
TBL1Y	Transducin (beta)-like 1, Y-linked
TC	Total cholesterol
TGFβ1	Transforming growth factor beta 1
TGIF2LY	Transforming growth factor-induced factor homeobox 2-like, Y-linked
TIMPs	Tissue inhibitor of matrix metalloproteinases
TMEM147	Transmembrane protein 147
TMSB4X	Thymosin beta 4 X linked
TMSB4Y	Thymosin beta 4, Y-linked
TNFα	Tumour necrosis factor alpha
tRNA	Transfer ribonucleic acid
TROVE2	TROVE domain family, member 2
TSPY	Testis specific protein, Y-linked
TPT1	Tumour protein, translationally-controlled 1
TTTY	Testis-specific transcript
UK	United Kingdom
USA	United States of America
USP9X	Ubiquitin specific peptidase 9 X linked
USP9Y	Ubiquitin specific peptidase 9, Y-linked
UTX	Ubiquitously transcribed tetratricopeptide repeat gene X linked
UTY1	Ubiquitously transcribed tetratricopeptide repeat gene, Y-linked
VAV1	Vav 1 guanine nucleotide exchange factor
VCAM-1	Vascular cellular adhesion protein
VCY	Variably charged, Y-linked
WOSCOPS	West of Scotland Coronary Prevention Study
WPKP	Western Poland Kidney Project
XKRK	XK, Kell blood group complex subunit-related, Y-linked
XKRX	X Kell blood group precursor-related, X-linked
Y	Years
YMCA 1	Young Men Cardiovascular Association Studies 1
YMCA 2	Young Men Cardiovascular Association Studies 2
Yp	Short arm of the Y chromosome
Yq	Long arm of the Y chromosome
YRI	Yoruba in Ibadan, Nigeria
ZFX	Zinc finger protein X linked
ZFY	Zinc finger protein, Y-linked
ZNF410	Zinc finger protein 410

Publications resulting from this work (including poster/oral conference presentations)

Publications

Bloomer LDS, Bown MJ, Tomaszewski M. Sexual dimorphism of abdominal aortic aneurysms: a striking example of "male disadvantage" in cardiovascular disease. ***Atherosclerosis*** 2012;[Epub ahead of print]

Charchar FJ, **Bloomer LDS**, Barnes TA, Cowley MJ, Nelson CP, Wang Y, Denniff M, Debiec R, Christofidou P, Nankervis S, Dominiczak AF, Bani-Mustafa A, Balmforth AJ, Hall AS, Erdmann J, Cambien F, Deloukas P, Hengstenberg C, Packard C, Schunkert H, Ouwehand WH, Ford I, Goodall AH, Jobling MA, Samani NJ, Tomaszewski M. New insights into ininheritance of coronary artery disease in men-the role of the Y sex chromosome. ***The Lancet*** 2012;**379:**915-922

Other publications

Büsst CJ, **Bloomer LDS**, Scurrah KJ, Ellis JA, Barnes TA, Charchar FJ, Braund P, Hopkins PN, Samani NJ, Hunt SC, Tomaszewski M, Harrap SB. The epithelial sodium channel γ-subunit gene and blood pressure: family based association, renal gene expression, and physiological analyses. ***Hypertension*** 2011;**58:**1073-1078

Conference presentations and publications

Bloomer LDS, Denniff M, Bradley D, Christofidou P, Debiec R, Nelson CP, Scott DJ, Sohrabi S, Jonson A, Hughes A, Badger S, Charchar F, Thompson J, Sayers R, Samani NJ, Bown M, Tomaszewski M. *Phylogenetic analysis of the Y chromosome in abdominal aortic aneurysms: a meta-analysis of three British cohorts.* 24th Scientific Meeting of the International Society of Hypertension, Sydney, Australia (30/09/2012-04/10/2012) - poster presentation

Bloomer LDS, Denniff M, Christofidou P, Debiec R, Nelson CP, Zukowska-Szczechowska E, Thompson J, Charchar F, Samaini NJ, Tomaszewski M. *Androgens and aggressiveness do not mediate the effect of the Y chromosome on cardiovascular risk.* 24th Scientific Meeting of the International Society of Hypertension, Sydney, Australia (30/09/2012-04/10/2012) - oral presentation

Bloomer LDS. *The Y chromosome in coronary artery disease.* Third NIHR BRC/BRU/CLAHRC Experimental Medicine Research Training Camp**,** Ashridge, UK (04/07/2012-06/07/2012) - poster presentation

Bloomer LDS, Christofidou P, Debiec R, Denniff M, Bown MJ, Sayers R, Samani NJ, Tomaszewski M. *Genetic variation within the male-specific region of the Y chromosome, immunity and cardiovascular disease.* Leena Peltonen Summer School of Human Genomics, Cambridge, UK (21/08/2011-25/08/2011) - poster presentation

Bloomer LDS, Denniff M, Bradley D, Christofidou P, Debiec R, Nelson CP, Scott DJ, Sohrabi S, Jonson A, Hughes A, Badger S, Charchar F, Thompson J, Sayers R, Samani

NJ, Bown M, Tomaszewski M. *Phylogenetic analysis of the Y chromosome in abdominal aortic aneurysms: a meta-analysis of three British cohorts.* Cardiovascular genetics incorporating the 15th international SHR symposium & 48th Japanese SHR meeting, Melbourne, Australia (27/09/2012-28/09/2012) - poster presentation

Bloomer LDS, Denniff M, Christofidou P, Debiec R, Nelson CP, Zukowska-Szczechowska E, Thompson J, Charchar F, Samaini NJ, Tomaszewski M. *Androgens and aggressiveness do not mediate the effect of the Y chromosome on cardiovascular risk.* Cardiovascular genetics incorporating the 15th international SHR symposium & 48th Japanese SHR meeting, Melbourne, Australia (27/09/2012-28/09/2012) - poster presentation

Bloomer LDS, Denniff M, Christofidou P, Debiec R, Nelson CP, Cardiogenics Consortium, Thompson J, Goodall A[1], Samani NJ, Charchar F, Tomaszewski M. *The mechanisms of association between the Y chromosome and coronary artery disease.* ISH New Investigators' Satellite, Sydney, Australia (29/09/2012) – oral presentation

Grants, awards and other esteem factors

- International Society of Hypertension New Investigator of the Month (03/2012) (http://www.ish-world.com/Nin/Pages/Spotlight.aspx)
- International Society of Hypertension Stuart Spencer Award (September 2012)
- Kroner Fellowship (27/06/2012)
- "The association between the Y chromosome and abdominal aortic aneurysms" Leicester NHS Charitable Funds (2012-2013), £9,889 (Co-applicant)
- "Genetic variation in the human Y chromosome and cardiovascular risk factors" Leicester NHS -Charitable Funds (2012-2013), £9,773 (Co-applicant)
- Reviewer for "Gender Medicine"

1 Introduction

1.1 Cardiovascular disease (CVD)

Cardiovascular disease (CVD) is a multi-factorial disease whereby many genes act in conjunction with environmental factors to result in terminal clinical manifestations such as coronary artery disease (CAD), stroke and abdominal aortic aneurysm (AAA) (Bhupathy *et al.* 2010). CVD kills up to 20 million people a year (Bhupathy *et al.* 2010) and is a leading cause of death in the UK (Bhupathy *et al.* 2010). CAD results from the build-up of atherosclerotic plaques in the arteries supplying blood to the heart (Watkins & Farrall 2006; Soler & Ruiz 2010). AAA is the localised, permanent widening and dilation of the infrarenal aortic vessel wall (Grootenboer *et al.* 2003). The data shown here are primarily focused on CAD and AAA.

1.2 Coronary artery disease (CAD)

CAD accounts for approximately 50% of all CVD (Roger *et al.* 2012) and is caused by atherosclerosis, a degenerative disease caused by deposition of lipids and fibrous matrix in the blood vessel wall (Watkins & Farrell 2006). Symptoms occur when the oxygen supply to the cardiac muscle is either reduced or blocked. CAD is the most common cause of death in the Western world (Watkins & Farrell 2006). Risk factors for CAD include; family history, age, male sex, dyslipidaemia, hypertension, diabetes mellitus, obesity and smoking (Mozaffarian *et al.* 2009). CAD is a common, complex, multi-factorial polygenic disease that results from interactions between environmental and genetic factors.

1.2.1 Heritability

CAD is commonly described as a heritable disease (i.e. a disease where a proportion of the observable differences in the disease between individuals is due to genetic differences) due to the common familial clustering of disease. Seminal work completed using ~10,000 Swedish twin pairs (both monozygotic and dyzygotic) found that significantly more monozygotic twin pairs were concordant for death resulting from CAD compared to dizygotic twins. Indeed, of 1,649 monozygotic twin pairs investigated, 144 were concordant for death from CAD; whilst of 3,865 dizygotic twin pairs only 99 were concordant for death from CAD (Marenberg *et al.* 1994). The relative hazard ratio (RHR) of death from CAD was significantly higher in monozygotic twins than in dizygotic twins (male monozygotic twins RHR: 8.1, male dizygotic twins RHR 3.8, female monozygotic twins RHR: 15.0, female dizygotic twins RHR: 2.6) (Marenberg *et al.* 1994). Furthermore, the heritability of fatal CAD events was estimated at 57% in men and 38% in women (Zdravkovic *et al.* 2002). These estimates are typical for many common complex cardiovascular disorders.

CAD has been described as both a monogenic (i.e. occurring from one mutation in a single gene) and polygenic (where disease is caused by more than one variant; each with small effect on the phenotype) disease. The end points of both forms of disease are identical; resulting in fatty plaque build up in the arteries leading to MI. Common CAD is known to be complex (resulting from interactions between both genetics and the environment) and heterogeneous (where the variability in genetic background and environment in individuals results in a variable phenotype).

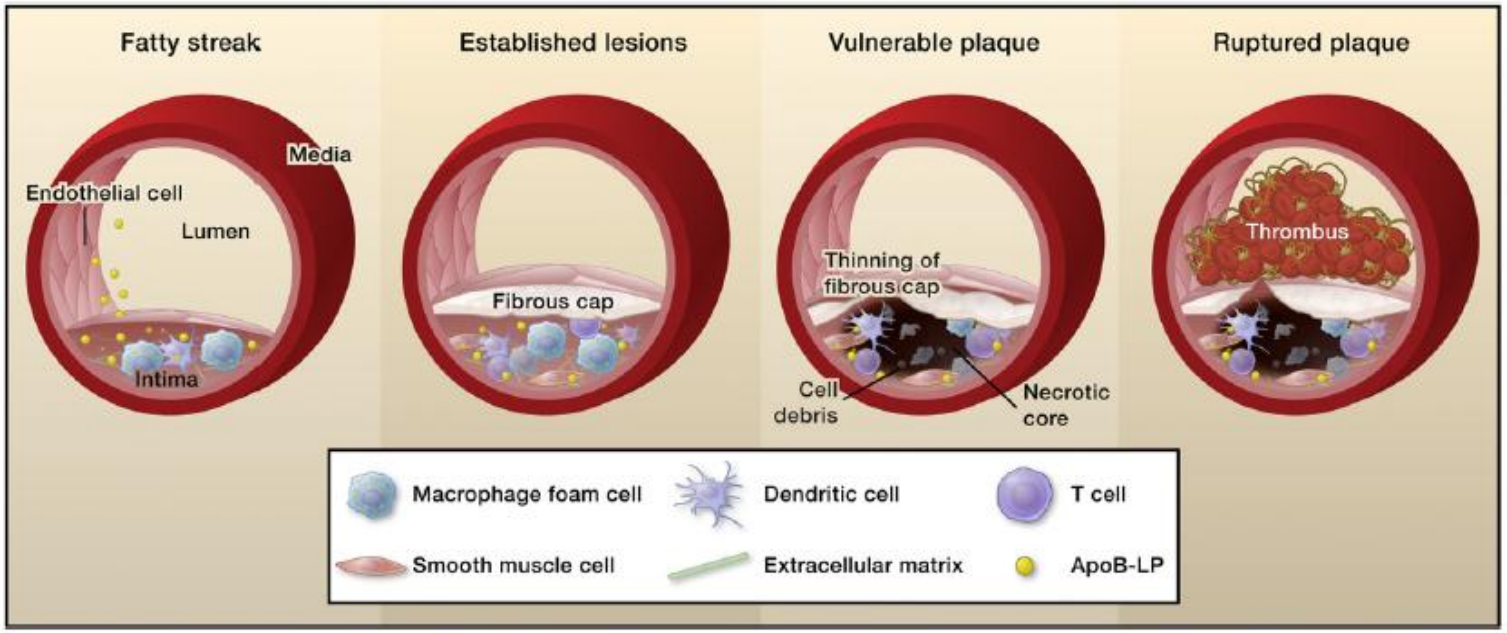
1.2.2 Pathogenesis

Historically, CAD was perceived purely as a cholesterol storage disease due to the production of lipid-laden foam cells. However, over the years it has become clear that many cell types, ranging from those in the artery wall to the immune system, are implicated in the pathogenesis of this disease (Libby & Theroux 2005) (Figure 1.1). Atherosclerosis is initiated when the normal barrier function of the arterial endothelium is lost. Causes of endothelial surface injury include bacterial products (such as lipopolysaccharide [LPS]), hyperlipidaemia/dyslipidaemia, vasoactive mediators, hyperglycaemia or proinflammatory cytokines released from adipose tissue. The initial formation of an inflammatory infiltrate (Libby & Theroux 2005; Watkins & Farrell 2006) is followed by increased adhesion of leukocytes (predominantly mononuclear cells, T lymphocytes and monocytes) to the inner surface of the arterial wall, causing the endothelium to lose functionality and the expression of adhesion molecules to increase. The action of chemo-attractants results in leukocyte infiltration into the arterial intima, where they release lipid mediators (such as protanoids and leukotrienes) acting on smooth muscle cells and causing cell migration to the intima (from the tunica media) as well as proliferation (Libby & Theroux 2005). Alongside the development of the inflammatory infiltrate, local proteoglycans bind lipoproteins, prolonging their presence in the intima and rendering them more susceptible to oxidative modification and glycation. The resultant oxidised phospholipids and advanced glycation end products sustain and develop the inflammatory response via the release of growth factors, cytokines and further up-regulation of adhesion molecules. These, in turn, recruit further leukocytes to the area of damaged endothelium (Libby & Theroux 2005; Watkins & Farrell 2006). Phagocytosis of oxidised lipids by macrophages precedes their conversion into foam cells. Accumulation of foam cells and further proliferation of

smooth muscle cells result in the growth of the plaque (Watkins & Farrell 2006). Apoptosis of foam cells and smooth muscle cells causes the classic lipid-rich necrotic core and fibrous cap of the plaque (Libby & Theroux 2005; Watkins & Farrell 2006). Sequential episodes of plaque rupture, thrombus formation and healing allow the plaque to grow (Watkins & Farrell 2006). Early lesions appear as fatty streaks due to the presence of foam cells. Once a plaque has formed, one of three events may occur:

1. The plaque will remain silent (Watkins & Farrell 2006).
2. The plaque will progressively grow, slowly narrowing the lumen and restricting blood flow. Generally these plaques are fibrotic and have thick caps. They result in angina pectoris (Libby & Theroux 2005, Watkins & Farrell 2006).
3. The plaque will completely occlude the lumen through acute thrombosis and cause myocardial infarction (MI). Generally these plaques are lipid-rich with thin fibrous caps (Libby & Theroux 2005; Watkins & Farrell 2006).

Figure 1.1 Pathogenesis of CAD



Early fatty streak lesions are characterised by the accumulation of lipoproteins which causes the recruitment of dendritic cells and macrophages to the area. Smooth muscle cells and T cells infiltrate the intima where macrophages engluf oxidised LDL-C (Moore *et al.* 2011).

1.2.3 Monogenic forms of CVD leading to CAD

Monogenic disorders that result in premature CAD are rare and predominantly stem from mutations in genes involved in regulation of lipid metabolism. One of the most common forms of Mendelian hyperlipidaemia, and human monogenic diseases in general, is familial hypercholesterolaemia. This disorder leads to high levels of circulating low-density lipoprotein cholesterol (LDL-C) and fast progression of atherosclerosis if left untreated. Mutations in four genes have been recognised as causes of familial hypercholesterolemia (Table 1.1 and 1.2). The mode of ininheritance of these monogenic disorders has been described as both autosomal dominant and recessive (Watkins & Farrell. 2006).

Table 1.1 Monogenic forms of hypercholesterolaemia that result in premature CAD

Disease (OMIM)	**Affected gene**
Familial hypercholesterolaemia (143890)	*LDLR*
Familial defective apolipoprotein B-100 (144010)	*APOB*
Familial hypercholesterolemia 3 (603776)	*PCSK9*
Autosomal recessive hypercholesterolemia (603813)	*LDLRAP1*

LDLR- low density lipoprotein receptor gene, *APOB*- Apolipoprotein B, *PCSK9*- proprotein convertase subtilisin/kexin type 9, *LDLRAP1*- low density lipoprotein receptor adaptor protein 1.

Several monogenic forms of decreased high-density lipoprotein cholesterol (HDL-C) have also been described (Table 1.2), although some, such as Fish-eye disease, do not result in premature CAD (Kral & Becker, 2007).

Table 1.2 Monogenic disorders of decreased high-density lipoprotein-cholesterol (HDL-C) (Kral & Becker, 2007).

Disease (OMIM)	**Affected gene**
Familial hypoalphalipoproteinaemia (604091)	*APOA-I*
Familial LCAT deficiency (245900)	*LCAT*
Fish-eye disease (136120)	*LCAT*
Tangier disease (205400)	*ABCA-I*
LPL deficiency (246650)	*LPL*
CETP deficiency (245900)	*CETP*
Hepatic lipase deficiency (614025)	*LIPC*

APOA- Apolipoprotein A, *LCAT*- Lecithin-cholesterol acyltransferase, *ABCA-I*- ATP-binding cassette transporter 1, *LPL*- lipoprotein lipase, *CETP*- cholesteryl ester transfer protein, *LIPC*- hepatic lipase.

It is hoped that future research into the genetic basis of monogenic forms of CAD will lead to the discovery of novel candidate pathways and the development of new therapeutic approaches (Luft *et al.* 1995; Tomaszewski *et al.* 2006).

1.2.4 Non-monogenic causes of CAD

The clustering of CAD in families and increased individual risk of CAD when a positive family history is present suggests that CAD has a sizable genetic component (Watkins & Farrell 2006). Candidate gene studies, linkage studies and genome-wide association studies (GWASs) have all been used in an effort to identify the underlying genetic mechanisms. Candidate gene studies are hypothesis-driven and search for genetic differences within biologically relevant genes and pathways (mostly related to atherosclerosis such as signalling molecules, inflammation and immunity and apolipoproteins) [Watkins & Farrell 2006]). Linkage studies investigate the entire genome to find chromosomal regions which segregate with a disease in affected families. GWASs investigate thousands of single nucleotide polymorphisms (SNPs) from across the genome in an attempt to associate a genetic variant with the disease. Neither linkage studies nor GWASs are hypothesis-driven.

1.2.4.1 Candidate gene studies

To date, approximately 100 candidate genes studies have been completed in an effort to identify the genetic cause of CAD (Patel & Ye 2011). The majority of these investigations have been completed in case-control studies on genes contained in, or associated with, major cardiovascular pathways. These include the renin angiotensin system, lipid metabolism pathways, oxidative stress pathways, vascular remodelling and immunological pathways (Patel & Ye 2011). While a number of these investigations have revealed data of importance to CAD (Yang *et al.* 2003; Arroyo-Espliguero *et al.* 2004; Edfeldt *et al.* 2004; Helgadottir *et al.* 2006), significant inconsistencies exist among their findings. A frequently discussed example is the association of

polymorphisms in the angiotensin-converting enzyme gene (*ACE*) and CAD. A recent meta-analysis investigating 46 studies found that there was a significant but modest relationship between polymorphisms of *ACE* and the risk of CAD in Chinese populations (Zhou *et al.* 2012). However, Zintzaras *et al*. identified only a modest effect of the same polymorphisms in a meta-analysis of 118 studies of differing ethnicities implying that genetic effects outside easily associated genes may contribute to CAD (Zintzaras *et al.* 2008).

Candidate gene studies have generally failed to provide the robust candidate loci for predisposition to CAD due to issues with study design, definition of phenotype (for example MI before a certain age or presence of a CAD plaque) and statistical power. However, the failure of candidate gene studies has shown that major alleles and genes underlying susceptibility to CAD may reside outside the classical systems of cardiovascular regulation, including the five pathways described above (Manolio 2007; Charchar *et al.* 2008; Patel & Ye 2011).

1.2.4.2 Linkage studies

Linkage studies investigate chromosomal regions which track with predisposition to disease in families (Patel & Ye 2011). While this method of investigation was particularly successful when applied to monogenic diseases, it is of limited value in complex diseases such as CAD (Patel & Ye 2011). To date, 11 linkage studies have been completed on a number of differing ethnicities. One of the most important contributions is the British Heart Foundation Family Heart Study that recruited 4,175 individuals from 1,933 families in the UK. Each family had at least two siblings concordant for validated early onset CAD (before age of 66 years). In genome-wide

analysis, CAD was linked to a region on chromosome 2 with a likelihood of linkage (LOD) score of 1.86. The LOD score increased to 2.7 in families with a minimum age of diagnosis of 56 years (Samani *et al.* 2005). Importantly, this finding was in concordance with two other small studies. Wang *et al.* reported a LOD score of 3.82 for the same locus in a study of 928 US families (Wang *et al.* 2005) while Pajukanta *et al*. identified the same region in a Finnish population of 156 families (LOD of 3.2) (Pajukanta *et al.* 2000). Other studies have revealed linkage to 16p13 (Francke *et al*. 2001), 2p36 (Harrap *et al.* 2002), 14 (Broeckel *et al*. 2002), 3p13 (Hauser *et al*. 2004), 13q.12 (Helgadorttir *et al.* 2004), 12q22 (Helgadottir *et al.* 2006) and 17p (Farrall *et al.* 2006) with CAD.

The major lesson learnt from genome-wide linkage scans is that CAD is polygenic and heterogenous in nature (multiple quantitative trait loci [QTLs] in different populations) (Patel & Ye 2011). One of the major problems with making further progress in this form of research was that the identified QTLs were very large (due to being based on LD), spanning several Mb and harbouring tens of potentially causative genes. Very few QTL-focused studies have managed to successfully refine the linkage peaks and identify the driver of linkage to cardiovascular phenotype (Tomaszewski *et al.* 2007).

1.2.4.3 Genome-wide association studies

GWASs search for alleles and SNPs associated with the trait of interest through systematic, unbiased screening of human autosomes and/or the X chromosome (Patel & Ye 2011). "SNP chips" of up to one million SNPs are currently used in this form of work. Generally, common SNPs, with a minor-allele frequency (MAF) ≥5% in the

population under investigation are chosen. Each SNP acts as a tagging SNP for the LD block in which it is posititioned. By choosing SNPs in differing LD blocks, more of the genome can be investigated maximising the potential academic use of the investigation. To date, 31 SNPs have been identified as significantly associated with CAD through GWASs (Table 1.3) (Charchar *et al.* 2012). Most notably, the chromosomal region on human chromosome 9 (9p21) was associated with CAD. The Wellcome Trust Case-Control Consortium identified a significant association with rs1333049 and CAD. This SNP increased risk of CAD by 36% (OR 1.9 in subjects homozygous for the risk allele) (Myocardial Infarction Genetics Consortium *et al*. 2009). The region contains three genes; cyclin-dependent kinase inhibitor 2B antisense RNA (*ANRIL*), cyclin-dependent kinase inhibitor 2A gene (*CDKN2A*), cyclin-dependent kinase inhibitor 2B gene (*CDKN2B*) and methylthioadenosine phosphorylase gene (*MTAP*). This finding has been extensively replicated across different populations (McPherson *et al.* 2007; Helgadottir *et al.* 2008; Hinohara *et al.* 2008; Shen *et al.* 2008; Samani *et al.* 2007), culminating in a meta-analysis that estimated the OR of CAD to be 1.29 for the rs1333049 risk allele (Schunkert *et al.* 2008).

While over thirty loci have now been significantly associated with CAD through GWASs, these common SNPs (MAF≥5% in the study population) only account for around 10% of the heritability of CAD (Patel & Ye. 2011). This means a significant portion of CAD heritability remains unexplained.

Table 1.3 SNPs associated with CAD in recent GWASs

SNP	Genomic loci	Gene	Allele frequency	β score/ OR	SE	Platform information	No. cases (controls)	Reference
rs2228671	19p13.2	*LDLR*	0.88	0.2	0.15 - 9.3E-24	16 cohorts, Illumina HumanHap300-Duo platform	17,797 (total study population)	Aulchenko *et al*. 2009
rs10953541	7q22-q31	*BCAP29*	0.75	1.1	1.05 - 1.11	Meta-analysis (4 studies), 575,000 SNPs	15,420 (15,062)	Coronary Artery Disease (C4D) Genetics Consortium, 2011
rs2505083	10p11.23	*KIAA1462*						
rs2246942	10q23.31	*LIPA*	0.34	1.08	1.05 - 1.12			
rs974819	11q22.3	*PDGFD*	0.3	1.09	1.05 - 1.13			
rs11206510	1p32.3	*PCSK9*	0.81	1.12	1.02 - 1.23	727,496 SNPs (Affymetrix 6.0	2,967 (3,075)	Myocardial Infarction Genetics Consortium
rs6725887	2q33.1	*WDR12*	0.14	1.24	1.12 - 1.38			

rs12526453	6p24.1	*PHACTR1*	0.65	1.15	1.07 - 1.24	GeneChip)		*et al*. 2009
rs4977574	9p32	*CDKN2A/B, ANRIL*	0.56	1.25	1.16 - 1.34			
rs1746048	10q11.21	*CXCL12*	0.84	1.22	1.10 - 1.34			
rs17465637	1q41	*MIA3*	0.29	1.23	1.12 - 1.34	3 cohorts, GeneChip Human Mapping 500K Array Set (Affymetrix)	1,926 (2,938)	Samani *et al*. 2007
rs6922269	6q25.1	*MTHFD1L*	0.3	1.23	1.13 - 1.35			
rs17114036	1p32.2	*PPAP2B*	0.91	1.17	1.13 - 1.22	Meta-analysis (14 studies)	22,233 (64,762)	Schunkert *et al*. 2011
rs2306374	3q22.3	*MRAS*						
rs17609940	6p21.31	*ANKS1A*	0.75	1.07	1.05 - 1.10			
rs12190287	6q23.2	*TCF21*	0.62	1.08	1.06 - 1.10			
rs3798220	6q25.3	*LPA*	0.02	1.54	1.36 - 1.74			

rs579459	9q34.2	*ABO*	0.21	1.1	1.07 - 1.13			
rs12413409	10q24.32	*CYP17A1, CNNM2, NT5C2*	0.89	1.12	1.08 - 1.16			
rs964184	11q23.3	*ZNF259, APOA5-A4-C3-A1*	0.13	1.13	1.10 - 1.16			
rs4773144	13q34	*COL4A1, COL4A2*	0.44	1.07	1.05 - 1.10			
rs3825807	15q25.1	*ADAMTS7*	0.57	1.08	1.06 - 1.10			
rs216172	17p13.3	*SMG6, SRR*	0.37	1.07	1.05 - 1.09			
rs46522	17q21.32	*UBE2Z, GIP, ATP5G1*	0.53	1.06	1.04 - 1.08			

rs1122608	19p13.2	*LDLR (SMARCA4)*	0.77	1.14	1.09 - 1.18			
rs9982601	21q22.11	*MRPS6*	0.15	1.18	1.12 - 1.24			
rs4299376	2p21	*ABCG8*	0.45	0.85	0.87 - 0.9	Meta analysis (11 studies)	13,764 (13,630)	Teupser *et al*. 2010

1.2.5 Prevalence of CAD

CAD is widely documented as a disease of older males. Men develop CAD ~10 years earlier, and die more frequently from CAD than age-matched women (Yusuf *et al.* 2004). The prevalence of CAD is around 2:1 men to women and this male predominance is observed in all geographical locations and ethnicities studied (Lowry *et al.* 1983; Ness & Aronow 1999; Roger *et al*. 2012). Prevalence and incidence data are commonly described in person years so as to account for situations where the amount of observation time differs between people, or for when the population risk varies with time (Table 1.4). In the Western world, the death rate per 100,000 individuals ranges from 75.6- 706 in men to 22.2- 237.1 in women (Roger *et al*. 2012). These statistics immediately indicate the apparent sexual dimorphism in the prevalence and diagnosis of CAD. Men are at significantly more risk from CAD than women. CAD is a worldwide health and economic burden that affects every country; 35% of African Americans and 38% of Asians had CAD in a cohort of older people (Ness & Aronow 1999). Both prevalence and severity of CAD vary geographically; more indidivuals of European ancestry are diagnosed with CAD, although Asian patients tend to suffer from more severe forms of CAD (Lowry *et al.* 1983).

All available CAD prevalence data can be found in Table 1.4. These data show that irrespective of country examined, the male disadvantage at diagnosis is always observed. In 2009 there were ~80,000 deaths from CVD in Britain, of which 32% were due to CAD (Scarborough *et al.* 2011). Approximately 30% of all male and 20% of all female deaths are due to CAD (Scarborough *et al.* 2011). The incidence of CAD increases with age in both sexes; in those <50 years of age the incidence of CAD is 11/100,000 and 4/100,000 in men and women, respectively. This rate increases to 120/100,000 men and 60/100,000 women in those <75 years of age (Scarborough *et al.*

2011). This trend is also visible in statistics of total deaths from CAD. At ages <35 years, just 460 men and 267 women died of CAD in 2009. This estimate increased to 8,986 men and 3,461 men at ages 55-64 years and was highest in those aged >75 years: 55,891 men and 78,069 women. The overall death rate from CAD is significantly higher in men; these data may be confounded by women living considerably longer than men. 78,000 women aged >75 years died from CAD in 2009 compared to 55,000 men; however, the average population at this age is 2.9 million women compared with 1.9 million men (Scarborough *et al.* 2011; Folsom *et al.* 1998; Ness & Aronow. 1999; Roger *et al*. 2012).

Table 1.4 Worldwide prevalence of CAD

Location	**Total prevalence**	
	Men	**Women**
Russia	706	237◊
Hungary	384	144◊
Bulgaria	253	86◊
Poland	190	55◊
Greece	136	34◊
Germany	125	38◊
Italy	75	22◊
England and Wales	138	43◊
Northern Ireland	125	51◊
Eire	152	48◊
Scotland	195	67◊

Australia	89	27◊
New Zealand	138	47◊
USA	153	60◊
Rural China	41	29◊
Urban China	108	72◊
Korea	47	13◊
Japan	51	20◊
Israel	78	27◊
Argentina	140	40◊
USA Black	10	5^
USA White	13	4^
England	38,531	28,565*
Scotland	2151	876*
Wales	1166	400*
Northern Ireland	551	220*

◊ Total prevalence per 100,000 person years (Roger *et al.* 2012), ^total prevalence per 1000 person years (Dornas *et al.* 2008; Kawaguchi *et al.* 2011), * total deaths due to AAA in 2009 (Scarborough *et al.* 2011)

1.2.6 Sexual dimorphism in risk factors of CAD

The extensive epidemiological data on CAD shown here highlight the extreme sexual dimorphism in diagnosis of CAD and CAD-related death. Men are at a significantly higher risk of developing CAD and CAD-related mortality than age-matched women. Likewise, many of the CAD risk factors such as smoking, hyperlipidaemia,

hypertension and diabetes mellitus are known to be sexually dimorphic and have a sexually dimorphic inheritance pattern (Pan *et al.* 2007; Ober *et al.* 2008).

1.2.6.1 Effect of smoking on CAD

There are up to 1.1 billion smokers worldwide (Huxley & Woodward 2011). However, only a fifth of these are female (Huxley & Woodward 2011). This over-representation of men amongst smokers is apparent at all ages and across different ethnicities. The INTERHEART study showed that 57.3% of men had smoked at some point in their lives compared to just 21.9% of women (Teo *et al.* 2006). Likewise, the number of cigarettes consumed per day is significantly higher in males (Lederle *et al.* 1997) and the increase in level of usage with age is higher in men (Lederle *et al.* 1997). 17.6% of men consume over 20 cigarettes per day compared to just 2.8% of women (Teo *et al.* 2006). In non-Hispanic Americans of European ancestry, 57.6% of individuals smoke (relative risk of CAD is 2.2 with history of and current smoking) (Wei *et al.* 1996). Smoking is known as a serious risk factor for CAD. The cardiovascular risk associated with smoking is up to 50% higher in men than women (Prescott *et al.* 1998). A meta-analysis based on 3,912,809 people from 26 different studies showed that prevalence of smoking and amount of smoking was significantly higher in males (male prevalence 2-71%, female prevalence 1-44%) (Huxley & Woodward 2011) but women had a 25% greater relative risk of CAD from smoking independent of other CVD risk factors than men (Huxley & Woodward 2011). The trend of higher CAD risk in female smokers is found across ethnicities, though in Asia the relationship is only seen in the heaviest smokers (Woodward *et al.* 2005).

1.2.6.2 Lipids levels in CAD

The presence of dyslipidaemia is well accepted to predispose to CAD. The most extreme sexual dimorphism in normal populations is observed in HDL-C; many prospective and cross-sectional studies have shown higher HDL-C in women than men (Genovesi *et al.* 2005; McQueen *et al.* 2008). Interestingly, decreased levels of HDL-C are particularly prevalent in males who developed CAD (Howard *et al.* 1995). Furthermore, the heritability of HDL-C may differ between the sexes; a heritability of 0.6 was calculated in men while this was estimated at 0.7 in women (Pan *et al.* 2007; Ober *et al.* 2008). The sexual dimorphism in other lipid fractions including LDL-C is less consistent. A study of 27,098 pre-menopausal women and men showed that women had significantly lower LDL-C and total cholesterol (TC) (Sharrett *et al.* 2001). However, a study on 21,465 individuals, where some women were post-menopausal, revealed that women had higher TC than men (Genovesi *et al.* 2005). 15% of all adults have total serum cholesterol over 6.2 mmol/L (Roger *et al*. 2012) and a TC of >6.2 mmol/L confers a 2-fold increase in relative risk of CAD (Wei *et al.* 1996). Increased LDL-C and TC are associated with increased risk of CAD; most significantly in women (LDL-C >4.1 mmol/L showed a prevalence of CAD of 31.2/100 in women and 20.9/100 in men, while TC values <6.2 mmol/L gave a prevalence of CAD of 27/100 in women and 20/100 in men [Howard *et al.* 1995]). Interestingly, LDL-C shows a higher heritability in men than women (0.5 and 0.3 respectively, additive model) (Pan *et al.* 2007; Ober *et al.* 2008).

1.2.6.3 Hypertension prevalence in CAD

Prevalence of hypertension is similar across the sexes. Burt *et al.* identified that 24% of the general American population had hypertension (defined as blood pressure [BP] ≥140/90 mmHg), and men were slightly more commonly diagnosed with hypertension (24.7%) compared to women (23.4%) (Burt *et al.* 1995). British men have a significantly higher prevelence of hypertension than women (8.0 and 7.5 per 100,000 in women and men, respectively). Kei *et al.* found that hypertension was associated with a higher risk of CAD in women than men (relative risk of 1.97 in men and 2.63 in women) (Keil *et al.* 1993).

1.2.6.4 Diabetes mellitus as a risk factor of CAD

Diabetes mellitus is a well-documented risk factor for CAD. 8% of Americans of European ancestry have diabetes, 36.8% have pre-diabetes, and a further 34% (35.1% of men and 32.6% of women) have metabolic syndrome (Roger *et al*. 2012). Diabetes is also considered by some authors as a sexually dimorphic trait affecting more men than women (Howard *et al.* 1995; Maric 2009), and more diabetic men develop CAD than diabetic women. Indeed, a study by Howard *et al.* showed that 2.2% of diabetic women developed CAD compared with 5.8% of diabetic men (Howard *et al.* 1995).

1.3 Abdominal aortic aneurysms (AAA)

1.3.1 Normal structure of the aorta

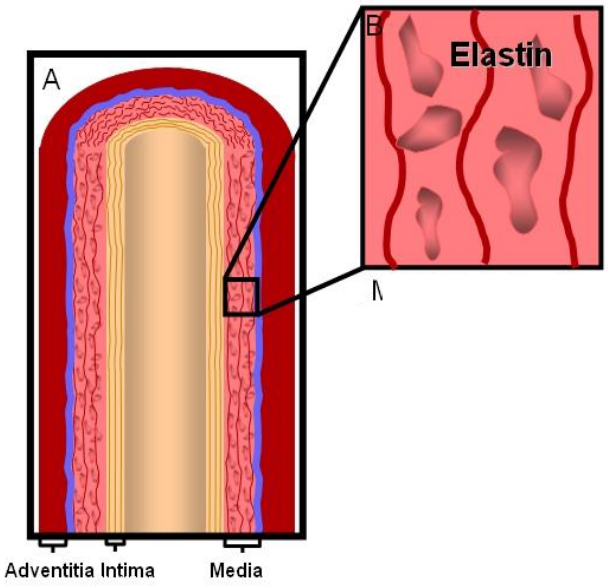
Abdominal aortic aneurysm is a disease of the aortic vessel wall, and to understand this disease it is first useful to describe the structure and physiology of the aorta. A physiologically normal aorta has an approximate diameter of 28 mm in the thorax, and 20 mm in the abdominal cavity (infrarenally) (Grootenboer *et al.* 2003). Aortic diameter is modulated by age, height, weight, sex and body surface area (Grootenboer *et al.* 2003). The major function of the aorta is to regulate the pulsatile pressure of blood originating from the heart and provide a consistent flow of blood to peripheral capillary beds (Shadwick 1999). This is achieved by maximising blood flow during diastole and minimising blood flow during systole (Wagenseil & Mecham 2009).

The aortic wall has three layers; the intima, media and adventitia, which are separated by the elastic lamina (comprised purely of elastin fibres) (Figure 1.2). The intima is formed of endothelial cells, (which control smooth muscle action throughout the artery), fibroblasts, smooth muscle cells and a small number of elastin fibres (Arnett 2007). The media is comprised of smooth muscle cells, elastin (which is organised into concentric elastic lamellae with microfibrils), and collagen fibres. The outermost layer of the aorta, the adventitia, is formed from collagen, microfibrils, elastin and fibroblasts (Arnett 2007).

The three protein constituents of the media each have distinct roles in maintaining the function and integrity of the vessel. Elastin, with low tensile strength and high dispensability, acts as a reservoir for elastic tension allowing the distribution of pressure from each heart contraction throughout the vessel wall (Wagenseil & Mecham 2009). Elastin is organised into thick concentric circles of elastin and microfibrils (elastic lamella) that form a highly extensible network of fibres. Collagen fibres are arranged into bundles and are dispersed throughout the circumference of the

aorta. These bundles stiffen and become circumferentially aligned during increased BP, preventing arterial damage and rupture (Wagenseil & Mecham 2009, Arnett 2007). The elastin:collagen ratio in the aortic arch is 2, while a ratio of 0.5 is observed infrareanally (Shadwick 1999). Smooth muscle cells secrete elastin and alter aortic diameter through vasodilatation and vasoconstriction, and are controlled by endothelial cells in the intima (Wagenseil & Mecham 2009).

Figure 1.2 The normal structure of the aorta



http://www.google.ca/imgres?q=structure+of+the+aorta+intima+media&um=1&hl=en&sa=N&tbo=d&biw=1091&bih=537&tbm=isch&tbnid=wtJeiWWZqYpRSM:&imgrefurl=http://academicdepartments.musc.edu/surgery/research/ctresearch/ikonomidis_research.htm&docid=CQVgjs0OJieDoM&imgurl=http://academicdepartments.musc.edu/sebin/h/u/ikon_figure.jpg&w=914&h=514&ei=N5LsULHkJMLorAHll4DYDw&zoom=1&iact=hc&vpx=381&vpy=255&dur=3752&hovh=168&hovw=300&tx=101&ty=120&sig=117898223437388299147&page=1&tbnh=136&tbnw=243&start=0&ndsp=19&ved=1t:429,r:10,s:0,i:118

1.3.2 Overview of AAA

AAA is the localised, permanent widening or dilation of the aortic vessel wall (Karkos *et al.* 2000; Grootenboer *et al.* 2003; Wanhainen 2008) (Figure 1.3). AAA is normally asymptomatic and often diagnosed during work-up of other, unrelated, disease states. Diagnosis is achieved primarily on the basis of clinical and ultrasound examinations (Naydeck *et al.* 1999; Karkos *et al.* 2000). AAA can be classified by four definitions;

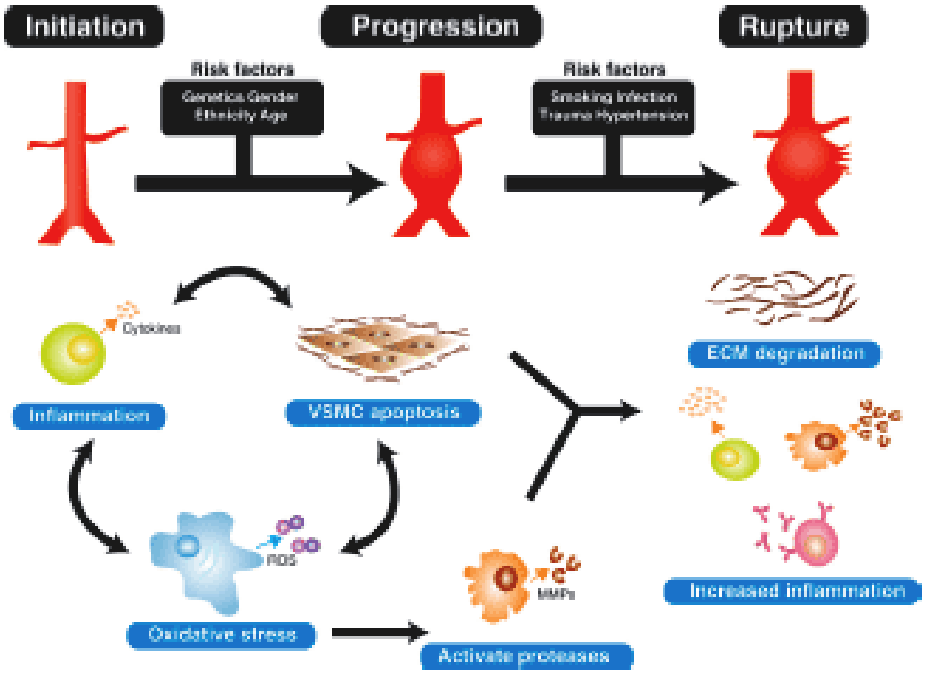
1. An aortic size of 1.5 times larger than the expected diameter of 20 mm.
2. An aortic diameter of over two standard deviations (SD) from the expected measurement of 20 mm.
3. An aorta measuring >30 mm (Wanhainen 2008, Karkos *et al.* 2000).
4. A ratio of 1.1 between the size of the infrarenal and super renal aorta (Wanhainen 2008)

Aortic screening has been implemented in some areas of the UK, as aneurysm rupture is fatal in 80-90% of cases due to haemorrhage (Grootenboer *et al.* 2003; Ogata *et al.* 2005). However, screening provides only a "snap-shot" of the vessels at one time-point and does not indicate who may be at risk of developing AAA in the future. The ability to predict those who are at risk from AAA could reduce mortality significantly (Ogata *et al.* 2005).

Though an aorta expanded to >30 mm is clinically relevant, surgical intervention is only recommended once the aneurysm diameter has increased to 55 mm or when symptoms, such as back pain, are reported, as this is the point at which aneurysm rupture is common (Wanhainen 2008; Harthun 2007). The most common treatment of AAA involves passing a stent-like structure from the arteries in the groin to the aneurysm, where it is released and serves to bypass the aneurysm. Alternatively, an

open aneurysm repair can be performed. This involves an incision in the abdomen and has increased mortality rates (Ogata *et al.* 2005). Open repairs, even with associated risks, are more commonly performed in women (Harthun 2007) due to the shorter and more diseased segments of the non-aneurysmal aorta proximal to the aneurysm that make stent procedures impossible (Harthun 2007). Mortality during both the endovascular procedure and open surgery may occur due to CAD or because of complications such as haemorrhage, lower extremity or intestinal ischaemia, renal failure, respiratory failure or stroke (Harthun 2007). Mortality in stent and open surgery is higher in women, possibly due to the increased average age of diagnosis and treatment (Ogata *et al.* 2005; Harthun 2007; Mehta *et al.* 2010).

Figure 1.3 Pathogenesis of AAA



(Boddy *et al.* 2008)

1.3.3 Causes of AAA

AAA may occur due to a variety of causes including connective tissue disorders, trauma, infection, inflammatory disorders, heritable factors or non-specific degenerative disease. The various aetiologies of AAA are briefly described below. However the primary focus of this work is based on AAA that results from non-specific and genetic causes; all other aetiologies (congenital, connective tissue disorders, trauma, infection and inflammatory disorders) are outside the scope of this work.

1.3.3.1 Congenital AAA

True AAA is extremely uncommon in infants and rarely documented (Mehall *et al.* 2001). Most occurrences of this disease in the very young can be ascribed to trauma or infection (Bell *et al.* 2003).

1.3.3.2 Connective tissue disorders

Mutations in fibrillin 1 gene (*FBN1*), transforming growth factor beta receptor I gene (*TGFB1*) and collagen type III alpha 1 gene (*COL3A1*) all cause monogenic forms of AAA (Marfan syndrome (OMIM 154700), Loeys-Dietz (OMIM 613795) and Ehlers-Danlos syndrome (OMIM 130050, respectively) (Mizuguchi *et al.* 2004; Bown *et al.* 2007). Mutations in collagen genes (causing reduced type III collagen fibres) and type II procollagen (*COL2A1*) were reported as causative in a subset of families with high AAA prevalence (Kuivaniemi *et al.* 2003).

1.3.3.3 Trauma

Trauma to the aorta, generally resulting from aortic bypass surgery, can cause false aneurysms, i.e. where the aneurysm does not involve all three layers of the aorta (Kronzon 1997).

1.3.3.4 Infection

Infective AAA is most commonly found in those with suppressed immune systems or following intravascular procedures (Gomes *et al.* 1992). The advent of antibiotics have reduced infective AAA from 3.4% of all AAA before 1960 to 1.3% today (Parkhurst & Decker 1955) through reduction in the prevalence of *Streptococcus, Staphylococcus aureus, Salmonella sp. Chlamydia pneumonia, Treponema pallidum* and tuberculosis (Juvonen *et al.* 1997; Choudhary *et al.* 2001; Miller *et al.* 2004).

1.3.3.5 Inflammatory disorders

Inflammatory AAA (IAAA) accounts for up to 10% of all AAA, though controversy remains over whether this form of AAA is truly a distinct disease. IAAA has been defined by a thickened aneurysm wall, extensive perianeurysmal and retroperitoneal fibrosis, and dense adhesion to abdominal organs (Walker *et al*. 1972; Rasmussen *et al.* 1997; Railo *et al.* 2005). Inflammatory disorders causing AAA include polyarteritis nodosa, and Kawasaki and Takayasu's arteritis.

1.3.3.6 Familial factors

AAA pathogenesis is known to have a genetic component and this is discussed in detail in *1.3.5 Genetics of AAA.*

1.3.4 Pathogenesis of AAA

Though the pathogenesis of AAA remains unclear, many proteins, genes and environmental factors, ranging from the balance between matrix metalloproteinases (MMPs) and tissue inhibitor of metalloproteinases (TIMPs) (Bown *et al.* 2007), to male sex, have been implicated (Daugherty & Cassis 2004; Ogata *et al.* 2005; Pearce & Shively 2006).

AAA is characterised by the localised destruction and remodelling of aortic extracellular matrix. The natural history of AAA can be split into three distinct phases: formation, growth and rupture (Grootenboer *et al.* 2003; Daugherty & Cassis 2004). Leukocytes recruited to areas of arterial wall damage mark the initiation of AAA, wherein smooth muscle and elastin are degraded due to the action of MMPs 2, 9 and 12 released from inflammatory cells (Daugherty & Cassis 2004). Subsequently, changes in the ratio of MMPs:TIMPS occur due to increasing numbers of inflammatory cells and degraded elastin, which together cause further arterial wall damage (Pearce & Shively 2006). Increased shear stress and vascular pressure are the hallmarks of AAA formation, and result from degradation of medial and adventitial elastin causing arterial stiffness and loss of pulsatile strength (Naydeck *et al.* 1999). Growth of the aneurysm is characterised by arterial stiffness (causing further arterial damage) and subsequent disorganised collagen turnover resulting from a transfer of stress to the aortic collagen (collagen regimentation) (Grootenboer *et al.* 2003; Bown *et al.* 2007; Sandford *et al.* 2007). Increased turnover of collagens I and III, due to the action of MMPs 1, 8 and 12,

causes reduced structural integrity of the aortic wall and generalised wall weakening (Sandford *et al.* 2007; Bown *et al.* 2007). Continued thinning of the media from the loss of collagen, elastin and neovascularisation leads to aortic wall dilation, expansion, and finally, rupture (Grootenboer *et al.* 2003; Pearce & Shively 2006; Bown *et al.* 2007; Sandford *et al.* 2007).

Research into AAA pathogenesis has revealed a significant role for the immune system, although the precise mechanism remains unclear. Both pro- and anti-inflammatory cytokines, including interleukins (IL) 1β, 1, 6, 8, 10, tumour necrosis factor alpha (TNFα), IL 8 receptor and chemokine (C-X-C motif) receptor 2 (CXCR2) have been identified at high levels within the AAA lesion (Bown *et al.* 2007; Sandford *et al.* 2007). There are many hypotheses as to the origin and significance of the inflammatory infiltrate. One proposes that AAA is an autoimmune disease that may be activated through molecular mimicry following infection with pathogens that express similar epitopes to those found within aortic tissue (Daugherty & Cassis 2004; Bown *et al.* 2007). Bystander activation, whereby degradation of the aortic matrix exposes aortic antigens, has also been contemplated (Daugherty & Cassis 2004). Opposing hypotheses place the inflammatory infiltrate as a symptom, rather than cause, of AAA (Daugherty & Cassis 2004). It has also been suggested that AAA may represent a subclinical form of either Marfan syndrome or Ehlers-Danlos syndrome (Bown *et al.* 2007). The most widely accepted hypothesis is that AAA results from wall weakening due to atherosclerosis (Pearce & Shively 2006). The basis for this assumption is that the aortic wall of the aneurysm is frequently affected by atherosclerosis and many risk factors are shared between atherosclerosis and AAA.

1.3.5 Genetics

AAA shows significant familial clustering, with between 15% and 28.6% of AAA patients reporting a positive family history of disease (Kuivaniemi *et al.* 2003). This was initially demonstrated by Clifton who described a family containing three brothers with AAA (Clifton 1977). The mode of AAA inheritance remains unclear; early segregation analysis has indicated that AAA could be an autosomal dominant or recessive disease (van Vlijmen *et al.* 2002; Kuivaniemi *et al.* 2003). Candidate gene studies, linkage studies and GWASs have all been used in an effort to discover the genetic elements of AAA.

1.3.5.1 Candidate gene studies

Many candidate genes studies have been conducted in search of genetic determinants of AAA, some examples of which are discussed below. Polymorphisms in the regulatory and coding regions of TIMPs and MMPs have been implicated in AAA pathogenesis (Ogata *et al.* 2005). Two SNPs identified in the regulatory regions of *MMP3* and *MMP9* (nt-1612 and nt-1562, respectively) were hypothesised to cause increased transcriptional activity of these genes. Specifically, the C to T substitution at nt-1562 in *MMP9* caused a 1.5-fold increase in promoter activity and was significantly more abundant in those with AAA than controls (Jones *et al.* 2003; Sandford *et al.* 2007). *MMP10* also harbours a SNP associated with male-specific AAA (nt+180) (Ogata *et al.* 2005). Two SNPs in *TIMP1* (nt +434 and rs2070584) have been associated with disease in males without a family history of AAA (Ogata *et al.* 2005). These polymorphisms may act through increasing the action of various MMPs (therefore increasing degradation of the extracellular matrix) and/or through reduced TIMP expression (Sandford *et al.* 2007). However, meta-analyses have shown no association between AAA and polymorphisms

in either *MMP9* or *TIMP3* (Pearce and Shively 2006). The T-allele of a polymorphism (nt+994) in platelet activating factor acetylhydrolase gene (*PAF-AH*) is over-represented in AAA patients, and may cause increased induction of MMPs 1, 2 and 9 from epithelial cells and fibroblasts in Asian populations (Sandford *et al.* 2007).

A 287-bp deletion in *ACE* was associated with high levels of protein and increased risk of AAA independently of BP (Pola *et al.* 2001; Sandford *et al.* 2007). However, this association was not replicated in other studies (Obukofe *et al.* 2010). Both AAA susceptibility and protective variants have been identified in 23 genes controlling the homocysteine and methionine metabolism pathway. This pathway is hypothesised to be involved in AAA since it results in proteins which remodel the arterial wall via the action of various MMPs (Pola *et al.* 2001; Sandford *et al.* 2007). One of the most commonly studied polymorphisms of this pathway, nt+677 in the methylenetetrahydrofolate reductase (*MTHFR*) gene, has been associated with AAA, although this has not been independently confirmed (Jones *et al.* 2005; Sandford *et al.* 2007).

An increase in the size of a short tandem repeat and SNP G894T in endothelial nitric oxide synthase gene (*eNOS*) has been associated with AAA (Pearce and Shively 2006; Sandford *et al.* 2007). It is hypothesised that these variants may cause aortic wall thinning due to changes in vasodilatation and smooth muscle migration and proliferation (Rasmussen *et al.* 2001).

Various human leukocyte antigen (*HLA*) subtypes have been associated with AAA. Specifically, HLA subtypes B1*02, B1*04, HLA-A2 and HLA-B61 are over-represented in patients with AAA, while B1*01, B1*08 and B1*14 are more common in healthy controls (Sugimoto *et al.* 2003; Sandford *et al.* 2007; Giusti *et al.* 2008). The

importance of the HLA system in AAA progression and protection highlights the immunological basis of this disease.

1.3.5.2 Linkage studies

Just one significant linkage study have been completed in search of the genetic basis to AAA. Details can be found in Table 1.5.

Table 1.5 Linkage studies in AAA

Genomic region	Candidate genes	Sample number	Reference
19q13.3	*LRP3,TGF-β1, PCDD5, PEPD*	233 families (653 cases)	Kuivaniemi *et al.* 2006,
4q3	*PDGFRA, SPINK2, BAB1, EDNRA*	233 families (653 cases)	Kuivaniemi *et al.* 2006

LRP3- low density lipoprotein receptor-related protein 3 gene, *TGF-β1*- transforming growth factor beta 1 gene, *PCDD5*- programmed cell death 5 gene, *PEPD*- peptidase D gene, *PDGFRA*- platelet-derived growth factor receptor alpha polypeptide gene, *SPINK2*- serine peptidase inhibitor Kazal type 2 gene, *BAB1*- bric a brac 1 gene, *EDNRA*- endothelin receptor type A gene

1.3.5.3 Genome-wide association studies

Three genomic regions have been associated with AAA through GWASs. A SNP located approximately 200 kb upstream of a gene which encodes a lipid anchored cell adhesion protein, contactin 3 plasmacytoma associated (*CNTNAP1*), was significantly associated with AAA (rs7635818) in 235 men of European ancestry (Elmore *et al.* 2009). rs1333049 resides in a locus initially associated with CAD, but now also associated with AAA. It contains no known coding genes, though recent investigations have identified the large non-coding ribonucleic acid (RNA), *ANRIL*, which is

expressed the walls of AAA (Bown *et al.* 2008). rs133049 showed a moderate association with AAA amongst British individuals (Bown *et al.* 2008). A region within chromosome 9q31 (rs7025486) was significantly associated to AAA in GWASs of 1,292 Icelandic individuals (Gretarsdottir *et al.* 2010). Disabled homolog 2 gene (*DAB2*) and disabled homolog 2 gene interacting protein (*DAB2IP*) are the strongest candidate genes within the locus; both genes are thought to suppress cell survival and proliferation through suppression of phosphatidylinositol-3 kinase Act (PI3K-Act) and RAS pathways which, in turn, may induce apoptosis, (Gretarsdottir *et al.* 2010). Finally, a SNP in the first intron of low-density-lipoprotein receptor-related protein 1 (*LRP1*) (rs1466535) also showed significant, replicated association with AAA (Bown *et al.* 2011).

1.3.6 Prevalence of AAA

Research from cross-sectional, prospective and autopsy studies across various ethnicities have found that the incidence of AAA is 11.3-117.2 per 100,000 person years in men and 2.2-33.9 per 100,000 in women (Castleden & Mercer 1985; Spark *et al.* 2001; Yii 2003; Salem *et al.* 2009; Bergqvist *et al.* 2010; Bloomer *et al.* 2012). 1.9-18.5% of men has an abdominal aortic diameter of >29 mm in industrialised countries (Grootenboer *et al.* 2003; Ogata *et al.* 2005).

Prevalence data shown in Table 1.6 and Figure 1.4 highlight the male disadvantage in AAA. Just 65 AAA were identified in a cross-sectional cohort of 2,756 Brazilian subjects (abdominal aorta >29 mm). However, the male disadvantage in AAA remained; of the 65 AAA patients, just 9 were female (Puech-Leão *et al.* 2004). Conversely, a reversal of the male disadvantage was identified in a Kenyan study; a

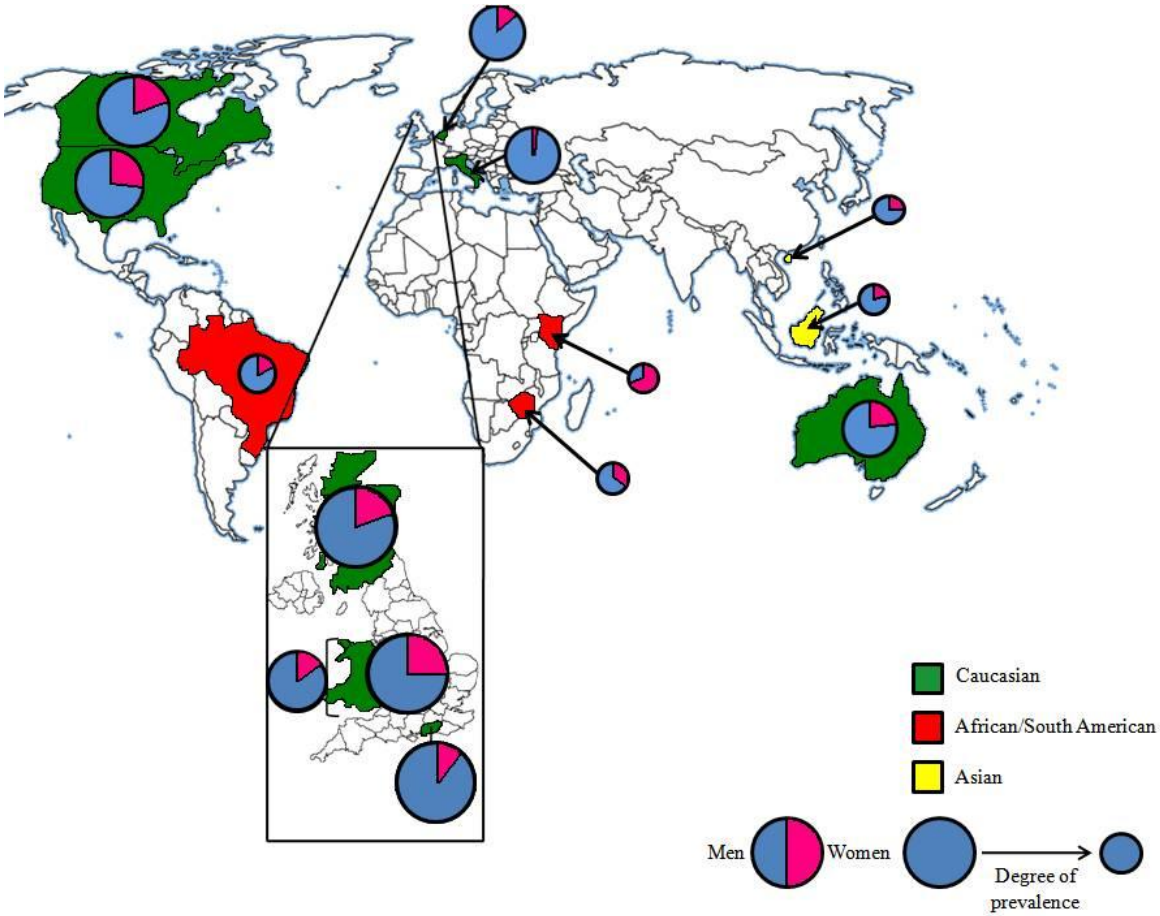
ratio of AAA diagnosis between men and women of 1:1.9 was found (Ogeng'o *et al.* 2010). Indeed, a prevalence of 13.7 per 100,000 person years was found in Hong Kong in 2000 (male to female ratio of 2:1) (Cheng *et al.* 2003). Likewise, a prospective study identified only 120 AAA cases in Hong Kong between 1975 and 1983 (Yii *et al.* 2003), whilst between 1996 and 1999 only 123 AAA diagnoses (aortic diameter >30 mm) were made in Indonesia (Yii *et al.* 2003).

Table 1.6 Total prevalence of AAA worldwide

		Total prevalence		
Location	**Time period**	**Men**	**Women**	**Reference**
Western Australia	1971-1981	377◊	125◊	Castleden *et al.* 1985
USA	1951-1980	45◊	19◊	Melton *et al.* 1984
Borneo	1996-1999	25◊	7◊	Yii *et al*. 2003
Scotland	1971-2000	526◊	144◊	Naylor *et al.* 1988
Worthing	1979-1986	23◊	6◊	Morris *et al.* 1994
Rotterdam	1989-1993	22∞	4∞	Pleumeekers *et al.* 1994
Chichester	1989-1994	8∞	1∞	Scott *et al.* 1995
Japan		3	1∞	Ishikawa *et al.* 1999
Italy	1991-1994	9∞	1∞	Simoni *et al.* 1995
Netherlands	1972-1992	8570*	1399*	Naylor *et al.* 1988; Reitsma *et al.* 1996
England and Wales	1979-1999	59194*	20301*	Filipovic *et al.* 2005

◊Total prevalence per 100,000 person years, ∞ percentage total prevalence, *total deaths due to AAA.

Figure 1.4 Worldwide prevalence of abdominal aortic aneurysms (Bloomer *et al.* 2012).



1.3.7 Sexual dimorphism in risk factors of AAA

In accordance with the sexually dimorphic nature of AAA, certain risk factors of this disease are also affected by sex. Risk factors not affected by sex include chronic obstructive pulmonary disease (COPD) and diabetes.

1.3.7.1 CAD as a risk factor for AAA

Given that the co-existence of CAD and AAA is common, it is not surprising that AAA is commonly considered to be a clinical manifestation of CVD (Madaric *et al.* 2005). In a cohort of individuals aged between 65-75 years, 58% were diagnosed with both CAD and AAA (Wanhainen *et al.* 2005). Madaric *et al.* identified the same over-representation of CAD and AAA co-existence (Madaric *et al.* 2005), and found that 14.7% of those with CAD had AAA while only 3.3% without CAD were diagnosed with AAA (Madaric *et al.* 2005). A review of 2,238 peer-reviewed articles reported that MI increased the risk of AAA by 2.3-fold (Cornuz *et al.* 2004), while Wilmink & Quick assigned a relative risk score of 2.2 to CAD (Wilmink & Quick 1998). These data suggest that the presence of CAD is a risk factor for AAA development, predominantly in men where the prevalence of both diseases is higher (Wilmink & Quick 1998; Bengtsson *et al.* 1991; Blanchard *et al.* 2000; Makomaski-Illing & Kaiserman 2004; Karp *et al.* 2005). However, this association could be also explained by the overlap in risk factors between CAD and AAA.

1.3.7.2 Peripheral vascular disease (*PVD*)

PVD is associated with an odds ratio of 2.5 and relative risk of 2.4 for developing AAA (Cheavechai *et al.* 2004). Galland *et al.* found that 14% of those with occlusive PVD were diagnosed with AAA (Galland *et al.* 1991), the majority of whom were male (Shapria *et al.* 1990). Likewise, 5.9% of patients with atherosclerotic PVD were found to have AAA (Galland *et al.* 1991). Indeed, in a cohort of 1,166 patients with peripheral arterial disease (PAD), 26% also had AAA (Galland *et al.* 1991). The co-existence of PAD and AAA was significantly higher in males (13.6%) compared with females (4.1%) (Barba *et al.* 2005).

1.3.7.3 Smoking level in AAA

Smoking is a well-documented risk factor in AAA development (Karkos *et al.* 2000; Thompson *et al.* 2006; Harthun 2007; Sandford *et al.* 2007). As previously stated, significantly more men than women smoke; of >1.1 billion smokers world-wide, the vast majority are male (Teo *et al.* 2006). Smoking is significantly over-represented in patients with AAA when compared to healthy controls (30.6% and 17.6%, respectively [Blanchard *et al.* 2000]). The prevalence of smoking in both men and women with AAA is significantly higher than amongst those who do not have AAA (9.8% of smokers had AAA while just 1.7% of non-smokers were diagnosed with AAA [Lederle *et al.* 1997]). The effect of smoking on the risk of AAA appears to be sex-dependent: only 2.2% of female smokers were diagnosed with AAA compared to 7.8% of male smokers (Lederle *et al.* 1997; Singh *et al.* 2001). Although active and history of smoking are strong risk factors for AAA in men than women, men are at higher risk of AAA; the odds of AAA are 7.4 and 5.8 for male and female smokers, respectively (Bass *et al.* 1993).

1.3.7.4 Lipid level in AAA

A meta-analysis of all investigations on HDL-C and predisposition to AAA found that individuals with AAA had significantly lower HDL-C levels (approximately 0.15 mmol/L lower) than healthy controls (Takagi *et al.* 2010). Men without AAA diagnosis have higher HDL-C levels compared with those who were diagnosed with AAA diagnosis (1.3 mmol/L vs. 1.2 mmol/L) (Yusuf *et al.* 2004). Increased LDL-C and TC are also considered to be risk factors for AAA (Wanhainen *et al.* 2005). Both male and female AAA patients have significantly higher average LDL-C than controls (Wanhainen *et al.* 2005). Higher average TC was also identified in patients with AAA than controls (Wanhainen *et al.* 2005).

1.3.7.5 Blood pressure level in AAA

High BP is a long disputed risk factor of CVD and AAA (Karkos *et al.* 2000; Thompson *et al.* 2006; Harthun 2007; Sandford *et al.* 2007). Hypertension has been implicated in pathogenesis of AAA, growth rate and the risk of rupture (Sharratt *et al.* 2001; Paradis *et al.* 2004; Genovesi *et al.* 2005; Takagi *et al.* 2010). Elevated diastolic blood pressure (DBP) (>90 mmHg) was associated with a 1.5-fold increase in risk of AAA in a cohort of 5,356 individuals (4.8% and 3.3% in AAA patients and normal controls, respectively, had a DBP >90 mmHg [Wilmink & Quick 1998; Karp *et al.* 2005]). It was estimated that every 10 mmHg increase in DBP and a 2.3 mmHg increase in systolic blood pressure (SBP) were associated with 1.9- and 1.15-fold increase in odds of AAA, respectively. Interestingly, women are more at risk when SBP is raised (OR of AAA was calculated at of 1.39 in women and 0.97 in men) (Singh *et al.* 2001). Elevated DBP increases the risk of AAA equally in both sexes (Singh *et al.* 2001).

1.4 Causes of sexual dimorphism in CAD and AAA

Differences in lifestyle, the immune system, sex-specific actions of sex hormones, sex chromosomes and other genetic effects are all possible drivers of the "male disadvantage" in CAD and AAA.

1.4.1 Lifestyle differences between the genders

Lifestyle choices differ between the sexes. Smoking is more common in men at all ages compared to women in the general population, and the male predominance in smoking is even more evident in AAA (Mckee *et al.* 1998; Genovesi *et al*. 2005; McQueen *et al.* 2008). Bengtsson *et al.* found that patients with AAA smoked significantly more than individuals without AAA (55 g of tobacco in cases and 27 g in controls per person per week). A similar phenomenon is observed in CAD whereby 67% of CAD (compared with 29% of healthy controls) smoked (Genest *et al.* 1991). Though more men smoke and more male smokers are diagnosed with CAD, smoking poses a higher risk of disease in women (female to male relative risk ratio of 1.25) (Huxley & Woodward 2011).

Likewise, men consume more alcohol more regularly than women in the general population (29.3% of men consume alcohol compared to just 9% of women (Yusuf *et al.* 2004). Heavy drinkers, predominantly males, are more likely to develop AAA (hazard ratio of 1.21) (Wong *et al.* 2007). Men were found to be significantly more overweight and obese in a cross-sectional study of 16,657 Polish individuals (48% of men and 39.2% of women were overweight while 39.9% of men and 31.0% of women were obese) (Kaess *et al.* 2010). Galland *et al.* found that body mass index (BMI) was higher in men with AAA than normal controls (OR of 1.14) (Galland *et al.* 1991). Lack

of physical activity is an important risk factor in both CAD and AAA; up to 36% of Americans take no physical activity (Roger *et al*. 2012) and women are less active during leisure time than males (18.2% of men are active compared to 13.3% of women [Yusuf *et al.* 2004]). However, inactivity in males was significantly associated with AAA; just 57.4% of male cases were active compared to 66% of controls (Galland *et al.* 1991). Similarly, men expending >4200 kJ/week reduced their risk of CAD by up to 20%, and less active men have higher incidence of CAD (Sesso *et al.* 2000).

1.4.2 Aggression as sexually dimorphic cause of CAD and AAA

Men are widely accepted to exhibit significantly higher rates of Type A behaviour pattern including aggressiveness, competitiveness, hostility and time urgency, than women (Barrett-Connor 1997). Indeed, being competitive and aggressive is a central component of the masculine personality characteristics and is atypical in women (Barrett-Connor 1997). Levels of aggression are known to be stable over time and both prospective and cross-sectional studies have shown that aggression levels are associated with CVD including CAD (Kop 1999; Rozanski *et al.* 1999; Krantz & McCeney 2002; Smith & Ruiz 2002; Gallo & Matthews 2003; Suls & Bunde 2005). Interestingly, traditional masculinity is thought to increase risk of CVD (Barrett-Connor 1997, Bem 1974, Helgeson 1990). Boyle *et al.* investigated the relationship between CAD risk factors and anger, anxiety and hostility. Both hostility and anger - recognised dimensions of aggressiveness (Helgeson 1990; Boyle *et al.* 2006) - were correlated with cigarettes smoked each day. Anxiety and hostility were associated with diabetes and hostility was related to a decrease in HDL-C (Boyle *et al.* 2006).

1.4.3 Sexual dimorphism of the immune system in AAA and CAD development

Differences in the regulation of the immune system between the sexes are apparent from the first year of life (Tollerud *et al.* 1990; Gourdy *et al.* 2005; Sinha *et al.* 2006) and become increasingly evident at sexual maturity (Gourdy *et al.* 2005). Tollerud *et al.* found that while human leukocyte antigen-D related (HLA-DR) cells are increased in young healthy adolescent males (17-19 years), CD4+ cells and the CD4+:CD8+ ratio are higher in age-matched females (Tollerud *et al.* 1990). It was hypothesised that these fundamental changes in cellular and adaptive immunity make males more likely to develop AAA and CAD since proteins which regulate immune cell migration, cytokine production, growth factors and chronic inflammation are known to be drivers of both diseases (Daugherty & Cassis 2004; Ogata *et al.* 2005).

Sexual disparity in the functioning of the immune system may, in part, explain why male apoE-/- mice develop angiotensin II (AngII)-induced AAA more readily than female mice (Mittendorfer 2005). Female mice show increased T-cell proliferative responses and increased interferon gamma (IFNγ) levels compared to male mice (Gourdy *et al.* 2005). Likewise, Sinha *et al.* found that in tissue from female rat tissue IL- 1, 2, 3, 5, 7, 8α, 8β, 11 and 12 showed a significant (up to 6.3-fold) reduction in expression when compared to male rat tissue (Sinha *et al.* 2006). The same trend was also observed in the TNFα, vascular endothelial growth factor A and C-C chemokine ligand and receptor families (Sinha *et al.* 2006). Similarly, an overall reduction in neutrophil and macrophage abundance was observed in female AAA tissue (Sinha *et al.* 2006). The ability of macrophages to absorb oxidised lipoproteins (through type A scavenger receptors) leading to the formation of foam cells was also significantly higher in males. The formation of foam cells is integral in the aetiology of CAD and the male preponderance to foam cell development due to differences in immune cells may

underlie some of the sexual dimorphism observed in CAD (Wu & von Eckardstein 2003).

1.4.4 Sexual dimorphism in androgen levels in predisposition to AAA and CAD

Androgens (such as testosterone) circulate at significantly higher levels in men (Barrett-Connor 1997), and have been hypothesised to contribute to development of CAD and AAA. Testosterone is known to affect cellular maintenance of vascular smooth muscle and extra-cellular matrix, which when altered may lead to AAA and CAD (Henriques *et al.* 2008). Testosterone is also heavily implicated in CVD risk factors. Elevated testosterone is known to increase HDL-C and reduce triglycerides, TC and LDL-C (Wu & von Eckardstein 2003). Thus, changes in testosterone levels may affect both BP and lipid levels, which indirectly contribute to increased inflammation, possibly leading to both CAD and AAA (Henriques *et al.* 2008; Webb & Collins 2010).

Male testosterone levels have been strongly associated with AAA in animal models, though lowered testosterone has also been implicated in human studies. It is hypothesised that androgens may drive (along with sexually dimorphic immune system functioning) the increased development of AAA in male mice compared to female when treated with AngII (Mittendorfer 2005). The androgen theory is based on the reduced development of AAA in AngII-treated male mice after orchidectomy (which causes a significant reduction in circulating androgens) (Mittendorfer 2005). Henriques *et al.* found that androgens increased expression of every constituent of the renin-angiotensin system which, in turn, increased the prevalence of AAA in this murine model (Henriques *et al.* 2008; Wu *et al.* 2009).

However, human studies do not show a clear association between AAA and testosterone levels. A study completed by Yeap *et al.* on 3,620 elderly men showed no correlation between free testosterone and increased risk of AAA; in fact the risk was reduced with increasing free testosterone levels (Yeap *et al.* 2010). Both total testosterone and free testosterone levels were significantly higher in controls than in patients with AAA (Yeap *et al.* 2010). However, free testosterone and sex hormone binding-globulin (SHBG) did show an association with aortic diameter when aortic diameter was treated as a continuous trait (Wu *et al.* 2009). This highlights a possible role for testosterone depletion in human AAA development rather than the reverse association identified in animal models (Mittendorfer 2005; Henriques *et al.* 2008).

A similarly complex relationship is also evident between androgens and CAD (Webb & Collins 2010). By 2003, 32 cross-sectional studies had investigated the association between testosterone and CAD. 16 studies found that reduced testosterone increased CAD (Barth *et al.* 1983; Mendoza *et al.* 1983; Breier *et al.* 1985; Hromadova *et al.* 1985; Aksut *et al.* 1986; Swedarsen *et al.* 1986; Chute *et al.* 1987; Hämäläinen *et al.* 1987; Lichtenstein *et al.* 1987; Swartz *et al.* 1987; Sewdarsen *et al.* 1988; Sewdarsen *et al.* 1990; Rice *et al.* 1993; Phillips *et al.* 1994; Zhao & Li 1998; English *et al.* 2000) and 16 found no association (Labropoulos *et al.* 1982; Luria *et al.* 1982; Zumoff *et al.* 1982; Heller *et al.* 1983; Phillips *et al.* 1983; Small *et al.* 1985; Franzen & Fex 1986; Baumann *et al.* 1988; Slowinska-Srzendnicka *et al.* 1988; Cengiz *et al.* 1991; Hauner *et al.* 1991; Mitchell *et al.* 1994; Marquez-Vidal *et al.* 1995; Feldman *et al.* 1998; Kabakci *et al.* 1999; Schuler-Luttmann *et al.* 2000). Similarly, no association was identified in any of the prospective studies completed (Cauley *et al.* 1987; Barrett-Connor & Khaw 1988; Phillips *et al.* 1988; Hautenen *et al.* 1994; Harman *et al.* 2001). There is an inconsistent and unconvincing body of evidence effect on the association between

dehydroepiandrosterone sulfate (DHEA-S) (a weaker androgen than testosterone) and CAD. It can be concluded that that though reduced testosterone may have a small effect on predisposition to CAD, the action of testosterone and DHEA-S probably do not account for the majority of sexual dimorphism in CAD.

1.4.5 Oestrogens

It has repeatedly been hypothesised that oestrogens may offer a protective influence against CVD in females. This association has been investigated extensively in animal models of both AAA and CAD. Ailawadi *et al.* showed that aortic transplantation between two male or female rats does not affect the rate of AAA development. However, a female aorta transplanted into a male rat will develop AAA at the same rate as male, rather than female, rats. This suggests that protection of the aorta given by the female sex was lost (Thompson *et al.* 2009), indicating the possible role of female sex hormones. Ovariectomised female rats had a significantly higher AAA growth rate than control animals (Webb & Collins 2010). Likewise, rats injected with oestradiol showed a significantly lower AAA growth rate than controls (Ailawadi *et al.* 2004). Although endogenous oestrogens appear to offer protection in cross-sectional studies of CAD (Tolbert & Oparil 2001), no investigations have been completed into their role in human AAA development.

It is hypothesised that oestrogens may act through the reduction of MMP 9 (a collagenase for collagen type IV) production from macrophages, thereby reducing the collagen destruction, smooth muscle rearrangement and the inflammatory infiltrate seen in AAA (Grootenboer *et al.* 2003; Ailawadi *et al.* 2004). Similarly, oestrogens alter the action of oxidative stress and inflammation. Specifically, oestrogens reduce the action of nuclear factor kappa-light-chain-enhancer of activated B cells (*NFκb*), oxidative

stress and p38 mitogen-activated protein kinase (*MAPK*) (which regulates the action of *TNFα, IL 1, 4, 6 and 8* (Tollerud *et al.* 1990), and activate cardio-protective genes such as peroxisome proliferator-activated receptor gamma co-activator 1-alpha gene (*PGC-1α*), modulatory calcineurin interacting protein 1 gene (*MCIp1*) and heat shock protein 72 gene (*HSP72*) (Bhupathy *et al.* 2010). Though few studies have been completed in humans, it can be hypothesised that the same reduction in MMP 9 production and protection of the aorta due to oestrogens may partially explain why women are at less risk from AAA (Thompson *et al.* 2009). Levels of endogenous oestrogens may affect risk factors for both CAD and AAA i.e. high endogenous levels of oestradiol and estrone in healthy young men are identified as a negative risk factor of lipid profile (Tomaszewski *et al.* 2008). Specifically, oestradiol was associated with decreased HDL-C, while estrone was positively correlated with LDL-C (Tomaszewski *et al.* 2008).

1.4.6 Sexual dimorphism in cardiovascular risk factors

The cumulative effect of risk factors may explain the male predominance in diagnosis of AAA and CAD. Though each of the risk factors outlined here is sexually dimorphic, no one factor explains the extent of male disadvantage in AAA or CAD. It is plausible that the relationship between each risk factor, i.e. the cumulative effect of increased smoking, BP, testosterone and altered lipid profile on male physiology, may predispose men to AAA or CAD. For AAA, this was shown in a cohort of 504 individuals (248 men and 256 women) where 90% of individuals diagnosed with AAA had at least one of the following risk factors: history of atherosclerotic disease, history of smoking or family history of AAA (Wanhainen *et al.* 2005). The risk of AAA increased with number of risk factors present; just 4% of AAA patients had none of the risk factors,

12% had one risk factor, 22% had two of the risk factors and 57% had all three risk factors (Wanhainen *et al.* 2005). A similar phenomenon was proposed in CAD (Genest *et al.* 1991). These observations show that increased number of risk factors in an individual may increase risk of AAA and CAD (Wanhainen *et al.* 2005). Although just three risk factors were investigated, it can be hypothesised that other sexually dimorphic risk factors, such as hypertension and hyperlipidaemia may also act in a collective manner.

1.4.7 Genetics and epigenetics

Various aspects of genetics are sexually dimorphic, ranging from epigenetic to gene expression differences. Males are known to have higher levels of methylation at long interspersed nuclear element 1 (LINE-1) sites (El-Maarri *et al.* 2011) and changes in epigenetics are known to predispose towards CVD (Gluckman *et al.* 2009; Handy *et al.* 2011). There are also significant differences in gene expression profiles between the sexes. Of particular relevance are the differences observed in heart tissue (Isensee *et al.* 2008). However, the most fundamental difference between the sexes is the complement of sex chromosomes (XY in males and XX in females). This major genetic difference will be explored in detail in this thesis.

1.5 The Y chromosome

The human X and Y chromosomes (approximately 155,Mb and 60,Mb, in the reference sequence respectively) evolved from a normally recombining autosomal pair. Differentiation of the Y chromosome was initiated due to a mutation giving rise to the

male-specific sex determining region, Y gene (*SRY*) giving it the key dominant role in male sex determination (Sinclair *et al.* 1990). Recombination in this area was then suppressed, probably through segmental inversion. Recruitment of further sex-specific genes extended the area of suppressed genetic recombination to prevent potential loss of function (Graves 1998). Recombination with the X chromosome was eventually suppressed across 95% of the Y chromosome – this part of the Y chromosome is called the male-specific region (MSY) (Skaletsky *et al.* 2003). Normal recombination was restricted to segments found at the extreme ends of the sex chromosomes called pseudoautosomal regions (PARs). In the absence of recombination events across the majority of the Y chromosome, gene conversion (Chen *et al*. 2007) has been observed to occur frequently (Rozen *et al*. 2003). In duplicated regions containing spermatogenic genes, this process may be acting to retard the degeneration of genes by using a non-mutated copy to 'correct' mutations in another copy; the mechanism for discriminating mutated from non-mutated versions is not understood (Rozen *et al*. 2003).

The Y chromosome accounts for less than 2% of the genome (Vogt 2005); the short arm of the Y chromosome (Yp) contains just 8 Mb of DNA while the long arm (Yq) contains 14.5 Mb of euchromatic DNA (Skaletsky *et al.* 2003). The MSY contains highly repetitive structural features including eight large palindromes (inverted repeats) encompassing 5.7 Mb (palindrome 1 [1,450 kb], palindrome 2 [122 kb], palindrome 3 [283 kb], palindrome 4 [190 kb], palindrome 5 [496 kb], palindrome 6 [110 kb], palindrome 7 [8.7 kb], palindrome 8 [6 kb]) (Skaletsky *et al.* 2003). The arms of the palindromes contain unique, non-duplicated spacer sequences at their centre range from 9 kb to 1.45 Mb (Skaletsky *et al.* 2003). Further genetic characteristics include five more widely-spaced inverted repeats (Skaletsky *et al.* 2003). Short tandem repeats (STRs), tandem arrays of up to 30 copies of adjacent repeats of 1-7 bp length are

common and polymorphic on the Y chromosome (Kayser *et al.* 2004), as they are elsewhere in the genome. Many Y chromosome SNPs have been identified, of which ~600 have been used to construct a Y chromosome phylogenetic tree encorporating 311 haplogroups (Karafet *et al.* 2008).

1.5.1 Pseudoautosomal regions 1 and 2

The pseudoautosomal regions of the Y chromosome are located at the extreme ends of Yp and Yq (PAR1 in Yp and PAR2 in Yq). These small regions recombine normally with the X chromosome. Indeed, during male meiosis at least one recombination event must take place within PAR1 (Delbridge & Graves 1999). PAR recombination rates differ greatly with sex; recombination rates in PAR1 and PAR2 in females are analogous to those found on the autosomes (outside of the telomeres and cetromere), whilst the rate in males is significantly higher (Flaquer *et al.* 2008). The PAR regions contain just 29 genes, with 24 genes in PAR1 and 5 genes in PAR2 (Flaquer *et al.* 2008). The PAR regions of the Y chromosome are not studied in this work (which is based on the MSY).

1.5.2 The male-specific region (MSY)

The MSY encompasses approximately 95% of the length of the Y chromosome (positions 2,699,520- 59,169,534 in the reference sequence) (Skaletsky *et al.* 2003) and does not cross over with the X chromosome. There are three large regions of heterochromatin in the MSY stretching across nearly 42 Mb, though are extremely variable in length. Two of these regions are situated at the centromere (approximately 1Mb) and on distal Yp (~40 Mb, though extremely variable among males [Repping *et al*. 2006]); both are constructed from massively amplified tandem repeats of low

sequence complexity. The third block spans 400 kb, is formed from >3,000 tandem repeats of 125 bp, and is situated in a long block of euchromatic sequence on Yq. Four different repetitive sequences make up the heterochromatic sequences, and include; 171 bp alphoid, GGAAT, 125 bp repeats and one non-characterised repeat thought to be AT rich. No part of the heterochromatin is known to be transcribed (Skaletsky *et al.* 2003).

The euchromatic sequence of the Y chromosome spans approximately 23 Mb (8 Mb on Yp, 14.5 Mb on Yq) (Skaletsky *et al.* 2003). These regions contain at least 156 transcriptional units, approximately half of which encode proteins. Of the 78 protein-coding genes, 60 are organised into one of nine gene families, while the remaining 18 are present in single copy (Table 1.7). There are three forms of Y chromosome genes. Firstly, X-transposed genes which show very high sequence homology (~99%) in both exonic and intronic sequences with gene counterparts found on the X chromosome (Skaletsky *et al.* 2003). These genes, containing both pseudogenes and protein-coding genes, were transposed from the long arm of the X chromosome 3-4 million years ago, while a subsequent Y chromosome inversion split this block into two separate locations. Just two protein-coding genes, the testis specific transforming growth factor-induced factor homeobox 2-like, Y-linked (*TGIF2LY*) and brain specific protocadherin 11 Y-linked (*PCDH11Y*), are known to lie in this segment. The remaining X-transposed sequence is composed from high density repeat elements such as LINE-1. Secondly, 27 coding X-degenerate genes and a number of pseudogenes represent surviving relics from the autosomes from which the modern day X and Y chromosomes originated. These genes show 60-95% sequence homology with their X chromosome counterparts. Fourteen genes, including all 12 ubiquitously expressed genes, are thought to produce functional proteins which are similar, but not identical to, those located on the X chromosome (Skaletsky *et al.* 2003). The difference in protein isoforms between those

derived from the X and Y chromosomes is particularly apparent in ribosomal protein S4, Y-linked 1 (*RPS4Y1*) and ribosomal protein S4, Y-linked 2 (*RPS4Y2*), which are full length homologs of the X-linked gene ribosomal protein S4, X-linked (*RPS4X*) (Table 1.7) (Figure 1.5). *CYorf15A* and *CYorf15B* are homologs to the 5´ and 3´ regions of *CXorf15*. A total of 126 SNPs were identified in X transposed and degenerate genes through direct re-sequencing of the reference sequences (mainly haplogroup R1b1b2) (Krausz *et al.* 2004; Rozen *et al.* 2009). Finally, the remaining euchromatic sequences are arranged into 7 ampliconic segments (palindromes) (10.2 Mb) and are identified through their high similarity (99.9%) to other MSY sequences. These sequences contain nine protein-coding families with copy numbers (in the reference sequence) ranging from two (*VCY, XKRY, HSFY, PRY),* 3 (*BPY2*), four (*CDY, DAZ*), six (*RBMY*) and up to

Table 1.7 The list of MSY genes

Full gene name	Symbol (Copy number)	Gene size (bp)	Category	Expression
Variable charge, Y-linked	*VCY* (2)	744	Ampliconic	Testis
TGFβ-induced factor homeobox 2-like, Y-linked	*TGIF2LY* (1)	987	X transposed	Testis
Chromodomain protein, Y-linked	*CDY* (4)	2,786	Ampliconic	Testis
Testis specific protein, Y-linked 1	*TSPY* (35)	2,795	Ampliconic	Testis
Basic charge, Y-linked, 2	*BPY2* (3)	21,203	Ampliconic	Testis
PTPN13-like, Y-linked	*PRY* (2)	24,241	Ampliconic	Testis
RNA binding motif protein, Y-linked	*RBMY* (6)	37,949	Ampliconic	Testis
Heat shock transcription factor, Y-linked 1	*HSFY* (2)	42,293	Ampliconic	Testis
Deleted in azoospermia	*DAZ* (4)	69,740	Ampliconic	Testis
XK, Kell blood group complex subunit-related, Y-linked	*XKRK* (2)	46,372	Ampliconic	Testis
Amelogenin, Y-linked	*AMELY* (1)	8,110	X degenerate	Tooth
Thymosin beta 4, Y-linked	*TMSB4Y* (1)	2,458	X degenerate	Ubiquitous
DEAD (Asp-Glu-Ala-Asp) box polypeptide 3, Y-linked	*DDX3Y* (1)	16,372	X degenerate	Ubiquitous
Eukaryotic translation initiation factor 1A, Y-linked	*EIF1AY* (1)	17,430	X degenerate	Ubiquitous
Chromosome Y open reading frame 15A	*CYorf15A* (1)	23,075	X degenerate	Unknown

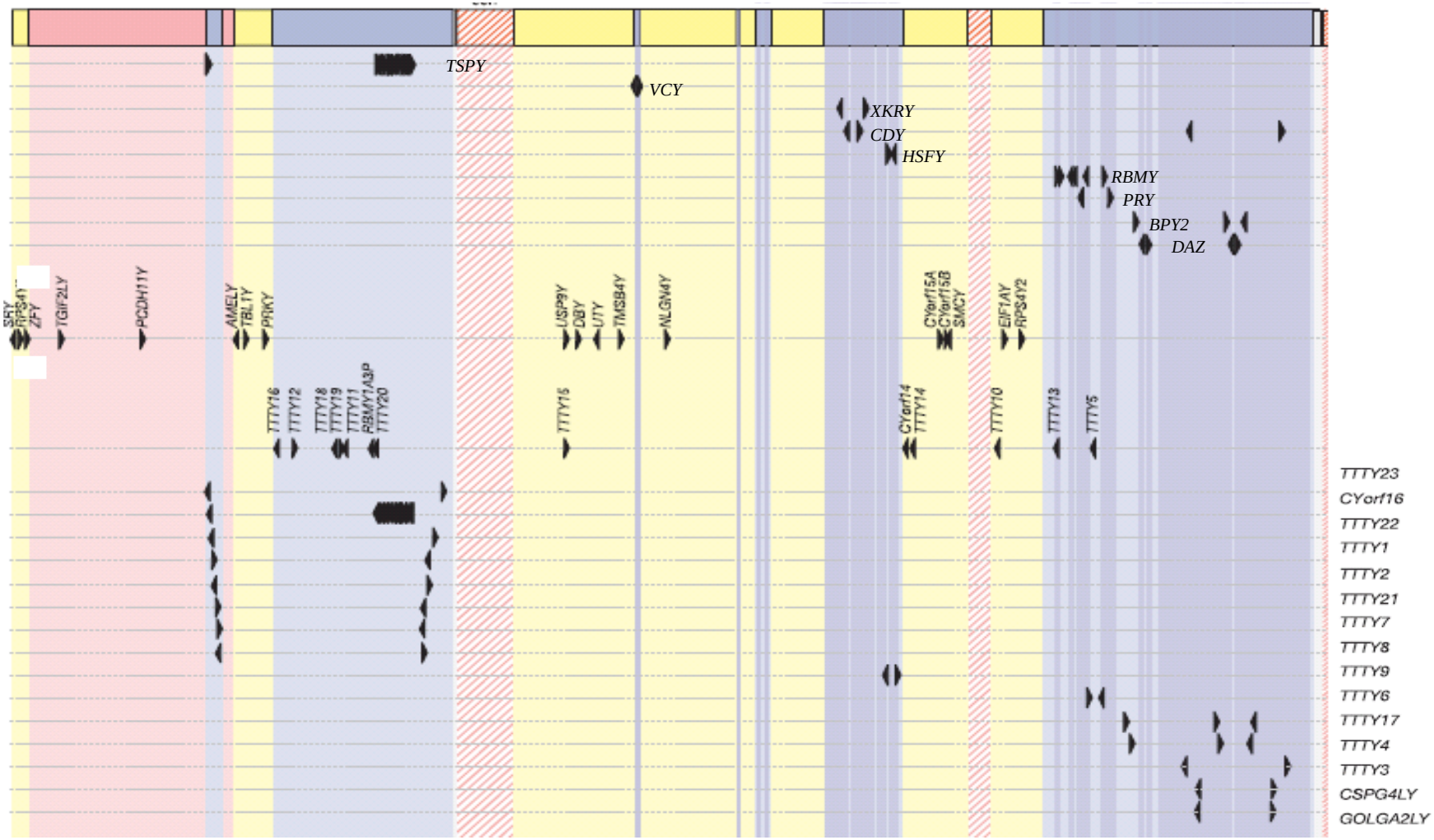
Ribosomal protein S4, Y-linked 2	*RPS4Y2* (1)	24,965	X degenerate	Ubiquitous
Ribosomal protein S4, Y-linked 1	*RPS4Y1* (1)	25,471	X degenerate	Ubiquitous
Lysine (K)-specific demethylase 5D	*KDM5D* (1)	41,074	X degenerate	Ubiquitous
Zinc finger protein, Y-linked	*ZFY* (1)	47,436	X degenerate	Ubiquitous
Protein kinase, Y-linked	*PRKY* (1)	145,838	X degenerate	Ubiquitous
Ubiquitin specific peptidase 9, Y-linked	*USP9Y* (1)	159,609	X degenerate	Ubiquitous
Ubiquitously transcribed tetratricopeptide repeat gene, Y-linked	*UTY* (1)	232,295	X degenerate	Ubiquitous
Sex determining region Y	*SRY* (1)	897	X degenerate	Unknown
Transducin (beta)-like 1, Y-linked	*TBL1Y* (1)	180,998	X degenerate	Unknown
Neuroligin 4, Y-linked	*NLGN4Y* (1)	323,013	X degenerate	Unknown
Chromosome Y open reading frame 15B1	*CYorf15B* (1)	15,522	X degenerate	Unknown
Protocadherin 11 Y-linked	*PCDH11Y* (1)	741,999	X transposed	Brain

Gene copy numbers and sizes are based on the reference sequence (mostly belonging to haplogroup R1b1b2, with a ~1-Mb segment encompassing the *AZFa* region deriving from a haplogroup G chromosome). Gene size includes introns and 3′ and 5′ UTRs. Data from UCSC (http://genome.ucsc.edu/)

35 in *TSPY* (Krausz *et al.* 2004). All nine protein-coding families are expressed purely in the testis. Ampliconic genes, such as *DAZ* which have two gene copies in both palindrome 1 and 2, are mostly found in palindromic regions. Likewise, *CDY* has two

copies in palindrome 1 and 5. One protein-coding gene, *TSPY*, exists in 35 copies in the reference sequence (Skaletsky *et al.* 2003), forming an array of 20.4-kb repeats account for around 700 kb of the ampliconic region; it is polymorphic in copy number (Repping et al. 2006). A number of non-coding gene families of unknown function are found in the five inverted repeats, including testis-specific transcript (*TTTY*) *1*, *2*, *4*, *7*, *8*, *18*, *19*, *21* and *22* (Skaletsky *et al.* 2003).

Figure 1.5 Positions of MSY genes (Adapted from Skaletsky *et al.* 2003)



Top bar represents the Y chromosome from Yp on the left to Yq on the right. *SMCY* is now known as *KDM5D, DBY* is now known as *DDX3Y*.

Gene families: *TSPY, VCY, XKRY, CDY, HSFY, RBMY, PRY, BPY2, DAZ.* Single copy genes: *SRY, RPS4Y1, ZFY, TGF2LY PCDH11Y, AMELY, TBL1Y, PRKY, USP9Y, DDX3Y, UTY1, TMSB4Y, NLGN4Y, CYorf15A, CYorf15B, KDM5D, EIF1AY, RPS4Y2.* Single copy transcription units: *TTTY16, TTTY12, TTTY8, TTTY9, TTTY1, RBMY1A3P, TTTY20, TTTY15, CYorf14, TTTY14, TTTY10, TTTY15.* Transcriptional unit families: *TTTY23, CYorf16, TTTY22, TTTY1, TTTY2, TTTY21, TTTY7, TTTY8, TTTY9, TTTY6, TTTY17, TTTY4, TTTY3, CSPG4LY, GOLGA2LY.*

Coloured blcoks represent: pink- X transposed, yellow- X degenerate, blue- ampliconic, red hash- heterochromatin, white- other.

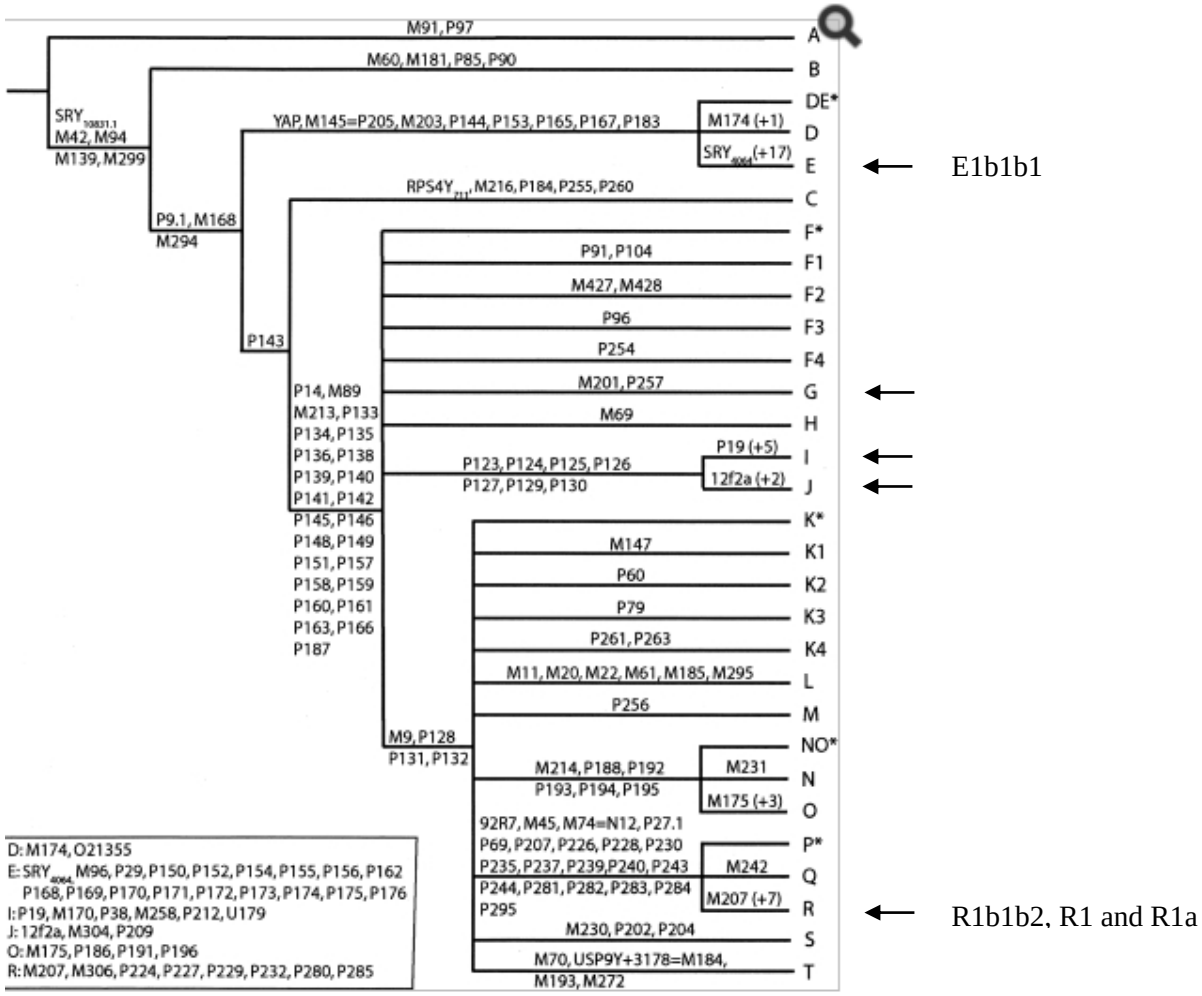
1.5.3 Y chromosome haplogroups

The MSY essentially acts as one block, being inherited by sons from fathers as an indivisible unit. Sequence variants, ranging from SNPs and indels to STRs, in the MSY can be used to reconstruct an evolutionary tree of the Y chromosome where each type of Y chromosome can be traced back to a common ancestor, sometimes referred to as Y-chromosomal Adam (The Y Chromosome Consortium 2002). Most commonly, the Y chromosome evolutionary tree is constructed using specific binary markers (due to their slow back and parallel mutation rates) which define stable Y chromosome 'haplogroups' (The Y Chromosome Consortium 2002; Karafet *et al.* 2008). The most detailed and useful published haplogroup tree uses 599 SNPs and results in 311 Y chromosome haplogroups (Karafet *et al.* 2008). STRs mutate at higher rates and can be used to define very specific haplotypes within a haplogroup (de Knijff 2000). Two current naming systems of Y chromosome haplogroups exist; the "Y Chromosome Consortium" labelled the 18 major haplogroups A to R (The Y Chromosome Consortium 2002) to which haplogroups S and T have now been added (Karafet *et al.* 2008). Sub-groups and halpotypes within each haplogroup were then named numerically, for example E1, E1a and E1a1. The secondary nomenclature identifies each haplogroup via its terminal marker, for example, E-M81 (The Y Chromosome Consortium 2002). All work in this thesis will utilise the naming system established by the Y Chromosome Consortium.

Haplogroup prevalence varies globally; haplogroups A and B are predominantly found in Africa, D and E in Africa and Asia, C in East Asia, Oceania and North America, F and H in Asia, K in East Asia, Q in North East Asia and America, while R is mostly found in Europe and West Asia (Figures 1.5-1.7). Haplogroups I and R are most common in Europe with haplogroup I being found exclusively in Europeans (Rootsi *et*

al. 2004; Karafet *et al.* 2008). Haplogroups K, F and P are superhaplogroups from which other haplogroups are derived. Investigation of twelve specific Y chromosome SNPs (SRY 10831, M35, M89, M201, M170, M304, M9, M45, M207, M173 and M269) can characterise the twelve most common major European haplogroups (BR[xE1b1b1,F], E1b1b1, G, I, J, R1a, and R1b1b2), and will also identify rare lineages within F*, K*, P* and R* (Semino *et al*. 2000; Underhill & Kivisild 2007), and will be used in all research performed in this study.

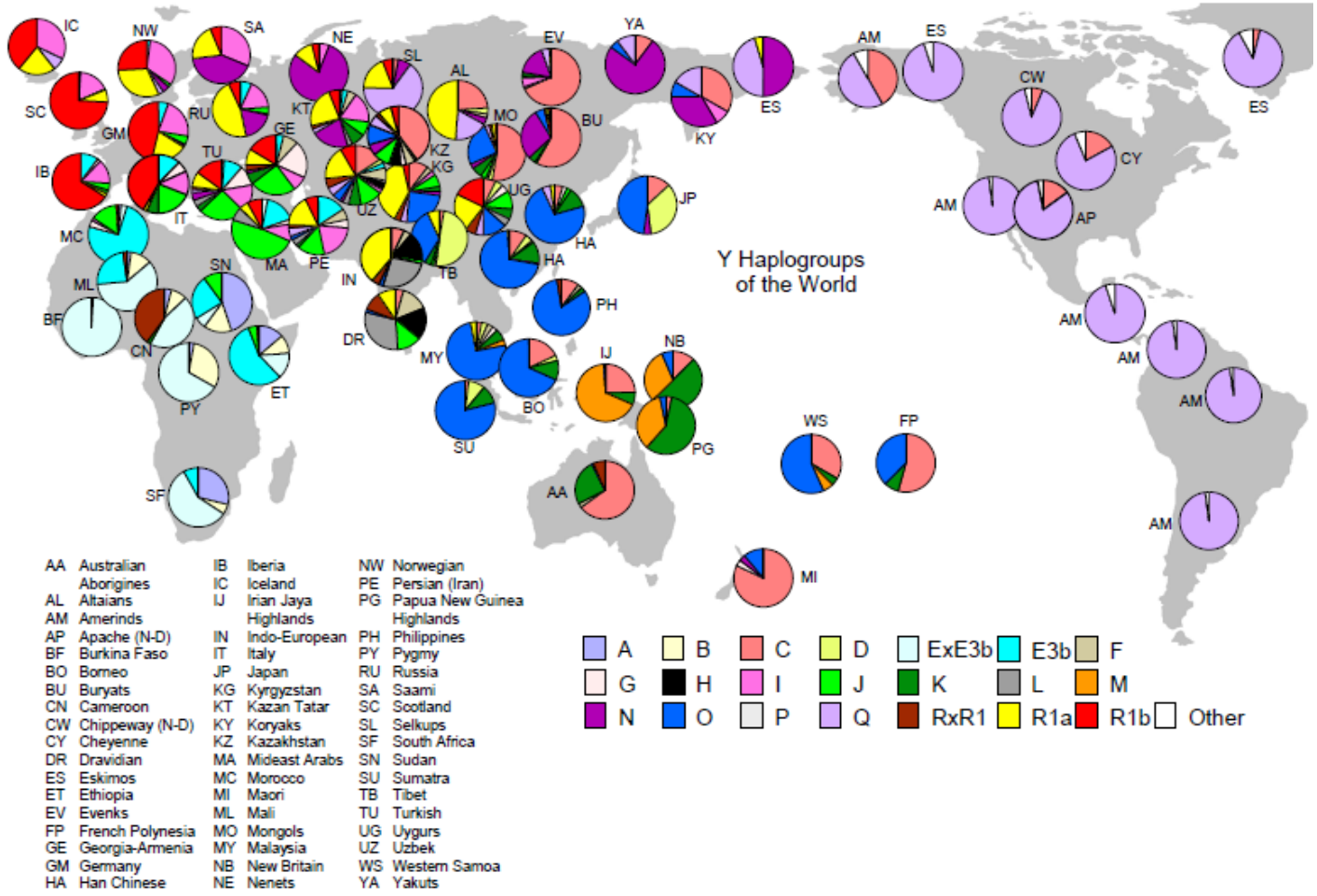
Figure 1.6 Worldwide Y chromosome haplogroup tree (Karafet *et al.* 2008). Lineages found in Europe are highlighted by arrows and specific lineage defined.



E1b1b1

R1b1b2, R1 and R1a

Figure 1.7 Worldwide Y chromosome haplogroup prevalence (http://www.scs.illinois.edu/~mcdonald/WorldHaplogroupsMaps.pdf)*Note: E3b is now known as E1b1b1



1.5.3.1 Common European haplogroups

Typically >50% of European men are designated as belong to haplogroup R, and most commonly to sub-clade R1 (defined by M173) (Myres *et al.* 2011). In Europe, the two most common sub-lineages of R1 are R1a in Eastern Europe (defined by M17) and R1b in Western Europe (defined by M343). The most common R1 sub-haplogroup in Europe is R1b1b2 (defined by M269) and is carried by approximately 110 million men, predominantly Western Europeans (Balaresque *et al.* 2010; Myres *et al.* 2011; Sjödin & François 2011) (Figure 1.8). The method of spread of this young haplogroup may have occurred through either demic diffusion (i.e. the movement and expansion of Near-Eastern farmers) or transmission of cultural innovation, including agriculture, to existing populations that then expanded (Balaresque *et al.* 2010; Myres *et al.* 2011; Sjödin & François 2011). The correlation of microsatellite variance with longitude rather than latitude means the most likely hypothesis is through Neolithic expansion (Balaresque *et al.* 2010), though this has been contested (Busby *et al.* 2012), and the area is controversial.

Approximately 18% of European men belong to haplogroup I (defined by M170), a haplogroup specific to Europe (Rootsi *et al.* 2004). This haplogroup is thought to have originated in Europe sometime before the last glacial maximum (approximately 24,000 years before present) (Semino *et al.* 2000), and at its height had a frequency of 40-50% in Scandinavian and Southern European populations (Rootsi *et al.* 2004). Haplogroup I has a number of sub-clades (I1a, I1b, I1c, I1d, I1e, I1f, I2) (Figure 1.6) which show variable prevalence across Europe. Lineage I1b is most common in France (Rootsi *et al.* 2004) (Figure 1.8). Interestingly, the sister group of haplogroup I, haplogroup J, is relatively rare in Europe (Figure 1.9).

Figure 1.8 Y chromosome haplogroup prevalence in Europe (http://www.scs.illinois.edu/~mcdonald/WorldHaplogroupsMaps.pdf) Note: E3b is now known as E1b1b1

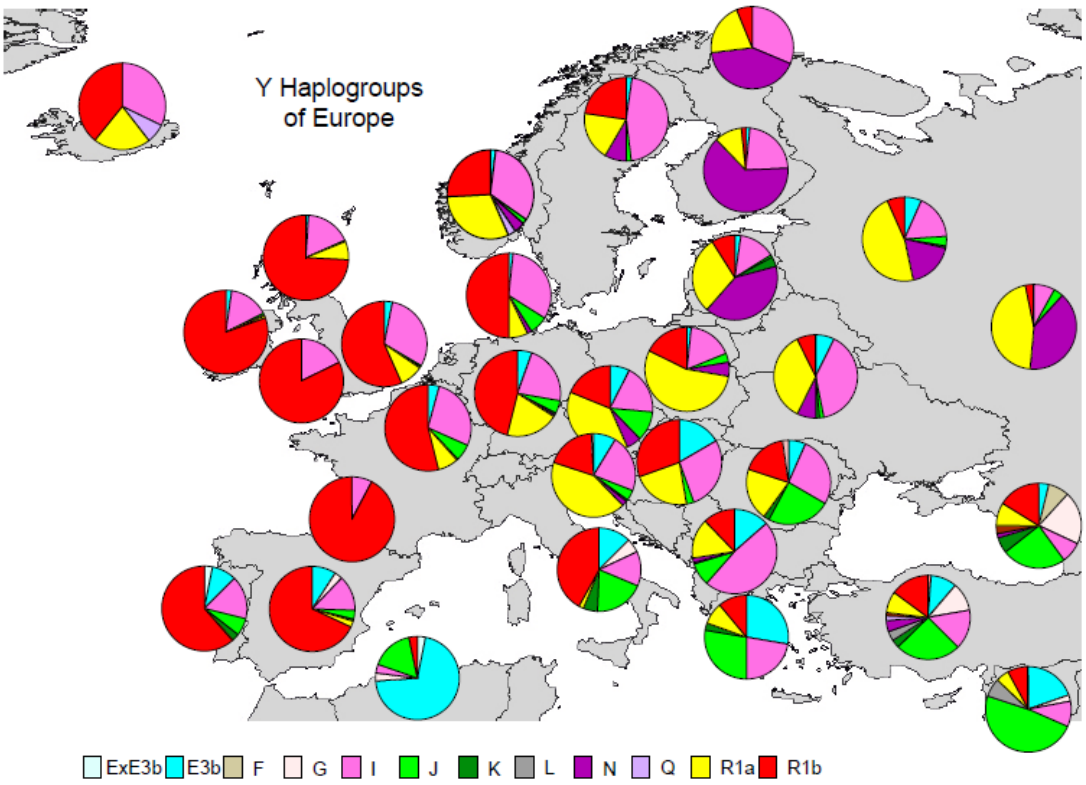
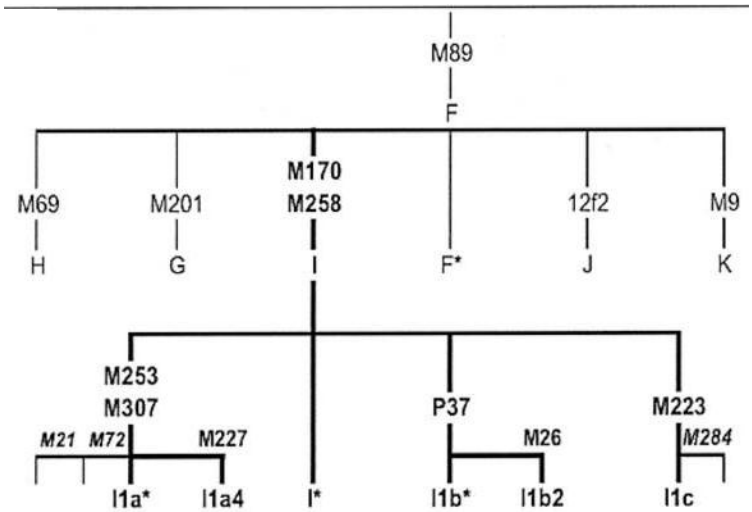


Figure 1.9 Y chromosome haplogroup I phylogeny (Rootsi *et al.* 2004)



1.5.4 The Y chromosome in disease

The Y chromosome has been implicated in a variety of different pathogeneses ranging from hypertension (Charchar *et al.* 2002) to autism (Serajee & Mahbubul Huq 2009). Studies typically used methods such as haplogroup analysis and of specific variants (e.g. the biallelic *Hin*dIII(+/-) restriction fragment length polymorphism).

The *Hin*dIII polymorphism, present in the alphoid repeat (*DYZ3*) near the Y chromosome centromere, indicates the presence or absence of a 6kb repreated higher-order alphoid unit that contains a diagnostic *Hin*dIII restriction site (Mathias *et al.* 1994). *Hin*dIII (+) approximately correlates with haplogroup P of the Y chromosome (Mathias *et al.* 1994). Ellis *et al.* found that men with the *Hin*dIII (+) variant had significantly higher DBP than men with *Hin*dIII (–) (81.2 mmHg vs. 83.2 mmHg) in an Australian population of European origin (Ellis *et al.* 2000). Polish men with *Hin*dIII (+) had significantly higher SBP and DBP (SBP +5.27 mmHg and DBP +3.14 mmHg) then those with the other allele. The same association was observed in Scottish men (Charchar *et al.* 2002). Shankar *et al.* showed that the association between this polymorphism and BP is present early in life - *Hin*dIII (-) was associated with increased BP in young men (Shankar *et al.* 2007). However, Russo *et al*. found no evidence of association with *Hin*dIII (+/-) and BP or hypertension (Russo *et al.* 2006) in three European populations (British, Belgian and Italian). These data highlight the complicated relationship between the Y chromosome and BP.

A similarly complex relationship exists between the *Hin*dIII(+/-) polymorphism and lipid levels. In this case only three studies have been completed; the first demonstrated a relationship between *Hin*dIII(-) and increased levels of TC and LDL-C in two Polish populations (Charchar *et al.* 2003). Furthermore, a protective haplotype encompassing this polymorphism was associated with lipid profile in African men (Russo *et al.* 2008). No association between any measure of serum lipids and the

*Hin*dIII polymorphism was identified in three European populations (British, Belgian and Italian) (Russo *et al.* 2006).

Association between the MSY *Hin*dIII restriction site and BP and lipids suggests that the Y chromosome may play a role in CVD. Haplogroup analysis of MSY is currently viewed as the "gold standard" for Y chromosome association analysis. For example, Sezgin *et al.* found that haplogroup I was associated with the rate of human immunodeficiency virus (HIV) progression in 2,292 European Americans (Sezgin *et al.* 2009). Specifically, haplogroup I caused faster disease progression from HIV to acquired immune deficiency syndrome (AIDS). Males with haplogroup I depleted CD4+ cells more quickly than those with other haplogroups (Sezgin *et al.* 2009). Certain Y chromosome haplogroups have also been found to predispose carriers towards developing autism, indicating that a specific signature present only on the "risk" haplogroup of the Y chromosome underlies at least part of the 4:1 male to female ratio in diagnosis of autism (Serajee & Mahbubul Huq 2009). Arredi *et al.* have shown that haplogroup E predisposes carriers to Y chromosome micro-deletions (b2/b4 deletion), implying that men with haplogroup E are more prone to chromosomal rearrangements (Arredi *et al.* 2007). Significantly more men with haplogroup I1c than other haplogroups were diagnosed with prostate cancer in 4,902 Swedish men (Lindstrom *et al.* 2008). However, a secondary study (also in Swedish men) did not find evidence of association (Lindstrom *et al.* 2008).

1.5.5 The role of the Y chromosome in CVD

Men have a significantly higher prevelence of CVD than age-matched women across a variety of ethnicities. The presence of the Y chromosome exclusively in males underlies the hypothesis that it may be responsible for male predisposition to CVD. Indeed, Y

chromosome polysomy (e.g. 47, XYY) has been associated with increased cardiovascular mortality in British men (Higgins *et al.* 2007). An analysis in the "British Heart Foundation Family Heart Study" (BHF-FHS) has shown that haplogroup I is over-represented in British men with early-onset CAD than controls free from disease (20.5% in cases, 12.6% in controls) and that carriers of haplogroup I had increased risk of CAD (OR 1.75; p= 0.04) (Charchar *et al.* 2012). There was no population stratification in the BHF-FHS - multi-dimensional scaling of approximately 500,000 SNPs from previously conducted GWASs showed no significant differences in genetic background between these with haplogroup I and those with all other haplogroups (Charchar *et al.* 2012). This work was then replicated in a large prospective study "West of Scotland Coronary Prevention Study" (WOSCOPS) (which is less susceptible to confounding factors than cross-sectional studies) - haplogroup I was an independent predictor of age-adjusted CAD (OR 1.45). Importantly, this association was independent of all traditional CVD risk factors, including hypertension, LDL-C, HDL-C, triglycerides, BMI, diabetes, glucose, alcohol consumption, smoking, socioeconomic factors, or loci previously associated with CAD (OR 1.6). Meta-analysis of both cohorts (n~3,000) revealed that haplogroup I increased risk of CAD by approximately 50% (OR 1.55) (Charchar *et al.* 2012). Unfortunately, no power calculation was completed for any aspect of the study (Charchar *et al.* 2012). These data show that the Y chromosome may significantly affect male predisposition to CAD. Further studies into the mechanisms of this association could provide insights into the role of the MSY in health and disease. Conventional CVD risk factors and male-associated/ sex-specific phenotypes such as aggression and sex steroids represent possible drivers of this association. As previously discussed, a complex relationship between the Y chromosome and CVD risk factors such as hypertension and hyperlipidaemia exists.

Thus, it is plausible that carriers of haplogroup I are predisposed to hypertension and/or hyperlipidaemia and therefore are also predisposed to CAD. Aggression is widely accepted as a major component of masculine behaviour and high levels of aggression were associated with increased risk of CVD. Thus, if haplogroup I is associated with increased aggression levels, the carriers of this lineage may be also more prone to CAD. The same logic could apply to androgens. To date the possibility of a Y chromosome effect on CVD risk factors, aggression or sex steroids has not been investigated.

1.5.6 The role of the Y chromosome in AAA

CAD, AAA and stroke are thought to possess similar molecular pathogenesis and aetiology, and all fall under the umbrella term of CVD. This idea is supported by the considerable (though not complete) overlap in risk factors between the diseases. It is plausible that should haplogroup I of the Y chromosome predispose towards CAD, haplogroup I may also predispose towards AAA and stroke. To date the role of the Y chromosome has not been investigated in relation to AAA or stroke.

1.6 Hypothesis

I hypothesise that the association between the MSY and CAD may be, at least in part, explained by a haplogroup I effect on conventional CVD risk factors and/or male-associated/ sex-specific phenotypes, including aggression and sex steroids. Furthermore, I hypothesise that haplogroup I of the Y chromosome is associated with significantly altered expression of genes and pathways in cells relevant to CVD (monocytes, macrophages, and kidney), possibly through changes in one of the ubiquitously expressed MSY X-degenerate genes. Finally, I hypothesise that haplogroup I of the Y chromosome may predispose its carriers not only to CAD, but also to AAA.

1.7 Aims

Specific aims of this thesis include;

1. Analysis of the association between haplogroup I of the Y chromosome and CVD risk factors (BMI, SBP, DBP, TC, HDL-C, LDL-C, glucose and creatinine) (Chapter 3), aggression scores (anger, hostility, physical aggression, verbal aggression and total aggression) and sex steroids (testosterone, androstenedione, DHEA-S, oestrogen and oestradiol) in phenotypically normal young males (Chapter 4).
2. A hypothesis-free transcriptome-wide investigation using RNA from human monocytes and macrophages in search of changes in expression between haplogroup I and other MSY lineages (Chapter 5).
3. Analysis of the differences in ubiquitously expressed MSY gene expression between males with haplogroup I and all other haplogroups in both human macrophage and kidney RNA (Chapter 6).
4. Investigation of the prevalence of haplogroup I in men with AAA and controls free from disease (Chapter 7).

2 General methodologies

2.1 DNA Extraction

Peripheral blood was obtained from all participants by venipuncture. In all cases, DNA was isolated from peripheral blood mononuclear cells. 4.5 ml of red blood cell lysis solution (Qiagen, USA) was added to 4.5 ml of whole blood and inverted ten times before being incubated at room temperature for 10 min. The lysate was then centrifuged at 2000g for 10 min at room temperature. All but 5µl of the supernatant was discarded and the same volume of lysis solution (Qiagen, USA) was added and the pellet dislodged using a pipette tip. This mixture was incubated at 37^{o}C for 1 hr, during which the tube was inverted at 10 min intervals. The solution was allowed to homogenise for between 12 hrs and 72 hrs. 1 ml of protein precipitation solution (Qiagen, USA) was added and vortexed for 20 sec, before being centrifuged at 2000g for 10 min at room temperature. The supernatant was subsequently poured into a tube containing 3 ml of isopropranol (Qiagen, USA), inverted ~50 times and centrifuged at 2000g for 3 min at room temperature. The supernatant was poured away and the tubes left to drain for 10 min. Once dry, 3 ml of 70% (v/v) ethanol was added and tubes were inverted ten times before centrifugation at 2000g for 1 min at room temperature. Again, the supernatant was discarded and the tubes left to drain for 10 min. 250 µl of DNA hydration solution (Gentra Systems, USA) was added once dry, and incubated for 1 hr (65^{o}C) during which the tubes were agitated at 20 min intervals. Finally, the DNA solution was left to cool overnight before being centrifuged at 2000g for 1 min at room temperature and stored at -20^{o}C.

2.1.1 Measurement of DNA quantity and quality

DNA quality and quantity was assessed using a Nanodrop 8-sample spectrophotometer ND-8000 (Fisher Scientific, UK). This technology measures DNA concentration between 5 and 2000 ng/µl using different wavelengths of light projected through the sample (depending on the absorbance of DNA compared to contaminants). A 260/280 ratio of 1.8 and a 260/230 ratio of 1.8-2.2 indicates a pure sample. The Nanodrop was blanked using 2 µl of DNA hydration solution before 1 µl of each sample was loaded for analysis. The concentration and purity of each DNA sample was measured. Those with DNA concentrations <20ng/µl, a 260/280 ratio <1.6 or >2.0, or a 260/230 ratio outside of the 1.7-2.3 range were discarded. The remaining samples were diluted to 15 ng/µl using DNA hydration solution before storage at -80^{o}C.

2.2 Genotyping

The Taqman genotyping assay is used to differentiate between different allelic states of SNPs through the use of coloured fluorophores. Commercially designed Taqman probe mixes (such as those used here) contain two probes, with one probe designed to anneal to each allelic state of a given SNP (e.g. one probe [VIC labelled] may anneal to the A state and one [FAM labelled] to the G state of a bi-allelic SNP). Each probe is initially attached to a quencher in order to prevent its fluorescence. Once the probe has annealed to the DNA (SNP state allows the binding of one or both probes to DNA; homozygous major allele, heterozygous, homozygous minor allele), polymerisation and DNA amplification takes place using forward and reverse PCR primers provided with the probes. The process of polymerisation removes the quencher from the probe allowing

the probes to emit fluorescence that can be detected (Figure 2.1). In all cases, probe sequences were BLASTed to ensure to off-target binding was unlikely to occur.

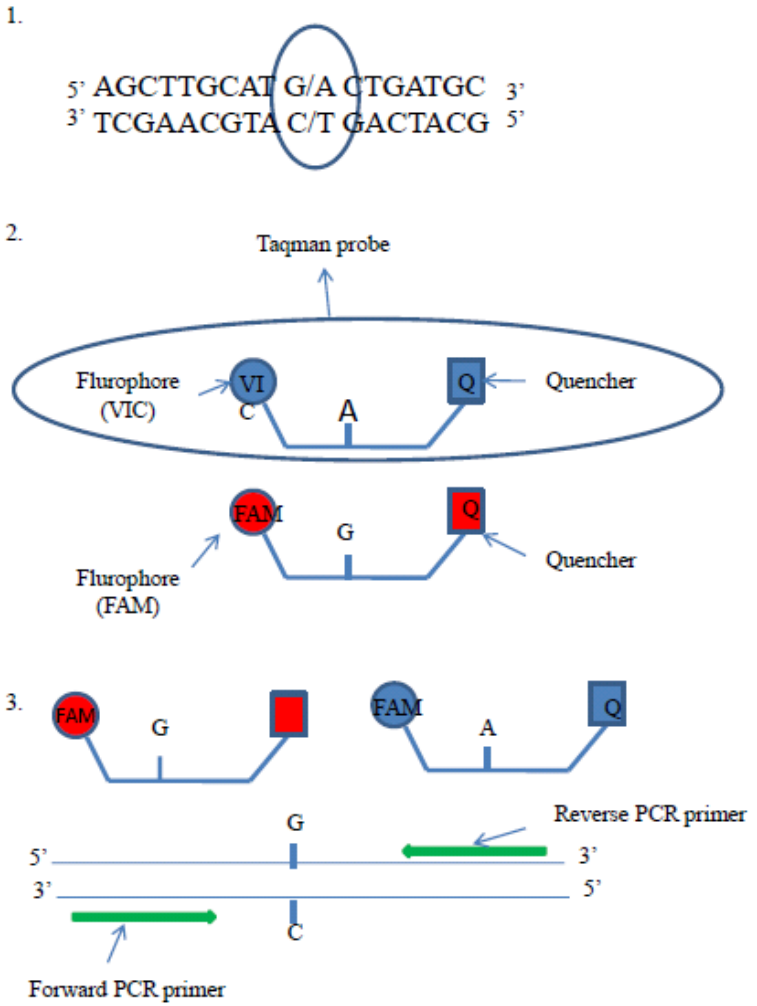
The "rs" number, "M" number (Y chromosome SNP identifier), SNP alleles and probe sequences for each SNP genotyped in this project are listed in Table 2.1. A master-mix containing 1,200 µl of Taqman genotyping master-mix (Quanta Biosciences, USA), 800 µl of deionised water and 10 µl of specific probe mix (Applied Biosystems, USA) was vortexed for 1 min. 5 µl of this master-mix was added to 1 µl of each DNA sample placed on a 384-well plate (one DNA sample per well). At least four negative controls (containing no DNA) and two positive controls (where the allelic state was known) were placed onto each 384-well plate. The plate was then covered with a film slip and centrifuged at 2000g for 1 min at room temperature. The assay reaction was completed on a GeneAmp PCR System 9700 (Applied Biosystems, USA) using the following program: 95^{o}C 10 min then 45 cycles of 92^{o}C for 15 sec each, and 60^{o}C for 1 min, 72^{o}C for 10 min. Depending on the SNP analysed, optimisation of the reaction consisted of either the alteration of the cycle number or annealing temperature (Table 2.1). Fluorescence was detected using a 7900HT ABI Prism Sequence Detection System (Applied Biosystems, USA), where genotype calls were automatically called and checked by two independent, experienced scientists.

Optimisation of five genotyping probes (Table 2.1) was required. This involved optimisation of annealing temperatures, denaturing temperatures, cycle number, and master-mix magnesium concentration. Due to the high AT content of the Y chromosome probes used, the decision was taken to reduce the annealing temperature by two degree increments (from 60^{o}C) to a minimum of 52^{o}C to increase probe efficiency. Both M89 (57^{o}C) and M173 (52^{o}C) worked sufficiently after this optimisation. The remaining three probes required modulation of both annealing

temperature and cycle number. In each case an extra annealing step (55^{o}C) was added alongside the standard annealing temperature of 60^{o}C, whilst cycle number was increased to 50.

In each cohort analysed, 10% of samples were independently typed to ensure correct allocation of base calls. On average, a SNP call was altered every 400 samples.

Figure 2.1 Taqman genotyping assay.

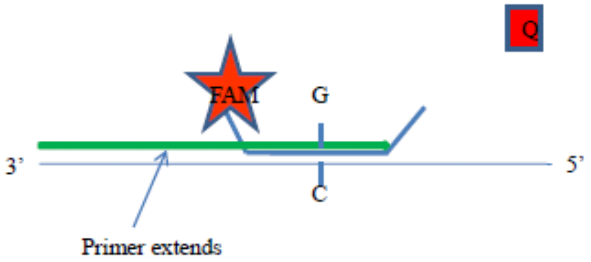


DNA template with G/A SNP to be investigated.

Two forms of Taqman probe are provided to differentiate between the two allelic states. One probe is fluorescently labelled with FAM and one with VIC. Each anneals to one version of the SNP. Probes also contain a quencher which prevents fluorescence of VIC and FAM until the probes are annealed to DNA.

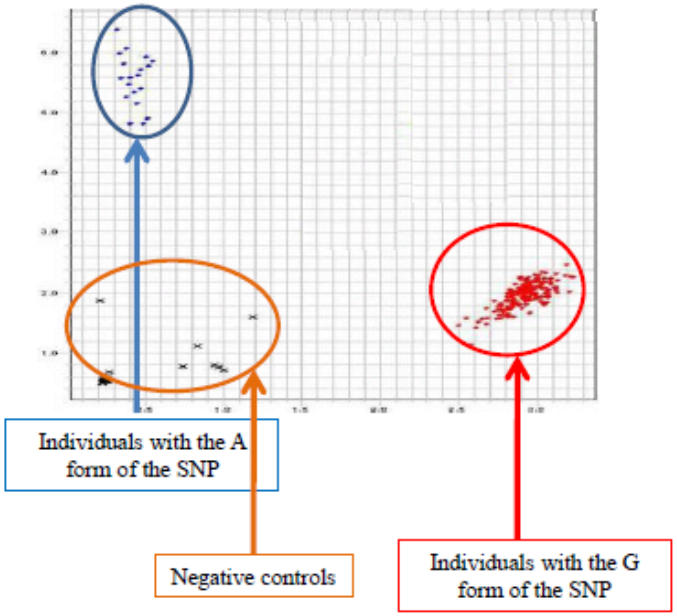
The DNA template (here containing the G-allele) is denatured and the assay components (forward and reverse primers and appropriate probe) bind.

4.



Polymerisation occurs removing the quencher and allowing FAM fluorescence, showing the allelic state here to be G.

5.



VIC and FAM emit at different wave lengths allowing differentiation between SNP state.

(Whitecomb *et al.* 1998)

Table 2.1 Details of Taqman probes used in genotyping for Y chromosome haplogrouping

SNP (rs number)	Alleles	Forward primer	Reverse primer	Amplification temperature (^{o}C) (cycle number)
M9 (rs3900)	C>G	ACCCTGAAATACAG AACTGCAAAGA	CTAAGTATGTAAGA CATTGAACGTTTGA	60 (45)
M35 (no rs number)	G>C	CAATACTCAGTGTC CCAATTTTCCTTT	GAACAACTAATCCA TGCAGACTTTCG	55 (55)
M45 (rs2032631)	G>A	AGAGAGGATATCAA AAATTGGCAGTGAA	GGCCTGGACCTCAG AAGGA	60 (45)
M89 (rs2032652)	C>T	TGGATTCAGCTCTCT TCCTAAGGTT	AGCAAAGGTAGCTG CAACTCA	57 (45)
M170 (rs2032597)	A>C	GCTCTTATTAAGTTA TGTTTTCATATTCTG	GTGAGACACAACCC ACACTGA	60 (45)
M173 (rs2032624)	A>C	GCCATATAAATTTA CTGTAACTTCCTAGA	GATCCTGAAAACAA AACACTGGCTTAT	52 (45)
M201 (rs2032636)	G>T	GCAATAGTTACTAC TTGAGTTACTATATT	CAGCTTCATCCAAC ACTAAGTACCT	55 (55)
M207 (rs2032658)	A>G	TCCCTGAAGAAGGA AAAAACGTTACA	ACTTCAACCTCTTGT TGGAAGATTATTCA	55 (55)
M269 (rs9786153)	C>T	ATTTAAAGTGGATT CTGTTACATGGTATC	CACTATACTTCTTTT GTGTGCCTTCTG	60 (45)
M304 (rs13447352)	A>C	CAGTATGTGGGATT TTTTTAGATGTGTTC	AGAAATAAACAAAT AACTTTCAAAACGT	60 (45)
SRY10831 (rs2534636)	G>A	AGTATCTGGCCTCTT GTATCTGACT	GACACAAGGCACCA CAT	60 (45)

2.3 RNA extraction

RNA extraction from kidney tissue was conducted using RNeasy mini kits (Qiagen, USA). RNeasy kits isolate RNA based on the selective binding properties of a silica-based membrane. All molecules longer than 200 nucleotides, including most mRNAs, bind to the silica-based membrane (excluding rRNAs and tRNAs) from which they can then be purified.

The apparatus was thoroughly cleaned using RNAase (Ambion, USA) and the experiments were completed on ice to prevent RNA degradation. A maximum of 20 µg of starting tissue was used providing a yield of ~30 µg of RNA. Tissue was suspended in lysis buffer (700 µl of β-mercaptoethanol [Sigma-Aldrich, UK] in 70 ml of RTL buffer [provided]) and homogenised using a rotor-stator homogeniser. The lysate was centrifuged for 3 min at full speed, removed and used in all future steps. An equal volume of 70% ethanol was added to the lysate and mixed. 700 µl of this mixture was added to the membrane of the RNeasy spin column (which was placed in a 2 ml collection tube) and centrifuged for 30 sec at 8,000g, room temperature (flow-through discarded). Three wash stages were completed, in each of which the flow-through was discarded:

1. 700 µl of RW1 buffer (provided) was added to the spin column and centrifuged for 30 sec at 8,000g room temperature
2. 500 µl RPE buffer (provided) was added to the spin column and centrifuged for 30 sec at 8,000g room temperature
3. 500 µl RPE buffer (provided) was added to the spin column and centrifuged for 2 min at 8,000g room temperature

The RNeasy spin column was placed in a 1.5 ml collection tube and 40 µl of RNase-free water was added to the spin column membrane. After incubation for 1 min (room temperature), the column was centrifuged for 1 min at 8,000g room temperature. The resultant RNA-containing liquid was stored at -80^{o}C.

2.3.1 Measurement of RNA quality and quantity

RNA quality and quantity was assessed using NanoDrop (Labtech) spectrophotometer in a manner identical to that described in section 2.2.1.2. The 260 nm wavelength is absorbed by RNA while the 280 nm wavelength is absorbed by sample contamination (in particular proteins). A ratio of around 2.1 (260:280 nm) indicated a pure RNA sample.

2.4 Reverse transcription PCR

Each reverse transcription PCR (RT-PCR) experiment converted 200 ng of mRNA to cDNA for use in gene expression studies. All surfaces and equipment were cleaned with RNase away (Ambion, USA) to prevent RNA degradation during the experimental setup. A master-mix containing buffer, $MgCl_2$, dNTPs, DTT (0.1µM), oligo sT_{16}, random hexamers, RNase inhibitor and MultiScribe RTase (concentrations Table 2.2) (all reagents sourced from Applied Biosystems, USA except 0.1 M DDT which was sourced from Invitrogen UK), was vortexed for 1 min and centrifuged at 12.2g for 1 min at room temperature. Relative volumes of each reagent used for one sample are shown in Table 2.2. 1 µl mRNA was added to 7.6 µl deionised water (Applied Biosystems, USA) and 11 µl of master-mix. Two negative controls were used; the first

contained deionised water rather than mRNA, while the second contained all reagents and mRNA other than RTase. Samples were vortexed for 1 min and centrifuged at 12.2g for 1 min at room temperature. Samples were subsequently incubated at 25^{o}C for 10 mins, 42^{o}C for 12 mins, and 85^{o}C for 15 mins. This involved optimisation of annealing temperatures, denaturing temperatures, cycle number, and master-mix magnesium concentration. Samples were stored at -20^{0}C on a short-term basis.

Table 2.2 Reagents of RT-PCR master-mix used to obtain 200 ng of mRNA from renal tissue

Reagent	Volume (µl)	Concentration
Buffer	4.0	10 x
$MgCl_2$	2.0	25 mM
dNTPs	2.0	10 mM
DTT	2.0	0.1 µM
Oligo dT_{16}	0.3	50 µM
Random hexamers	0.3	50 µM
RNase inhibitor	0.5	20 units/µl
MultiScribe RTase	0.3	50 units/µl

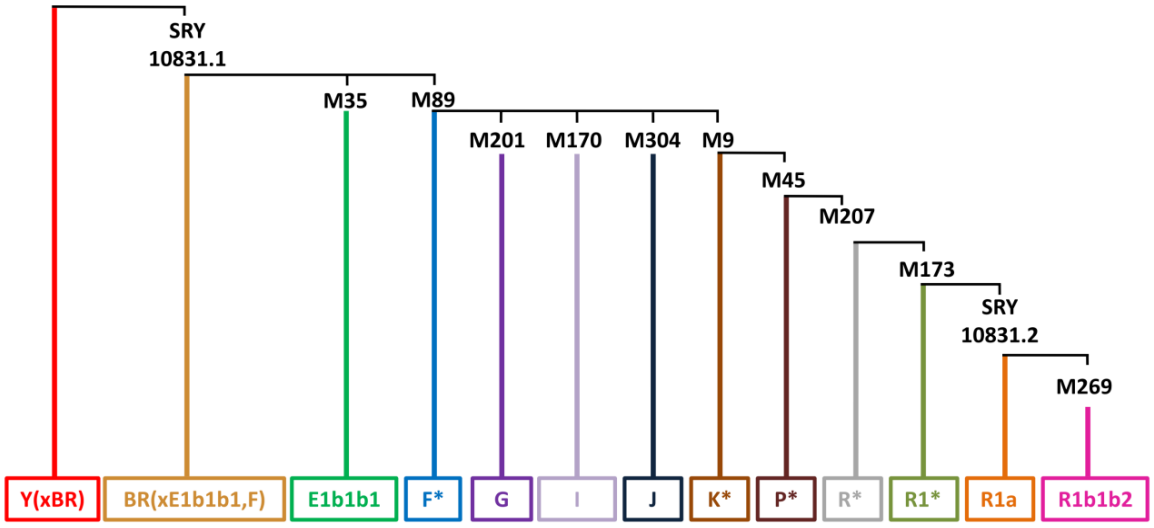
Relative volumes and concentrations shown.

2.5 The phylogenetic analysis of the Y chromosome

Phylogenetic analysis allows each Y chromosome to be assigned to a particular haplogroup. Based on the hierarchical relationship of 11 specific MSY SNPs (Figure 2.2), the allelic state of each SNP was analysed in turn, beginning with SRY10831, then M35, and so on. Where the derived state of a SNP was identified, the Y chromosome was placed in the appropriate group. For example, if the ancestral form of SRY 10831.1, M35, M89 and M201 but the derived form of M170 was identified, the Y

chromosome would be classified as haplogroup I. Once each Y chromosome was assigned a haplogroup, the percentage prevalence of each haplogroup was calculated.

Figure 2.2 Phylogenetic tree showing markers analysed to define the common European Y chromosomes. Haplogroups are arranged from [Y(xBR)] on the left to (R1b1b2) on the right with colours representing different haplogroups (e.g. pink is haplogroup R1b1b2). M and SRY numbers represent the terminal SNP dictating that haplogroup. SRY10831 is a recurrent mutation.



(Rootsi *et al.* 2004)

3 Haplogroup I of the Y chromosome and CAD risk factors

3.1 Introduction

Haplogroup I of the Y chromosome has been strongly implicated in the male predisposition to CAD (Chapter 1 Charchar *et al.* 2012). One of the major questions that remain to be answered is whether the effect of the human Y chromosome on the risk of CAD is mediated, at least to some extent, by traditional cardiovascular and metabolic risk factors, such as adiposity, BP, lipids or glucose. The analysis of WOSCOPS data, where information on some of main cardiovascular risk factors was available, showed that the effect of haplogroup I on CAD was independent of hypertension, LDL-C, HDL-C, triglycerides, C-reactive protein (CRP), BMI, diabetes, glucose, alcohol consumption, smoking and socioeconomic status (Charchar *et al.* 2012). However, this analysis was limited only to one cohort of middle-aged Scottish men, many of whom were on different types of pharmacological medications with a potential to affect measurements of the examined phenotypes. In addition, several major risk factors were only examined as categorised phenotypes (i.e. hypertension) rather than more informative continuous traits (such as BP). Finally, it is difficult to extrapolate these data into the young, apparently healthy male population prior to the development of overt cardiovascular disease.

Interestingly, genetic variation within the MSY of the human Y chromosome has previously been associated with cardiovascular risk factors in the general population. Firstly, the alphoid *Hin*dIII(+/-) polymorphism of the MSY has been associated with BP in men from general Scottish (MIDSPAN Family Study) (Charchar *et al.* 2002) and Australian (Victorian Family Heart Study) (Ellis *et al.* 2000) populations. Secondly, the same polymorphism was associated with levels of LDL-C, TC/HDL-C ratio (a surrogate

of small dense LDL-C particles) and paternal history of MI amongst young apparently healthy Polish men from the general population (Charchar *et al.* 2004). An MSY haplotype encompassing the *Hin*dIII(+/-) restriction fragment length polymorphism was also associated with HDL-C and triglycerides in men of African descent (Russo *et al.* 2008). Although not all reports showed association of the Y chromosome with BP and/or lipid fractions (Russo *et al.* 2006), the collected body of evidence suggests that the MSY may modulate cardiovascular risk.

The evidence discussed above means that it is reasonable to hypothesise that the effect of haplogroup I on the predisposition to CAD may be meditated through a known CAD risk factor. We therefore investigated the role of haplogroup I of the Y chromosome in major cardiovascular risk factors, using samples from three young, apparently healthy cohorts of European men recruited from the general population.

3.2 Materials and methods

3.2.1 Materials

3.2.1.1 Young Men Cardiovascular Association (YMCA) Studies 1 and 2

The "Young Men Cardiovascular Association Study 1" (YMCA 1) consists of 1,157 biologically unrelated and apparently healthy males (mean age: 19.1 years) recruited in Silesia (Southern Poland) (Charchar *et al.* 2004; Tomaszewski *et al.* 2007). The "Young Men Cardiovascular Association Study 2" (YMCA 2) - an extension of YMCA 1- recruited an additional sample of unrelated 597 young (mean age: 19.0 years) men in Silesia (Tomaszewski *et al.* 2007). Clinical and biochemical phenotyping protocols of each study were described in detail in previous publications (Charchar *et al.* 2004; Tomaszewski *et al.* 2007). In brief, clinical history was collected using anonymised, coded questionnaires. Recorded anthropometric measurements included height and weight, as well as three consecutive BP measurements (Charchar *et al.* 2004; Tomaszewski *et al.* 2007). Each subject underwent thorough biochemical phenotyping (under fasting conditions) including analysis of fasting lipids (TC, HDL-C, triglycerides), glucose and creatinine. LDL-C was calculated using the Friedewald formula (which was appropriate for use as no readings were extremely high). Only 1.6% and 0.3 % men in YMCA 1 and YMCA 2, respectively, were on antihypertensive medication (Charchar *et al.* 2004; Tomaszewski *et al.* 2007), and none were prescribed lipid-lowering medications. BP values from those on antihypertensive treatment were adjusted for BP-lowering effect of therapy using a previously reported method (Charchar *et al.* 2004; Tomaszewski *et al.* 2007).

3.2.1.2 Genetic Regulation of Arterial Pressure of Humans in the Community (GRAPHIC)

The "Genetic Regulation of Arterial Pressure in Humans in the Community" (GRAPHIC) cohort consists of 2,037 healthy individuals from 520 European British families (two parents, two offspring) recruited from the general population of Leicester, UK (Tobin *et al.* 2008; Tomaszewski *et al.* 2010). A total of 1,028 men of European ethnicity were recruited (516 fathers [average age: 53.8 years] and 512 sons [average age 25.0 years]). Of 512 male offspring, 391 biologically unrelated individuals were selected for this analysis; in families with two male offspring only the elder one was included in this study. Clinical and biochemical phenotyping protocols were described in detail elsewhere (Tomaszewski *et al.* 2010). In brief, clinical history was collected through standardised questionnaires alongside anthropometric measurements (height, weight and waist circumference). Each individual had his BP measured using both clinic and 24-hr ambulatory monitoring (Tomaszewski *et al.* 2010). Blood samples were obtained for the measurement of TC, HDL-C, glucose, and creatinine. LDL-C was measured enzymatically. Amongst GRAPHIC men, 0.8% of sons and 52.6% of fathers were on antihypertensive medication.

3.2.1.3 Bioethical approval

Written, informed consent was obtained from all subjects in all studies used here, in accordance with the Declaration of Helsinki. All studies had approval from relevant institutional ethical committees.

3.2.2 Methodology

Details of DNA extraction and quality assessment, genotyping and Y chromosome haplogrouping can be found in Chapter 2.

3.2.2.1 Assessment of normality

Each continuous trait (age, BMI, BP, lipid fractions, creatinine, glucose) was plotted and assessed by eye for normality. Age, triglycerides and CRP showed skewed distributions and as such were natural log-transformed before analysis. Each trait was summarised using arithmetic means and SD (when normally distributed) or geometric mean, which was calculated by multiplying each data-point and dividing the total by the n^{th} root and 95% confidence intervals (CI) (for traits showing a skewed distribution).

3.2.2.2 Assessment of population stratification

The presence of population stratification was assessed in GRAPHIC using approximately 1,700 ancestral SNPs genotyped using the HumanCVD BeadChip platform (50K IBC array) (Keating *et al.* 2008). SNPs were chosen based on differing allele frequencies between populations (European, Asian and African). SNPs were used to calculate the genome-wide, identity-by state, distance matrix. This matrix was further transformed with a non-metric, multidimensional scaling technique to visualise genetic similarity between participants GRAPHIC against three HapMap populations (CEU [European population]), JPT+CHB [Japanese and Chinese populations combined], YRI [African population]). Assuming all test individuals were European, three separate groups would be evident (with the test population positioned close to the CEU

population). If the test population were not European, the coloured dots relating to this population would be positioned elsewhere on the graph. GRAPHIC was also stratified by Y chromosome haplogroup (I vs. all others) and the non-metric, multi-dimensional scaling was then used to visualise the genetic similarity between the two groups. Where the two groups overlapped, no population stratification was identified; a separation of the two groups implies the presence of population stratification.

The presence of population stratification in YMCA 1 and YMCA 2 was assessed using the extensively validated STRUCTURE program (http://pritch.bsd.uchicago.edu/software.html) (Pritchard *et al.* 2000; Falush *et al.* 2003; Falush *et al.* 2007; Hubisz *et al.* 2009). This method is based on the genotyping of 34 ancestrally informative SNPs (chosen for their differing allele frequencies between ethnicities) (rs10141763, rs1024116, rs1084344, rs12913832, rs1321333, rs1335873, rs142665, rs149844, rs1573020, rs16891982, rs182549, rs1886510, rs1978806, rs2026721, rs2040411, rs2065160, rs2065982, rs2303798, rs2304925, rs239031, rs2572307, rs2814778, rs3785181, rs4540055, rs5033240, rs5997008, rs722098, rs727811, rs730570, rs773658, rs7897550, rs881929, rs896788, rs917118) and predicts the likelihood of any one sample belonging to a predefined set of subpopulations (in this case Asian, African and European). Reference population genotypes were downloaded from http://spsmart.cesga.es/snpforid.php and examined cohorts were commercially genotyped (KBiosciences, UK). The STRUCTURE software was run using length of burnin period 200,000, number of MCMC reps after burnin 200,000 and expected number of populations of 3 (Asian, African and European). These data were then visualised using the "triplot" command in Stata.

3.2.2.3 Genetic associations

Crude and age-adjusted linear regression was used to investigate the role of haplogroup I in each CAD risk factor. Crude analyses were completed in these healthy cohorts so as to allow the investigated traits to be researched without prior assumption of any effectors (such as age). Age was found to be a significant confounder in all traits and as such was added to the regression model before meta-analysis. The method of haplogroup I vs. all other MSY haplogroups was used. Meta-analysis was conducted when data were available for more than one cohort (BMI, SBP, DBP, TC, HDL-C, LDL-C, glucose and creatinine was available for all three cohorts, triglycerides were available in YMCA 1 and YMCA 2, CRP was available for GRAPHIC only) using a random-effect, inverse-variance model in Metan script (Deeks *et al.* 2001). In all cases a p-value if below 0.05 after Bonferroni correction indicated signficant association. Effect sizes from each study were weighted inversely proportional to the variance of the effect size. The β coefficient, standard error and p-values reflect the magnitude of the effect size and statistical significance for the subjects from all cohorts.

Power calculations based on 5% and 10% change in phenotypes were completed using http://www.statisticalsolutions.net/pss_calc.php.

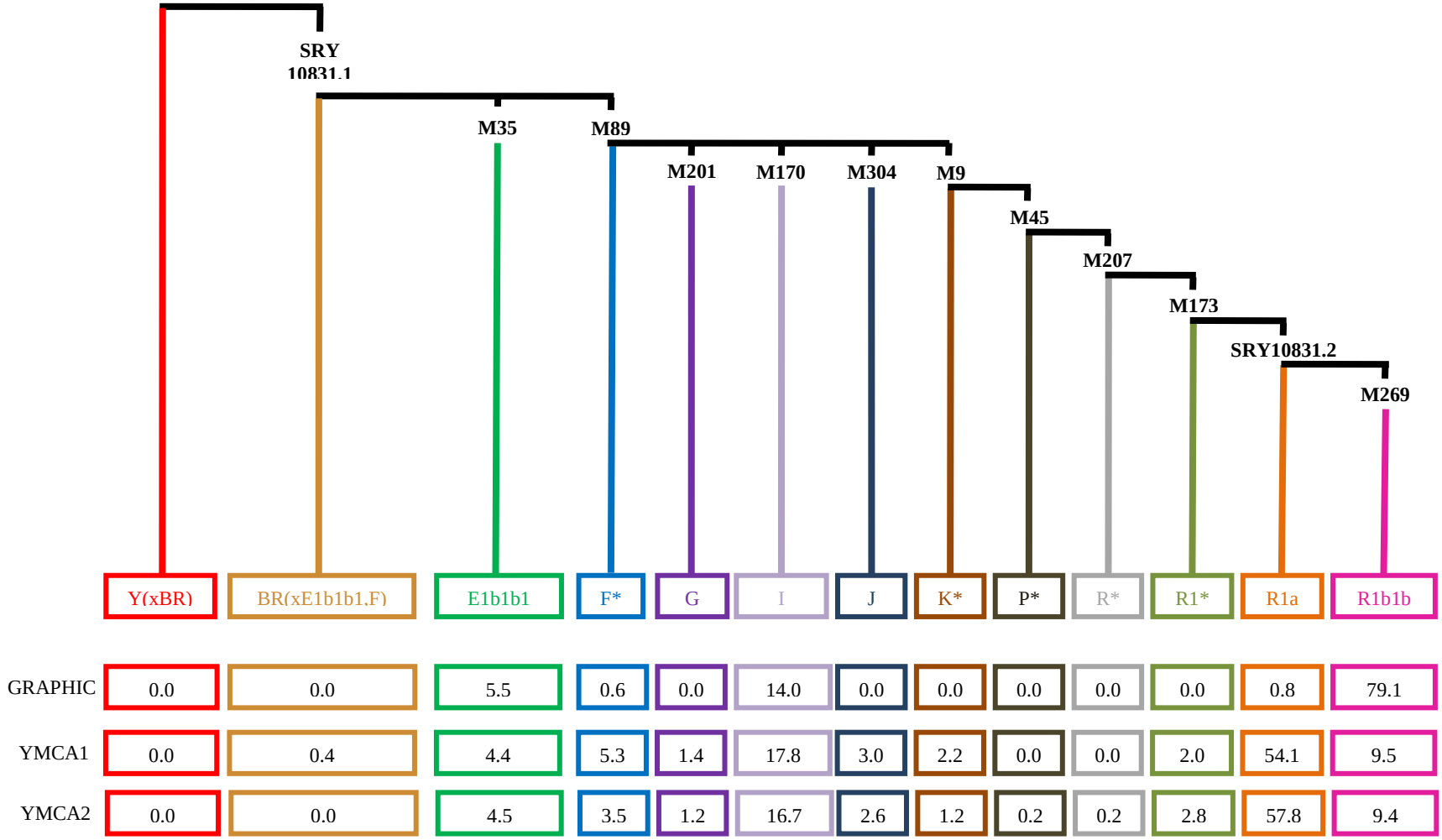
3.3 Results

3.3.1 Y chromosome haplogroup distribution in GRAPHIC, YMCA 1 and YMCA 2

Of 2,037 individuals recruited to GRAPHIC, 1,028 were male (Büsset *et al.* 2011). Of those, 516 represented fathers and 512 were sons (representing families with one son and various families containing two sons) (Büsset *et al.* 2011). For this analysis, 363 biologically unrelated sons were selected and gave a full set of genotypes necessary for haplogroup resolution. In families with two male offspring, the elder brother was selected for this analysis. A total of 1,068 men from 1,157 individuals recruited into YMCA 1 and 509 of 597 men recruited into YMCA 2 provided a full set of genotypes necessary for haplogrouping.

Of the 12 common European haplogroups, 5 MSY lineages were identified in GRAPHIC (frequency ranging from 0.6 – 79.1%). Haplogroups I and R1b1b2 accounted for ~93% of the lineages in the British population (GRAPHIC). A total of 10 and 11 Y chromosome lineages were identified in YMCA 1 and YMCA 2, respectively. Haplogroups I and R1a accounted for ~72% and 75% of the lineages in these cohorts, respectively (Figure 3.1). The prevalence of haplogroup I and R1a were different between the cohorts (GRAPHIC and both YMCA cohorts) and in keeping with published data (Semino *et al.* 2000; Figure 3.1).

Figure 3.1 Phylogenetic tree of the Y chromosome and frequency of haplogroups in GRAPHIC, YMCA 1 and YMCA 2. %- percentage prevalence. Haplogroups are arranged Y(xBR) on the left to R1b1b2 on the right. M and SRY numbers represent the terminal SNP dictating that haplogroup. SRY10831 is a recurrent mutation.



3.3.2 Analysis of population stratification

Analysis of population stratification in GRAPHIC was completed based on ancestrally informative SNPs from across the genome. Figure 3.2A shows the GRAPHIC cohort plotted in the context of three HapMap populations (CEU, JPT+CHB and YRI). A very low level of separation was detected between the CEU HapMap individuals and those recruited to the GRAPHIC cohort. The lack of significant separation shows that all those included in this project were of European ancestry. Figure 3.2B illustrates the level of genetic difference between carriers of haplogroup I and all others, based on the same ancestrally informative SNPs. Significant population stratification would have been indicated by a clear separation of carriers of haplogroup I and all others. However, there was no separation between the two groups, indicating that there are no systematic genetic differences between the two groups outside of the MSY.

STRUCTURE confirmed that all individuals recruited into YMCA 1 and YMCA 2 are of European ancestry (Figure 3.3). Differing ancestry would have been shown by a wider dispersal of data points away from the European corner of the triangle plot.

Figure 3.2 Assessment of population stratification in the GRAPHIC study based on 17,00 polymorphisms from across the genome, genotyped using 50 IBC array. A. A matrix of identity-by-state distance was created in the GRAPHIC cohort and the three HapMap populations (CEU, JPT+CHB, YRI). Coloured dots represent each population: blue- JPT+CHB cohort, green- YRI cohort, red- CEU cohort, yellow- GRAPHIC cohort. The first two components that were obtained from the subsequent multi-dimensional scaling of the matrix and used as co-ordinates to compare the relative distances between the four populations. B. Assessment of population structure stratified by Y chromosome haplogroup (I vs. all others) of GRAPHIC cohort. Colours represent haplogroups: green- carriers of haplogroup I, red- carriers of all other haplogroups. A matrix of identity-by-state distances was created and the first two components obtained from the subsequent multi-dimensional scaling of the matrix were used as co-ordinates to generate a scatter plot.

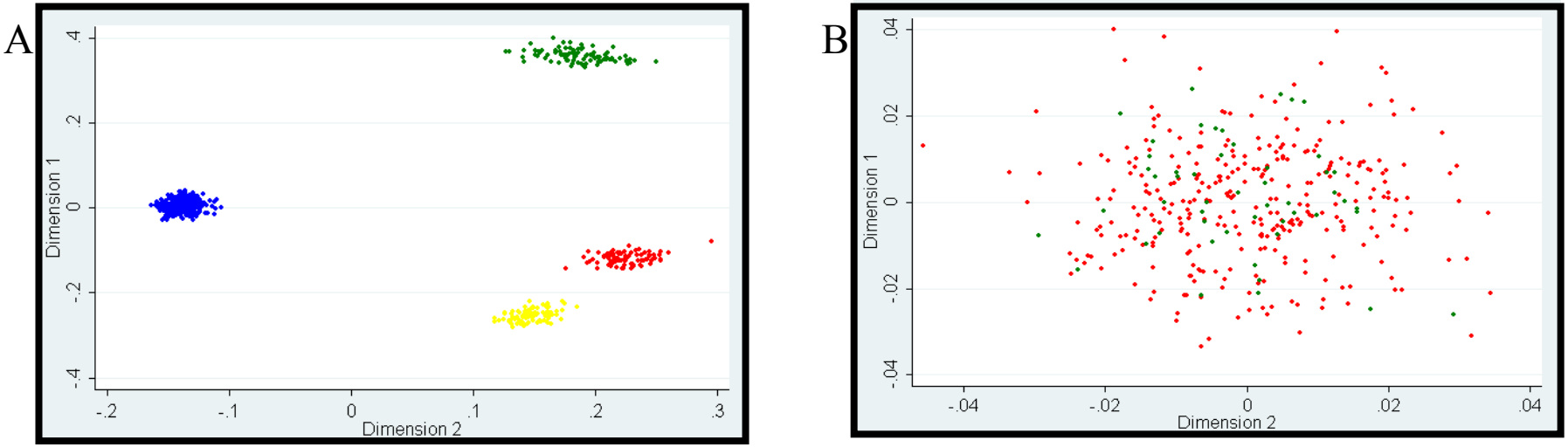
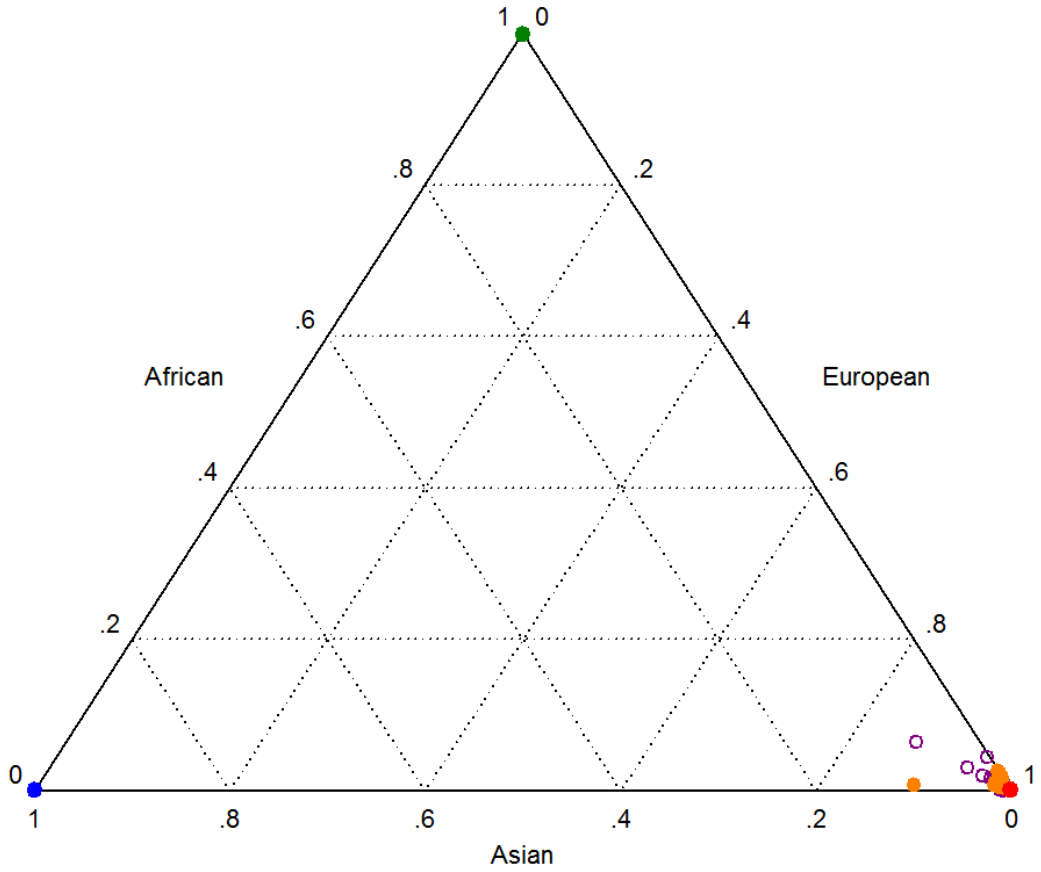


Figure 3.3 Analysis of population stratification in YMCA 1 and YMCA 2 based on the STRUCTURE program (using 34 validated ancestrally informative SNPs); blue- Asian populations, green- African populations, red- European populations, orange- YMCA 1, purple- YMCA 2.



3.3.3 Phenotype distribution and clinical characteristics

Summarised data from each cohort can be found in Table 3.1. With the exception of age, triglycerides and CRP, all phenotypes were approximately normally distributed in each cohort (Figure 3.4-3.6). In research such as this, it is widely accepted that age is unlikely to be normally distributed and no action is required. Triglycerides and CRP were log-transformed to provide a distribution closer to the expected "normal" distribution. All characteristics measured showed levels analogous to those expected for the age of the populations studied (data not shown).

Table 3.1 Clinical and demographic data in GRAPHIC, YMCA 1 and YMCA 2

Phenotype	YMCA 1	YMCA 2	GRAPHIC sons
N	1068	509	363
Age (years)	19.3 (3.5)	18.9 (3.3)	25.7 (5.0)
BMI (kg/m^2)	22.9 (3.0)	22.6 (3.0)	25.1 (4.1)
Clinic SBP (mmHg)	118.1 (13.2)	118.8 (13.2)	128.4 (13.0)
Clinic DBP (mmHg)	74.2 (7.9)	74.1 (7.9)	76.4 (9.6)
TC (mmol/L)	4.3 (0.9)	3.7 (1.0)	4.6 (0.9)
HDL-C (mmol/L)	1.2 (0.3)	1.1 (0.3)	1.3 (0.3)
Triglycerides (mmol/L)	1.1 (1.0-1.1)	0.9 (0.8-0.9)	-
LDL-C (mmol/L)	2.6 (0.9)	2.2 (0.8)	2.6 (0.7)
Glucose (mmol/L)	4.8 (0.7)	4.4 (0.8)	4.9 (0.8)
CRP (mg/L)	-	-	0.8 (0.70-0.81)
Creatinine (µmol/L)	82.8 (11.1)	74.9 (9.8)	82.2 (10.1)

Data are means and standard deviations or geometric means and 95% confidence intervals (triglycerides and CRP); n –number of subjects, BMI – body mass index, SBP – systolic blood pressure, DBP – diastolic blood pressure, TC – total cholesterol, HDL-C – high density lipoprotein cholesterol, LDL-C – low density lipoprotein cholesterol, CRP – C-reactive protein.

Figure 3.4 Distribution of cardiovascular risk factors in GRAPHIC cohort (sons). BMI- body mass index, SBP- systolic blood pressure, DBP- diastolic blood pressure, HDL-C- high density lipoprotein cholesterol, LDL-C- low density lipoprotein cholesterol, CRP- C reactive protein.

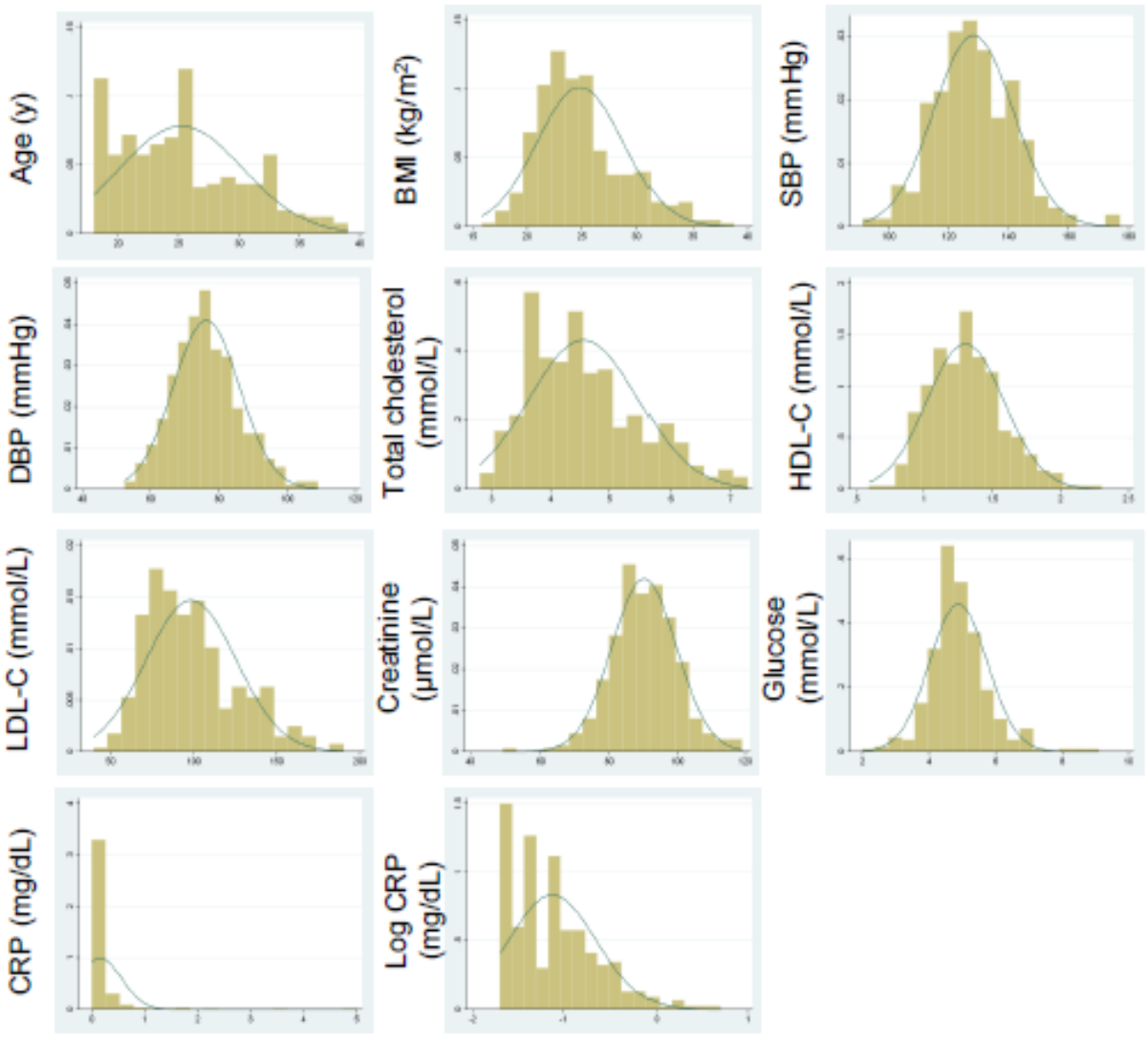


Figure 3.5 Distribution of cardiovascular risk factors in YMCA 1. BMI- body mass index, SBP- systolic blood pressure, DBP- diastolic blood pressure, HDL-C- high density lipoprotein cholesterol, LDL-C- low density lipoprotein cholesterol.

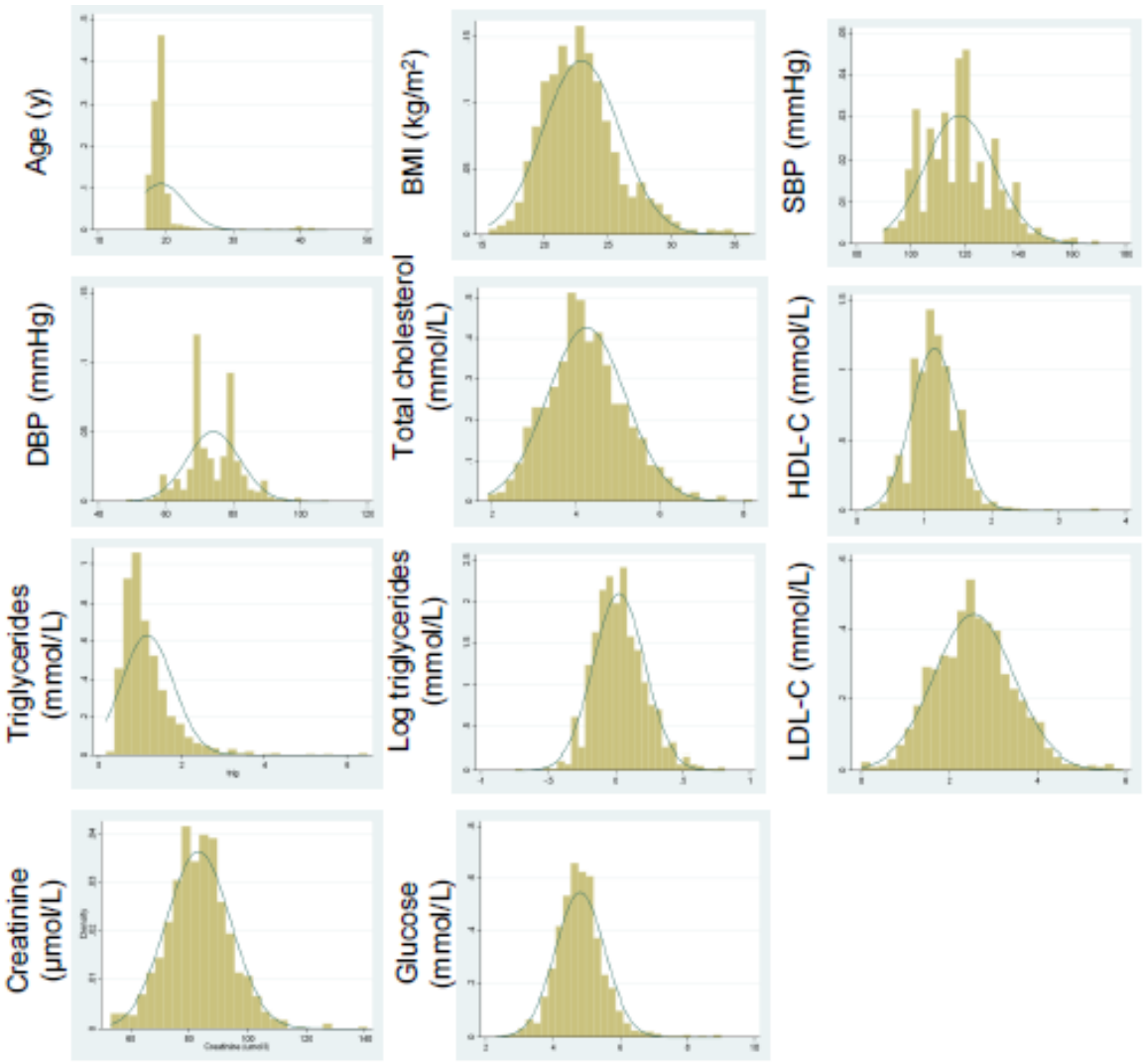


Figure 3.6 Distribution of cardiovascular risk factors in YMCA 2. BMI- body mass index, SBP- systolic blood pressure, DBP- diastolic blood pressure, HDL-C- high density lipoprotein cholesterol, LDL-C- low density lipoprotein cholesterol.

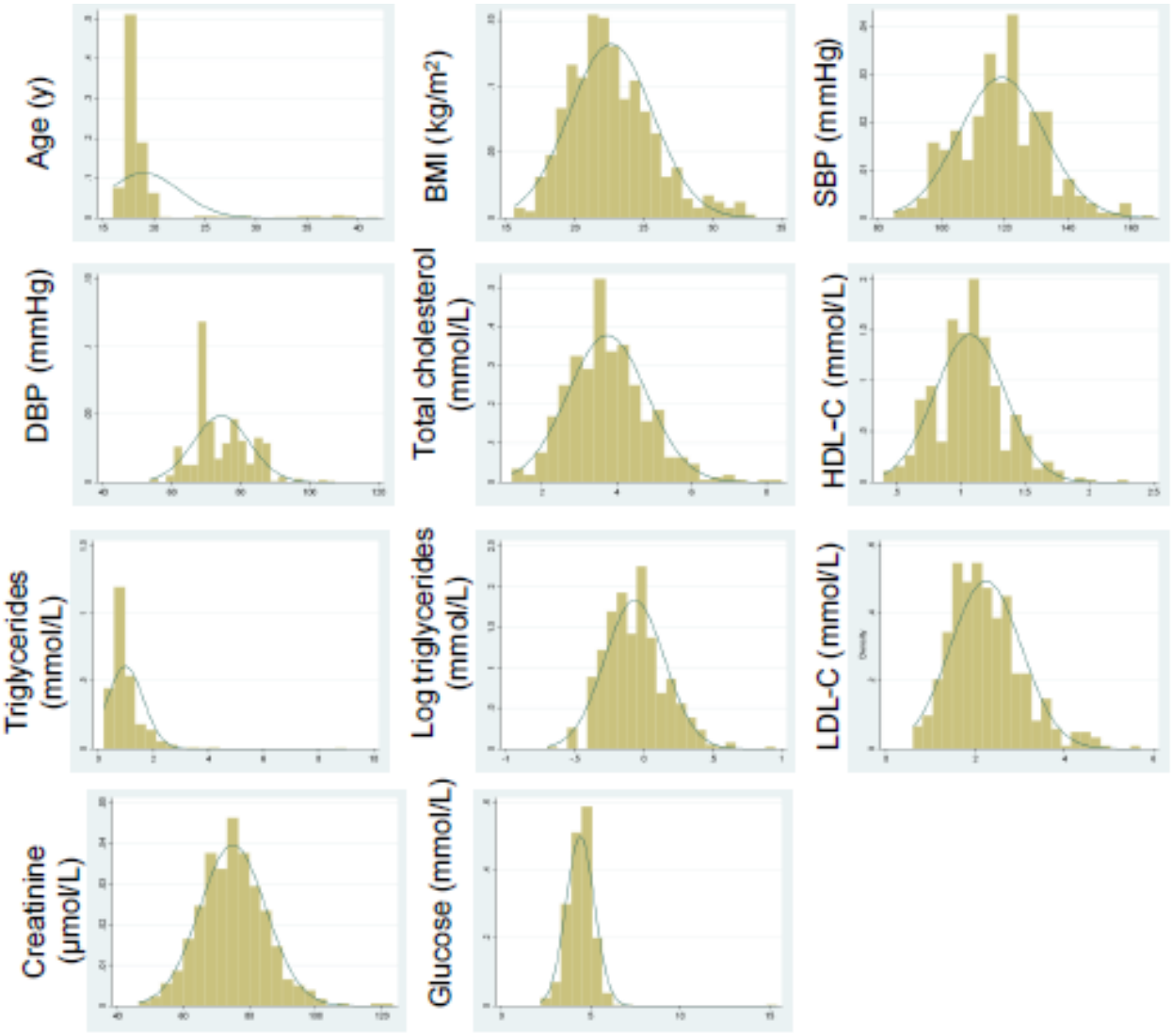


Table 3.2 Statistical power calculations for cardiovascular risk factor data.

Phenotype	Statistical power	
	5% Change	10% Change
BMI	1.00	1.00
SBP	1.00	1.00
DBP	1.00	1.00
HDL-cholesterol	0.81	1.00
Triglycerides	0.31	0.84
LDL-cholesterol	0.73	1.00
Glucose	1.00	1.00
Creatinine	1.00	1.00
CRP	0.04	0.05

Statistical power calculations for detection of a 5% or 10% change cardiovascular risk factor values between carriers of haplogroup I versus all other haplogroups (at the 5% significance level). Calculations were performed using an online tool (http://www.statisticalsolutions.net/pss_calc.php).

3.3.4 Effect of the Y chromosome on CAD risk factors

No CAD risk factor was significantly associated with haplogroup I in the GRAPHIC cohort (lowest P= 0.285 for glucose). Haplogroup I of the Y chromosome was associated with LDL-C levels in both YMCA 1 and YMCA 2 (P= 0.034 and 0.054, respectively). The direction of effect differed in each cohort; haplogroup I decreased levels of LDL-C in YMCA 1 while increasing levels in YMCA 2. Interestingly, LDL-C was the only phenotype to show any level of heterogeneity between studies. CRP was available in the GRAPHIC cohort only and was not associated with haplogroup I (P= 0.291).

Meta-analysis of all three populations showed that the Y chromosome was not associated with any of the traits analysed (lowest P= 0.397 DBP) (Tables 3.3- 3.11). There was no heterogeneity in any of the traits between the three studies (other than

LDL-C [P= 0.017]) (p values for BMI- 0.614, SBP- 0.896, DBP- 0.338, TC- 0.957, 0.098, HDL-C- 0.463, triglycerides- 0.955, LDL-C- 0.017, creatinine- 0.937, glucose- 0.318).

Table 3.3 Age-adjusted fixed-effect inverse-variance meta-analysis of association between haplogroup I of the Y chromosome and BMI

Cohort	N	β coefficient (SE)	P value	Weight (%)	Heterogeneity P value
GRAPHIC	363	0.498 (0.6)	0.395	11.00	-
YMCA 1	1068	-0.127 (0.2)	0.601	66.74	-
YMCA 2	509	-0.055 (0.4)	0.888	25.26	-
Meta-analysis	1940	-0.040 (0.3)	0.830	100	0.830

SE- standard error.

Table 3.4 Age-adjusted fixed-effect inverse-variance meta-analysis between haplogroup I of the Y chromosome and SBP

Cohort	N	β coefficient (SE)	P value	Weight (%)	Heterogeneity P value
GRAPHIC	363	0.460 (1.9)	0.814	18.14	-
YMCA 1	1068	-0.584 (1.1)	0.588	59.72	-
YMCA 2	509	-0.257 (1.8)	0.885	22.14	-
Meta-analysis	1940	-0.320 (1.2)	0.699	100	0.869

SE- standard error.

Table 3.5 Age-adjusted fixed-effect inverse-variance meta-analysis between haplogroup I of the Y chromosome and DBP

Cohort	N	β coefficient (SE)	P value	Weight (%)	Heterogeneity P value
GRAPHIC	363	1.454 (1.4)	0.292	13.70	-
YMCA 1	1068	-0.726 (0.6)	0.257	63.32	-
YMCA 2	509	-0.744 (1.1)	0.485	22.98	-
Meta-analysis	1940	-0.430 (1.2)	0.397	100	0.338

SE- standard error.

Table 3.6 Age-adjusted fixed-effect inverse-variance meta-analysis between haplogroup I of the Y chromosome and total cholesterol

Cohort	N	β coefficient (SE)	P value	Weight (%)	Heterogeneity P value
GRAPHIC	363	0.074 (0.1)	0.544	23.27	-
YMCA 1	1068	-0.112 (0.1)	0.147	57.47	-
YMCA 2	509	0.196 (0.1)	0.14	19.26	-
Meta-analysis	1940	-0.010 (0.1)	0.873	100	0.957

SE- standard error.

.

Table 3.7 Age-adjusted fixed-effect inverse-variance meta-analysis between haplogroup I of the Y chromosome and HDL-C

Cohort	N	β coefficient (SE)	P value	Weight (%)	Heterogeneity P value
GRAPHIC	363	1.181 (1.6)	0.466	0.02	-
YMCA 1	1068	0.035 (0.1)	0.213	63.57	-
YMCA 2	509	-0.12 (0.1)	0.738	36.41	-
Meta-analysis	1940	0.02 (0.1)	0.426	100	0.463

SE- standard error.

Table 3.8 Age-adjusted fixed-effect inverse-variance meta-analysis between haplogroup I of the Y chromosome and LDL-C

Cohort	N	β coefficient (SE)	P value	Weight (%)	Heterogeneity P value
GRAPHIC	363	0.014 (0.1)	0.884	26.50	-
YMCA 1	1068	-0.155 (0.1)	0.034	43.94	-
YMCA 2	509	0.172 (0.1)	0.054	29.56	-
Meta-analysis	1940	-0.014 (0.1)	0.779	100	0.017

SE- standard error.

Table 3.9 Age-adjusted fixed-effect inverse-variance meta-analysis between haplogroup I of the Y chromosome and triglycerides

Cohort	N	β coefficient (SE)	P value	Weight (%)	Heterogeneity P value
YMCA 1	1068	0.018 (0.1)	0.736	21.41	-
YMCA 2	509	0.031 (0.1)	0.688	9.76	-
Meta-analysis	1577	0.012 (0.1)	0.609	100	0.880

SE- standard error.

Table 3.10 Age-adjusted fixed-effect inverse-variance meta-analysis between haplogroup I of the Y chromosome and creatinine

Cohort	N	β coefficient (SE)	P value	Weight (%)	Heterogeneity P value
GRAPHIC	363	-0.448 (1.4)	0.746	19.00	
YMCA 1	1068	0.025 (0.9)	0.978	49.59	-
YMCA 2	509	0.175 (1.1)	0.874	31.41	-
Meta-analysis	1940	-0.022 (0.6)	0.972	100	0.880

SE- standard error.

Table 3.11 Age-adjusted fixed-effect inverse-variance meta-analysis between haplogroup I of the Y chromosome and glucose

Cohort	N	β coefficient (SE)	P value	Weight (%)	Heterogeneity P value
GRAPHIC	363	0.135 (0.1)	0.285	14.66	
YMCA 1	1068	-0.026 (0.1)	0.669	64.63	-
YMCA 2	509	-0.113 (0.1)	0.289	20.71	-
Meta-analysis	1940	-0.02 (0.1)	0.677	100	0.880

SE- standard error.

3.4 Discussion

It was hypothesised that an established, sexually dimorphic CAD risk factor was associated with haplogroup I of the Y chromosome, and as such, may underlie the increased risk of CAD in carriers of haplogroup. The data shown here refute this hypothesis.

3.4.1 Y chromosome haplogroup distribution and analysis of population stratification in the GRAPHIC cohort and YMCA 1 and YMCA 2 cohorts

The distribution of Y chromosome haplogroups in all cohorts studied were in keeping with published Y chromosome data (Capelli *et al.* 2003; Charchar *et al.* 2012). It is well established that in British cohorts (such as GRAPHIC), haplogroups I and R1b1b2 account for between 70 and 95% of the haplotypic variance (Semino *et al.* 2000; Charchar *et al.* 2012). Conversely, in Polish cohorts (such as YMCA 1 and YMCA 2) the majority of individuals are those from haplogroup I or R1a (Semino *et al.* 2000). These differences in haplotypic variation were confirmed here. Meta-analysis of the differing cohorts is applicable as the overall prevalence of haplogroup I (the research haplogroup) is consistent across the ethnicities.

Analysis of population stratification was completed in all three cohorts using two differing methods. Both forms of analysis confirmed that all individuals used here were of European origin. While analysis investigating genomic differences between carriers of haplogroup I and all others was not possible in YMCA 1 and YMCA 2 due to the method of population stratification analysis used, no evidence of systematic genetic differences were identified in the GRAPHIC cohort. These data have confirmed that in

this cohort the only observable systematic genetic difference between carriers of haplogroup I and all others is the lineage of the Y chromosome.

3.4.2 The association of haplogroup I with CAD risk factors

At the 5 and 10% change level all analyses other than CRP were sufficiently powered. There was no association between haplogroup I of the Y chromosome and any of ten sexually dimorphic CAD risk factors (BMI, SBP, DBP, TC, HDL-C, LDL-C, triglycerides, creatinine, glucose and CRP) , either in the analysis of individual cohorts or during their meta-analysis (after correction for multiple testing).

These data provide important support for the previously proposed hypothesis that the association between haplogroup I of the Y chromosome and CAD is independent of traditional cardiovascular risk factors (Charchar *et al.* 2012). Indeed, using cohorts un-confounded by medications and continuous measures of cardiovascular risk (BP, LDL-C, HDL-C, glucose or CRP) it was shown that no risk factor was associated with haplogroup I of the Y chromosome in apparently healthy men.

The absence of association between the Y chromosome and CRP in this analysis is particularly important, given that the immunity hypothesis of atherosclerosis is becoming increasingly accepted and that CRP is a recognised surrogate of cardiovascular risk (Ridker 2004).

Our data also reveal that other correlates of CAD not included in the original observation (Charchar *et al.* 2012) such as renal function, are not associated with Y chromosome lineages amongst young apparently healthy men. Indeed, circulating levels

of creatinine, a clinically recognised measure of renal function, was not associated with any of the Y chromosome lineages. Creatinine was used as a measure of renal function instead of estimated glomerular filtration rate (eGFR) because a significant proportion of YMCA 1 and YMCA 2 men were aged <18 years and the MDRD eGFR equation is not validated as a measure of renal function in this age category. Irrespective of the measure used to estimate it, renal function is associated with the risk of CAD and cardiovascular mortality in the general population (Hallan *et al.* 2007; Di Angelantonio *et al.* 2010). There is also apparent evidence of sexual dimorphism in susceptibility to renal disease with men showing higher incidence of chronic kidney disease and faster progression to end-stage renal disease than women (Iseki *et al.* 1996; Evans *et al.* 2005; Tomaszewski *et al.* 2009). The absence of association between the common Y chromosome lineages and circulating levels of creatinine suggests that the Y chromosome is not likely to drive the biological mechanisms underlying this phenomenon.

3.4.3 Conclusions

Taken together, our results show that the I haplogroup of the Y chromosome, which has been previously linked to increased risk of CAD, is not associated with conventional cardiovascular risk factors prior to development of overt CVD.

3.4.4 Future work

Future work should include an analysis of cohorts with differing ethnic backgrounds. Indeed, studies involving a variety of ethnicities would further clarify if the Y chromosome (rather than simply European haplogroups) has a regulatory role in CAD.

Furthermore, efforts should be made to investigate the role of the Y chromosome in other sexually dimorphic CAD risk factors. Many of these risk factors were not available in this study, including smoking, alcohol intake and other markers of inflammation.

3.4.5 Major limitations

3.4.5.1 Haplogroup I vs. all other haplogroups

Work outlined here represents a follow-up study to previously published work (Charchar *et al.* 2012). As such, the method of analysis used here focused on comparison of haplogroup I against all other lineages of the Y chromosome, and was found to be consistent with previous publications (Charchar *et al.* 2012). Furthermore, this approach reduces the possibility of over-interpreting these data and reduces the risk of type 1 errors. However, as a relatively conservative method, the effects of other Y chromosome haplogroups on CAD risk factors may have been overlooked. While less conservative methods, such as analysis of variance or grouping together of haplogroups (such as grouping the entire R branch into one group) would help increase statistical power, such methods are not in line with previously published data (Charchar *et al.* 2012), are less conservative and have major limitations. The major problems with this form of analysis include the possibility that the cause of a statistically significant result (i.e. which haplogroup) would remain unclear, and means that over-interpretation of these data is imminently possible.

It is plausible that an investigation of more precisely defined Y chromosome haplogroups (for example haplogroup I1, I2a1a, I2a1b, I2a2a, I2a2b and I2c rather than I) may highlight a role for the Y chromosome in cardiovascular risk factors (for

example, one sub-lineage may drive the detected association between haplogroup I and CAD). While possible, the loss of statistical power brought about by dividing haplogroup I into its sub-lineages in the cohorts used here makes this investigation impossible. This limitation may be addressed through the use of larger cohorts, as proposed in future work. While this may provide interesting data, historically, the association between a specific haplogroup and a disease state is lost upon analysis of sub-lineages (Sezgin *et al.* 2009).

3.4.5.2 Method of meta-analysis

Meta-analysis based on β coefficient and standard error was deemed most appropriate here. While an argument can be made to complete meta-analysis based on p-value, which can be useful when a significant level of heterogeneity is detected, this method does not take into account the direction of effect, and as such, is a less robust method (Iaonnidis & Lau 1999).

The two cohorts used here were of Polish origin, however, the initial association between haplogroup I of the Y chromosome and CAD was identified in two British cohorts. While the comparison of Y chromosome analyses across European cohorts is generally accepted (Stouffs *et al.* 2011; Charchar *et al.* 2012), it is possible that the association made between haplogroup I of the Y chromosome and CAD might have been UK-specific. Therefore, the mechanism of action may not be identified in cohorts of differing ancestry. To test this hypothesis, further British cohorts with appropriate phenotyping would be required. Unfortunately, such cohorts (apart from GRAPHIC) were not available during the duration of this project. Nonetheless, it should be noted that the haplogroup I prevalences in Britain and Poland are largely similar (Semino *et*

al. 2000; Rootsi *et al.* 2004) and will have undergone many of the same evolutionary pressures. It is therefore reasonable to assume that haplogroup I would predispose all carriers to CAD (regardless of country of origin). Further discussion on the ancestry of populations used can be found in Section 8.5.2.

3.4.5.3 Population stratification

All the populations used here underwent some level of analysis of population stratification. Due to the availability of ~1,700 ancestrally informative SNPs in the GRAPHIC cohort, a multi-dimensional scaling approach was used. This represents a well-established and validated method of assessing the presence of population stratification (Charchar *et al.* 2012). Conversely, the STRUCTURE computer package was used to assess population stratification in YMCA 1 and YMCA 2. The length of burnin period and number of MCMC repetitions were both set at 200,000. Both parameters were selected because previous work has demonstrated those variables result in the most consistent data, given the computing resources that were available (data not shown). While this method does not indicate if there are significant genetic differences between carriers of haplogroup I and all others in these cohorts, it does confirm that all individuals were of one ethnic background (European). Future work should include analysis of differences in autosomal signatures in these cohorts.

4 Haplogroup I of the Y chromosome and "male-associated" and/or "sex-specific" phenotypes

4.1 Introduction

The results from the analysis of association between the Y chromosome and traditional cardiovascular risk factors presented in the previous chapter clearly indicate that these phenotypes are unlikely to act as major mediators of the haplogroup I effect on CAD risk. One of the biologically most plausible, and perhaps most intuitive, potential mechanisms underlying this association are highly heritable, intermediate phenotypes strongly linked to male sex and/or sex-specific action. Perhaps the strongest biological candidates are aggressiveness and sex steroids, which have significant genetic component and strong sex specificity (Barrett-Connor 1995). Aggression is well accepted to be a male-dominated trait; significantly more men are described as of Type A personality (defined as having higher aggressiveness, competitiveness, hostility and time urgency), while women more are more commonly of Type B personality (defined as reflective and nurturing) (Barrett-Connor 1997). Previous studies have found that male traits such as anger, hostility, time urgency and anxiety were associated with increased risk of MI (Haynes *et al.* 1978; Haynes & Feinleib 1980; Barrett-Connor 1997). Furthermore, meta-analyses of twin studies showed that approximately 50% of the inter-individual variation of aggressiveness is genetic (Miles & Carey 1997; Rhee & Waldman 2002). Likewise, it is well accepted that the concentrations of sex steroids (including testosterone and oestradiol) are significantly different at all stages of life between men and women (Ober *et al.* 2008). At birth testosterone is at its percentage maximum in both sexes and drops dramatically. This drop is followed by a significant increase in testosterone in males (to ~40% of maximum), however, no increase in testosterone is witnessed in females. Similarly, oestradiol increases in females (to ~30%

of maximum) while remaining low in males (~10% of maximum in males) (Ober *et al.* 2008). During puberty, oestradiol increases to maximum in both the sexes, while testosterone increases significantly more in males (to ~100% of maximum) (Ober *et al.* 2008). Testosterone remains at ~5% of maximum in females. This difference in sex steroid concentration remains across the reproductive life. At menopause (~50 years) oestradiol falls significantly in females (levels are maintained in men), while testosterone falls slowly in men to ~50% of maximum by age 85 years (testosterone remains very low, ~5% of maximum, in females across this time) (Ober *et al.* 2008). It is well established that sex steroid levels vary significantly between the sexes, however, it is also well defined that inter-individual variation in sex steroid levels is common. Bogaert *et al.* found that up to 65% of the inter-individual variation in testosterone is accounted for by genetic factors (Bogaert *et al.* 2008; Vanbillemont *et al.* 2009). Furthermore, both aggression and sex steroids have been strongly implicated in the pathogenesis of CAD (further details can be found in Chapter 1).

Though no analysis of the human Y chromosome has ever been completed in the context of these "male-related" or "sex-specific" phenotypes, evidence from animal models supports a hypothetical role for the Y chromosome in regulation of both aggression and sex hormones. Firstly, the Y chromosome has been repeatedly implicated in aggression in murine models; Gatewood *et al.* found that female mice carrying an *SRY*-deleted Y chromosome had significantly higher aggression levels than wild-type female mice, similar to those found in males (Gatewood *et al.* 2006). Similarly, the Y chromosome was associated with sex steroids through its function in sex determination (Wilhelm *et al.* 2007).

Thus, it is reasonable to postulate that aggressiveness and/or sex steroids may be the causative mechanism(s) underlying the observed association between haplogroup I

of the Y chromosome and CAD. Through the analysis of two cohorts containing phenotypically normal, apparently healthy young men, the potential association between haplogroup I and five measures of aggressiveness and sex steroids levels were investigated.

4.2 Materials and methodology

4.2.1 Materials

4.2.1.1 Aggression scoring and sex steroid measurement in YMCA 1 and 2

Aggression was assessed using the validated Buss-Perry aggression questionnaire (Buss & Perry 1992). This improved version of the well-documented "Buss and Durkee inventory" (Buss & Durkee 1957) has been used in a number of investigations (Meesters *et al.* 1996; Nakano 2001; Garcia-Leon *et al.* 2002; Tsorbatzoudis 2006; Shah *et al.* 2009), and has been extensively evaluated for validity and reliability (Harris 1997). The questionnaire assesses four dimensions of aggression using a 29-question analysis, including an assessment of physical aggression, verbal aggression, anger and hostility (Buss & Perry 1992). Each facet of aggression represents a different reaction being measured. For example, hostility is an internal feeling while anger has external consequences. For each facet of aggression, the participants were asked to score the degree to which they agreed or disagreed with a statement on a scale of 1-5 (1= this is uncharacteristic of me, 5= this is very characteristic of me). Physical aggression scores from the questionnaire can range from 9-45, verbal aggression from 5-25, anger from 7-35, and hostility from 8-40. Total aggression score was calculated as a sum of the scores from each facet of aggression. In each case, a higher score represented higher levels of aggression (Buss & Perry 1992). A total of 606 men from YMCA 1 and 477 men from YMCA 2 completed the Buss and Perry questionnaires anonymously. A Cronbach's α coefficient (Vernon & Roberts 1981) was calculated as 0.77 and 0.84 for YMCA 1 and 2, respectively and confirmed these data had good internal consistency. The questionnaire used can be found at http://psychology-tools.com/buss-perry-aggression-questionnaire/.

Sex steroid levels, including total testosterone, DHEA-S, androstenedione, estrone and oestradiol were measured in 933 men from YMCA 1 (Tomaszewski *et al.* 2009a, Tomaszewski *et al.* 2009b). Circulating concentrations of sex steroids were measured using a 1,470 Wallac Wizard Gamma and commercially available radio-immunoassays (TESTO-RIA-CT, ESTRONE-RIA-CT, ANDROSTENEDIONE-RIA-CT, DHEAS-RIA-CT, all Biosource, Belgium). Each assay was highly specific and the cross-reactivity of a measured hormone with other sex steroids was minimal (generally below 1.0%). Consistent with the previous studies that examined oestradiol concentrations in male populations (Orwoll *et al.* 2006; Tivesten *et al.* 2006), an ultra-sensitive RIA kit designed specifically to measure oestradiol in men and children was used. The thresholds of detection of the assay used to evaluate oestradiol, estrone, total testosterone and androstenedione were 7.34 pmol/L, 11.84 pmol/L, 0.17 nmol/L, 0.03 ng/ml, respectively. On average, intra- and inter-assay coefficients of variation were 4.1% and 5.5% (total testosterone), 5.4% and 8.2% (total oestradiol), 6.3% and 8.6% (estrone), and 3.9% and 7.5% (androstenedione), respectively. These coefficients were comparable with indicators of intra- and inter-assay variation in other large-scale population studies that used RIA assays to measure sex steroids in men (Feldman *et al.* 2002; Orwoll *et al.* 2006).

4.2.2 Methodology

Details of DNA extraction and quality assessment, genotyping and Y chromosome haplogrouping can be found in Chapter 2.

4.2.2.1 Assessment of normality

All facets of aggression and each sex steroid were plotted and assessed by eye for normality. While all aggression measures were normally distributed, all sex steroids showed skewed distributions and were natural log-transformed before analysis. Potential differences in demographic/clinical profile between those who completed the aggression questionnaire and those who did not were compared using Student's *t* test. The consistency of aggression scores between YMCA 1 and YMCA 2 was also assessed using Student's *t* test.

4.2.2.2 Assessment of population stratification

Details on assessment of population admixture in both YMCA 1 and YMCA 2 can be found in Chapter 3.

4.2.2.3 Descriptive statistics

All normally distributed traits were summarised using arithmetic means and SD, while log-transformed data were described using geometric mean, which was calculated by multiplying each data point and dividing the total by the n^{th} root and 95% CI.

4.2.2.4 Genetic associations

Both crude and age-adjusted linear regression was used to assess association between haplogroup I and each facet of aggression (comparison of carriers of haplogroup I vs. all other haplogroups). The meta-analysis was conducted using a fixed-effect, inverse-variance model in Metan script (Deeks *et al.* 2001). Effect sizes from each study were

weighted inversely proportional to the variance of the effect size. The β coefficient, standard error and P-values reflect the magnitude of the effect size and statistical significance for the subjects from all cohorts of the examined associations.

Crude, age- and fully-adjusted (age, BMI, TC and HDL-C) linear regression was used to assess the association of haplogroup I of the Y chromosome and natural log-transformed sex steroids.

Power calculations based on 5% and 10% change in phenotypes were completed using http://www.statisticalsolutions.net/pss_calc.php.

Bonferroni correction was used to correct for multiple testing.

4.3 Results

4.3.1 Y chromosome haplogroup distribution

607 and 477 men recruited into YMCA 1 and YMCA 2, respectively, completed and returned the questionnaires with aggression assessment. Of these, 562 (YMCA 1) and 436 (YMCA 2) had DNA of sufficient quality and quantity for this analysis. Sex steroids (total testosterone, androstenedione, DHEA-S, oestradiol, estrone) were measured in a total of 933 individuals recruited into YMCA 1, of which 861 had DNA of sufficient quality and quantity for genetic resolution of the Y chromosome haplogroups.

Ten haplogroups were identified in YMCA 1 (percentage prevalence 0.4–52.7%), and ~73% of the lineages were accounted for by haplogroups I and R1a. A total of 11 haplogroups were found in YMCA 2 (percentage prevalence 0.2-59.6%), with ~77% of men assigned into either haplogroup I or R1a (Figure 4.1).

4.3.2 Assessment of phenotypic differences between those with and without available aggression phenotype

After correction for multiple testing, there were no clinical differences between men who completed the aggression questionnaire and those who did not in either YMCA 1 or YMCA 2 (Table 4.1).

Figure 4.1 Y chromosome haplogroup distributions in the subsets of YMCA 1 and YMCA 2 individuals used in Chapter 4. %- percentage prevalence. Haplogroups are arranged from [Y(xBR)] on the left to (R1b1b2) on the right. M and SRY numbers represent the terminal SNP dictating that haplogroup. SRY10831 is a recurrent mutation.

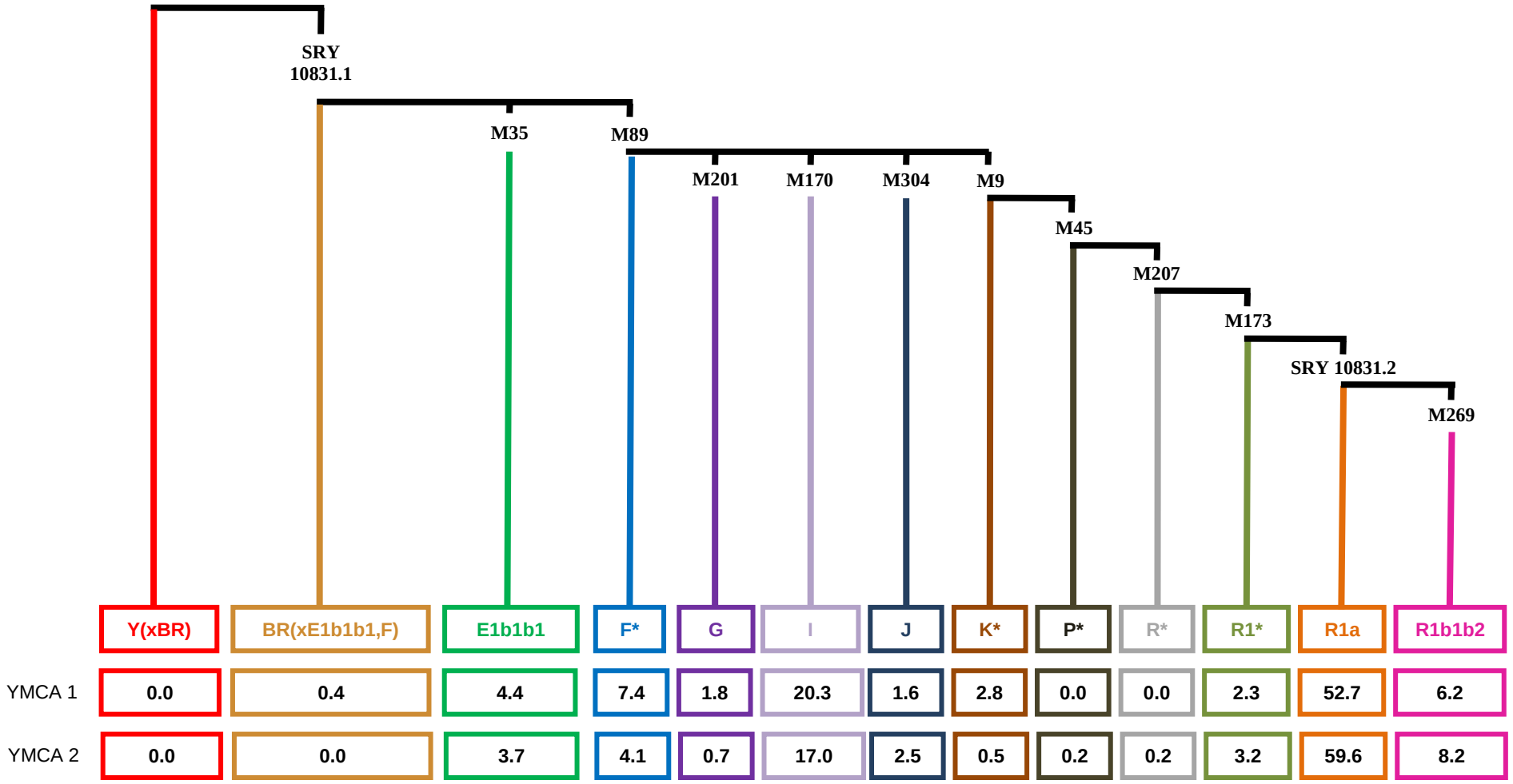


Table 4.1 Demographic and clinical characteristics of those who completed Buss and Perry questionnaire and those who did not in YMCA 1 and YMCA 2

Phenotype	Available	Unavailable	P value
YMCA 1			
N	607	550	-
Age (years)	18.9 (2.9)	19.7 (4.1)	0.089
BMI (kg/m^2)	22.6 (3.0)	23.1 (3.1)	0.215
Clinic SBP (mmHg)	118.7 (13.3)	117.3 (12.9)	0.345
Clinic DBP (mmHg)	74.6 (7.9)	73.8 (7.9)	0.517
TC (mmol/L)	4.1 (0.9)	4.3 (0.9)	0.033
HDL-C (mmol/L)	1.2 (0.3)	1.1 (0.3)	0.614
Creatinine (µmol/L)	82.0 (11.7)	83.8 (10.5)	0.753
Glucose (mmol/L)	4.9 (0.7)	4.7 (0.7)	0.402
YMCA 2			
N	477	120	-
Age (years)	18.9 (3.6)	19.4 (3.3)	0.226
BMI (kg/m^2)	22.5 (3.0)	23.1 (3.1)	0.861
Clinic SBP (mmHg)	118.6 (13.7)	122.4 (11.8)	0.369
Clinic DBP (mmHg)	74.4 (8.3)	74.3 (7.3)	0.388
TC (mmol/L)	3.7 (1.1)	3.8 (0.7)	0.932
HDL-C (mmol/L)	1.1 (0.3)	1.1 (0.2)	0.734
Creatinine (µmol/L)	74.2 (10.1)	79.6 (8.3)	0.548
Glucose (mmol/L)	4.4 (0.8)	4.5 (0.6)	0.390

Data are means and SD; p value-statistical significance of a difference between two groups calculated by a two tailed Student's *t* test, n – number of individuals, BMI – body mass index, SBP – systolic blood pressure, DBP – diastolic blood pressure, TC – total cholesterol, HDL-C – high-density lipoprotein cholesterol.

4.3.3 Difference in aggression between YMCA 1 and YMCA 2

There were no statistically significant differences in aggression scores between those recruited to YMCA 1 and YMCA 2 (Table 4.2).

Table 4.2 Mean values for five parameters of aggression in YMCA 1 and YMCA 2

Aggressive trait	YMCA 1	YMCA 2	P value
N	562	436	-
Physical aggression	24.5 (6.3)	24.9 (6.0)	0.285
Verbal aggression	16.0 (3.7)	16.2 (3.3)	0.387
Hostility	22.9 (5.4)	23.1 (5.2)	0.531
Anger	18.6 (5.7)	18.8 (4.7)	0.391
Total aggression	82.0 (15.4)	83.1 (14.1)	0.239

Data are means and SD, p value – statistical significance of a difference between two groups calculated by a two-tailed Student's *t* test.

4.3.4 Analysis of population stratification

There was no evidence of population stratification in the subset of YMCA 1 or YMCA 2 individuals available for this analysis. The details of this analysis can be found in Chapter 3 (Figure 3.3, and section 3.3.2).

4.3.5 Phenotype distribution

All aggression measures in both YMCA 1 and YMCA 2 showed approximately normal distributions (Figure 4.2 and 4.3). All sex steroid distributions were skewed and therefore were log-transformed before analysis (Figure 4.4).

Figure 4.2 Distribution of aggression scores (YMCA1)

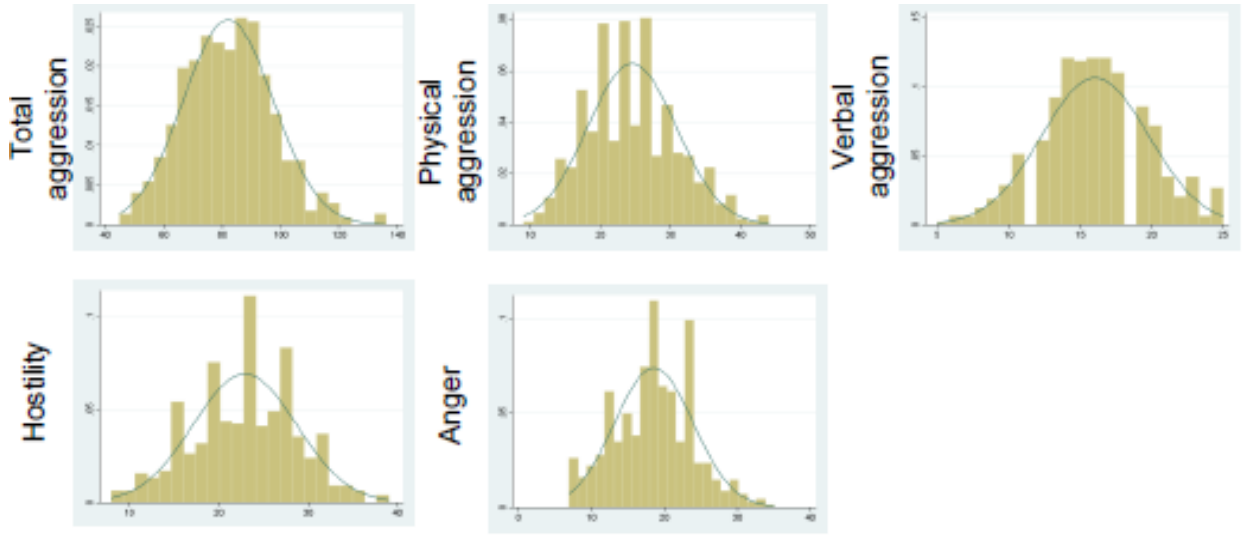


Figure 4.3 Distribution of aggression scores (YMCA2)

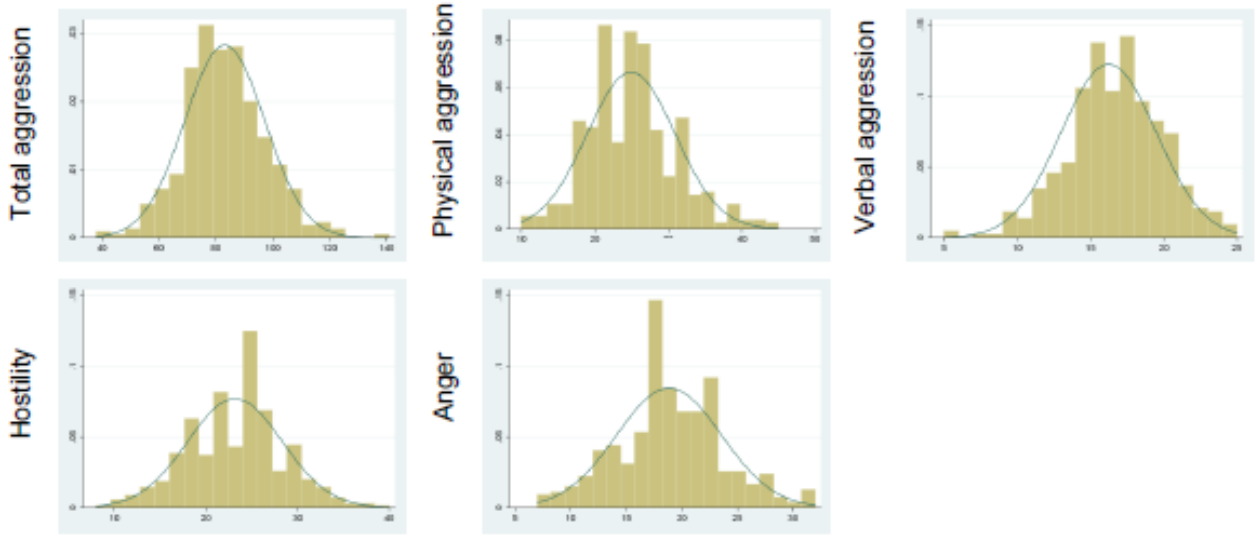


Figure 4.4 Distribution of sex steroids levels in YMCA 1

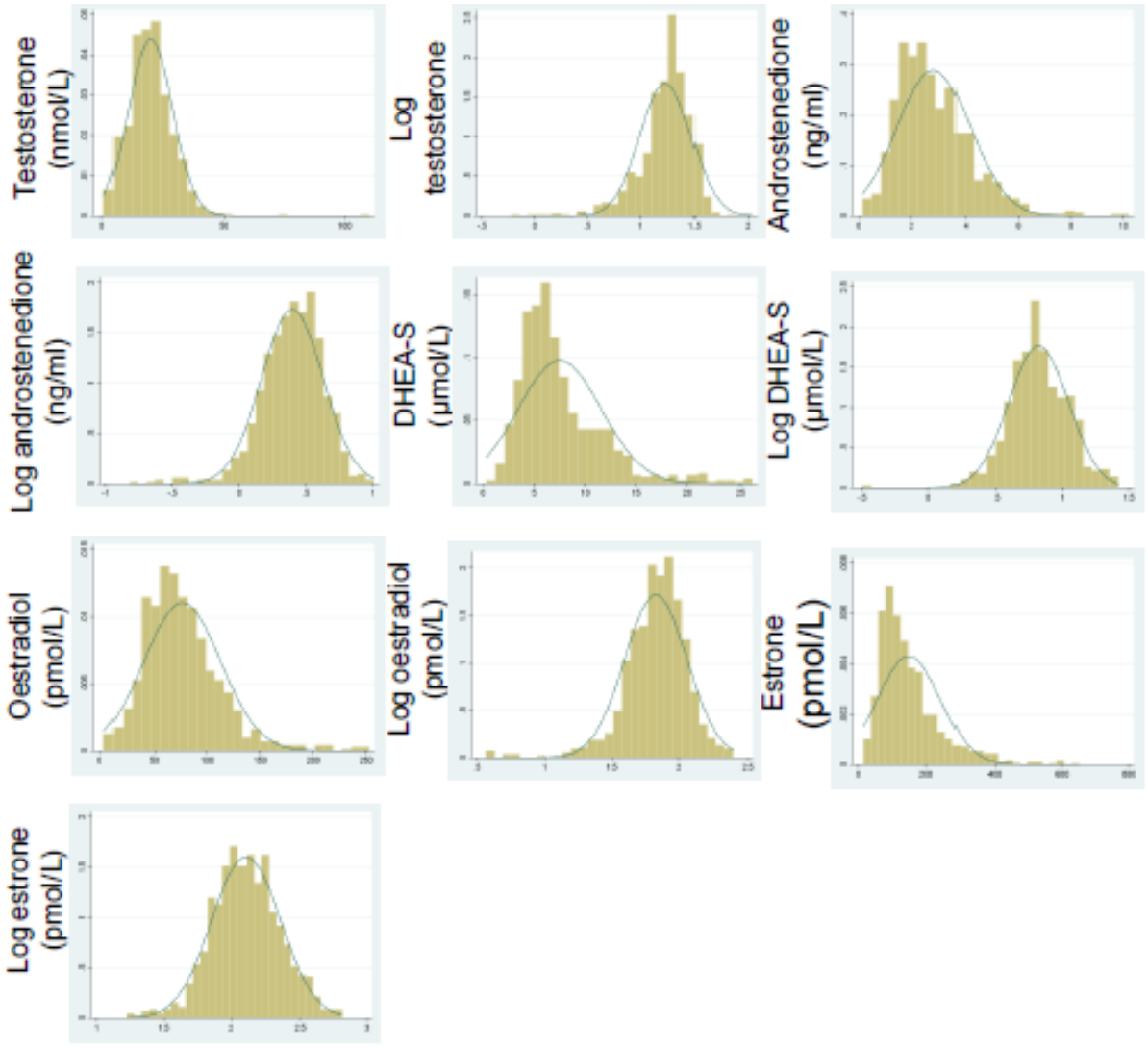


Table 4.3 Statistical power calculations for aggression and sex steroids data

Phenotype	**Statistical power**	
	5% Change	10% Change
Aggression		
Physical aggression	0.83	1.00
Verbal aggression	0.90	1.00
Anger	0.75	1.00
Hostility	0.85	1.00
Total aggression	0.98	1.00
Sex steroids		
Oestradiol	0.25	0.73
Androstenedione	0.25	0.73
Estrone	0.17	0.52
DHEA-S	0.21	0.63
Total testosterone	0.28	0.79

Statistical power calculations for detection of a 5% or 10% change in aggression (n=1,124) and sex steroids (n=861) values between carriers of haplogroup I versus all other haplogroups (at the 5% and 10% significance level). Calculations were performed using an online tool (http://www.statisticalsolutions.net/pss_calc.php)

4.3.6 Effect of the Y chromosome on aggression scores

While no evidence of association was identified between haplogroup I of the Y chromosome and any facet of aggression in YMCA 1, two suggestive associations were identified in YMCA 2 [P= 0.078 and P= 0.067 in physical aggression (Table 4.5) and verbal aggression, respectively (Table 4.6)]. The meta-analysis of values from YMCA 1 and YMCA 2 showed that haplogroup I was associated (at the nominal level of statistical significance) with reduced verbal aggression scores (P= 0.016). However, this finding did not survive correction for multiple testing (Table 4.6). No evidence of heterogeneity was identified between both cohorts for total aggression (P= 0.674), physical aggression (P= 0.365), verbal aggression (P= 0.762), hostility (P= 0.530) or anger (P= 0.548) (Tables 4.4- 4.8).

Table 4.4 Age-adjusted fixed-effect inverse-variance meta-analysis between haplogroup I of the Y chromosome and total aggression

Cohort	n	β coefficient (SE)	P value	Weight (%)	Heterogeneity P value
YMCA 1	562	-1.370 (1.7)	0.433	57.72	-
YMCA 2	436	-2.498 (2.0)	0.221	42.26	-
Meta-analysis	998	-1.85 (1.9)	0.164	100	0.674

SE- standard error.

Table 4.5 Age-adjusted fixed-effect inverse-variance meta-analysis between haplogroup I of the Y chromosome and physical aggression

Cohort	n	β coefficient (SE)	P value	Weight (%)	Heterogeneity P value
YMCA 1	562	-0.520 (0.7)	0.459	60.35	-
YMCA 2	436	-1.529 (0.9)	0.078	39.65	-
Meta-analysis	998	-0.92 (0.8)	0.092	100	0.365

SE- standard error.

Table 4.6 Age-adjusted fixed-effect inverse-variance meta-analysis between haplogroup I of the Y chromosome and verbal aggression

Cohort	n	β coefficient (SE)	P value	Weight (%)	Heterogeneity P value
YMCA 1	562	-0.672 (0.4)	0.111	56.48	-
YMCA 2	436	-0.863 (0.5)	0.067	44.52	-
Meta-analysis	998	-0.760 (0.4)	0.016	100	0.762

SE- standard error.

Table 4.7 Age-adjusted fixed-effect inverse-variance meta-analysis between haplogroup I of the Y chromosome and hostility

Cohort	n	β coefficient (SE)	P value	Weight (%)	Heterogeneity P value
YMCA 1	562	0.225 (0.6)	0.728	57.22	-
YMCA 2	436	0.845 (0.7)	0.258	42.78	-
Meta-analysis	998	0.490 (0.6)	0.315	100	0.530

SE- standard error.

Table 4.8 Age-adjusted fixed-effect inverse-variance meta-analysis between haplogroup I of the Y chromosome and anger

Cohort	n	β coefficient (SE)	P value	Weight (%)	Heterogeneity P value
YMCA 1	562	-0.403 (0.6)	0.505	56.12	-
YMCA 2	436	-0.950 (0.7)	0.164	43.88	-
Meta-analysis	998	-0.640 (0.6)	0.155	100	0.548

SE- standard error.

4.3.7 Effect of the Y chromosome on sex steroids – YMCA 1

There was no evidence of association between haplogroup I and sex steroids in men from YMCA 1. The lowest fully adjusted nominal association (P= 0.170) was calculated for oestradiol (Table 4.9).

Table 4.9 Association between haplogroup I of the Y chromosome and 5 major sex steroids – YMCA 1

Sex steroids	I	Others	P value $^{\alpha}$	P value $^{\gamma}$	P value $^{\Lambda}$
n	153	708	-	-	-
Total testosterone	17.1 (16.5- 17.8)	17.8 (16.5- 19.2)	0.351	0.388	0.506
Androstenedione	2.5 (2.4- 2.6)	2.5 (2.3- 2.7)	0.834	0.912	0.989
DHEA-S (µmol/L)	6.6 (6.3- 6.8)	6.7 (6.1- 7.3)	0.701	0.665	0.601
Oestradiol (pmol/L)	67.6 (65.2- 70.0)	70.2 (65.4- 75.5)	0.323	0.293	0.170
Estrone (pmol/L)	127.7 (122.8- 132.7)	126.9 (114.9- 140.2)	0.895	0.977	0.856

Data are geometric means and 95% confidence intervals, p value$^{\alpha}$ – level of crude statistical significance, p value$^{\gamma}$– age-adjusted level of statistical significance calculated from regression analysis, P value $^{\Delta}$– age, BMI, total cholesterol and HDL-C adjusted level of statistical significance, DHEA-S- dehydroepiandrosterone sulphate.

4.4 Discussion

To identify the mechanism of association between haplogroup I of the Y chromosome and CAD, the effect of this haplogroup on "sex-associated" and "sex-specific" phenotypes was investigated in a series of experiments highlighted in this chapter. No relationship between the Y chromosome and any facet of aggression or sex steroid was identified. These data indicate that these traits are unlikely to drive the association between haplogroup I and increased predisposition to CAD.

4.4.1 Y chromosome haplogroup distribution

The prevalence and distribution of Y chromosome haplogroups in both YMCA 1 and YMCA 2, including the high prevalence of haplogroup R1a, were in keeping with published data.There was no evidence of population stratification in YMCA 1 or YMCA 2 (Knowler *et al.* 1988; Lander & Schork 1994; Ewens & Spielman 1995). This is reassuring because population stratification is known to inflate the false positive rate. Furthermore, no statistically significant differences in demographic or clinical phenotypes by haplogroup were identified in either YMCA 1 or YMCA 2 (further details in Chapter 3).

There were also no systematic differences between individuals with available aggression data and those without such data in YMCA 1 or YMCA 2. Cronbach's α coefficient showed that the assessment of aggression had good internal consistency in both cohorts. There were also no differences in aggression scores of those recruited into YMCA 1 compared with YMCA 2. These data imply that the cohorts were not subject to recruitment bias and were phenotypically homogenous.

4.4.2 The effect of haplogroup I on aggression levels

The meta-analysis of five facets of aggression across the two cohorts revealed only a suggestive association between haplogroup I of the Y chromosome and verbal aggression, which was not significant after correction for multiple testing. The role of the Y chromosome in personality has been indicated in a number of studies (Lovell-Badge 2005). For example, men with an 47,XYY karyotype had more criminal convictions, were diagnosed more frequently with antisocial personality disorder or were alcoholic more commonly than men with a normal karyotype (Götz *et al.* 1999). This finding suggested a role for the Y chromosome in personality phenotypes such as aggression.

The results from YMCA 1 and YMCA 2 are is in agreement with Shah *et al.* who investigated 159 Pakistani men and found no association between the Y chromosome and aggression levels (Shah *et al.* 2009). Though haplogroup I of the Y chromosome was not individually investigated in this study (due to its absence in the study population), the lack of association between other Y chromosome haplogroups and aggression adds to evidence that the genetic variation in the human Y chromosome is not associated with human aggression (Shah *et al.* 2009). Similarly, a study using Y chromosome STRs near the *DYZ3* loci also failed to demonstrate its association with antisocial personality disorder in men with alcoholism (Kittles *et al.* 1999).

There are significant differences in methods between these studies that may account for some of the discrepancy in association between the Y chromosome and aggression. Firstly, a number of different phenotypes were used across the studies. Aggression was used as the main phenotype in the negative study shown here alongside studies by Shah *et al.* and Gatewood *et al.*. Conversely, Kittles *et al.* and Götz *et al.*

used antisocial personality disorder as the main phenotype. Studies in murine models using reported association with the Y chromosome and aggression (Gatewood *et al.* 2006). Indeed, it is plausible that the Y chromosome is involved in determining aggression level in rodents but not humans.

It should also be stressed that studies that reported a positive association investigated either the whole Y chromosome or Y chromosomes with specific deletions (human studies; Y chromosome polysomy [Götz *et al.* 1999]; murine models - *SRY* negative Y chromosomes [Gatewood *et al.* 2006, Toot *et al.* 2004]). Conversely, all negative studies investigated only the MSY (through haplogroups such as the approach used here and by Shah *et al.* or through analysis of certain aspects of Y chromosome variation including STRs) [Kittles *et al.* 1999]). It is plausible that a gene or transcript in the PAR regions may underlie the association between the Y chromosome polysomy and antisocial behaviour (Götz *et al.* 1999) and aggression (Gatewood *et al.* 2006, Toot *et al.* 2004). Previous work in murine models has indicated that steroid sulfatase (*STS*), a gene which resides in the PAR region, correlated with aggression scores in mice (Le Roy *et al.* 1999; Gatewood *et al.* 2006). The orthologous gene cannot underlie the association between Y chromosome polysomy and antisocial/aggressive phenotypes in men, however, due to the degradation of *STS* to a pseudogene in humans (Yen *et al.* 1987).

4.4.3 The effect of haplogroup I on sex steroid levels

No evidence of association between haplogroup and sex steroids was identified. This finding is important for several reasons. Firstly, there is general assumption in the wider scientific community that the human Y chromosome is a major regulator of sex

hormones, particularly androgens, through its role in sex differentiation (Angelopoulou *et al.* 2006; Velasco *et al.* 2011). There are data that, to some extent, support this hypothesis. Aksglaede *et al.* found that after puberty, the complement of sex chromosomes in boys may play a role in regulation of testosterone levels (Aksglaede *et al.* 2007) since individuals with XXY and XX combinations had significantly lower average testosterone compared to men with a normal sex chromosome constellation (XY) (Aksglaede *et al.* 2007). Interestingly, a microdeletion in the azoospermia factor c *(AZFc)* region is associated with increased levels of testosterone (alongside follicle stimulating hormone [FSH] and luteinizing hormone [LH]) and infertility (Pandey *et al.* 2010). These investigations analysed the sex chromosomes using a significantly different method to that used here; Aksglaede *et al.* investigated the effect of XY polysomy, while Pandey *et al.* investigated specific Y chromosome variation (Pandey *et al.* 2010). The study presented here represents the first phylogenetic analysis of the Y chromosome haplogroup in relation to sex steroid levels.

4.4.4 Conclusions

Phenotypes previously perceived as "male-associated" or "sex-specific", including aggression and sex steroids, are not associated with haplogroup I of the Y chromosome in the cohorts used here. As such these phenotypes are unlikely to underlie the association between haplogroup I of the Y chromosome and increased risk of CAD (Charchar *et al.* 2012).

4.4.5 Future work

The findings from this analysis derive from cohorts recruited in one country (Poland). As such they will require replication in independent cohorts of different origin.

Furthermore, an investigation into other common European haplogroups with regard to these phenotypes may prove fruitful.

It is also plausible that other "male-associated"/ "sex-specific" phenotypes not investigated here may underlie the association between haplogroup I of the Y chromosome and CAD, and should be investigated. Examples of other such sexually dimorphic biological phenotypes include, for example, height and muscle mass.

4.4.6 Major limitations

4.4.6.1 Y chromosome analysis

It is plausible, though unlikely, that a smaller lineage of haplogroup I may affect aggression or sex steroid levels (further details can be found in Chapter 3.4.5.1). However, due to limited sample numbers in each individual haplogroup I lineage, this analysis was not able to be completed here.

4.4.6.2 Cohorts

The decision was made to investigate the potential role of the Y chromosome on "male-associated"/ "sex-specific" traits in young, healthy individuals. Our analysis provides significant analysis into the role of the Y chromosome in various traits which were not subject to confounding by external forces such as antiypertensive treatment. While this study was sufficiently powered to detect small changes in aggression score, only larger changes in sex steroids would have been identified. To rectify this limitation, larger sample sets are required, however, at the time of writing these were not available.

4.4.6.3 Ethnicity

The original association between the Y chromosome and CAD was identified in British cohorts. Therefore, the use of Polish cohorts may be perceived as a limitation. An investigation using British cohorts with aggression and/ or sex steroids data would be more optimal. Unfortunately, such cohorts were not available during the duration of this project. As explained in Chapter 3, the prevalences of haplogroup I in Britain and Poland are very similar and thus the difference in cohort ethnicity is not expected to be a significant limitation (Semino *et al.* 2000; Rootsi *et al.* 2004). Further information can be found in Section 8.5.2.

4.4.6.4 Sex steroid analysis

The five sex steroids investigated here represent the major biochemical products of the sex steroid pathway. It is plausible that haplogroup I may exert its effect on sex steroids (and therefore predisposition to CAD) elsewhere in the pathway and therefore not be captured in this analysis. The unavailability of sex hormone binding globulin (SHBG) levels prevented the analysis of free, biologically active hormonal fractions. The analysis of the free testosterone and oestradiol, would increase the reliability of the analysis (Tomaszewski *et al.* 2009). Similarly, the effect of fat mass on hormone level is widely accepted, and while the use of BMI provides a good estimate of fat mass in young, apparently healthy men (Pietrobelli *et al.* 1998; Tomaszewski *et al.* 2009), waist circumference would have been a better estimate of abdominal obesity (relevant to sex steroids analysis as a site of aromatase synthesis) (Pietrobelli *et al.* 1998; Tomaszewski *et al.* 2009).

General limitations of radio-immunoassays (which were used to measure all sex steroids used here) include hook effects (resulting from high antigen concentration), heterophilic antibodies, human anti-mouse antibodies and cross reactivity. However, this method of sex steroid measurement is widely accepted within the field (Feldman *et al.* 2002; Orwoll *et al.* 2006; Tivesten *et al.* 2006).

5 Effect of haplogroup I of the Y chromosome on the monocyte and macrophage transcriptome

5.1 Introduction

It has been demonstrated that haplogroup I increases the risk of CAD (Charchar *et al.* 2012). This association is unlikely to act through traditional cardiovascular risk factors or "male-typical" or "sex specific" phenotypes (Chapters 3 and 4), yet the mechanisms underlying this association remain elusive. Genome-wide transcriptome data provide a snap-shot of the expression level of multiple genes in a given cell type (Allocco *et al.* 2004). Through the parallel measurement of expressed genes, genome-wide transcriptome data allow hypothesis-free experiments (including differential expression analysis, disease association and pathway analysis) to be conducted (Allocco *et al.* 2004). Genome-wide transcriptome technology was originally developed from Southern blotting, in which fragmented DNA is attached to a substrate and probed (Makos *et al.* 1992). The first reported use of this technology investigated 378 bacterial colonies in analysis of gene expression levels of normal and cancer genes (Augenlicht *et al.* 1982). This was swiftly followed up by the analysis of over 4,000 human sequences comparing colonic cancer cells with healthy cells (Augenlicht *et al.* 1987; Augenlicht *et al.* 1991). While early arrays were formed by spotting cDNAs onto filter paper using a pin spotting device, the first reported use of miniaturised "chip" arrays was in 1995 (Schena *et al.* 1995). This method of study became widely used; the search term "transcriptome analysis" in PubMed retrieves over 75,000 papers at the time of writing. Importantly, analysis of genome-wide transcriptome data has the potential to identify if haplogroup I of the Y chromosome alters the expression of a single gene or the expression of an entire pathway in a given cell type. Pathway analysis integrates normalised data from transcriptomic arrays with gene annotations in order to detect changes in the expression

of networks of molecules (Curtis *et al.* 2005). Pathway analysis ranges from the simple use of gene ontology to those based on complex statistical analysis.

The aim of this chapter is to conduct a hypothesis-free investigation, using genome-wide transcriptome data from cells relevant to CAD (monocytes and macrophages), in order to identify which genes or pathways (through the use of linear regression analysis and gene set enrichment analysis [GSEA]) show differential expression in carriers of haplogroup I compared to carriers of all other Y chromosome lineages.

5.2 Materials and methodology

5.2.1 Materials

5.2.1.1 CARDIOGENICS Transcriptomic Study (*CTS*)

The "CARDIOGENICS Transcriptomic Study" (CTS) was a multi-centre collaboration that recruited 917 (458 healthy and 459 CAD/MI patients) individuals from five European centres (Schunkert *et al.* 2011). Patients recruited in Leicester (n=161), Lübeck (n=102) and Regensburg (n=122) all had validated MI within the previous 6 months. Cases recruited in Paris (n=74) were angiographically diagnosed with CAD. All healthy individuals were blood donors recruited in Cambridge (UK) as a part of the Cambridge Bioresource. Medical history of patients with CAD/MI was taken alongside BP measurements using digital BP monitors (with an appropriately sized cuff after a minimum of 10 min rest and with an interval of at least 3 min between readings). A sample of peripheral blood was collected for all individuals for monocyte isolation and culture (for further DNA and RNA extraction and baseline blood chemistry [cases only]).

5.2.1.2 Bioethical approval

Written, informed consent was obtained from all subjects in all studies used here, in accordance with the Declaration of Helsinki. All studies had approval from relevant institutional ethical committees.

5.2.2 Methodology

5.2.2.1 Genotyping and haplogrouping

All details regarding genotyping and Y chromosome haplogrouping of all samples can be found in Chapter 2.

5.2.2.2 Descriptive statistics

Age, BMI, weight, height, SBP and DBP were summarised using arithmetic means and SD. Differences between traits between carriers of haplogroup I and all other haplogroups were analysed using Student's *t* test.

5.2.2.3 Population stratification

Population stratification was examined in a manor akin to that described in Chapter 3 using the GRAPHIC cohorts. Population stratification was investigated using multi-dimensional scaling of genome-wide SNP data in PLINK. Genome-wide SNP data were obtained using the Human 610 Quad Custom array (Illumina, USA). SNPs failing standard quality filters [minor allele frequency - MAF<1%), Hardy-Weinberg equilibrium $\chi 2$ $p<0{\cdot}001$ and call rate <95%] were removed. This was followed by pruning SNPs in LD ($r^2 \geq 0{\cdot}2$) to leave approximately 35,000 SNPs. The pruned set of SNPs was used to construct a matrix of identity-by-state an used in multi-dimensional scaling (further detail in Chapter 3).

5.2.2.4 Measurement of genome-wide transcriptome levels

Transcriptome-wide analysis was conducted using RNA extracted from macrophages and monocytes for each individual. All experiments described in this study were completed using only male patients recruited in Leicester and controls from Cambridge (n=255). Genomic DNA was collected from peripheral monocytes using standardised Qiagen kits and genotyped either at the Wellcome Trust Sanger Institute on the Human 610 Quad Custom Array (594,398 SNPs and 66,049 CNVs), or the SNP&SEQ Technology Platform at Uppsala University, using the Sentrix Human Custom 1.2 M array (1,115,839 SNPs and 80,128 CNVs) (Rotival *et al.* 2011; Wallace *et al.* 2012). RNA extraction was also standardised; monocytes were isolated from 30 ml of peripheral blood using positive selection with CD14+ microbeads (Miltenyi), as specified in the manufacturer's instructions (purity was measured using flow cytometry). Half of the monocytes underwent RNA extraction (monocytes were lysed in Trizol, RNA was extracted in chloroform and ethanol, washed in RNeasy columns, and incubated with DNase I before extracting in RNase-free water) (Johnson *et al.* 2011; Rotival *et al.* 2011; Wallace *et al.* 2012). The other half of the monocytes were activated to macrophages using macrophage colony stimulating factor (M-CSF) and RNA extracted using the same method (Johnson *et al.* 2011; Rotival *et al.* 2011; Wallace *et al.* 2012). Each sample was analysed using the Illumina Human Ref 8 v3.0 bead chip array. The array contains 79-nucleotide oligonucleotide sequences immobilised to beads held in wells on the surface of the array (van Gelder *et al.* 1990; Pease *et al.* 1994; Nuwaysir *et al.* 2002). The 5′ 50 bases of the oligonucleotides had sequences complementary to known gene sequences (available in public databases). The remaining 3′ region of the probe represented a molecular "address" allowing unambiguous identification of each oligonucleotide during analysis (van Gelder *et al.*

1990). Probes were designed so as to ensure; a lack similarity to other genes, an absence of highly repetitive sequences, correct sequence complexity, self-complementarily for hairpin structure prediction, appropriate melting temperature and uniform distance from the 3´ region of the transcript under analysis. Each probe was present on the array in many copies, allowing significant redundancy (van Gelder *et al.* 1990). After hybridisation, array images were scanned using the Illumina BeadArray Reader, and probe intensities were extracted using the Gene expression module (version 3.3.8) of the Illumina BeadStudio software (version 3.1.30) (Johnson *et al.* 2011; Rotival *et al.* 2011; Wallace *et al.* 2012). Raw intensities were processed using the lumi (Du *et al.* 2008; Lin *et al.* 2008) and beadarray (Welcome Trust Case-control Consortium 2007) packages in R (R Development Core Team 2010). All array outliers were excluded and only arrays with high concordance in terms of gene expression measures (pair-wise Spearman correlation coefficients within each cell type 0.85) were included in the analyses (Johnson *et al.* 2011; Rotival *et al.* 2011; Wallace *et al.* 2012).

5.2.2.5 Single-gene analysis of genome-wide transcriptome data

The effect of haplogroup I on genome-wide gene expression levels in both monocytes and macrophages was investigated using linear regression (adjusted for age, CAD status and centre of recruitment). The correction for multiple testing was calculated using false-discovery rate (FDR) [QVALUE software (http://genomine.org/qvalue)] with a threshold of 0.2 to identify suggestive associations.

5.2.2.6 Validation of gene(s) from transcriptome-wide analysis

Gene(s) showing significant association with haplogroup I were replicated in 48 available macrophage samples. The significant drop in sample number was unavoidable because of severely restricted RNA availability.

Gene expression assays measure the relative concentration of cDNA (from mRNA as described in Chapter 2) in any given tissue. In all cases, FAM-labelled TaqMan probes were used for each gene of interest and normalised by the expression of the housekeeping gene β-microglobulin (*B2M*) (control gene) (VIC-labelled). Analysis was completed on the 7900HT Sequence Detection System (Applied Biosystems, USA). Control genes (also known as endogenous controls) should have a constant RNA transcription level across the different experimental conditions (in this case those with haplogroup I and all other haplogroups), and be sufficiently abundant across the different research tissues (in this case monocyte, macrophage) (Popovici *et al.* 2009). While any stably expressed gene can be selected as a control gene, the most common choice is from a range of housekeeping genes (involved in pathways such as glycolysis, protein folding and the cytoskeleton) or ribosomal RNAs (including 18S rRNA). While there are other common choices of control genes (including ribosomal protein, beta actin, peptidylpropyl isomerase and glucuronidase β), *B2M* is widely accepted to have good expression levels in a number of tissues, including those used here.

Due to the sensitivity of gene expression assays to minor changes in experiment set-up, it is recommended that all samples be analysed in triplicate. However, due to the scarcity of RNA input all samples were analysed in duplicate.

The contig sequence (sequence where the probe binds) for the *IFI6* probe used here was: CTGCAGAGCCTCGGGGCTGGTGGCA. Probes of interest were assessed for correct functioning using a serial dilution of cDNA input. 11 µl of TaqMan gene expression master-mix (Applied Biosystems, USA) was added to 1 µl probe of interest or control gene and 9 µl of cDNA (one reaction at each concentration of the cDNA serial dilution; 80 ng/µl, 40 ng/µl, 20 ng/µl, 10 ng/µl, 5 ng/µl). This was incubated following the protocol: denaturing at 50^{o}C for 2 min, further denaturing at 95 ^{o}C for 10 min, 50 cycles of 95 ^{o}C for 15 sec each and 60 ^{o}C for 1 min. A standard line graph of delta cycle threshold (dCT) against log of cDNA input was constructed. All probes were validated and showed an efficiency of >80%.

5.2.2.9 Analysis of the effect of haplogroup I on expression levels of interferon 6 (IFI6) in macrophage tissues

Gene expression analysis of *IFI6* followed the same protocol outlined above although at one cDNA concentration (20 ng/µl), in duplex (both probe of interest and control gene in one cDNA sample) and in duplicate. A negative control (containing water rather than cDNA) and two positive controls of known RNA abundance were used. Data were expressed as cycle threshold (CT) and dCT (average probe of interest minus average control gene). All individuals were split into two groups based on Y chromosome haplogroup status (I vs. all others), and the fold-difference in mean expression was calculated using the formula $=2^{-\text{difference in dCt}}$.

5.2.2.10 Pathway analysis- Gene Set Enrichment Analysis

GSEA was used to identify pathways differentially regulated in carriers of haplogroup I compared to all other haplogroups. GSEA evaluates a gene list, ordered by effect size (e.g. based on β coefficients), at the pathway level. GSEA aims to establish if an input gene set tends towards the top (up-regulated) or bottom (down-regulated) of a pathway. Input files were generated using effect size scores (β coefficient from linear regression), one score per probe. These input data were used in GSEA "pre-ranked" mode (Subramanian *et al.* 2005). Data were analysed using Kyoto Encyclopaedia of Genes and Genomes (KEGG) contained within the Molecular Signatures Database (v3.0) (a collection of annotated gene sets) and 1,000 permutations. This analysis provides a number of measures of association including; an enrichment score, normalised enrichment score, P-value and FDR. The enrichment score is calculated as a running statistic; as GSEA reaches a gene included in the pathway the score is increased, as GSEA reaches a gene not in the pathway the score is decreased (the increment increased or decreased depends on the correlation of the gene with the phenotype (Subramanian *et al.* 2005). The normalised enrichment score is the enrichment score normalised for the size of the pathways investigated (normalised enrichment score= actual enrichment score/ mean enrichment score against all permutations of these dataset) (Subramanian *et al.* 2005). P-values illustrate the level of nominal significance of the difference in expression of a given pathway between two groups.

Enrichment Map (v1.0) (Merico *et al.* 2010) plugin for Cytoscape (v2.8.0) (Cline *et al.* 2007) was used to analyse and pictorially represent the degree of gene overlap in those pathways significantly deregulated in carriers of haplogroup I (P= 0.99 FDR<0.2, overlap coefficient >0.3).

5.3 Results

5.3.1 Y chromosome haplogroup distribution in CTS

A subset of the CTS cohort that consisted of men recruited in Leicester and Cambridge were used in this investigation (n=255) (Charchar *et al.* 2012). Of the ten haplogroups represented in the CTS subset (haplogroup prevalence ranging from 0.4-63%), haplogroups I and R1b1b2 accounted for approximately 85% of the observed lineages (21% haplogroup I, 63% haplogroup R1b1b2) (Figure 5.1).

5.3.2 Clinical and demographic characteristics of CTS by Y chromosome haplogroup

There were no statistically significant differences in clinical or demographic data between those with haplogroup I and those from all other Y chromosome lineages (Table 5.1).

Figure 5.1 Y chromosome haplogroup distributions in CTS. Percentages-prevalence of each haplogroup represented in CTS. Haplogroups are arranged from [Y(xBR)] on the left to (R1b1b2) on the right. M and SRY numbers represent the terminal SNPs that define specific haplogroups. SRY 10831 is a recurrent mutation.

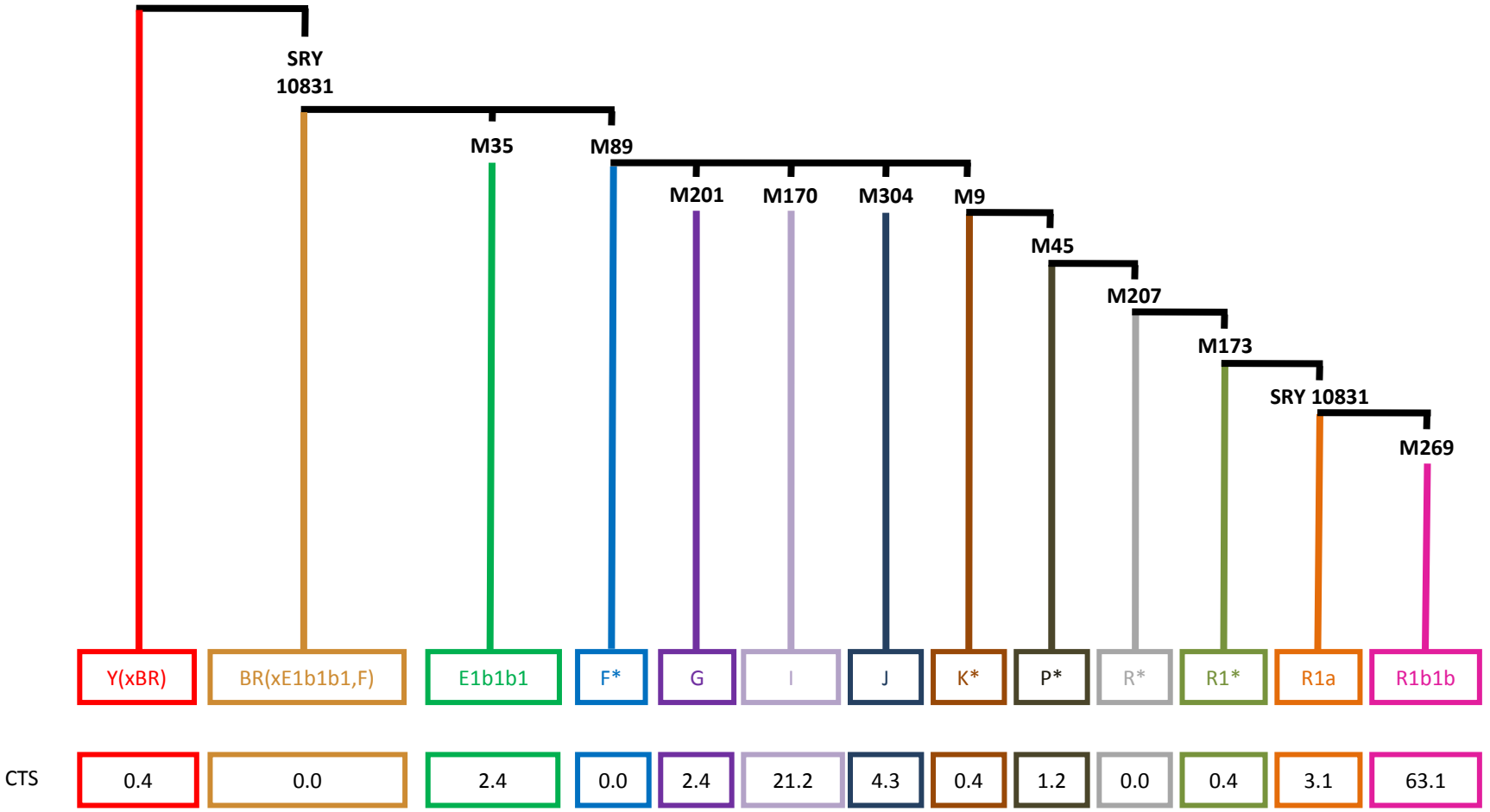


Table 5.1 Clinical and demographic data from individuals in CTS stratified by haplogroup

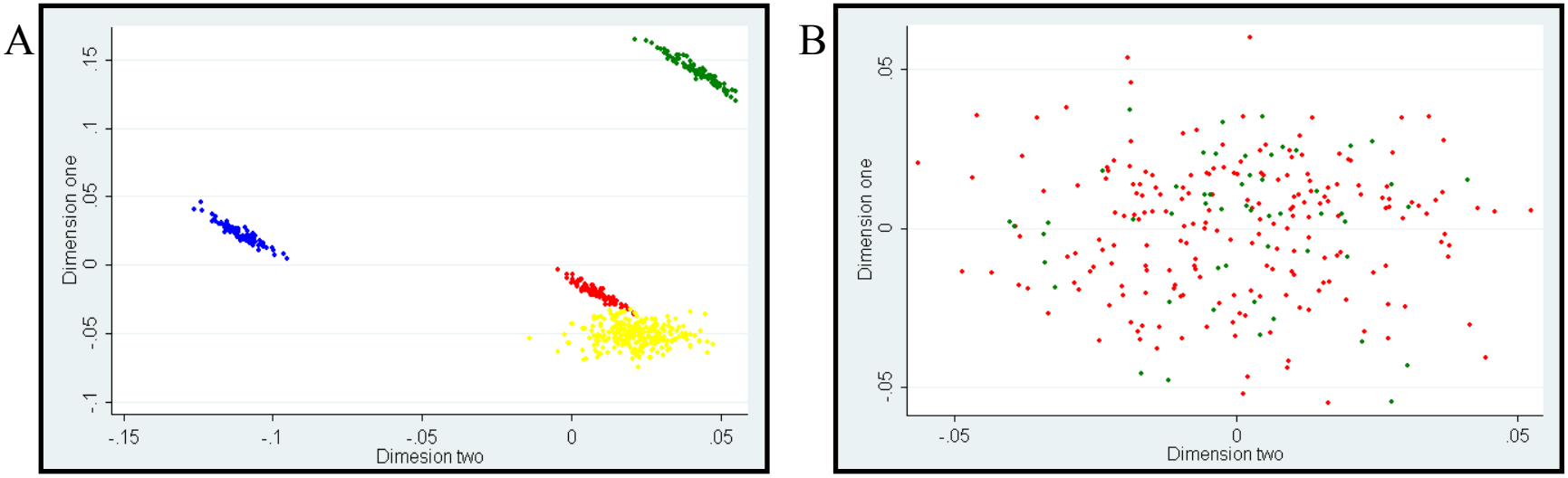
Phenotype	I	Others	P value
N	54	201	-
Age (years)	54.2 (7.5)	54.3 (7.1)	0.138
BMI (kg/m^2)	27.4 (3.9)	27.5 (4.5)	0.971
Weight (kg)	86.5 (12.7)	86.3 (13.8)	0.966
Height (m)	177.9 (6.3)	177.5 (6.1)	0.438
SBP (mmHg)	127.6 (17.5)	135.6 (18.7)	0.939
DBP (mmHg)	77.4 (7.9)	82.7 (10.6)	0.138

Data are means (SD). P value-level of statistical significance for a difference in a phenotype between carriers of haplogroup I and all other haplogroups calculated by a two-tailed Student's *t* test, I – carriers or haplogroup I, Others – carriers of all other haplogroups, n – number of subjects, BMI – body mass index, SBP – systolic blood pressure, DBP – diastolic blood pressure.

5.3.3 Analysis of population stratification

All CTS subjects were of European ancestry, as shown by the appearance of the CTS subjects in the same cluster as the CEU HapMap population (Figure 5.1A). There were no systematic difference in autosomal SNP distribution between carriers of haplogroup I and men from all other lineages, as shown by the complete interspersion of data points from these groups (Figure 5.2B).

Figure 5.2 Analysis of population stratification in CTS. A. A matrix of identity-by-state distance was created in the CTS subset and the three HapMap populations (CEU, JPT+CHB, YRI). Coloured dots represent each population: blue- JPT+CHB, green- YRI, red- CEU, yellow- CTS. The first two components obtained from the subsequent multi-dimensional scaling of the matrix were used as co-ordinates to compare the relative distances between the four populations. B. Assessment of population structure stratified by Y chromosome haplogroup (I vs. all others). Colours represent haplogroups: green- carriers of haplogroup I, red- carriers of all other haplogroups. A matrix of identity-by-state distances was created and the first two components obtained from the subsequent multi-dimensional scaling of the matrix were used as co-ordinates to generate the scatter plot.



5.3.4 Genome-wide transcriptome analysis

Of 24,526 probes present on the bead chip, 12,480 were either absent in both monocytes and macrophages or failed quality checks, leaving 12,145 probes (representing 10,082 genes) available for study. Both single-gene and pathway analysis was completed using this dataset.

5.3.4.1 Genome-wide single-gene transcriptome analysis in monocytes

No single gene showed a differential expression between carriers of haplogroup I and all other haplogroups in monocytes at the genome-wide level of significance (10^{-7}). However, one gene, cytoskeleton associated protein 5 (*CKAP5*), showed a nominal level of differential expression (Table 5.2).

5.3.4.2 Genome-wide single-gene transcriptome analysis in macrophages

While no genes reached genome-wide significance, 33 genes showed a nominal level of differential expression between carriers of haplogroup I and carriers of all other haplogroups in macrophages (Table 5.3). *IFI6* showed the most statistically significant up-regulation in carriers of haplogroup I (P= 0.000059). Approximately 20% of the differentially regulated genes were associated with the immune system (i.e. integrin αL *[ITGAL]*, butyrophilin subfamily 3A1 *[BTN3A1]* major histocompatibility complex class 1A *[HLA-A]*, heat shock protein 1 *[HSPD1]*) (Table 5.3). A number of genes relating to protein synthesis and the cell-cycle were also differentially regulated (Table 5.3).

Eleven genes up- or down-regulated in macrophages from men with haplogroup I survived correction for multiple testing (FDR<20%) (Table 5.4). Three genes were up-regulated in carriers of haplogroup I while 8 were down-regulated. Four of these 11 genes were associated with various immunological diseases (*IFI6* in psoriasis (Szegedi *et al.* 2010), eukaryotic translation initiation factor 2 and subunit 1 alpha [*EIF2S1*] in hepatitis B (Zhang *et al.* 2006), cytochrome P450, family 27, subfamily A and polypeptide 1 [*CYP27A1*] in asthma (Bossé *et al.* 2009; Li *et al.* 2011), and *ITGAL* in inflammatory bowel disease (Koch *et al.* 2010). Four genes were associated with CAD: purinergic receptor P2X, ligand-gated ion channel, 4 [*P2Rx4*] through association with central fat mass (Wilson *et al.* 2006), *CYP27A1* through cholesterol balance (Luoma. 2008), protein tyrosine phosphatase type IVA and member 3 [*PTP4A3*] in the progression of cardiac hypertrophy (Matter *et al.* 2001), and *ITGAL*, which shows increased expression in those with CAD (Rahimi *et al.* 2003). Detailed information on differentially regulated genes surviving correction for multiple testing can be found in Table 5.4.

5.3.5 Validation of IFI6

In an effort to validate the effect of haplogroup I on the macrophage expression level of *IFI6*, RT-PCR was completed on 48 of the samples used in the transcriptome analysis. No association between haplogroup I of the Y chromosome and expression of *IFI6* (P= 0.924) was identified.

Table 5.2 Differentially regulated genes in monocytes: nominal level of statistical significance

Gene symbol	Gene name	P value	β coefficient (SE)	Functional overview
CKAP5	Cytoskeleton associated protein 5	0.00011	0.097 (0.02)	Organisation of spindle poles.

SE – standard error, P value-level of statistical significance for a difference in expression between haplogroup I and others (adjusted for age, BMI and centre of recruitment). Functional overview from Gene cards (www.genecards.org).

Table 5.3 Differentially regulated gene in macrophages: nominal level of statistical significance

Gene symbol	Gene name	P value	β coefficient (SE)	Functional overview
IFI6	Interferon, α-inducible protein 6	0.000059	0.363 (0.09)	Shows cytokine activity and is involved in the immune response and anti-apoptosis.
EIF2S1	Eukaryotic translation initiation factor 2, subunit 1α	0.000061	0.112 (0.03)	Catalyses the first regulated step of protein synthesis initiation.
P2RX4	Purinergic receptor P2X, ligand-gated ion channel, 4	0.00011	-0.271 (0.07)	Ligand-gated ion channel.
CAD	Carbamoyl-phosphate synthetase 2	0.00013	-0.131 (0.03)	Trifunctional protein associated with the enzymatic activities of pyrimidine biosynthesis.
MGC71993	Ribonuclease, RNase K	0.00017	-0.117 (0.03)	An endoribonuclease which preferentially cleaves ApU and ApG phosphodiester bonds.

CYP27A1	Cytochrome P450, family 27, subfamily A, polypeptide 1	0.00019	-0.213 (0.06)	Monooxygenases which catalyse reactions involved in drug metabolism and synthesis of various lipids including cholesterol and steroids.
PTP4A3	Protein tyrosine phosphatase type IVA, member 3	0.00019	-0.044 (0.01)	Stimulates progression from G1 into S phase. Enhances cell proliferation, cell motility and invasive activity.
PDIA5	Protein disulfide isomerase family A, member 5	0.00019	-0.188 (0.05)	Protein metabolism.
ALKBH7	AlkB, alkylation repair homolog 7	0.00021	-0.141 (0.04)	Unknown.
RRAS2	Related RAS viral (r-ras) oncogene homolog 2	0.00021	0.122 (0.03)	Plasma membrane-associated GTP-binding protein with GTPase activity.
ITGAL	Integrin, alpha L (antigen CD11A (p180)	0.00022	-0.191 (0.05)	Has roles in leukocyte-endothelial cell interaction, cytotoxic T-cell mediated killing, and antibody-dependent killing by granulocytes and monocytes, adhesion of cytotoxic T cells to their target cells, leukocyte intercellular adhesion through interactions with its ligands, lymphocyte co-stimulatory signalling, intercellular binding and delivers co-stimulatory signals to T cells.
C12ORF45	Open reading frame chromosome 12	0.00035	-0.165 (0.05)	Unknown.
AK1	Adenylate kinase	0.00041	0.205 (0.05)	Adenine nucleotide composition.
CBS	Cystathionine-β-synthase	0.00044	-0.350 (0.05)	Catalyses the conversion of homocysteine to cystathionine.

BTN3A1	Butyrophilin, subfamily 3, member A1	0.00053	-0.119 (0.10)	Involved in the MHC- associated genes that encode type I membrane proteins with 2 extracellular immunoglobulin domains and an intracellular B30.2 domain.
FBXO10	F-box protein 10	0.00056	-0.161 (0.03)	Protein- ubiquitin ligase.
NUDT16	Nudix (nucleoside diphosphate linked moiety X)-type motif 161	0.00058	-0.115 (0.05)	RNA de-capping enzyme.
TPT1	Tumour protein, translationally-controlled 1	0.00058	0.060 (0.03)	Calcium binding and microtubule stabilisation.
C22ORF36	Open reading frame chromosome 22	0.00064	-0.115 (0.02)	Unknown.
PEPD	Peptidase D	0.00065	-0.122 (0.03)	Recycling of proline and control of the production of collagen.
TMEM147	Transmembrane protein 147	0.00070	-0.126 (0.04)	Unknown.
FAM134A	Family with sequence similarity 134, member A	0.00074	-0.122 (0.04)	Unknown.
NDFIP2	Nedd4 family interacting protein 2	0.00076	0.145 (0.04)	Activates homologous to the E6-AP Carboxyl Terminus domain-containing E3 ubiquitin-protein ligases and may modulate eGFR signalling.
HLA-A	Major histocompatibility complex, class I, *A*	0.00079	-0.173 (0.04)	Class I molecule which plays a central role in the immune system by presenting peptides derived from the endoplasmic reticulum.
ADCK5	aarF domain containing kinase 5	0.00081	-0.070 (0.05)	Unknown.
SCG5	Secretogranin V (7B2 protein)	0.00081	0.415 (0.02)	Molecular chaperone for proprotein convertase, preventing its premature activation in the regulated secretory

				pathway.
TROVE2	TROVE domain family, member 2	0.00081	0.108 (0.10)	Binds small cytoplasmic RNAs.
GFM1	G elongation factor, mitochondrial 1	0.00082	0.108 (0.03)	Mitochondrial translation elongation factor.
ITGA6	Integrin, α6	0.00082	0.225 (0.07)	Cell adhesion and cell surface mediated signalling.
ZNF410	Zinc finger protein 410	0.00084	-0.115 (0.03)	Transcription factor activating transcription of matrix-remodelling genes during fibroblast senescence.
PSMA2	Proteasome (prosome, macropain) subunit, αtype, 2	0.00084	0.101 (0.03)	Cleaves peptides in an adenosine triphosphate (ATP)/ubiquitin-dependent process.
HSPD1	Heat shock 60kDa protein 1 (chaperonin)	0.00085	0.164 (0.05)	Involved in the folding and assembly of imported proteins to the mitochondria and may have a role in the innate immune system.
FAM116B	Family with sequence similarity 116, member B	0.00086	0.169 (0.05)	Unknown.

SE – standard error, P value-level of statistical significance for a difference in expression between haplogroup I and others (adjusted for age, BMI and centre of recruitment). Functional overview from Gene cards (www.genecards.org).

Table 5.4 Function of genes showing differential regulation in macrophages (FDR<0.2)

Gene symbol	Gene name	Up/down regulation	Gene size (bp)	Relevance to immunological diseases and/or CAD
IFI6	Interferon, α-inducible protein 6	↑	6,158	Over-expressed in psoriasis (Szegedi *et al.* 2010). Inhibits hepatitis C sub-genomic replication (Itsui *et al.* 2009).
EIF2S1	Eukaryotic translation initiation factor 2, subunit 1α	↑	26,489	rs3759756 influences IFN response in hepatitis B patients (King *et al.* 2002).
P2RX4	Purinergic receptor P2X, ligand-gated ion channel, 4	↓	24,250	Associated with central fat mass (Wilson *et al.* 2006).
CAD	Carbamoyl-phosphate synthetase 2	↓	26,403	Unknown.
MGC71993	Ribonuclease, RNase K	↓	2,054	Unknown.
CYP27A1	Cytochrome P450, family 27, subfamily A, polypeptide 1	↓	33,545	Associated with asthma and atopy (Bosse *et al.* 2009). Involved in physiological maintenance of cholesterol balance (Luoma *et al.* 2009).
PTP4A3	Protein tyrosine phosphatase type IVA, member 3	↓	9,614	Related to the progression of cardiac hypertrophy by inhibiting intracellular calcium mobilisation in response to angiotensin II (Matter *et al.* 2001).
PDIA5	Protein disulfide isomerase family A, member 5	↓	158,166	Inhibition of lymphocyte surface- associated protein-disulfide isomerases which blocks HIV: cell fusion (Barbouche *et al.* 2003).

ALKBH7	AlkB, alkylation repair homolog 7	↓	2,597	Unknown.
RRAS2	Related RAS viral (r-ras) oncogene homolog 2	↑	81,264	Unknown.
ITGAL	Integrin, αL (antigen CD11A (p180)	↓	50,524	Associated with inflammatory bowel disease (Frenzel *et al.* 2002). Involved in the development of experimental autoimmune encephalomyelitis (Dugger *et al.* 2009). Required in virological synapse formation involved in HIV-1 spread (Jolly *et al.* 2007). Increased expression in advanced CAD (Rahimi *et al.* 2003).

All genes survive correction for multiple testing (FDR<20) after age, body mass index and centre adjusted differential expression analysis (I vs. all others). Bp – base pairs, CAD – coronary artery disease, IFN – interferon γ, HIV- human immunodeficiency virus. All genes are differentially expressed in macrophages only.

5.3.6 Gene Set Enrichment Analysis

No individual KEGG pathways were differentially regulated between carriers of haplogroup I and all other haplogroups in monocytes. However, 30 pathways were differentially regulated in macrophages. Detailed information on the function of each pathway is in Table 5.6. ~30% of the differentially regulated pathways were involved in the immune system (KEGG IDs: 05330, 04612, 05310, 05320, 05332, 04672, 05140, 05322, 04940, 05416) and primarily acted in B cells, T cells or macrophages. Specifically, six down-regulated pathways were involved in autoimmunity, while four up-regulated pathways were related to inflammation. Furthermore, 54% of the differentially regulated pathways were previously associated with CAD (KEGG IDs: 04612, 05310, 04514, 05332, 04640, 04672, 04142, 00980, 05322, 04940, 04060, 04510, 04670, 04310, 05340, 00500, 05212).

Table 5.5 Pathways showing significant differential regulation in carriers of haplogroup I in macrophages

KEGG Pathway name	KEGG functional pathway description and important points	FDR	↑/↓ regulated	Genes in pathway (n) and genes conveying the major contribution to the enrichment score	Lead cell
05330: Allograft rejection	The consequence of the recipient's alloimmune response to non-self-antigens expressed by donor tissues.	0.0001	↓	25: *HLA-F, HLA-DPB1, HLA-DMB, HLA-A, HLA-DRA, HLA-DPA1, HLA-DRB3, HLA-DQA1, HLA-DRB1, HLA-DRB5*	T cells
04612: Antigen processing and presentation	Capturing, processing and presentation of antigens to enable their recognition by T cells are the basis of adaptive immunity. Processing and presentation of antigens (i.e. oxidised LDL-C) by antigen presenting cells is a well-recognised process in the early stages of atherosclerosis (Nilsson & Hansson. 2008).	0.0001	↓	52: *CTSS, HSPA5, HLA-G, LGMN, HSP90AB1, IFI30, HLA-F, HLA-DPB1, HLA-DMB, HLA-A, HLA-DRA, HLA-DPA1, HLA-DRB3, HLA-DQA1, HLA-DRB1, HLA-DRB5*	Macrophages, dendritic cells, B cells
05310: Asthma	Chronic inflammation of airways initiated by T cells and propagated by B cells, mast cells and eosinophils. Asthma increases the risk of CVD independently of traditional risk factors (Appleton *et al.* 2009).	0.0001	↓	19: *HLA-DPB1, HLA-DMB, HLA-DRA, HLA-DPA1, HLA-DRB3, HLA-DQA1, HLA-DRB1, HLA-DRB5*	T cells

5320: Autoimmune Thyroid disease	Activation of immune effector mechanisms against self-antigens leads to the damage of thyroid epithelial cells.	0.0001	↓	26: *HLA-F*, *HLA-DPB1*, *HLA-DMB*, *HLA-A*, *HLA-DRA*, *HLA-DPA1*, *HLA-DRB3*, *HLA-DQA1*, *HLA-DRB1*, *HLA-DRB5*	T cells
04514: Cell adhesion molecules	Integrin family, the immunoglobulin super family, selectins, and cadherins are glycoproteins expressed on the cell surface and playing a critical role in haemostasis, the immune response, inflammation, embryogenesis, and development of neuronal tissue. Soluble CD54 levels are independently associated with coronary artery calcium, aortic wall thickness and aortic compliance. Soluble adhesion molecules are related to risk of death from cardiovascular causes among patients with documented CAD (Blankenberg *et al.* 2001).	0.011	↓	69: *CD6*, *NCAM1*, *MPZL1*, *HLA-F*, *HLA-DPB1*, *ICAM2*, *CD99*, *HLA-DMB*, *HLA-A*, *HLA-DRA*, *ITGAL*, *HLA-DPA1*, *HLA-DRB3*, *HLA-DQA1*, *HLA-DRB1*, *HLA-DRB5*	Ubiquitous
00480: Glutathione metabolism	Glutathione is a major endogenous anti-oxidant with an important role in defence against reactive oxygen species. It contributes to regulation of several fundamental biological processes including DNA and protein synthesis as well as cytokine production and immune response.	0.048	↓	36: *MGST1*, *GPX3*, *MGST2*, *GSTK1*, *IDH1*, *GPX4*, *GGCT*, *GSTP1*, *MGST3*, *ANPEP*, *PGD*, *SMS*, *GGT1*, *GSTT1*	Ubiquitous
00531: Glycosaminoglycan degradation	Major components of the extracellular matrix and cell surface of most cell types are degraded in lysosome by the concerted action of exohydrolase activities following partial catabolism by endoenzymes (Fuller *et al.* 2004).	0.197	↓	15: *NAGLU*, *HEXA*, *HYAL2*, *SGSH*, *HYAL3*, *IDUA*	Ubiquitous

05332: Graft versus host disease	Complication of allogenic transplantation driven by immunocompetent donor T cells that attack the genetically disparate host cells.	0.0001	↓	26: *HLA-F, HLA-DPB1, HLA-DMB, HLA-A, HLA-DRA, HLA-DPA1, HLA-DRB3, HLA-DQA1, HLA-DRB1, HLA-DRB5*	T cells
04640: Hematopoietic cell lineage	Blood cells originate from haematopoietic stem cells that differentiate into multi-lineage committed progenitor cells. Haematopoietic cell lineage pathway is enriched in genes with prior evidence of association with CAD (Liu *et al.* 2011).	0.06	↓	55: *CD37, ANPEP, HLA-DRA, HLA-DRB3, HLA-DRB1, HLA-DRB5*	Hematopoietic stem cells
04672: Intestinal immune network for IGA production	Secreted Ig (immunoglobulin) A promotes immune exclusion by entrapping dietary antigens and microorganisms in the mucus and neutralisation of toxins. Salivary IgA is positively associated with CAD (Janket *et al.* 2010).	0.001	↓	28: *HLA-DPB1, HLA-DMB, HLA-DRA, HLA-DPA1, HLA-DRB3, HLA-DQA1, HLA-DRB1, HLA-DRB5*	B cells
05140: Leishmania infection	A disease with visceral and cutaneous manifestations caused by an intracellular protozoan parasite of macrophages. Silenced T cell activation leads to abnormal immune response. This pathway contains 13 genes showing prior evidence for association with coronary artery disease (Liu *et al.* 2011).	0.01	↓	57: *PTPN6, IFNGR1, JUN, MYD88, HLA-DPB1, HLA-DMB, HLA-DRA, HLA-DPA1, HLA-DRB3, HLA-DQA1, HLA-DRB1, HLA-DRB5*	Macrophages

00980: Metabolism of xenobiotics by cytochrome P450	One of the major pathways of xenobiotic biodegradation. Carriers of certain CYP polymorphisms (e.g. CYP2C19*2) are at increased risk of adverse cardiovascular events (Jin *et al.* 2010).	0.062	↓	21: *CYP2S1, MGST1, MGST2, GSTK1, GSTP1, MGST3, EPHX1, AKR1C3, AKR1C2, GSTT1*	Ubiquitous
05322: Systemic Lupus Erythematous	Autoimmune disease associated with production of auto-antibodies against self-antigens (DNA, nuclear proteins and cytoplasmic components) leading to inflammation, vasculitis, immune complex deposition and activation of complement system as well as macrophage- and neutrophil-mediated tissue injury. Lupus is strongly associated with atherosclerosis (Handa 2010).	0.0001	↓	63: *HLA-DPB1, TRIM21, HLA-DMB, C1QB, HLA-DRA, H2AFY2, HLA-DPA1, HLA-DRB3, HLA-DQA1, HLA-DRB1, HLA-DRB5*	Plasma B cells
04940: Type 1 Diabetes Mellitus	Autoimmune destruction of the insulin-producing β-cell of the pancreatic islets by macrophages and cytotoxic T cells as well as non-specific inflammatory mediators. Young people with Type 1 diabetes often show early signs of atherosclerosis (Trigona *et al.* 2010).	0.0001	↓	28: *HLA-F, HLA-DPB1, HLA-DMB, HLA-A, HLA-DRA, HLA-DPA1, HLA-DRB3, HLA-DQA1, HLA-DRB1, HLA-DRB5*	
05416: Viral myocarditis	Myocarditis is a cardiac disease associated with inflammation and injury of the myocardium. It results from various aetiologies, both non-infectious and infectious, but coxsackievirus B3 (CVB3) is still considered the dominant aetiological agent. Myocarditis may be caused by direct cytopathic effects of virus, a pathologic immune response to persistent virus, or autoimmunity triggered by the viral infection.	0.0001	↓	46: *HLA-F, HLA-DPB1, HLA-DMB, HLA-A, HLA-DRA, ITGAL, HLA-DPA1, HLA-DRB3, HLA-DQA1, HLA-DRB1, HLA-DRB5*	Cardiomyocyte

00330: Arginine and proline metabolism	Co-metabolism of several amino-acids: arginine, ornithine, proline, citrulline and glutamate. Elevated levels of amino acids and their products increase cardiovascular disease. Conversely dietary supplementation with arginine may reduce cardiovascular disease (Wu 2009).	0.158	↑	31: *GAMT, LAP3, ABP1, GLS, AMD1, P4HA1, MAOA*	Ubiquitous
05412: Arrhythmogenic right ventricular cardiomyopathy	Inherited heart muscle disease caused by genetically mediated disruption of desmosomal function and myocardial injury accompanied by inflammation.	0.078	↑	28: *ITGB3, CTNNA1, ATP2A2, CACNB3, CACNA2D4, SGCB, ITGA5, ITGB1, DSC2, ITGAV, ITGB7, ITGB5, JUP, CTNNB1, ITGA6*	Cardiomyocyte
04060: Cytokine: cytokine receptor interaction	Cytokines and their receptors are expressed in human atherosclerotic plaques and modulate plaque development and stability (Tedgui & Mallat 2006). Persistent increases in cytokines are associated with vascular dysfunction and vascular disease such as atherosclerosis, abdominal aortic aneurysm, varicose veins and hypertension (Sprague *et al.* 2009). The pathway is enriched for genes with prior evidence of association with CAD (Liu *et al.* 2011).	0.125	↑	114: *TGFBR2, TNFRSF12A, IFNA10, IL28RA, TSLP, PPBP, TNFRSF9, BMPR2, IL7R, IL6R, CSF2RA, TNFRSF4, PDGFC, CCR5, CSF3R, ACVR1, CXCL10, ACVR2A, IL13RA1, TNFSF15, TNFRSF1B, IL18, CXCL5, CSF1R, IL10RB, TNFSF14, LEP, IL10, LEPR, CCL2, CCR6, CCL8, CCL5, IL1B, CCL7,*	Ubiquitous

				CCL22, CCL4L1, TNFRSF21	
05414: Dilated cardiomyopathy	Progressive heart muscle disease with a significant immune and autoimmune component (increased serum levels of auto-antibodies, cytokines, viral genomes); may develop as a consequence of viral myocarditis (Tang *et al.* 2010).	0.111	↑	36: *ADCY3, ITGB3, TPM3, ATP2A2, CACNB3, CACNA2D4, SGCB, PRKACB, ITGA5, ITGB1, TPM2, ITGAV, TPM4, ITGB7, ITGB5, ITGA6*	Cardiomyocyte
04512: ECM receptor interaction	Extracellular matrix is a mixture of structural and functional macromolecules that play an important role in tissue and organ morphogenesis and in the maintenance of cell and tissue structure and function. ECM receptor interaction pathway is strongly enriched in genes associated with CAD (Blackenberg *et al.* 2001).	0.098	↑	32: *ITGB1, CD47, COL6A1, SPP1, SDC2, ITGAV, LAMC1, ITGB7, ITGB5, CD36, ITGA6*	Ubiquitous

04510: Focal adhesion	Focal adhesion regulates cytoskeletal/adhesion dynamics and in consequence cellular shape and motility. Molecules signalling within this pathway are targets for drugs used in treatment of acute coronary syndromes (Kraemer 2000; Curtiss 2009; Cox *et al.* 2010; Liu *et al.* 2011). There is a very strong enrichment of this pathway in genes with prior evidence of association with CAD (Liu *et al.* 2011).	0.115	↑	108: *ITGA5, PPP1R12A, PDPK1, PDGFC, ILK, PTK2, BIRC2, MYL9, PIK3R1, ITGB1, COL6A1, RAPGEF1, PAK2, ROCK1, VASP, BIRC3, TLN1, VCL, SPP1, RHOA, PRKCA, PPP1CB, ITGAV, RAP1B, ROCK2, VAV3, PIK3CG, LAMC1, CCND2, ITGB7, ITGB5, CCND1, CTNNB1, ITGA6*	Ubiquitous
00010: Glycolysis gluconeogenesis	Glycolysis is the process of converting glucose into pyruvate and generating small amounts of ATP and nicotinamide adenine dinucleotide (NADH). Gluconeogenesis is a synthesis of glucose from non-carbohydrate precursors and essentially a reversal of glycolysis.	0.093	↑	41: *PGAM1, ALDOA, ALDH3A2, ALDH3B1, HK3, HK1, DLAT, PFKM, ALDOC, PGM2, GAPDH, LDHA, PDHA1, DLD, PGK1, LDHB, PKM2, HK2*	Ubiquitous
05410: Hypertrophic cardiomyopathy	One of the most common inherited cardiac disorders with cardiomyocyte hypertrophy, myofibrillar disarray, and interstitial fibrosis. Hypertrophic cardiomyopathy pathway is enriched in genes showing association with coronary artery disease (Liu *et al.* 2011).	0.109	↑	33: *ITGB3, TPM3, ATP2A2, CACNB3, CACNA2D4, SGCB, ITGA5, ITGB1, TPM2, ITGAV, PRKAA1, TPM4, ITGB7, ITGB5,*	Cardiomyocyte

				ITGA6	
04670: Leukocyte transendothelial migration	Leukocyte migration from the blood into tissues is vital for immune surveillance and inflammation. During this diapedesis of leukocytes, the leukocytes bind to endothelial cell adhesion molecules (CAM) and then migrate across the vascular endothelium. A leukocyte adherent to CAMs on the endothelial cells moves forward by leading-edge protrusion and retraction of its tail. In this process, alphaL /beta2 integrin activates through vav 1 guanine nucleotide exchange factor (Vav1), Ras homolog gene family, member A (RhoA), which subsequently activates the kinase p160ROCK. ROCK activation leads to MLC phosphorylation, resulting in retraction of the actin cytoskeleton. Moreover, leukocytes activate endothelial cell signals that stimulate endothelial cell retraction during localised dissociation of the endothelial cell junctions. Intercellular adhesion molecule 1 (ICAM-1)-mediated signals activate an endothelial cell calcium flux and protein kinase C, which are required for ICAM-1 dependent leukocyte migration. Vascular cellular adhesion protein (VCAM-1) is involved in the opening of the endothelial passage through which leukocytes can extravasate. In this regard,	0.11	↑	68: *PTK2, MYL9, PIK3R1, ITGB1, CLDN23, ROCK1, VASP, PTPN11, ITGAM, VCL, RHOA, PRKCA, RAP1B, CLDN14, ROCK2, VAV3, PIK3CG, NCF1, F11R, CTNNB1*	Leukocytes

	VCAM-1 ligation induces nicotine adenine disphosphonucleotide, reduced (NADPH) oxidase activation and the production of reactive oxygen species in a Rac-mediated manner, with subsequent activation of matrix metallopoteinases and loss of VE-cadherin-mediated adhesion. The rolling, adhesion and transmigration of leukocytes across the endothelial barrier into the intima is a major part of atherosclerosis development (Erriksson 2004).				
05130: Pathogenic Escherichia coli infection	Enteropathogenic *Escherichia coli (E. coli)* and enterohaemorrhagic *E. coli* mediated colonisation and damage of intestinal epithelial cells.	0.126	↑	42: *TUBB2A, TUBB6, ITGB1, ROCK1, YWHAQ, ARPC4, RHOA, PRKCA, ROCK2, TLR4, TUBA4A, ARPC5, CTNNB1*	Epithelial cells
04310: PPAR signalling pathway	PPARs are nuclear hormone receptors that are activated by fatty acids and their derivatives. PPAR has three subtypes (PPARα, β/Δ, and γ) showing different expression patterns in vertebrates. Each of them is encoded in a separate gene and binds fatty acids and eicosanoids. PPARα plays a role in the clearance of circulating or cellular lipids via the regulation of gene expression involved in lipid metabolism in liver and skeletal muscle. PPARβ/Δ is involved in lipid oxidation and cell proliferation. PPARγ promotes adipocyte differentiation to enhance blood glucose uptake. PPAR have effects on both metabolic risk factors and vascular inflammation and are expressed in	0.125	↑	44: *ACSL3, PDPK1, DBI, ILK, OLR1, ACOX2, FABP3, PLTP, PPARG, ACADM, NR1H3, ME1, CD36, ACSL1, ACSL4, LPL, FABP4*	Ubiquitous

	atherosclerotic lesions (Neve *et al.* 2000).				
05340: Primary immunodeficiency	A heterogeneous group of disorders, which affect cellular and humoral immunity, in addition to non-specific host defence mechanisms mediated by complement and cells such as natural killer (NK) cells. These disorders of the immune system cause increased susceptibility to infection, autoimmune disease, and malignancy. Angina and myocardial infarction associated with infusion in a patient with common variable immunodeficiency. Coronary artery aneurysms found in patients with hyper-IgE recurrent infection syndrome (Ling *et al.* 2007). Intima media thickness, an early predictor of atherosclerosis, may be associated with subclinical inflammation in children with familial Mediterranean fever (Bilginer *et al.* 2008).	0.119	↑	20: *PTPRC, IL7R, ADA, DCLRE1C, BTK, AIRE*	Phagocytes, natural killer cells

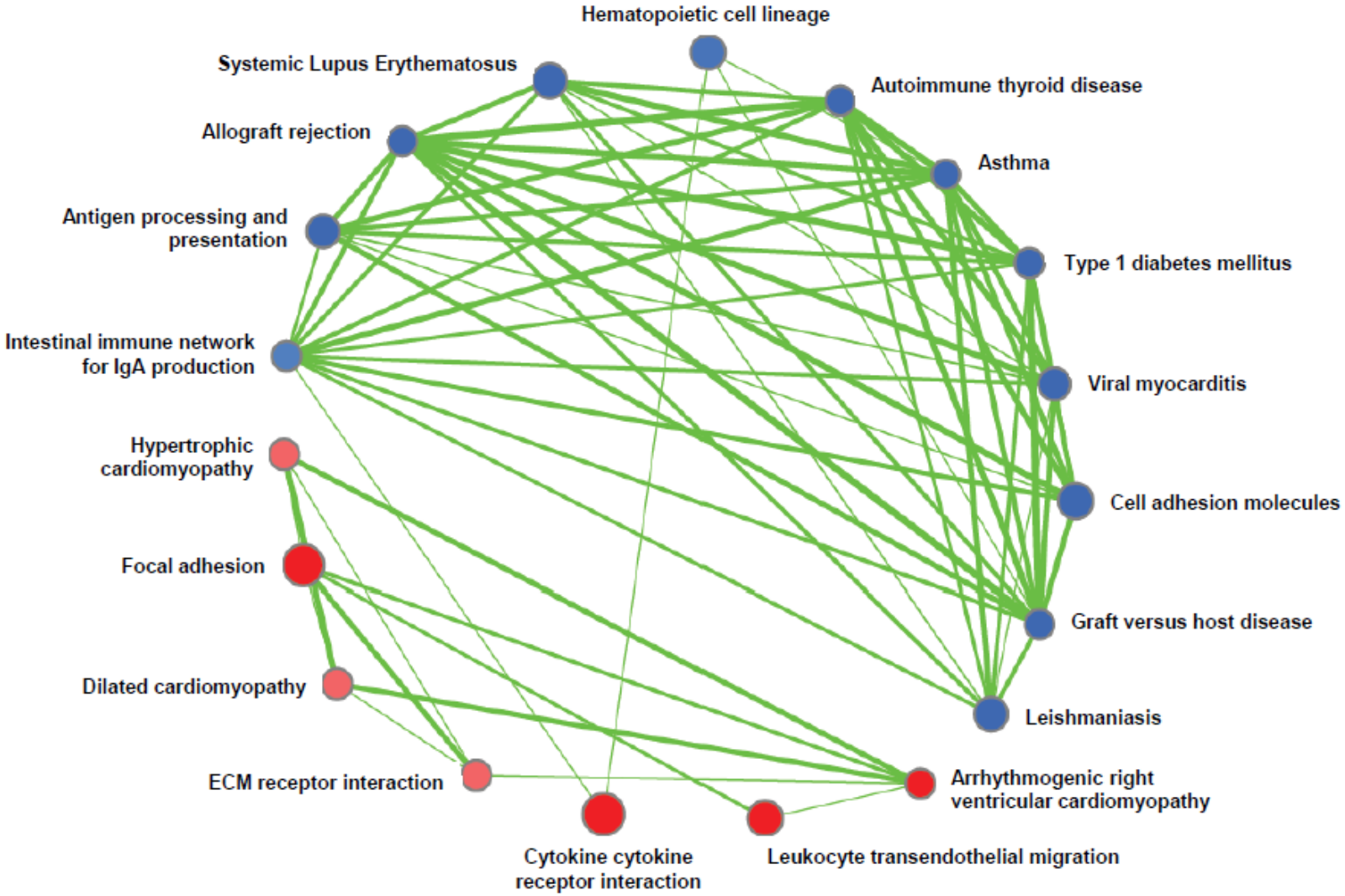
00500: Starch and sucrose metabolism	Two major dietary carbohydrates that influence postprandial glucose levels. This can produce undesirable changes in metabolic risk factors such as blood triglycerides, TC and its lipoprotein distribution, insulin and uric acid which are risk factors for cardiovascular disease (Reiser 1985).	0.109	↑	20: *AGL, HK3, PYGL, AMY2B, HK1, PGM2L1, UGDH, PGM2, GBE1, AMY1A, UGP2, HK2*	Ubiquitous
05212: TGFβ signalling pathway	The TGF-β family members, which include TGF-β, activins and bone morphogenetic proteins (BMPs), are structurally related secreted cytokines found in species ranging from worms and insects to mammals. A wide spectrum of cellular functions such as proliferation, apoptosis, differentiation and migration are regulated by TGF-β family members. TGF-β family members bind to the Type II receptor and recruit Type I, whereby Type II receptor phosphorylates and activates Type I. The Type I receptor, in turn, phosphorylates receptor-activated Smads (R-Smads: Smad1, Smad2, Smad3, Smad5, and Smad8). Once phosphorylated, R-Smads associate with the co-mediator Smad, Smad4, and the heteromeric complex then translocates into the nucleus. In the nucleus, Smad complexes activate specific genes through cooperative interactions with other DNA-binding and co-activator (or co-repressor) proteins. Activation of TGFβ pathway in the vascular wall may have both pro-atherogenic and anti-atherogenic effects (Xu *et al.* 2010).	0.115	↑	43: *TGFBR2, ID3, PPP2CA, BMPR2, ACVR1, RBX1, SMAD5, SMAD7, ZFYVE16, ROCK1, ACVR2A, MYC, RHOA, ROCK2, PPP2CB, ID2*	Ubiquitous

See http://www.genecards.org for definitions of gene abbreviations.

5.3.7 Analysis of overlapping genes in deregulated pathways

Nineteen pathways (FDR<0.2) showing differential expression in macrophages from carriers of haplogroup I and those from different MSY lineages were interconnected by common genes (including various major histocompatibility complex (MHC) genes and interleukins) involved in immunity and inflammation. A pictorial representation of how these pathways are connected through their differentially regulated gene sets is shown in Figure 5.3.

Figure 5.3 Gene overlap between differentially regulated pathways in carriers of haplogroup I in macrophages.



Immune pathways showing differential expression in macrophages from men of haplogroup I compared to carriers of all other haplogroups. Statistically different pathways (FDR <0.2) are shown as a network (nodes represent KEGG pathways [red is activation, blue is repression], colour intensity reflects the statistical significance, node size is a function of the number of genes in each pathway. Connections are drawn between pathways (green) with genes in common (line thickness proportional to the overlap coefficient).

5.4 Discussion

The mechanism of action predisposing carriers of haplogroup I of the Y chromosome to CAD has so far remained elusive. Data shown here demonstrate the role of the immune system as a mediator of the effect of haplogroup I of the Y chromosome on susceptibility to CAD amongst British men.

5.4.1 Y chromosome haplogroup distribution

The haplogroup distribution of British individuals recruited to CTS is in keeping with previously published data (Figure 5.1) (Capelli *et al.* 2003; Charchar *et al.* 2012). Indeed, amongst British men the prevalence of haplogroups I and R1b1b2 were reported at approximately 11% and 66%, respectively. There was no statistical difference in prevalence of available clinical and demographic characteristics between carriers of haplogroup I and all other haplogroups (Table 5.1). Analysis of population stratification confirmed that all CTS subjects were of European descent (Figure 5.2). No systematic differences in frequency of autosomal SNPs were identified between the individuals carrying haplogroup I of the Y chromosome and those with all other haplogroups. These data show that there are no systematic genetic or demographic differences between the two groups. This finding is reassuring given the potential for inflation of false positive results in data where population stratification is present.

5.4.2 Single-gene analysis of the monocyte transcriptome

The role of monocytes in the development and progression of CAD, together with the recent associations of the Y chromosome to immune-based diseases (Sezgin *et al.*

2009), have provided foundations to the hypothesis that immune system may explain the relationship between the Y chromosome and CAD (Charchar *et al.* 2012). Surprisingly, no single gene showed significant differential expression in monocytes between carriers of haplogroup I compared to carriers of all other Y chromosome lineages. These data imply that the mechanism through which the Y chromosome predisposes to CAD is unlikely to be routed in unstimulated monocytes.

5.4.3 Single-gene analysis of the macrophage transcriptome

Haplogroup I of the Y chromosome altered the gene expression levels of a number of genes in the macrophage transcriptome (Table 5.4). Though the significantly differentially regulated genes did not reach genome-wide significance, they did survive correction for multiple testing at the suggestive association level. Importantly, a number of differentially regulated genes had previously been associated with CAD (*P2RX4*, *CYP27A1*, *PTP4A3* and *ITGAL*) and imply that the Y chromosome may exert its effect on CAD through known pathways involving central fat mass and cholesterol. However, the role of haplogroup I of the Y chromosome in lipid fractions and BMI has been investigated in previous work (Chapter 3) and no association was identified. This implies that though the presence of haplogroup I may affect expression these genes; it may not be the major cause of the increased risk of CAD in carriers of haplogroup I.

A number of genes related to immunity or immunological diseases were either up- or down-regulated in men with haplogroup I of the Y chromosome (*EIF2S1*, *CYP27A1*, *PDIA5* and *ITGAL*). Potentially the most relevant of these genes is the significant up-regulation of *IFI6* in carriers of haplogroup I. *IFI6* (also known as *G1P3*) is a gene induced by type 1 interferon (Tahara *et al.* 2005). Very little is known of the

function of this protein; however, data have shown that this protein may have functions in cell survival, potentially through the inhibition of mitochondria-meditated apoptosis (Tahara *et al.* 2005). It is well accepted that interferon (both type 1 and 2) has an important role in the pathogenesis of atherosclerosis (Lopez-Pedrera *et al.* 2010); interferon is a well-known cytokine showing increased levels in atherosclerotic tissue (Sprague *et al.* 2009; López-Pedrera *et al.* 2010). Interferons, and their downstream effector molecules such as *IFI6*, are thought to drive cellular inflammatory responses through activation of endothelial cells, vascular smooth muscle cells, monocytes, macrophages, lymphocytes, dendritic cells and mast cells (Gordon *et al.* 2003; Mehra *et al.* 2005; Tedgui *et al.* 2006). The increased level of *IFI6* in carriers of haplogroup I may suggest that the downstream effects of interferon, through *IFI6*, may also be increased in carriers of haplogroup I. Interestingly, an increased expression of molecules induced by interferon, such as *IFI6*, is well documented in atherosclerosis (Ehrt *et al.* 2001; Mühl *et al.* 2003; Lee *et al.* 2007; McLaren *et al.* 2009; López-Pedrera *et al.* 2010). However, the association between haplogroup I and *IFI6* did not validate in independent replication using a small subset of discovery samples. The very small sample number (in comparison to the cohort in which this association was discovered) is the most likely explanation for the lack of replication.

These data indicate that several genes associated with haplogroup I of the Y chromosome are involved in the regulation of immune system and various immunological diseases. It can therefore be hypothesised that haplogroup I of the Y chromosome predisposes carriers to CAD through its effect on genes involved in the immune system, particularly inflammation and autoimmunity.

5.4.4 Pathway analysis of transcriptome data

5.4.4.1 Pathway analysis

Single-gene analysis has provided a glimpse of how haplogroup I of the Y chromosome may alter expression of various genes and predispose carriers to CAD. However, it has not provided a detailed insight into mechanisms that could drive the association between Y chromosome and CAD. In an effort to further elucidate the effect of haplogroup I of the Y chromosome on CAD risk, pathway analysis was completed. Analysis of genome-wide transcriptome levels allows the simultaneous analysis of the global gene expression and facilitates functional interpretation of the findings.

5.4.4.2 Pathways previously associated with CAD

While in monocytes no pathways were differentially expressed, in macrophages 30 pathways showed differences in expression between carriers of haplogroup I and carriers of all other Y chromosome lineages. Three differentially regulated pathways have previously been associated with CAD (leukocyte transendothelial migration [hsa04670], focal adhesions [hsa04510], cytokine cytokine receptor interaction [hsa04060]) (Eriksson 2003; Hägg *et al.* 2009). Leukocyte transendothelial migration (hsa04670) was significantly up-regulated in carriers of haplogroup I. The process of leukocyte movement through the endothelium has been heavily implicated in the pathogenesis of CAD in both its early (Eriksson 2003) and late (Hägg *et al.* 2009) stages. Specifically, Hägg *et al.* found that genes in the transendothelial migration of leukocytes pathway were over-expressed in atherosclerotic tissue (Hägg *et al.* 2009). Up-regulation of the "focal adhesions" pathway (hsa04510) highlights the importance of cytoskeletal and adhesion dynamics in CAD. Cytoskeletal and adhesion dynamics

significantly affect interactions between monocytes, platelets and the endothelium. Without monocyte/platelet and endothelium interaction, monocytes cannot migrate into the vessel to perpetuate CAD (Mitra *et al.* 2005, Curtiss 2009, van Gils *et al.* 2009). Furthermore, a well-known mechanism in the development of CAD is the immune cascade and balance between pro-inflammatory and anti-inflammatory cytokines (Hansson & Libby 2006). The "cytokine: cytokine receptor interaction" (hsa04060) pathway was up-regulated in carriers of haplogroup I, implying its dysfunction in these individuals and possibly predisposing to CAD. Alongside the three pathways heavily implicated in CAD, there are a number of other pathways hypothesised to be involved in CAD that were differentially regulated by the presence of haplogroup I of the Y chromosome. These include; lysosome (hsa04142) (Ryter *et al.* 2010), metabolism of xenobiotics by cytochrome P450 (hsa00980) (Jin *et al.* 2010), peroxisome proliferator-activated receptors (PPAR) signalling (hsa04310) (Zandberge & Plutzky 2007), primary immune deficiency (hsa05340) (Dave & Hagen 2007; Ling 2007, Bilginer 2008), starch and sucrose metabolism (hsa00500) (Reiser 1985), and the TGFβ signalling pathway (hsa05212) (Goumans *et al.* 2009). Similarly, four up-regulated pathways (focal adhesions [hsa04510], cytokine-cytokine receptor interaction [hsa04060], hypertrophic cardiomyopathy [hsa05410] and extra cellular matrix (ECM) receptor [hsa04512]) had the strongest enrichment for genes associated with CAD (Liu *et al.* 2001; Charchar *et al.* 2012) and have all been investigated for pharmaceutical inhibition to ameliorate or reverse CAD (Eriksson 2003, Braunersreuther & Mach 2006; Curtiss 2009; Charchar *et al.* 2012).

5.4.4.3 Immune pathways

Though the modulation of established CAD pathways intuitively represents a likely mechanism of action for the Y chromosome on CAD, a statistically more convincing effect was identified in the immune system. Indeed, a number of pathways involved in the adaptive immune system including; leishmaniasis (hsa05140), antigen processing and presentation (hsa04612), and intestinal immune network for IgA (hsa04672), were down-regulated in carriers of haplogroup I of the Y chromosome. Conversely, pathways involved in self-recognition and autoimmunity, including autoimmune thyroid disease (hsa05320) and type 1 diabetes (hsa09490), were significantly down-regulated in carriers of the "risk haplogroup" (further detail Table 5.6). Furthermore, 19 of the significantly deregulated pathways were interconnected by genes related to inflammation and immunity (Charchar *et al.* 2012). Finally, both arrhythmogenic right ventricular cardiomyopathy (hsa05412) and dilated cardiomyopathy (hsa05414) are immune related and were identified as being down-regulated here (Cokkinos *et al.* 1985; Tentolouris *et al.* 2004).

Regulation of the immune system has been repeatedly implicated in the pathogenesis of CAD (Hansson & Libby 2006; Charchar *et al.* 2012). Indeed, LDL receptor-negative mice missing certain genes vital to the functioning of the immune system resulted in poorer development of atherosclerosis compared to those with a fully functioning immune system (Piedrahita *et al.* 1992). A similar response was identified in severe combined inmmuno-deficient mice (Zhou *et al.* 2000). Evidence of a local, T-cell-mediated adaptive immune response has been identified in human atherosclerotic plaques. T cells are activated through their interaction with antigen-presenting cells (including macrophages), which present antigens either from local pathogens or self (Stemme *et al.* 1995; de Boer *et al.* 2000; Hansson & Libby 2006). These findings

heavily relate to pathways associated with haplogroup I of the Y chromosome including; leishmaniasis (hsa05140), viral myocarditis (hsa05416), antigen processing and presentation (hsa04612) and intestinal immune network for IgA (hsa04672).

To date, the Y chromosome has been associated with several disorders related to the immune system. Namely, carriers of haplogroup I are at a higher risk of AIDS development in men infected with HIV. Compared to carriers of other haplogroups, patients with haplogroup I also take longer to achieve HIV suppression on treatment (Sezgin *et al.* 2009). Data from Sezgin *et al.* provided the first evidence that variation within the Y chromosome may influence prevalence and predisposition to certain immune diseases. Interestingly, numerous genes associated with HIV progression were significantly deregulated in carriers of haplogroup I. Taken together, these data suggest that the Y chromosome is a systemic immuno-modulatory force potentially predisposing certain people to immune-based diseases through the action of haplogroup-specific variants.

5.4.4.4 Other pathways

Though the modulation of the immune system is the statistically and biologically most convincing potential mechanism of action emerging from the data shown here, seven further pathways were also differentially expressed in carriers of haplogroup I of the Y chromosome. Interestingly, three pathways relating to the metabolism of various amino acids were deregulated in carriers of haplogroup I of the Y chromosome (glutathione metabolism [hsa00480], glycine, serine and threonine metabolism [hsa00260], and arginine and proline metabolism [hsa00330]). There are data showing specific alterations in myocardial amino acid metabolism and oxidative stress in ischaemic heart

disease (Mudge *et al.* 1976; Cavalca *et al.* 2009) indicating the possibility that haplogroup I of the Y chromosome may also contribute to amino acid metabolism. However, this is probably not the major mechanism through which the Y chromosome predisposes to CAD.

5.4.5 Conclusions

Both single-gene and pathway analysis highlighted a role for the Y chromosome in the regulation of the immune system. Specifically, carriers of haplogroup I of the Y chromosome show increased expression of pathways related to inflammation and decreased expression of pathways related to the adaptive immune system and autoimmunity. Both of these mechanisms have been previously implicated in CAD. Interestingly, the identification of this association in macrophages and not monocytes implies that the Y chromosome exerts its effect on the predisposition of CAD at a stage subsequent to cell differentiation from monocyte to macrophage.

5.4.6 Future work

Primarily this work requires replication in independent cohorts. However, at the time of writing no such resource was available. Outside of replication, future research should focus on investigations into which transcripts(s) of the Y chromosome contains the genetic variant found only in carriers of haplogroup I and underlies the association identified here. Furthermore, having identified that haplogroup I of the Y chromosome does affect genome-wide expression levels, it is important to ascertain how this occurs,

for example through a specific gene or transcript of the Y chromosome or effects on the maintenance of DNA methylation.

Research into how the Y chromosome influences the regulation of the immune system must also be completed. The implicated pathways suggest that haplogroup I may modulate self-recognition and cell movement. As such, investigations into the rate at which oxidised LDL-C are absorbed by macrophages and the rate at which monocytes move across the endothelial barrier may prove interesting.

5.4.7 Limitations

5.4.7.1 Data analysis

As in the previous chapters the analysis here was completed on haplogroup I vs all other haplogroups. Advantages and disadvantages of this approach can be found in section Chapter 3.

5.4.7.2 Cohort size and gene number

Due to the variety of ethnicities recruited into CTS, approximately one third of samples were appropriate for use here. While this represents a significant drop in sample number, the power to detect small (and large) changes in gene expression should not be significantly altered. Indeed, research is successfully completed on cohorts significantly smaller than that used here. For example, Chalmel *et al.* utilised just 47 samples during their investigation (Chalmel *et al.* 2012), while Voineagu*et al.* used just 36 (Voineagu*et*

al. 2011). This implies the 255 samples used here would most likely be sufficient to reliably identify changes in the transcriptome in relation to the Y chromosome lineages.

Approximately 10,000 genes were investigated here. This is a significant limitation of this study as it is plausible that one (or many) of the other ~10,000 genes not investigated are associated with haplogroup I of the Y chromosome. While further research into the remaining human genome may add to our knowledge of how the Y chromosome influences cardiovascular risk, it is unlikely to significantly alter the findings presented here given the high level of statistical significance obtained during pathway analysis. Similarly, only one transcript of each gene was investigated. Transcript and splice variation are known to affect ~50% of genes (Modrek & Lee 2002; International Human Genome Sequencing Consortium 2004) in the human genome, and further investigation into individual splice variants may shed further light the systemic actions of the Y chromosome. However, resources were not available to complete this during this study.

5.4.7.3 Replication

The major limitation of this investigation is the lack of independent replication. Transcriptome data are known to be difficult to replicate in the same samples, between cohorts and between different platforms (Kuo *et al.* 2002). While some efforts can be made to ensure data reproducibility, such as standardised extraction procedures, homogenous populations (such as the CD14+ monocytes used here), duplication of samples and correct normalisation, and standard inclusion criteria, independent replication using a second cohort remains the "gold-standard" of validation. Due to the lack of a suitable replication cohort, validation was not completed as part of this study.

However, recruitment of a similar cohort is currently underway at the University of Leicester and this poses the ideal situation for data replication in the future.

5.4.7.4 Single-gene analysis

Any p-value below 10^{-7} is deemed statistically significant in genome-wide transcriptome analysis, while suggestive associations are defined by $P<10^{-5}$. While the use of these criteria reduces the risk of inflated type 1 errors, they also prevent the detection of small to medium effect sizes. The use of a larger cohort size, such as that used here, goes some way to ameliorate this limitation. Furthermore, pathway analysis is a widely accepted method for increasing statistical power and detecting smaller effects (Wang *et al.* 2010).

The investigation of gene expression, at both the single-gene and genome-wide levels, assumes a linear correlation between mRNA production and level of protein product. While this assumption is generally accepted, it should be kept in mind that this may not be true of all genes. Mechanisms such as post-translational modification, including methylation and acetylation, are well known to affect protein levels, and this potential for variability is not incorporated in this type of study (Day & Tuite 1998).

5.4.7.5 Pathway analysis

GSEA was deemed the most appropriate pathway analysis above programs such as GO-get, MAPP-finder and GO-cluster (Curtis *et al.* 2005). However, as with all statistical packages, GSEA does have some limitations. Most importantly, GSEA depends on good, highly reproducible data as the precise order and β coefficient assigned to each

gene is vital to the analysis. A gene list which differs, even slightly, in order may indicate a different set of pathways as significantly deregulated (Curtis *et al.* 2005). The reliability of data used here is high due to the standardisation of protocols, replication of samples and probes and the use of correct normalisation and gene inclusion criteria (Charchar *et al.* 2012).

A widely discussed limitation of GSEA involves the possible lack of statistical continuity during statistical evaluation. If the order of the input gene list is reversed (meaning all up-regulated genes are down-regulated and vice versa), GSEA may not indicate the same pathways as differentially regulated as identified in the original research (Curtis *et al.* 2005). Though this can be perceived as a major problem in data continuity, a full understanding of the software generating the enrichment score shows that this is GSEA working correctly. While a statistically significant pathway score implies enrichment in that pathway of differentially regulated genes, not all genes in the pathway need to show changes in expression. As such, reversing the order of the genes may cause normally expressed genes in the previously expressed pathway to be higher (or lower) in the list leading to a smaller enrichment score.

The decision was taken to execute GSEA in "pre-ranked" mode rather than the standard mode (meaning the gene list uploaded was ranked by the user rather than GSEA) as the use of the "pre-ranked" mode is generally accepted to be more statistically robust (Curtis *et al.* 2005). Similarly, only those pathways containing between 5 and 500 genes were analysed, as the normalisation process completed by GSEA is not accurate for very small or very large gene sets. However, analysis of pathways with between 5 and 900 genes did not alter statistical significance of the pathways shown here (data not shown). The KEGG pathway repository was chosen as it represents the most accurate and well maintained repository. Though using other

pathway repositories, such as chromosomal positions, alongside KEGG may have generated further hits, it is doubtful that knowledge would have been gained as KEGG is well accepted to be the most well maintained and validated repository. Finally, statistical significance was based on FDR, rather than simply p-value, to reduce the risk of false positive findings.

5.4.7.6 Tissue type

It is plausible that during the isolation of monocytes from whole blood and subsequent mRNA extraction was subject to an element of contamination from α and β globin mRNA (Vartanian *et al.* 2009). The effect of this contamination was limited as far as possible through the use of magnetic micro-beads in the isolation of monocytes (Vartanian *et al.* 2009).

All work was completed in cells known to be of high relevance to CAD, namely monocytes and macrophages. Both monocytes and macrophages are integral in the development of atherosclerosis, particularly initiation and progression (Ley *et al.* 2011), through increased inflammation and production of foam cells (Shantsila & Lip 2009). Modified LDL-C causes endothelial cells to express adhesion molecules and chemo-attractants which recruit monocytes (amongst other cells) to the damaged area (Yan & Hansson 2007; Ley *et al.* 2011). Upon recruitment, monocytes directly interact with both the endothelium and myocardium, causing further cardiac tissue damage and further promoting monocyte recruitment to the area (Shantsila & Lip 2009). Monocytes are heavily involved in the activation of the inflammatory cascade through the synthesis of pro-inflammatory cytokines. Monocytes also secrete a variety of MMPs which have been implicated in plaque destabilisation (Shantsila & Lip 2009). Once monocytes have

migrated through the arterial wall, they differentiate into macrophages. The primary systemic role of macrophages under normal physiological function is the handling/removal of altered self-components. In atherosclerosis, macrophages recognise oxidised LDL-C as non-self and phagocytose it to produce lipid-laden foam cells that ultimately form the initial stage of CAD - the fatty streak lesions (Yan & Hansson 2007). While it is accepted that these cells are highly relevant in development of CAD, it is also known that they represent a disparate collection of cells (including $CD68^{+}/CD14^{+}$ and $CD68^{+}/CD14^{-}$) (Waldo *et al.* 2008). It is plausible that only one subset of macrophages is affected by haplogroup I of the Y chromosome. A further study investigating specific subpopulations of macrophages may improve statistical significance of both single-gene and pathway analysis.

While monocytes and macrophages are both easily accessible (through venipuncture) and highly important in CAD development, the investigation of other cells related to this disease may further explain the role of the Y chromosome in CAD. Of particular interest is diseased coronary artery tissue and vascular endothelial cells (Hansson 2005).

6 Association between haplogroup I and MSY gene(s) – gene expression analysis

6.1 Introduction

Having identified that haplogroup I of the Y chromosome predisposes carriers to CAD through the modulation of the immune system, I have hypothesised that the driver of this association resides within the coding regions of the MSY. It has previously been shown that various Y chromosome sequence alterations, including SNPs and micro-deletions, can be haplogroup-specific (for example, two SNPs (*TBL1Y*+519-96 and *USP9Y*+3283-72), are mapped specifically to haplogroup I [Rozen *et al.* 2009], and that micro-deletions resulting in severe oligospermia were significantly more common in carriers of haplogroup E [Arredi *et al.* 2007]). In light of this, it is plausible that haplogroup I may be linked to variants in one or more MSY gene(s), which is/are not present in any other haplogroup and contribute to CAD through immunity/inflammation. As previously described in detail (Chapter 1), the MSY contains ten testis-specific gene families, ten ubiquitously expressed genes, two tissue-specific genes and five genes an unknown expression profiles. The MSY genes are traditionally classified into three groups; two X-transposed genes, sixteen X-degenerate genes, and nine protein-coding families of varied expression patterns (Skaletsky *et al.* 2003). While testis-specific genes are believed to play a primary role in spermatogenesis and testis development (Noordam & Repping 2006), it is unknown what, if any, systemic function the remaining MSY genes hold. Without further study, this knowledge limitation prevents the identification of which MSY gene(s) may underlie the previously identified effect of haplogroup I of the Y chromosome on the immune system and predisposition to CAD.

The ten ubiquitously expressed genes (*TMSB4Y, DDX3Y, EIF1AY, RPS4Y1, RSP4Y2, KDM5D, ZFY, PRKY, USP9Y, UTY1*), two tissue-specific genes (*PCDHY* and *AMELY*) and five genes with unknown expression profiles (*SRY, TBL1Y, NLGN4Y CYorf15a; CYorf15b*) were examined in human macrophages and kidney (as chronic kidney disease is a major risk factor for CAD and kidney function is a well established surrogate predictor of CAD [Hallan *et al.* 2007] risk implying both known and unknown pathways associated with the kidney maybe implicated in CAD pathogenesis) tissue to see if their expression differs between men from haplogroup I and those from other lineages of the Y chromosome. Due to very limited availability of RNA, only genes associated with haplogroup I in macrophages were investigated in monocytes.

6.2 Materials and methodology

6.2.1 Materials

6.2.1.1 Subset of CTS

RNA was available from 48 men recruited to CTS. Further details on CTS can be found in Chapter 5.

6.2.1.2 Silesian Renal Tissue Bank (*SRTB*)

The "Silesian Renal Tissue Bank" (SRTB) is an ongoing recruitment of individuals electing for unilateral nephrectomy due to non-invasive renal cancer, and currently contains 63 individuals (38 men) (Tomaszewski *et al.* 2011, Tomaszewski *et al.* 2007). All patients were recruited in Silesia (Poland) and all were of white European ancestry. Renal tissue samples were taken from the unaffected (by cancer) pole of the kidney immediately after surgery and placed in RNAlater (Ambion, Austin, Tex), before being preserved at –80°C (Tomaszewski *et al.* 2011, Tomaszewski *et al.* 2007). An average of three BP measurements (after 10 min of rest) was also recorded using a cuff size adjusted to the arm circumference. Each patient was classified as hypertensive or normotensive in accordance with the guidelines set out by the International Society of Hypertension (Whitworth *et al.* 2003). 30 men had DNA and RNA of sufficient quality and quantity for use in this study.

6.2.1.3 Western Poland Kidney Project (*WPKP*)

The "Western Poland Kidney Project" (WPKP) is an on-going study collecting human renal tissue from patients who undergo unilateral nephrectomy for non-invasive renal cancer. A detailed medical history and anthropometric measurements (height and weight) were recorded alongside BP measurement (using an automatic digital BP monitor after 10 min of rest, using cuff size adjusted to the arm circumference). Hypertension was defined as SBP and/or DBP >140/90 mmHg on 3 separate occasions and/or remaining on anti-hypertensive treatment (European Society of Hypertension Guidelines). Renal tissue from the unaffected pole of the kidney was preserved in RNAlater at –80°C (Ambion, Austin, Tex) immediately after surgery. 14 men had DNA and RNA of sufficient quality and quantity for use in this study.

6.2.1.4 Bioethical approval

Written, informed consent was obtained from all subjects in all studies used here, in accordance with the Declaration of Helsinki. All studies had approval from relevant institutional ethical committees.

6.2.2 Methodology

6.2.2.1 Sample genotyping and haplogrouping

All details regarding genotyping and Y chromosome haplogrouping can be found in Chapter 2.

6.2.2.2 Descriptive statistics

Age, BMI, SBP and DBP measurements from individuals recruited to CTS were summarised using mean and SD. Age, BMI, SBP, DBP and hypertension from both kidney cohorts in conjunction (SRTB and WPKW) using mean, SD and percentage prevalence. Student's *t* test was used to assess the difference in continouous clinical characteristics between those with haplogroup I and all other haplogroups.

6.2.2.3 Analysis of population stratification in CTS subset

Analysis of population stratification was completed in the CTS subset used here using the method outlined in Chapters 3 and 5.

6.2.2.4 Analysis of sequence similarity between 16 MSY genes and their X chromosome homologs

During this investigation it was of paramount importance that only RNA transcripts originating from the Y chromosome (and not their X chromosome homologous counterparts) were measured. To ensure that sufficient differences between the sequences of the MSY ubiquitously expressed genes on the Y chromosome and X chromosomes existed, an analysis of sequence similarity between X and Y was completed (Table 6.1). Secondly, after generation of the gene expression code, the contig sequence was BLASTed to ensure no off-target binding would likely occur.

Table 6.1 Sequence similarity between MSY genes and their X chromosome homologs

Y chromosome gene	Homologue on X chromosome	Similarity between X and Y gene transcript sequences (%)
UTY1	*UTX*	94
PRKY	*PRKX*	95
CYorf15A	*CXorf15*	96
USP9Y	*USP9X*	92
NLGN4Y	*NLGN4X*	99
AMELY	*AMELX*	93
SRY	*SOX3*	10
PCDH11Y	*PCDH11X*	98
TBL1Y	*TBL1X*	93
DDX3Y	*DDX3X*	96
RPS4Y2	*RPS4X*	83
RPS4Y1		
KDM5D	*SMCX*	87
CYorf15B	*CXorf15*	92
ZFY	*ZFX*	95
EIF1AY	*EIF1AX*	92
TMSB4Y	*TMSB4X*	94

The X chromosome homolog of *RSP4Y1* and *RSP4Y2* is one gene.

6.2.2.5 Analysis of the difference in gene expression of 16 MSY genes between carriers of haplogroup I and all others

All information related to gene expression, probe validation and experimental set–up can be found in Chapter 5. Relative expression levels of 16 MSY (*TMSB4Y, DDX3Y, EIF1AY, RPS4Y1, RSP4Y2, KDM5D, ZFY, PRKY, USP9Y, UTY1, PCDHY, AMELY, SRY, TBL1Y, NLGN4Y, CYorf15a, CYorf15b*) genes was compared between carriers of haplogroup I and all other haplogroups in macrophages and kidney tissue. Due to limited RNA availability, only those gene(s) showing association with haplogroup I in

macrophages were investigated in monocytes. Contig sequences of MSY gene probes can be found in Table 6.2. In each case female controls were used to ensure only Y chromosome transcripts were being targeted by each probe.

6.2.2.6 Analysis of co-regulation of significant MSY genes

Should gene expression levels of more than one gene be identified as associated with haplogroup I, an investigation into their co-regulation will be completed using Pearson's correlation coefficient.

Table 6.2 Gene expression probe sequences

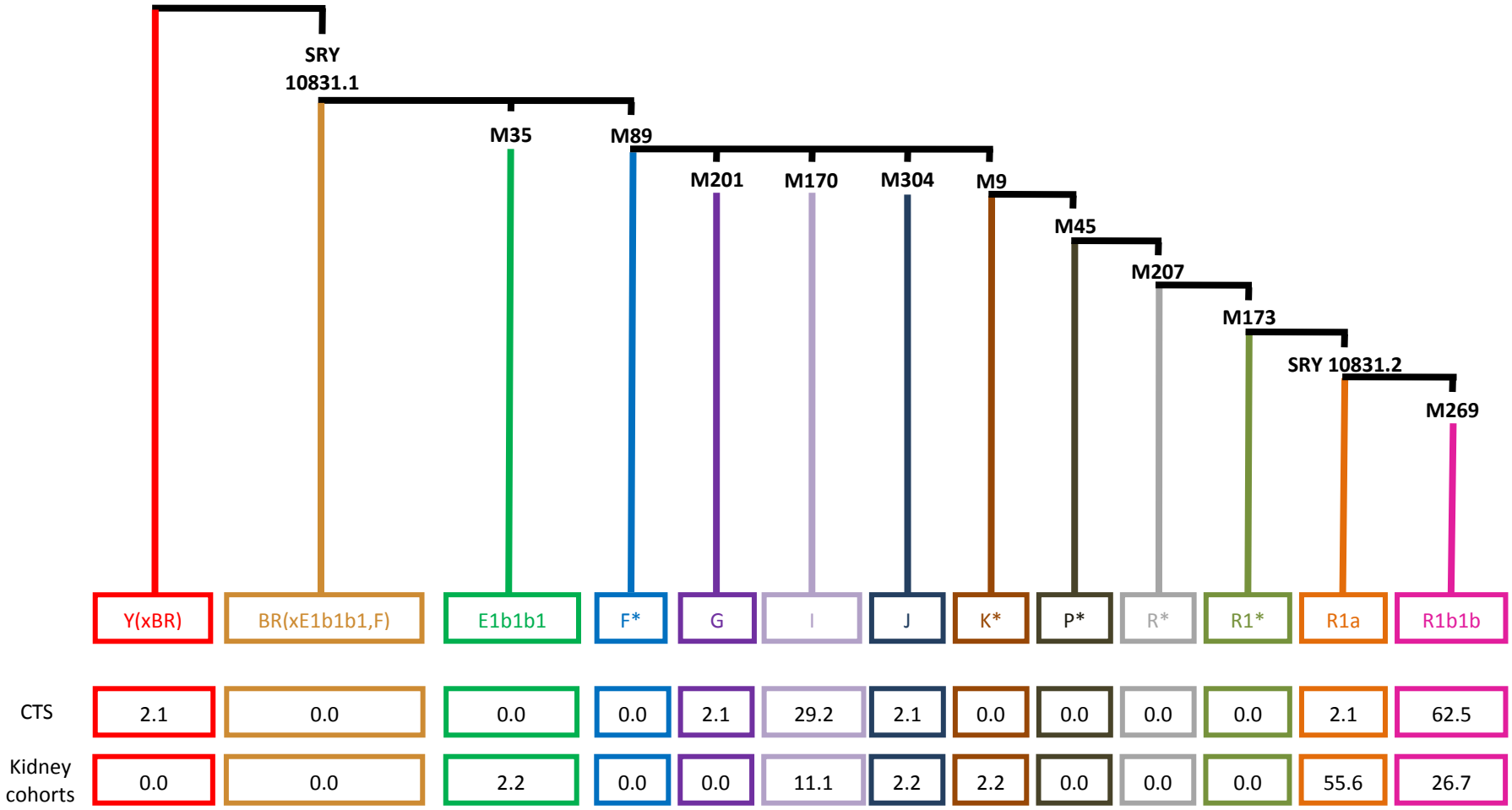
Gene	Contig sequence
UTY1	TTAAATTTCTACAGAATGGTTCTGA
PRKY	ACGCGCTGGTCACCATGGGCACTGG
CYorf15A	GAGGTGAATTTCCAAGATCAGCATA
USP9Y	GATTGTATGCTGGAGATCATGAAGA
NLGN4Y	ATGGGATTTTCTGACGTAAGATTTT
AMELY	AAGCAGGAGGAAGTGGATTAAAAGA
SRY	GGCGAAGATGCTGCCGAAGAATTGC
PCDH11Y	ACTACGAACTAATTAAGAGTCAAAA
TBL1Y	GGACTGGAACAGTGATGGAACACTA
DDX3Y	GACCAGCAATTTGTTGGTCTTGACC
RSP4Y2	TCAAATTTGACACAGGCAATGTATG
RSP4Y1	TTATCAAATTTGATACAGGCAATTT
KDM5D	GGTCAGGTGGCCAGGATGGACACTC
CYorf15B	CAAGCAGGAAACGGAAAAGCTGACA
ZFY	TACCTAATGATTTCGTTGGATGATG
EIF1AY	GCCCAAGAATAAAGGTAAAGGAGGT
TMSB4Y	CAAAGAAACTATCGAACAGGAGAGG

6.3 Results

6.3.1 Y chromosome haplogroup distribution in CTS, SRTB and WKPK

Only a subset of the CTS samples used in Chapter 5 had both macrophage and monocyte RNA of sufficient quantity and quality to be used here (n=48). Six Y chromosome haplogroups were represented in this subset of CTS with prevalence ranging from 2.1-62.5% (Figure 6.1). The analysis of SRTB and WKPK was conducted jointly in all available men recruited into both studies (n=45). Six Y chromosome haplogroups were represented in SRTB and WPKP and their prevalence ranged from 2.2-55.6%.

Figure 6.1 The prevalence of Y chromosome haplogroups in CTS, SRTB and WKPK. SRTB and WKPK analysed in conjunction. Haplogroups are arranged from Y(xBR) on the left to R1b1b2 on the right. M and SRY represent the terminal SNP that defines each haplogroup. SRY10831 is a recurrent mutation.



6.3.2 Distribution of clinical and demographic characteristics in CTS and kidney cohorts stratified by Y chromosome haplogroup

There were no statistically significant systematic differences in demographic or clinical data between carriers of haplogroup I and carriers of all other lineages in either CTS or in the kidney cohorts (Table 6.3).

Table 6.3 Clinical and demographic characteristics of the CTS (subset) and SRTB and WKPK (analysed in conjunction)

Phenotype	I	Other	P value
CTS			
n	14	34	-
Age (years)	55.8 (6.0)	54.1 (6.0)	0.380
BMI (kg/m^2)	27.3 (3.5)	29.2 (4.8)	0.221
SBP (mmHg)	127.8 (15.1)	136.1 (21.5)	0.215
DBP (mmHg)	78.2 (7.0)	84.2 (11.5)	0.083
SRTB and WPKP			
n	5	40	-
Age (years)	63.0 (7.2)	61.1 (10.1)	0.687
BMI (kg/m^2)	28.7 (6.0)	27.2 (3.4)	0.401
SBP (mmHg)	129.3 (14.7)	140.1 (11.3)	0.059
DBP (mmHg)	80 (10)	91.4 (38.9)	0.521
Hypertension(%)	66.0	80.1	0.518

Data are means and standard deviations or percentages. P-value – statistical significance of difference in a phenotype between carriers of haplogroup I and all other haplogroups, n- number of subjects, BMI- body mass index, SBP- systolic blood pressure, DBP- diastolic blood pressure.

6.3.3 Analysis of population stratification

There was no evidence of population stratification in 48 CTS men with biological material available for this analysis (Figure 6.2). There were no data available to examine the presence of population stratification in the kidney cohorts.

6.3.4 Analysis of 17 ubiquitously expressed MSY genes in macrophages

All probes were successfully validated and showed efficiency of 81-99% (data not shown). mRNA from 15 of 17 MSY genes analysed was detected in macrophage RNA; *PCDHY* and *SRY* showed no expression (Table 6.4). Amongst those detected was *AMELY*, a gene with expression profile previously described as tooth-bud-specific (Table 6.4). Of the 15 MSY genes expressed in macrophages, two genes (*UTY1* and *PRKY*) showed a significant difference in expression level between carriers of haplogroup I and men from all other lineages. Indeed, both *UTY1* and *PRKY* showed a significant reduction in expression in carriers of haplogroup I when compared to men who belonged to other MSY lineages. This difference was apparent in both crude and age-adjusted analysis (*UTY1* age-adjusted P= 0.0001, fold difference= 0.92, *PRKY* age-adjusted P= 0.001, fold difference= 0.92) (Figure 6.3).

Figure 6.2 Analysis of population stratification of the CTS subset. A. A matrix of identity-by-state distance was created in the CTS subset and the three HapMap populations (CEU, JPT+CHB, YRI). Coloured dots represent each population: blue- JPT+CHB, green- YRI, red- CEU, yellow- CTS. The first two components obtained from the subsequent multidimensional scaling of the matrix were used as co-ordinates to compare the relative distances between the four populations. B. Assessment of population structure stratified by Y chromosome haplogroup (I vs. all others). Colours represent haplogroups: green- carriers of haplogroup I, red- carriers of all other haplogroups. A matrix of identity-by-state distances was created and the first two components obtained from the subsequent multidimensional scaling of the matrix were used as co-ordinates to generate a scatter plot.

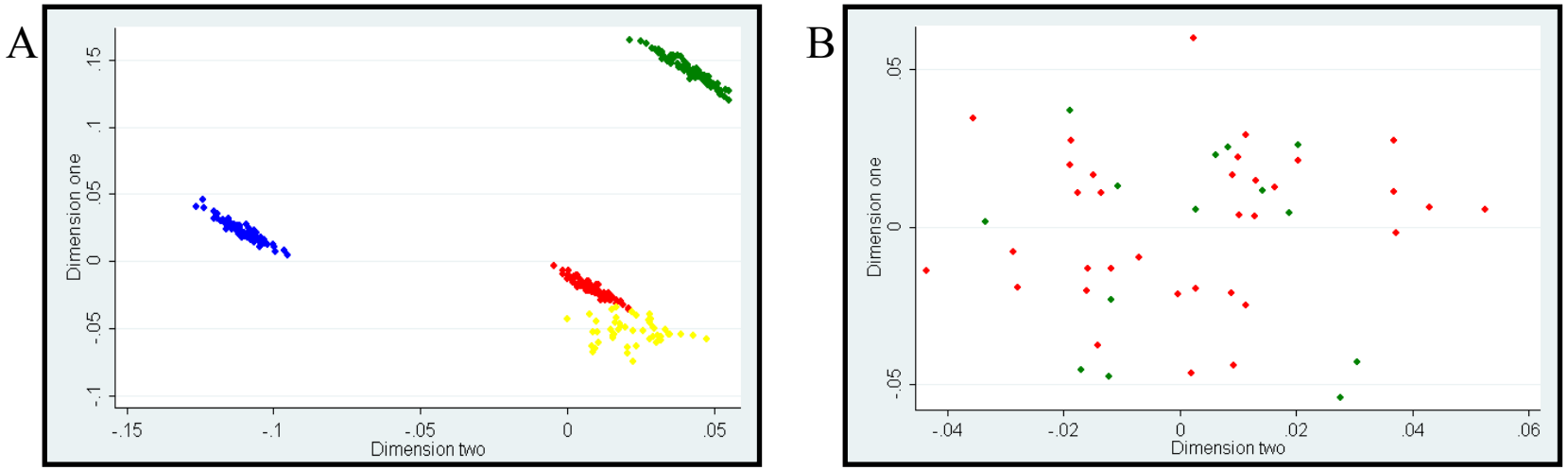
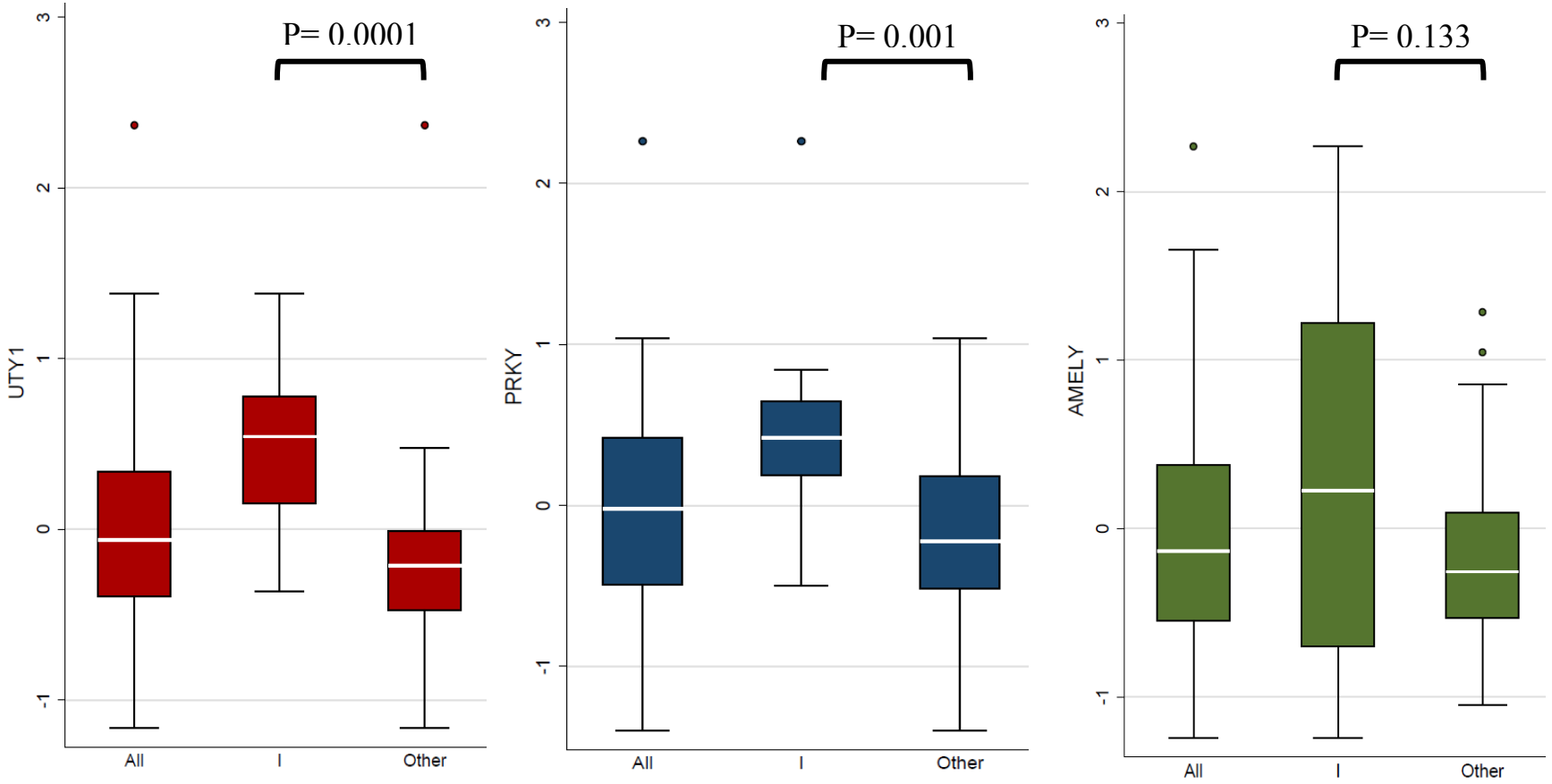


Table 6.4 Analysis of the association between haplogroup I and the expression of 17 MSY genes in macrophages

MSY gene	CT	dCT I	dCT Other	P value$^{\alpha}$	P value$^{\gamma}$
n	48	14	34	-	-
USP9Y	31.0 (1.3)	8.3 (1.0)	8.2 (0.8)	0.618	0.510
DDX3Y	29.3 (1.2)	6.1 (0.7)	5.8 (0.6)	0.085	0.147
PCDH11Y	*NOT EXPRESSED*				
NLGN4Y	36.53(1.7)	13.9 (1.7)	13.3 (3.9)	0.581	0.433
SRY	*NOT EXPRESSED*				
PRKY	31.1 (1.0)	8.2 (0.8)	7.6 (0.6)	0.004	0.001
UTY1	31.3 (1.3)	8.4 (0.6)	7.8 (0.6)	0.001	0.0001
TBL1Y	26.0 (5.0)	3.6 (1.0)	3.3 (0.9)	0.327	0.636
AMELY	24.6 (1.6)	3.6 (1.1)	3.1 (0.6)	0.066	0.113
TMSB4Y	35.3 (1.5)	10.0 (2.3)	11.1 (0.9)	0.014	0.012
EIF1AY	27.9 (0.8)	4.2 (0.5)	4.2 (0.5)	0.868	0.630
CYorf15A	32.9 (1.5)	9.4 (1.0)	9.4 (0.9)	0.895	0.664
CYorf15B	31.9 (1.5)	8.7 (1.4)	8.6 (1.5)	0.885	0.625
RPS4Y1	26.9 (1.8)	3.6 (0.5)	3.3 (0.6)	0.069	0.562
RPS4Y2	23.6 (3.6)	3.1 (0.4)	3.3 (0.8)	0.403	0.221
KDM5D	33.4 (1.3)	10.5 (1.9)	9.6 (2.2)	0.198	0.605
ZFY	30.5 (1.2)	7.1 (0.7)	7.1 (0.8)	0.929	0.402

Data are average CT, dCT and standard errors; P value$^{\alpha}$ – level of statistical significance in crude analysis, P value$^{\gamma}$– level of statistical significance in age-adjusted analysis.

Figure 6.3 Association between haplogroup I expression of *UTY1*, *PRKY* and *AMELY* in macrophages. Data are age-adjusted dCT residuals. P-value - age-adjusted statistical significance for a difference in expression level between carriers of haplogroup I and all other lineages.



6.3.5 Analysis of UTY1 and PRKY in monocytes

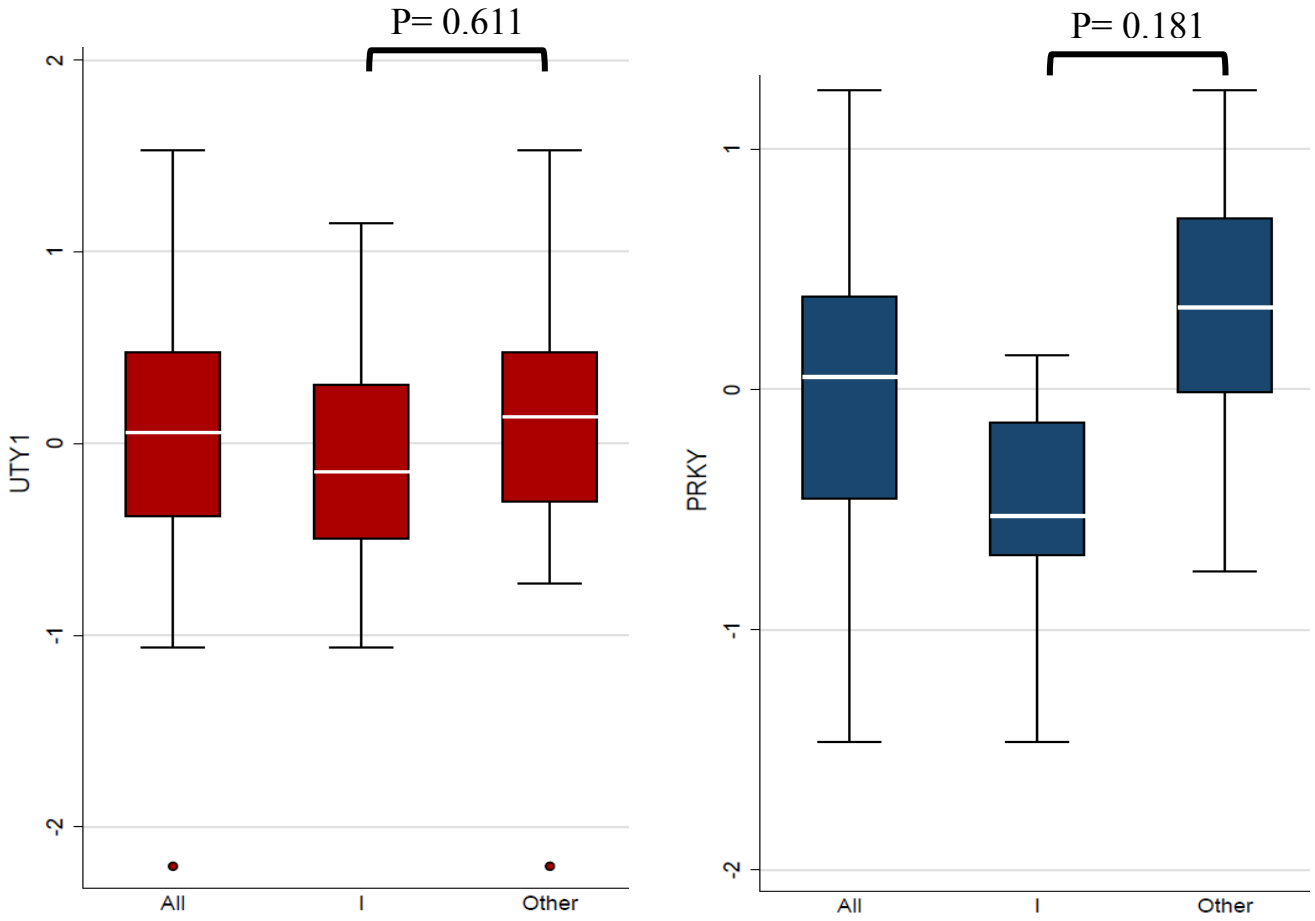
Due to limited availability of monocyte RNA, only *UTY1* and *PRKY* were investigated in this cell type. There was no evidence of association between haplogroup I and expression levels of either gene in monocytes after age adjustment (Table 6.5, Figure 6.4) (age-adjusted P= 0.611 and 0.181 for *UTY1* and *PRKY*, respectively).

Table 6.5 Effect of haplogroup I on expression level of *UTY1* and *PRKY* in monocyte RNA

MSY gene	CT	dCT I	dCT Other	P value$^{\alpha}$	P value$^{\gamma}$
n	48	14	38	-	-
UTY1	32.9 (0.9)	8.9 (0.6)	9.1 (0.8)	0.441	0.611
PRKY	32.6 (1.1)	9.1 (0.4)	10.0 (0.5)	0.0008	0.181

Data are average CT, dCT and standard errors; P value$^{\alpha}$ – crude analysis, P value$^{\gamma}$ – age-adjusted analysis.

Figure 6.4 Association between haplogroup I and expression of *UTY1* and *PRKY* in monocytes. Data are age-adjusted dCT residuals, P-values - age-adjusted statistical significance for a difference in expression level between carriers of haplogroup I and all other lineages.



6.3.6 Analysis of 17 ubiquitously expressed MSY genes in the human kidney

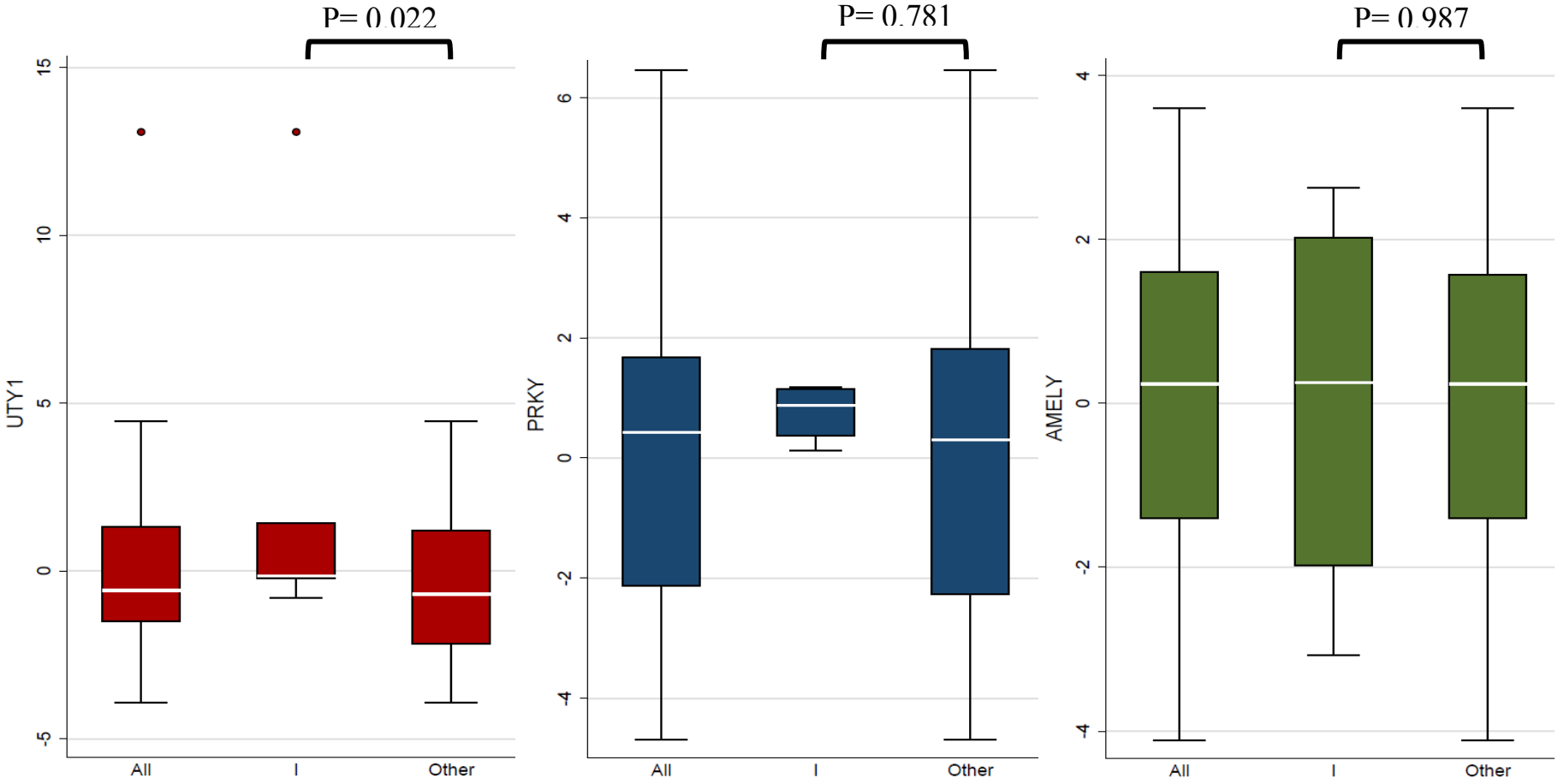
Of the 17 MSY genes investigated, mRNA from 16 genes was detected in the kidney tissue (*RPS4Y1* showed no expression). *UTY1* showed a borderline age-adjusted association with haplogroup I of the Y chromosome (P= 0.022). However, this association did not survive correction for multiple testing (Table 6.3 and Figure 6.5). Interestingly, mRNA from *AMELY* was also detected in this tissue (Table 6.6 and Figure 6.5).

Table 6.6 Association between haplogroup I and expression of 17 MSY genes in kidney tissue

MSY gene	CT	dCT I	dCT Other	P value$^{\alpha}$	P value$^{\gamma}$
n	45	5	40	-	-
USP9Y	31.0(3.5)	5.0 (1.1)	5.3 (1.4)	0.677	0.335
DDX3Y	30.2(4.0)	6.0 (2.5)	4.6 (1.4)	0.089	0.283
PCDHY	35.1(4.0)	8.8 (2.6)	9.2 (3.1)	0.807	0.512
NLGN4Y	31.4(3.3)	4.6 (1.7)	4.0 (2.5)	0.618	0.938
SRY	33.0(1.8)	11.0 (1.0)	9.0 (2.9)	0.144	0.177
PRKY	33.2(3.1)	8.1 (0.7)	7.4 (2.5)	0.621	0.781
UTY1	32.6(4.8)	8.8 (5.9)	5.9 (2.2)	0.033	0.022
TBL1Y	37.4(2.7)	11.6 (2.8)	13.2 (2.2)	0.143	0.242
TMSB4Y	34.1(3.3)	7.7 (1.0)	7.2 (1.8)	0.622	0.904
AMELY	36.6(1.4)	13.5 (2.3)	13.9 (2.0)	0.926	0.987
EIF1AY	28.1(3.1)	2.6 (0.4)	2.1 (1.1)	0.255	0.209
CYorf15A	32.1(4.2)	5.9 (1.1)	4.8 (2.2)	0.342	0.523
CYorf15B	33.6(4.2)	5.2 (1.0)	5.3 (2.3)	0.941	0.980
RPS4Y1	*NOT EXPRESSED*				
RPS4Y2	39.0 (0.8)	10.8 (0.5)	12.4 (1.7)	0.206	0.597
KDM5D	34.3 (2.1)	6.2 (2.0)	6.9 (1.8)	0.477	0.919
ZFY	33.5 (3.1)	5.3 (0.6)	6.4 (1.3)	0.059	0.245

Data are average dCT and standard errors; P value$^{\alpha}$ – level of statistical significance from the crude analysis, P value$^{\gamma}$ – level of statistical significance from age-adjusted analysis.

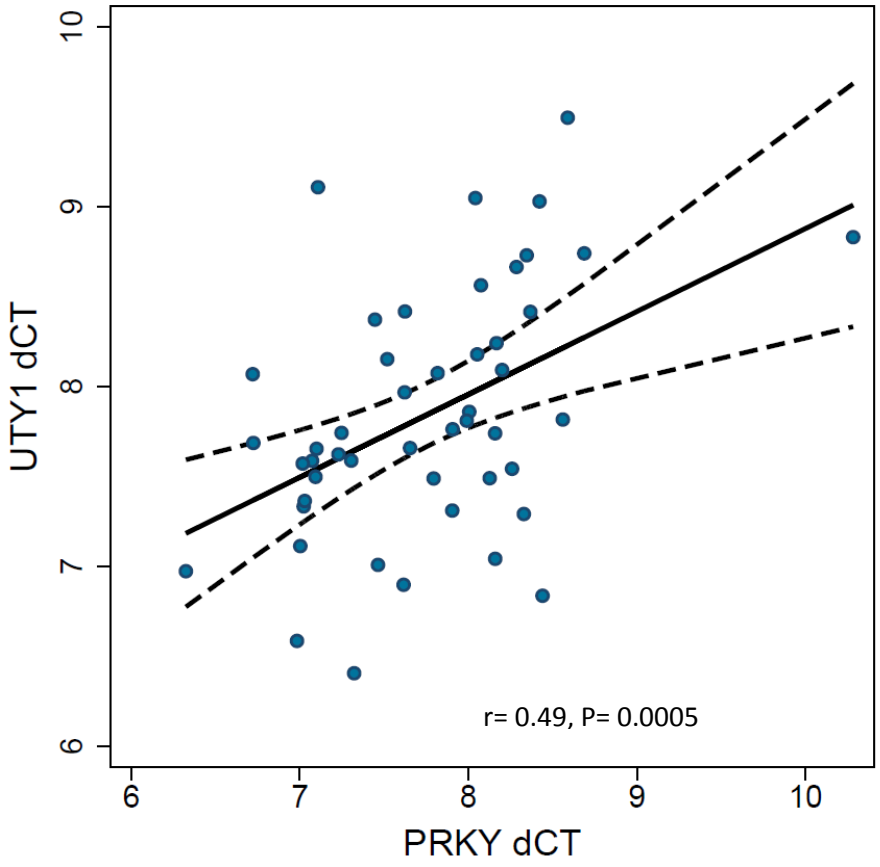
Figure 6.5 Association between haplogroup I and expression of *UTY1* and *PRKY* in the kidney. Data age-adjusted dCT residuals. P value shows age-adjusted statistical significance for a difference in gene expression level between carriers of haplogroup I and all other lineages.



6.3.7 Analysis of co-regulation of PRKY and UTY1 in macrophages

The dCT of *UTY1* and *PRKY* show a significant linear correlation in macrophages (r= 0.49, P= 0.0005).

Figure 6.6 Analysis of linear co-relation between *PRKY* and *UTY1* in macrophage RNA



6.4 Discussion

Haplogroup I of the Y chromosome predisposes carriers to CAD through its effect on the immune system, most likely at the stage of monocyte differentiation into macrophage. A significant reduction in expression levels of both *UTY1* and *PRKY* in macrophages identified here may underlie this association through an as yet undefined mechanism.

6.4.1 Y chromosome haplogroup distribution and clinical demographics

Haplogroups I, R1a and R1b1b2 accounted for approximately 93% of the Y chromosome lineages in both CTS and the kidney cohorts (93.8 and 93.4%, respectively). R1b1b2 was the predominant haplogroup in CTS (British men), while R1a was most common (62.5 and 55.6% respectively) in the kidney cohorts (Polish men) (Figure 6.1). The observed differences in haplogroup prevalence were expected and in keeping with published Y chromosome haplogroup data for both Poland and the UK (Charchar *et al.* 2012).

There were no statistically significant differences in clinical or demographic data between carriers of haplogroup I and all other haplogroups in either CTS or the kidney cohorts (Table 6.1).

There was no evidence of population stratification in the subset of CTS used here, indicating that the only genetic difference between the groups was Y chromosome haplogroup. This is in keeping with previous data from this cohort (Chapter 4) (Charchar *et al.* 2012). No data were available to assess population stratification in either SRTB or WKPK.

6.4.2 UTY1 and PRKY are differentially regulated in carriers of haplogroup I of the Y chromosome

Both *UTY1* and *PRKY* were significantly down-regulated in carriers of haplogroup I in comparison with carriers of all other haplogroups in macrophages. The lack of effect in monocytes is particularly intriguing given the findings from the previously completed pathway analysis (Charchar *et al.* 2012). Indeed, this study showed that haplogroup I was associated with immunity and inflammation at the stage of monocyte differentiation to macrophage (Charchar *et al.* 2012).

6.4.2.1 UTY1

UTY1 is one of three MSY genes known to encode a male-specific minor histocompatibility epitope (alongside *KDM5D* and *USP9Y*) (Vogt *et al.* 2000). The resultant peptide is thought to be heavily involved in protein-protein interactions and graft rejection (Simpson *et al.* 1997; Vogt *et al.* 2000). Minor histocompatibility epitopes are either processed via the MHC class I pathways or presented in the context of MHC class II proteins where they are recognised by HLA-B60 restricted cytotoxic T lymphocytes (Goulmy. 1997; Vogt *et al.* 2000). The 1,347 amino acids of *UTY1* encode a number of tetratricopeptide repeats potentially producing various polymorphic intracellular proteins (Vogt *et al.* 2000; Warren *et al.* 2000). This gene has a large capacity for splice variation and six *UTY1* transcripts have been definitively identified (Laaser *et al.* 2011). Transcript variation is generally found at the C terminal while the N terminal remains ubiquitous (Vogt *et al.* 2000). The tissue-specific expression profile of this gene is wide and includes; thymus, bone marrow, lymph node, brain, cerebellum,

muscle, kidney, lung, prostate, testis, heart, thyroid, skin, monocytes and macrophages (Vogt *et al.* 2000).

The association between reduced expression level of *UTY1* and CAD, such as that hypothesised here, is interesting assuming *UTY1* and *UTX* have the same function. The X chromosome homolog of *UTY1* (*UTX*), which shows 86% similarity (Vogt *et al.* 2000), escapes X inactivation in females implying that normal physiological functioning requires a "double-dose" of the resultant peptide (Greenfield *et al.* 1998). Mortality and morbidity resulting from CAD is significantly higher in those women with only one functioning X chromosome, such as in those with Turner Syndrome (Greenfield *et al.* 1998). Turner Syndrome (45X-) affects 1 in 2,500 live births (Aligeti & Horn 2007) and is associated with increased osteoporosis, type 2 diabetes mellitus and CAD. It has been repeatedly hypothesised that the resultant haploinsufficiency of those genes which routinely escape X inactivation, such as *UTY1*, may be the cause of these pathologies in those with Turner syndrome (Greenfield *et al.* 1998). While it would be tempting to assign an association between decreased levels of *UTY1* and increased risk of CAD in those with Turner Syndrome, these data must be interpreted in context of high prevelence of CVD risk factors in these patients (28% and 50% of patients with Turner syndrome have hypertension and hyperlipidemia, respectively) (Kozlowska-Wojciechowska *et al.* 2006). However, the data shown here add to the emerging body of evidence suggestive of reduction in a Y homolog of a gene which escapes X inactivation as a potential contributor to increased risk of CAD in men. Indeed, reduced levels of *UTY1* in human macrophages associated with haplogroup I of the Y chromosome may hypothetically contribute to down-regulation of adaptive immunity and heightened response to inflammation – mechanisms linked to CAD. The changes in expression of a minor histocompatibility complex, such as *UTY1*, may translate into

changes in the immune system (von Boehmer & Hafen 1986). This hypothetical association is in keeping with data from the transcriptome analysis – immune/inflammatory pathways were identified at differentially regulated between carriers of haplogroup I and men with other lineages of the MSY (Chapter 4). Thus, the down-regulation of *UTY1* in macrophages of carriers of haplogroup I may be interpreted as a biologically plausible mechanism for how haplogroup I of the Y chromosome could increase risk of CAD (Cokkinos *et al.* 1985; Piedrahita *et al.* 1992; Stemme *et al.* 1995; de Boer *et al.* 2000; Zhou *et al.* 2000; Tentolouris *et al.* 2004; Hansson & Libby 2006). Interestingly, non-synonymous polymorphisms in *UTY1* are common (Vogt *et al.* 2000). It is therefore also tempting to speculate that haplogroup I-specific sequence alteration in *UTY1* may indeed lead to a reduction in gene expression level in macrophages. Give the 14% sequence divergence between *UTY1* and *UTX* it is difficult to hypothesise if the two genes may have the same, or differing function. Research into functional conservation, including primate work, may shed light on this..

6.4.2.1.1 UTY1 and methylation

The most attractive hypothesis emerging from data shown here is that down-regulation of *UTY1* (a minor histocompatibility complex) in macrophages may link the immune system and the Y chromosome and translate into the increased risk of CAD. An alternative hypothesis revolves around the epigenetic effect of the Y chromosome and potentially *UTY1*. The X chromosome homolog of this gene (*UTX*), which escapes X inactivation, is a well-known histone 3 lysine 27 demethylase implicated in development, self-renewal and cell differentiation through the modulation of transcription factors including *HOX* (Liu *et al.* 2012). It is plausible that *UTY1* also

harbours epigenetic effects. The role of the Y chromosome in immune modulation through epigenetic modulation has been shown in various animal models (Teuscher *et al.* 2006; Lemos *et al.* 2008; Spach *et al.* 2009; Lemos *et al.* 2010; Case *et al.* 2012). Of particular interest is the significant modulation of immune-related genes resulting from natural variation in heterochromatin (Lemos *et al.* 2012). Previous data from animal models and the fact that a gene with potential histone demethylase activity (*UTY1*) shows differential regulation between carriers of haplogroup I and all other haplogroup implies a possible role for epigenetics (including DNA methylation and nucleosome placement) in the Y chromosome modulation of the immune system (Teuscher *et al.* 2006; Lemos *et al.* 2008; Spach *et al.* 2009; Lemos *et al.* 2010; Case *et al.* 2012).

6.4.2.2 PRKY

PRKY showed a significant reduction in expression level in carriers of haplogroup I. *PRKY* is a transcribed pseudogene as it lacks exon 6 and the last nucleotide of exon 5 (which causes a frame-shift mutation). This mutation causes the loss of the last 31 amino acids of the kinase domain and a complete loss of the PKC C-terminal domain. As a result of these genetic losses, *PRKY* transcripts are liable to nonsense-mediated decay. The expression of this gene is wide – it was found in the thymus, bone marrow, spleen, lymph node, brain, spinal cord, cortex, cerebellum, heart, muscle, kidney, lung, colon, monocytes and macrophages, amongst others. However, no functional properties of this transcript have been identified to date. The primary function of the protein kinase family, of which *PRKY* is a part, is to modify proteins through phosphorylation of OH groups found on serine and threonine (Manning *et al.* 2002; Hanks 2003). The lack of

documented biological function and likelihood of transcript decay make PRKY an unlikely driver of the association between haplogroup I and immunity.

6.4.3 Co-regulation of UTY1 and PRKY

A relatively high level of mRNA correlation was observed between *UTY1* and *PRKY* (r~0.5). It has been repeatedly documented that genes showing a high correlation in mRNA expression profile are likely to be regulated by a common element or transcription factor (Allocco *et al.* 2004, Ideker *et al.* 2001). Allocco *et al.* identified in yeast that a correlation coefficient of ~0.8 gave a 50% chance that two genes share a common regulator (Allocco *et al.* 2004). While the correlation observed here falls well below this "cut-off" value, this may be due to interspecies differences.

6.4.4 Conclusions

Data presented here indicate that a currently unidentified haplogroup I-specific sequence alteration significantly decreases the expression levels of both *UTY1* and *PRKY* in macrophages. It can be hypothesised that the sequence alteration affecting expression of these genes resides in either the regulation element of MSY common for both UTY1 and PRKY. Further work is required to identify this variant. A logical extrapolation of these data suggests that the identified gene down-regulation may link haplogroup I to immune system (identified in Chapter 4), predisposing carriers of haplogroup I to CAD.

6.4.5 Future work

The work presented here represents the first step into identifying which transcript(s) of the Y chromosome underlie the effect of haplogroup I on the immune system and predisposition to CAD. Future work is likely to involve the identification and quantification of specific transcripts originating from both *UTY1* and *PRKY* in an effort to isolate the transcript and the functional variant which decreases gene expression levels in carriers of haplogroup I. Given the high number of none coding RNAs originating from the Y chromosome (Jehan *et al.* 2006) research into the expression of these by haplogroup may prove interesting.

6.4.6 Novel expression profile of AMELY

Data shown here represent the novel identification of *AMELY* mRNA in cells and tissues other than the tooth bud. To date, only limited investigations into MSY gene expression profiles have been completed, none of which have identified the presence of *AMELY* outside of the tooth (Lau *et al.* 2000), where it is involved in bio-mineralisation during tooth enamel development (Fincham *et al.* 1991; Salido *et al.* 1992). *AMELY* mRNA was identified in 46 of the macrophage and 35 kidney samples used here. Interestingly, two macrophage samples (~4%) and ten kidney samples (~20%) did not show evidence of *AMELY* mRNA. This is potentially due to the *AMELY* deletion, a well-documented mutation involving the loss of the *AMELY* gene in some males; however, the recorded population frequencies of this are significantly lower than that seen here (Mitchell *et al.* 2006; Iwase *et al.* 2007).

Data shown here imply that *AMELY* may have a more systemic role than previously hypothesised and serves as a reminder of how little is known of the true function and expression profile of the MSY genes.

6.4.7 Major limitations

6.4.7.1 Cohort size

Only approximately 20% of the cohort used in the transcriptome-wide analysis presented in Chapter 4 had RNA of sufficient quantity and quality for use here. While this represents a significant drop in number and sample size should be maximised, it is in keeping with the size of cohorts used in this form of analysis in the wider scientific community (LeGoff *et al.* 2011, Thuny *et al.* 2012).

6.4.7.2 Statistical analysis

As with all analyses completed in this work, an approach of haplogroup I vs. all other haplogroups was used. This has a number of advantages and disadvantages discussed in detail in section 3.4.5.1.

6.4.7.3 Kidney tissue

Kidney tissue was gained from the none-diseased renal pole from those under going nephrectomy for renal cancer. Normal histology of samples used here was identified through various methods (including immunohistochemistry) and the effect of maliganacy on work completed here is expected to be limited.

6.4.7.4 Hypothetical presence of population stratification in the kidney cohorts

While the presence of population stratification was excluded in the subset of CTS used here, no data were available for study in the two Polish kidney cohorts. As the two most common Polish Y chromosome haplogroups (R1a and I) have significantly different genetic backgrounds due to differing evolutionary and geographic pressures (haplogroup I originated in the middle-east [Charchar *et al.* 2012, Semino *et al.* 2000], while haplogroup R1a most likely originated from the Eurasian continent [Underhill *et al.* 2010]) the presence of population stratification cannot be formally excluded. However, to date, there has been no evidence of population stratification in any of the cohorts analysed (Charchar *et al.* 2012). This implies that the risk and potential effect of population stratification in this cohort is small. The definitive exclusion of population stratification in both kidney cohorts should be completed in future work.

6.4.7.5 RNA extraction

While the RNA extraction kits used here are a widely accepted and validated method of RNA extraction (Ishii *et al.* 2007), two limitations must be acknowledged. Firstly, the kits do not capture mRNAs below 200 nucleotides in length. This limitation has limited effect on this research as the vast majority of Y chromosome mRNAs are over 200 nucleotides (Skaletskey *et al.* 2003). Should shorter RNAs be targeted in the future, a specific RNA extraction kit must be used. Secondly, isolation of rRNA and tRNA is not possible with this kit. This limitation has no effect on this study as investigation of rRNA (the RNAs used for decoding mRNA to amino acid sequence) and tRNA was outside the scope of this work. Furthermore, no ncRNAs were captured through this

method. Given the high number of ncRNAs originating from the Y chromosome this limitation should be rectified in future work.

6.4.7.6 Alternative methods of gene expression analysis

The use of Taqman probes in low-throughput gene expression analysis projects is generally accepted to be the "gold-standard" of gene expression analysis (Giulietti *et al.* 2001). This is due to high specificity and fidelity of binding between the probes and cDNA sequence, and the subsequent reduction in false positive rate (Giulietti *et al.* 2001). An alternative system, Sybr green, has a significantly lower binding specificity and fidelity and a higher false positive rate (TaqMan® Chemistry vs. SYBR® Chemistry for Real-Time PCR). Taqman was therefore used in this study.

6.4.7.7 MSY testis-specific genes

Only those genes showing gene expression profiles outside of the testis were investigated here (Skaletsky *et al.* 2003). This decision was based on the assumption that genes not expressed in tissues relevant to CAD are extremely unlikely to affect predisposition to disease. However, the identification of mRNA in macrophages resulting from *AMELY*, a gene previously assumed to only be expressed in the tooth (Skaletsky *et al.* 2003), casts some doubt over the reported expression profiles of MSY genes. It is plausible that the 10 genes assumed to be testis-specific show a wider expression profile which has not been documented. Based on this assumption, it is possible that one of these genes could affect male predisposition to CAD. The

possibility of testis-specific genes being expressed in macrophages, monocyte and kidney tissue should be explored in future work.

7 The Y chromosome and abdominal aortic aneurysms

7.1 Introduction

CAD and AAA have been proposed as different manifestations of the same disease (atherosclerosis) (Grootenboer *et al.* 2003), and are known to share many environmental and genetic risk factors. These include male sex, family history of disease (particularly male), increased BMI, hyperlipidaemia and smoking (as explained in detail in Chapter 1). Possibly the most significant similarity between CAD and AAA is the over-representation of the male sex at diagnosis. CAD has a ratio of ~2:1 between men and women at diagnosis while AAA shows up to 6:1 male predominance. That AAA is male dominated is best exemplified by screening programmes that target men exclusively (Grootenboer *et al.* 2003, Ogata *et al.* 2005). Alongside the similarity in risk factors between CAD and AAA, both diseases are thought to share a similar molecular pathogenesis that leads to either an atherosclerotic plaque or aneurysm. Indeed, a number of common pathways, including immunity and cell-cycle pathways, have been identified in both disorders (Charchar *et al.* 2012, Lenk *et al.* 2007). Though the function of the inflammatory infiltrate in the arterial wall is not completely understood; it is thought to perpetuate or even initiate both diseases through the increased production of proteins such as ILs, MMPs and TIMPs (Daugherty & Cassis 2004; Libby & Theroux 2005; Pearce & Shively 2006; Watkins & Farrell 2006) (these processes are explained in detail within Chapter 1). Furthermore, CAD and AAA share a significant similarity at the genetic level. Several genes were associated with both CAD and AAA including; *ANRIL* (Samani *et al.* 2007; Bown *et al.* 2008), *ABO* blood group locus (Carpeggiani *et al.* 2010, Mahmoodi *et al.* 2012), chemokine (C-X-C motif) ligand 12 (*CXCL12*) (Samani *et al.* 2007, Ocaña *et al.* 2008) and disabled homolog 2 interacting protein (*DAB2IP*) (Gretarsdottir *et al.* 2010; Harrison *et al.* 2012). This genetic overlap

implies that there may be a common genetic cause of both CAD and AAA. Indeed, the importance of paternal history in the strong sexual dimorphism of both diseases suggests that their sexual inequality may be due to a genetic overlap. As previously discussed in detail (Chapters 3, 4 and 5), haplogroup I of the Y chromosome was identified as an independent risk factor for CAD, increasing risk of CAD by ~50% in its carriers (Charchar *et al.* 2012). It can be hypothesised that this "male-specific" risk factor is common between CAD and AAA, and possibly underlies the sexual dimorphism of both diseases. I therefore used three British cohorts of men with AAA, together with disease-free controls, to investigate if the Y chromosome is associated with the male predisposition to AAA, in the same way as it contributes to CAD.

7.2 Materials and methodology

7.2.1 Materials

7.2.1.1 Leicester Abdominal Aortic Aneurysm Cohort

The Leicester AAA cohort was recruited in Leicester, UK. Patients with AAA were identified through three sources: an ongoing regional AAA screening program, a vascular outpatient clinic and the vascular admissions unit at Leicester Royal Infirmary between 2000 and 2009 (n=2,466). Unrelated healthy individuals were also recruited through the AAA screening program (Bown *et al.* 2008). Both cases and controls were screened for AAA using either ultrasonography or computed tomography. AAA was defined as an abdominal aortic aneurysm >3 cm while healthy controls had a maximal infrareanal aortic diameter of 2.5 cm. A complete medical history was obtained for each individual, including smoking history, familial history of AAA or hypertension (defined as >140/90 mmHg or existing BP lowering treatment), hyperlipidaemia (defined as TC >5.5 mmol/L or existing lipid lowering treatment), diabetes mellitus or presence of CAD. A peripheral blood sample was also taken for DNA analysis (Bown *et al.* 2008). 2,008 male individuals had DNA of sufficient quality and quantity available for use in this study.

7.2.1.2 Leeds Abdominal Aortic Aneurysm Cohort

Patients with AAA were recruited from those attending the Leeds Vascular Institute (Leeds General Infirmary). AAA-free controls were recruited from non-vascular outpatient clinics (total n=472). Presence or absence of AAA was confirmed by abdominal ultrasonography. Medical history and anthropometric measurements (height and weight) were recorded using a standardised questionnaire. A peripheral blood

sample for DNA extraction and blood biochemistry (including lipid profile) was secured from each participant (Bown *et al.* 2011). 369 male individuals had DNA of sufficient quality and quantity for use in this project.

7.2.1.3 Belfast Abdominal Aortic Aneurysm Cohort

Patients recruited into the Belfast AAA cohort were identified through in-patient and out-patient clinics, as well as AAA screening lists at Belfast City Hospital. Controls were recruited through the AAA screening program (total n=650). Presence or absence of AAA was confirmed by abdominal ultrasonography. Medical history, anthropometric data (height and weight) and a peripheral blood sample (for DNA extraction and baseline blood chemistry [including lipids]) were obtained from each individual (Bown *et al.* 2011, Bradley *et al.* 2011). 421 male individuals had DNA of sufficient quality and quantity for use in this study.

7.2.1.4 Bioethical approval

Written, informed consent was obtained from all subjects in all studies used here, in accordance with the Declaration of Helsinki. All studies had approval from relevant institutional ethical committees.

7.2.2 Methodology

7.2.2.1 Genotyping and Y chromosome haplogrouping

All details regarding genotyping and Y chromosome haplogrouping can be found in Chapter 2.

7.2.2.2 Descriptive statistics

Age, dyslipidaemia, hypertension, history of smoking and prevalence of diabetes were summarised using mean and SD and percentages in Leicester AAA cohort. Age, BMI, HDL-C, LDL-C, current and ex-smoking prevalence were summarised in Leeds AAA cohort using means, SD and percentages. Finally, age, BMI, dyslipidaemia, hypertension, diabetes and history of smoking were summarised using means, SD and percentages in Belfast AAA cohort. Difference in clinical characteristics was between either cases and controls or haplogroup I and all others as assessed using Student's *t* test (continuous traits) or Chi^2 test (discrete traits).

7.2.2.3 Analysis of prevalence of Y chromosome haplogroups by ethnicity

Due to the difference in ethnicity of the three cohorts (two English cohorts and one Irish cohort), a Chi^2 test was used to assess if the prevalence of six haplogroups (E1b1b1, I, K*, J*, R1a and R1b1b2) differed due to ethnicity.

7.2.2.4 Analysis of population stratification

The presence of population stratification in cases recruited to Leicester AAA cohort was assessed using genome-wide SNP data from 670K BeadChips (Illumina, USA) (Bown *et al.* 2011). The method of analysis was identical to that used in the assessment of CTS (Chapter 3 and 5).

Analysis of population admixture in Leicester AAA cohort control samples and Leeds AAA cohort was analysed in a manner identical to that used in the analysis of YMCA 1 and 2. Details of this method can be found in Chapter 3.

7.2.2.5 Analysis of the role of haplogroup I in predisposition to AAA

Crude, age-adjusted and fully-adjusted (adjusted for age, treated hypertension and history of smoking) linear regression was used to investigate the role of haplogroup I in predisposition to AAA in all three cohorts. Meta-analysis of all three cohorts was conducted using a fixed-effect, inverse-variance model in Metan script (Deeks *et al.* 2001) as described in Chapter 3.

7.2.2.6 Analysis of the role of the Y chromosome in predisposition to AAA

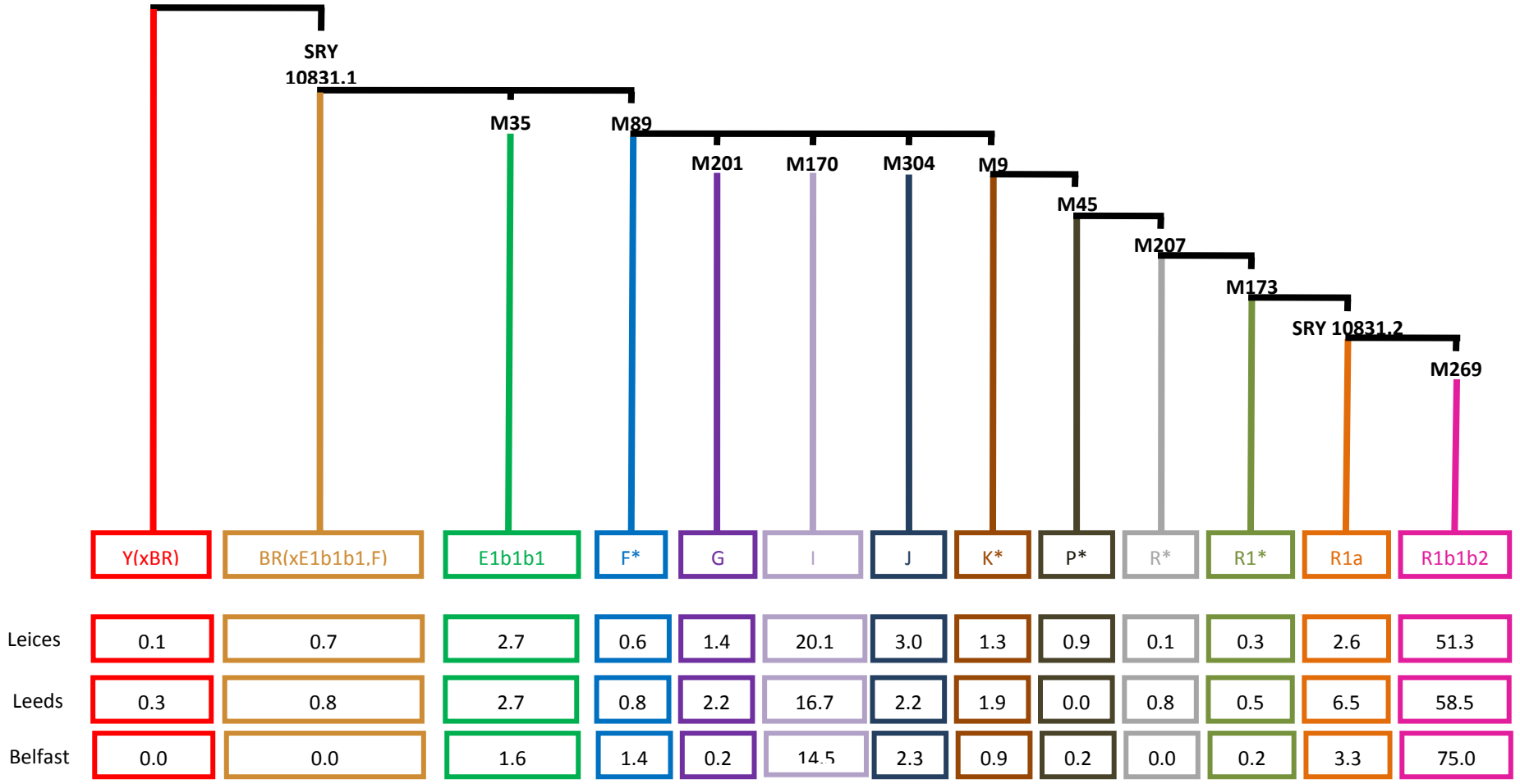
Five Y chromosome haplogroups (with prevalence over 2% in each cohort) were examined for association with AAA in a manor identical to that detailed in Chapter 6.2.2.5.

7.3 Results

7.3.1 Y chromosome haplogroup distribution in three AAA cohorts

Of the 13 common European Y chromosome lineages analysed here, all 13 were represented in Leicester AAA cohort study, 12 were identified in Leeds AAA cohort study and 10 were detected in the Belfast AAA cohort population (Figure 7.1). Of the haplogroups identified, E1b1b1, I, K*, J, R1a and R1b1b2 had a prevalence greater than ~2% in all three cohorts. Haplogroups R1b1b2 (51%, 58% and 75% in Leicester AAA cohort, Leeds AAA cohort and Belfast AAA cohort, respectively) and I (20%, 16% and 14% in Leicester AAA cohort, Leeds AAA cohort and Belfast AAA cohort, respectively) accounted for between 71% and 89% of the haplotypic variance across the three cohorts (Leicester AAA cohort: 71.4%, Leeds AAA cohort: 75.2%, Belfast AAA cohort: 89.5%).

Figure 7.1 Y chromosome haplogroup prevalence in Leicester AAA cohort, Leeds AAA cohort and Belfast AAA cohort. Leices-Leicester AAA cohort, Leeds-Leeds AAA cohort, Belfast-Belfast AAA cohort, %-percentage prevalence. Haplogroups are arranged from Y(xBR) on the left to R1b1b2 on the right. M and SRY numbers represent the terminal SNP dictating that haplogroup. SRY10831 is a recurrent mutation.



7.3.2 Distribution of clinical and demographic characteristics of populations, stratified by case-control status and haplogroup (*I vs. all others*)

As expected, there was a significant difference in prevalence of various AAA risk factors between AAA cases and controls. The prevalence of dyslipidaemia and hypertension was higher in Leicester AAA cohort cases when compared to controls. Dyslipidaemia, hypertension and smoking were more common in Belfast AAA cohort cases compared to controls (Table 7.1). No statistically significant difference in any clinical or demographic characteristic was observed between Leeds AAA cohort cases and controls (Table 7.1). Conversely, there was no statistically significant difference (after correction for multiple testing) in clinical or demographic measurements in any of the three cohorts when stratified by haplogroup (I vs. all other haplogroups) (Table 7.2).

Table 7.1 Distribution of clinical and demographic data of individuals recruited into Leicester AAA cohort, Leeds AAA cohort and Belfast AAA cohort stratified by case-control status

	Leicester AAA cohort		
	AAA cases	**AAA-free controls**	**P value**
n	1025	983	-
Age (years)	72.9 (7.3)	66.9 (5.0)	0.832
Dyslipidaemia (%)	342 (33)	201 (20)	0.0001
Hypertension (%)	567 (55)	374 (38)	0.0001
Current smoker (%)	134 (13)	101 (10)	0.052
Diabetes (%)	92 (9)	101 (10)	0.326
	Leeds AAA cohort		
n	183	186	-
Age (years)	74.2 (8.1)	69.8 (7.5)	0.994
BMI (kg/m^2)	27.1 (4.2)	27.8 (3.8)	0.163
HDL-cholesterol	1.2 (0.3)	1.3 (0.4)	0.937
LDL-cholesterol (mmol/L)	2.2 (0.9)	2.6 (0.9)	0.783
Current smokers (%)	42 (23)	32 (17)	0.193
Ex-smokers (%)	120 (66)	112 (60)	0.332
	Belfast AAA cohort		
n	184	237	-
Age (years)	72.1 (6.7)	69.1 (4.2)	0.612
BMI (kg/m^2)	26.4 (3.5)	26.3 (3.9)	0.342
Dyslipidaemia (%)	86 (47)	89 (38)	0.001
Hypertension (%)	116 (63)	129 (54)	0.015
Diabetes (%)	25 (14)	42 (18)	0.448
History of smoking (%)	54 (35)	27 (11)	0.001

Data are means and SD or counts and percentages. P value - statistical significance of a difference between cases and controls calculated by Student's *t* test or Chi^2 test, n-number of subjects, BMI-body mass index.

Table 7.2 Distribution of clinical and demographic data of individuals recruited to Leicester AAA cohort, Leeds AAA cohort and Belfast AAA cohort stratified by Y chromosome haplogroup

	Leicester AAA cohort		
	I	**Other**	**P value**
n	403	1605	-
Age (years)	70.1 (7.0)	69.8 (6.9)	0.324
Dyslipidaemia (%)	103 (26)	406 (25)	0.490
Hypertension (%)	183 (45)	715 (45)	0.539
Current smoker (%)	50 (12)	168 (10)	0.604
Diabetes (%)	31 (8)	156 (10)	0.157
	Leeds AAA cohort		
n	60	309	-
Age (years)	71.4 (7.8)	72.1 (8.1)	0.257
BMI (kg/m^2)	26.9 (4.4)	27.6 (3.9)	0.719
HDL-cholesterol (mmol/L)	1.2 (0.4)	1.1 (0.5)	0.356
LDL-cholesterol (mmol/L)	2.3 (1.2)	2.1 (1.2)	0.343
Current smokers (%)	14 (23)	60 (19)	0.484
Ex-smokers (%)	33 (55)	199 (65)	0.190
	Belfast AAA cohort		
n	62	359	-
Age (years)	69.9 (5.6)	70.5 (5.7)	0.475
BMI (kg/m^2)	26.1 (3.3)	26.4 (3.8)	0.095
Dyslipidaemia (%)	13 (21)	127 (35)	0.013
Hypertension (%)	26 (42)	171 (48)	0.491
Diabetes (%)	8 (13)	46 (13)	0.523
History of smoking (%)	9 (15)	46 (13)	0.838

Data are means and SD or counts and percentages. P-value - statistical significance of a difference between contrasting lineages of the MSY calculated by Student's *t* test or Chi^2 test, n-number of subjects, BMI-body mass index.

7.3.3 Prevalence of Y chromosome haplogroups by ethnicity

Of the six haplogroups represented in each cohort with a frequency >2%, only haplogroups I and R1b1b2 showed significant differences in prevalence between English and Belfast samples after correction for multiple testing (Table 7.3).

Table 7.3 Difference in prevalence of common Y chromosome haplogroups between English and Irish men

Haplogroup	English cohorts	Belfast cohort	P value
n	2367	421	-
E1b1b1	78 (3)	7 (2)	0.085
I	343 (15)	93 (22)	0.001
K*	36 (2)	4 (1)	0.501
J*	78 (3)	9 (2)	0.285
R1a	118 (5)	14 (3)	0.165
R1b1b2	1411 (59)	316 (75)	0.001

Data are counts and percentages (over ~2% prevalence) stratified by origin (English cohorts [Leicester AAA cohort and Leeds AAA cohort] and Belfast cohort [Belfast AAA cohort]). P value - level of statistical significance in difference in prevalence of each haplogroup calculated using Chi^2 test.

7.3.4 Analysis of population stratification

7.3.4.1 Population stratification in AAA cases recruited into Leicester cohort

Analysis of population stratification in AAA cases recruited to the Leicester AAA cohort was completed using genome-wide SNP data. No systematic genetic differences were identified between carriers of haplogroup I compared to carriers of all other haplogroups (Figure 7.2b). The three outliers detected (Figure 7.2b) differed only weakly from the remaining population and therefore were not excluded from further

analysis. The multi-dimensional scaling of Leicester AAA cohort cases with three HapMap populations as a reference confirmed that all recruited individuals were of European ancestry (Figure 7.2a).

7.3.4.2 Analysis of population stratification in Leicester AAA cohort cases and Leeds AAA cohort

The validated 34 SNP-based assessment of population substructure in Leicester AAA cohort controls and Leeds AAA cohort showed that all individuals used in this analysis were of European ancestry (Figure 7.3).

Figure 7.2 Analysis of population stratification of AAA cases recruited into Leicester AAA cohort. A. A matrix of identity-by-state distance was created in the Leicester AAA cohort cases and the three HapMap populations (CEU, JPT+CHB, YRI). Coloured dots represent each population: blue- JPT+CHB, green- YRI, red- CEU, yellow- Leicester AAA cohort cases. The first two components obtained from the subsequent multi-dimensional scaling of the matrix were used as co-ordinates to compare the relative distances between the four populations. B. Colours represent haplogroups: green- carriers of haplogroup I, red- carriers of all other haplogroups. A matrix of identity-by-state distances was created and the first two components obtained from the subsequent multi-dimensional scaling of the matrix were used as co-ordinates to generate a scatter plot.

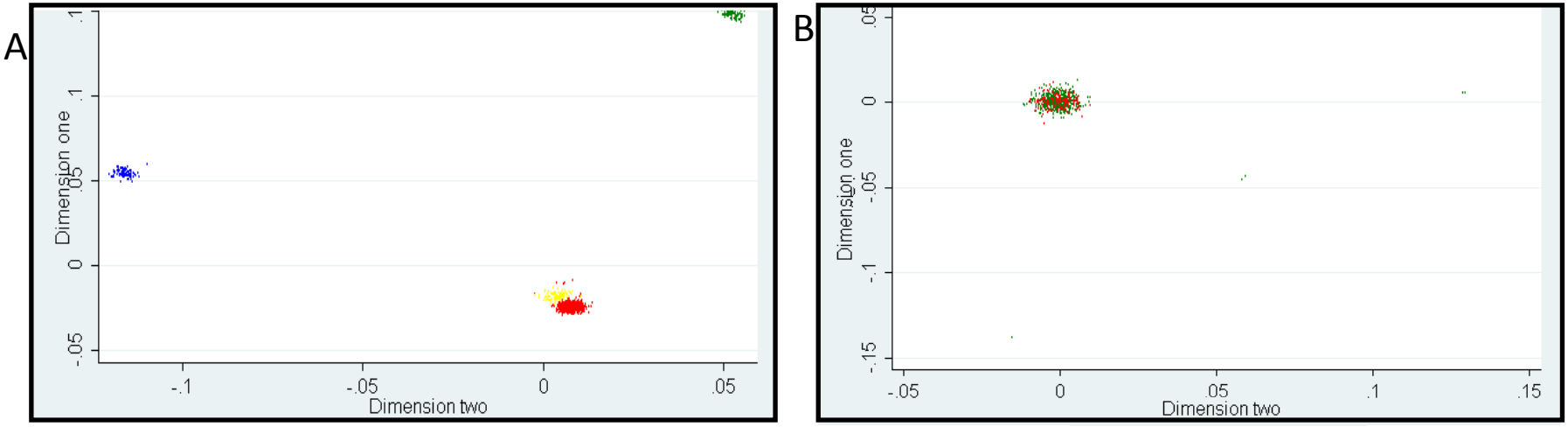
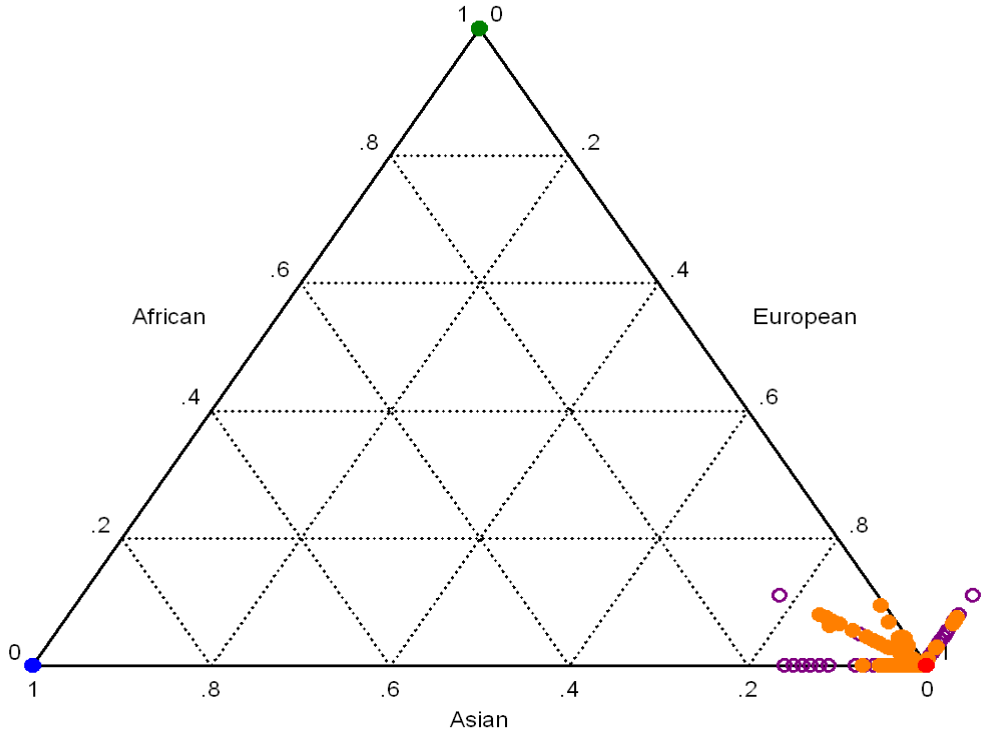


Figure 7.3 Analysis of cohort ancestry- Leicester AAA cohort controls and Leeds AAA cohort. Colour points represent individuals. Green- African reference samples, blue- Asian reference samples, red- European reference samples, orange- Leicester controls, purple- Leeds AAA cohort. Individuals were plotted outside of the triangle to ease interpretation of high sample numbers.



7.3.4.3 Analysis of association between haplogroup I of the Y chromosome and prevalence of AAA in three British cohorts

Prevalence of haplogroup I of the Y chromosome across the three cohorts ranged from 10-21% and 12-19% in cases and controls, respectively (Table 7.3). There were no statistically significant difference in prevalence of this haplogroup between those with AAA and those without in any of three cohorts (Table 7.4). The fixed-effect inverse variance age-adjusted meta-analysis confirmed that there was no relationship between haplogroup I and prevalence of AAA in these cohorts (P= 0.693) (Table 7.4).

Table 7.4 Analysis of association between haplogroup I of the Y chromosome and prevalence of AAA in three British cohorts

Cohort	AAA cases (%)	AAA-free controls	P value$^{\alpha}$	P value$^{\gamma}$	P value$^{\Delta}$
Leicester AAA cohort	21.2	18.9	0.220	0.650	0.632
Leeds AAA cohort	15.2	18.1	0.489	0.815	0.644*
Belfast AAA cohort	10.3	12.5	0.499	0.530	0.913

Data are percentage prevalences of haplogroup I of the Y chromosome in AAA cases and AAA-free controls in each of the three cohorts analysed. P value$^{\alpha}$- crude analysis, P value$^{\gamma}$- age-adjusted linear regression, P value$^{\Delta}$- fully-adjusted (age, treated hypertension, any smoking) linear regression. *Adjusted for age and smoking only due to limited availability of phenotypes.

Table 7.5 Age-adjusted fixed-effect inverse-variance meta-analysis of association between haplogroup I of the Y chromosome and prevalence of AAA

Cohort	n	β coefficient (SE)	P value	Weight (%)	Heterogeneity P value
Leicester AAA	2008	0.024 (0.1)	0.849	74.44	-
Leeds AAA cohort	369	-0.150 (0.3)	0.612	12.42	-
Belfast AAA	427	-0.055 (0.3)	0.850	15.14	-
Meta-analysis	2804	-0.100 (0.2)	0.693	100	0.852

SE- standard error.

7.3.4.4 Analysis of association between other Y chromosome haplogroups and the risk of AAA in three British cohorts

Five Y chromosome haplogroups (other than haplogroup I) had an approximate prevalence of >2% in all cohorts and sufficient power to detect small effects (Figure 7.1) and were chosen for further analysis (each individual haplogroup vs. all other haplogroups). Unadjusted analysis in each cohort showed an association between haplogroups R1a, R1b1b2 and prevalence of AAA in Leicester AAA cohort, while haplogroup R1b1b2 was significantly associated with odds of AAA in Belfast AAA cohort. However, neither of these associations retained significance in age and fully adjusted analyses. No significant association was identified between any haplogroup and AAA in the age-adjusted meta-analysis of all three cohorts (Tables 7.5- 7.9).

Table 7.6 Age-adjusted fixed-effect inverse-variance meta-analysis of association between haplogroup E1b1b1 of the Y chromosome and prevalence of AAA

Cohort	n	β coefficient (SE)	P value	Weight (%)	Heterogeneity P value
Leicester AAA	2008	0.405 (0.3)	0.189	75.70	-
Leeds AAA cohort	369	-1.133 (0.7)	0.314	14.31	-
Belfast AAA	427	-0.854 (0.8)	0.113	9.99	-
Meta-analysis	2804	0.060 (0.4)	0.825	100	0.073

SE- standard error.

Table 7.7 Age-adjusted fixed-effect inverse-variance meta-analysis of association between haplogroup J* of the Y chromosome and prevalence of AAA

Cohort	n	β coefficient (SE)	P value	Weight (%)	Heterogeneity P value
Leicester AAA	2008	-1.77 (0.3)	0.246	72.44	-
Leeds AAA cohort	369	0.074 (0.7)	0.925	12.42	-
Belfast AAA	427	-0.062 (0.7)	0.913	15.14	-
Meta-analysis	2804	-0.130 (0.5)	0.620	100	0.946

SE- standard error.

Table 7.8 Age-adjusted fixed-effect inverse-variance meta-analysis of association between haplogroup K* of the Y chromosome and prevalence of AAA

Cohort	n	β coefficient (SE)	P value	Weight (%)	Heterogeneity P value
Leicester AAA	2008	-0.574 (0.5)	0.560	61.40	-
Leeds AAA cohort	369	0.371 (0.8)	0.762	24.05	-
Belfast AAA	427	0.308 (1.0)	0.634	14.55	-
Meta-analysis	2804	-0.220 (0.7)	0.573	100	0.511

SE- standard error.

Table 7.9 Age-adjusted fixed-effect inverse-variance meta-analysis of association between haplogroup R1a of the Y chromosome and prevalence of AAA

Cohort	n	β coefficient (SE)	P value	Weight (%)	Heterogeneity P value
Leicester AAA	2008	0.064 (0.3)	0.811	63.47	-
Leeds AAA cohort	369	-0.637 (0.5)	0.199	22.48	-
Belfast AAA	427	0.723 (0.6)	0.162	14.25	-
Meta-analysis	2804	0.001 (0.4)	0.995	100	0.156

SE- standard error.

Table 7.10 Age-adjusted fixed-effect inverse-variance meta-analysis of association between haplogroup R1b1b2 of the Y chromosome and prevalence of AAA

Cohort	n	β coefficient (SE)	P value	Weight (%)	Heterogeneity P value
Leicester AAA	2008	0.076 (0.1)	0.502	68.12	-
Leeds AAA cohort	369	0.133 (0.2)	0.361	16.07	-
Belfast AAA	427	-0.182 (0.2)	0.562	15.81	-
Meta-analysis	2804	0.040 (0.1)	0.636	100	0.560

SE- standard error.

7.4 Discussion

To date haplogroup I of the Y chromosome has been associated with ~50% increase in risk of CAD (Charchar *et al.* 2012). Given the strong sexual dimorphism and molecular and genetic overlap in pathogenesis of CAD and AAA, it was hypothesised that haplogroup I of the Y chromosome could predispose its carriers to AAA.

7.4.1 Y chromosome haplogroup distribution and clinical demographics

R1b1b2 (50-80%) and I (10-20%) are the most common Y chromosome haplogroups in England (Capelli *et al.* 2003). The prevalence of these haplogroups in both Leicester AAA cohort and Leeds AAA cohorts is analogous to the published data (Figure 7.1). Due to differing evolutionary pressures, the haplogroup distribution in Ireland shows differences from that in England; namely reduced prevalence of haplogroup I and increased prevalence of haplogroup R1b1b2 (Moore *et al.* 2005). These expected differences were confirmed in the analysis here (Figure 7.1, Table 7.5).

As expected, there were significant differences in the prevalence of AAA risk factors (including dyslipidaemia, hypertension and smoking) between cases with AAA and controls. This is in keeping with published data on prevalence of AAA risk factors (Grootenboer *et al.* 2009). Surprisingly, there was no statistical difference in diabetes between cases and controls. Diabetes is a known protective factor in predisposition to AAA (Grootenboer *et al.* 2003). That there was no difference in prevalence of diabetes between cases with AAA and the controls used here is most likely due to the overall decreased rate of diabetes in these cohorts compared to the expected value (expected value ~30% [Howard *et al.* 1995]). There was no significant difference in clinical or

demographic factors between carriers of haplogroup I compared with carriers of all other haplogroups.

7.4.2 Association of haplogroup I of the Y chromosome and AAA

As previously discussed in detail (Chapters 1, 4 and 5), haplogroup I of the Y chromosome acts as a "male-specific" risk factor for CAD and increases risk of CAD by ~50% in its carriers. This effect is most likely mediated through the modulation of pathways associated with autoimmunity and inflammation (Charchar *et al.* 2012). Interestingly, both inflammation and autoimmunity have previously been associated with development and perpetuation of AAA (Golledge *et al.* 2009). This is exemplified in a study by Lenk *et al.* where 17 pathways heavily involved in the immune system were found to be differentially regulated in aortic tissue of those with AAA compared with non-aneurysmal aortic tissue (Lenk *et al.* 2007). Furthermore, there was ~50% overlap in pathways identified as associated with AAA (Lenk *et al.* 2007), and those identified as up- or down-regulated in carriers of haplogroup I (cell adhesion molecules, focal adhesions, leukocyte transendothelial migration, type 1 diabetes mellitus, ECM receptor interaction, hematopoietic cell lineage, antigen processing and presentation and cytokine-cytokine receptor interaction (Lenk *et al.* 2007; Charchar *et al.* 2012). AAA has frequently been hypothesised to represent an autoimmune disease (Hirose & Tilson 2001; Pearce & Shively 2006). These data strongly suggest that haplogroup I of the Y chromosome may play a role in the male predisposition to AAA through modulation of the immune system. However, data shown here indicate that there is no association between haplogroup I and the risk of AAA in any of the three cohorts individually or in their meta-analysis (Figure 7.4). These data also suggest that there are significant

differences in the pathogenesis of CAD and AAA, and that the effect haplogroup I has on the male predisposition to CAD is absent in AAA. This may be surprising given the overlap in pathways associated with AAA and those in men with haplogroup I, yet these data are in keeping with emerging evidence on differences in the genetics of AAA and other forms of CVD. Indeed, there is a very small overlap in genetic loci identified as affecting risk of CAD in various GWASs (Charchar *et al.* 2012) and in AAA. Just five loci (*ANRIL, ABO, CXCL12, LDLR, DAB2IP*) were found to be associated with AAA (Bown *et al.* 2008; Ocaña *et al.* 2008; Gretarsdottir *et al.* 2010; Hinterseher *et al.* 2011; Owens *et al.* 2011; Harrison *et al.* 2012; Mahmoodi *et al.* 2012) and CAD.

7.4.3 Association of other European Y chromosome lineages and AAA

While haplogroup I of the Y chromosome does not predispose its carriers to AAA, it is possible that another Y chromosome haplogroup may be associated with this disease. Therefore, five common haplogroups with prevalence over ~2% (E1b1b1, J*, K*, R1a and R1b1b2) were investigated in relation to AAA. This analysis showed no evidence of association between haplogroups E1b1b1, J*, K*, R1a or R1b1b2 and AAA (meta-analysis P= 0.825, P= 0.620, P= 0.573, P= 0.995 and P= 0.636, respectively). These data suggest that, at least in populations from the UK, common European Y chromosome haplogroups do not have a major role in the male predisposition to AAA.

7.4.4 Possible causes of the sexual dimorphism in AAA

While data shown here do not support the notion of a major role of the Y chromosome in the male predisposition to AAA, a number of potential contributors could account for

the sexual dimorphism in AAA. Arguably, the most significant of these is the over-representation of smoking in males. Smoking represents the biggest risk factor for AAA and significantly increases the risk of the disease (Teo *et al.* 2006). More men have a history of smoking, and smoke more cigarettes than women; this means that the prevalence of AAA in men may simply be a consequence of increased smoking in men (Teo *et al.* 2006). A similar finding is observed in relative alcohol consumption (Singh *et al.* 2001). Sex steroids may also have a marked effect on prevalence to AAA, although the contribution of both androgens and oestrogens to AAA is disputed (Tollerud *et al.* 1990, Henriques *et al.* 2008, Webb & Collins 2010, Wu *et al.* 2009, Yeap *et al.* 2010). Similarly, the differences in the functioning of the immune system between the sexes may affect male predisposition to AAA (possible causes of sexual dimorphism are discussed in detail in Chapter 1). Furthermore, it is plausible that male-dominated workplaces, including refuse collection and goods distribution, may increase AAA prevalence in men by unknown mechanisms (Yusuf *et al.* 2004). However, there is currently insufficient data available to either support or disprove this hypothesis. It is most likely that the cumulative effect of sexually dimorphic risk factors cause the male predominance in AAA diagnosis (Yusuf *et al.* 2004; Bloomer *et al.* 2012). Finally, it is possible that a Y-chromosome-specific effect common to all European haplogroups predisposes European men to AAA; this would explain the sex-bias, and the failure to find differences between European haplogroups. Studies of Y haplogroups (including C, D and O) found at high frequency in Asian males could in principle be informative. Finally, the role of the X chromosome in male-dominated diseases is widely accepted. It is plausible that X-linked "recessive" alleles may increase risk of AAA in men.

7.4.5 Conclusion

Data shown here do not support the hypothesis that the male-specific CAD risk factor, haplogroup I of the Y chromosome, plays a role in predisposition to AAA.

7.4.6 Future work

Further studies into the potential role of the Y chromosome and predisposition to AAA in other ethnicities (including Asian, where AAA is very rare) may prove interesting.

Due to lack of funds, analysis of population stratification in the Belfast AAA cohort was not completed here. While this is unlikely to have significantly affected the data here, the presence of this confounding factor should be formally discounted.

7.4.7 Major limitations

7.4.7.1 Cohort phenotyping

Recruitment of cases in all three cohorts was executed in both vascular clinics (criteria: aortic diameter >3 cm) and emergency department (generally ruptured aneurysm assumed to be >5.5 cm) (Bown *et al.* 2011). It is plausible that these systematic differences in selection criteria may have resulted in a wide phenotypic spectrum (Biros *et al.* 2011). Indeed, small aneurysms (those below 5.5 cm [Powell *et al.* 2011]) and large/ruptured aneurysms may represent differing phenotypes (Powell *et al.* 2010), and their joint analysis may dilute the power to identify the causative genetic factors. Though possible, the same phenotypic spectrum has been successfully used in the identification of other genetic effectors of AAA (Bown *et al.* 2008). While the separate

analysis of small and ruptured aneurysms may prove fruitful, the numbers of rupture AAA cases are low and thus difficult to collect.

7.4.7.2 Case-control studies

There are significant drawbacks which must be kept in mind when using case-control studies. Firstly, the risk of confounding effects from both selection bias (where cases and controls are recruited from separate groups) and recall bias (where information is remembered with differing accuracy in case and control individuals) is an inherent limitation of this type of analysis (Lewallen & Courtright 1998, Austin *et al.* 1994). The efforts made to minimise such problems across the three cohorts included recruiting control samples from the same screening programs (those who were validated as healthy) and gaining medical history from medical notes. The fidelity of this recruitment was based upon accurate and up to date medical records. The use of larger cohorts, such as those used here, partially alleviates this problem, although very weak effect associations may not be identified. The use of prospective studies would obviously be more informative, however, the availability of these cohorts are very limited (Li & Chiu 1996) and no such cohort with AAA as main phenotype was available for this study.

As previously discussed, Y chromosome analysis using cohorts of differing ethnicity can cause both false positive findings and prevent the identification of associations due to hidden population stratification. Prevalence data of haplogroup I and R1b1b2 show significant differences between the two origins studied here, although these are in line with previous studies.

7.4.7.3 Statistical analysis

7.4.7.3.1 Haplogroup analysis

As previously discussed (Chapter 3.4.5.1), it remains possible that a higher resolution European lineage of the Y chromosome (e.g. I1) is associated with AAA. Further statistical analysis into the association of Y chromosome sub-haplogroups in these cohorts was not possible due to the small sample numbers which would be assigned to each specific lineage. Similarly, it is plausible that a Y chromosome haplogroup not analysed here predisposes its carriers to AAA (including F*, G, P*, R* and R1*); again the same limitations of sample size prevented its investigation. However, given the very low prevalence of these haplogroups, they are unlikely to account for a significant proportion of the sexual dimorphism observed in AAA. It is also plausible that the Y chromosome as a whole predisposes men to AAA, and given the haplogroup-specific analysis completed here this effect would not have been identified.

While the use of covariate analysis is very important in genetic association studies, its use can elicit problems. In this case, the use of smoking as a covariate, while important due to the significant increase in AAA in smokers, could be problematic. Generally, patient reporting of smoking levels and duration is unreliable and fraught with inconsistencies (for example, having quit smoking yesterday may count as an ex-smoker to one individual, whilst a second individual may require the subject to be cigarette free for a number of months to be deemed an ex-smoker). A more reliable measure of smoking uses blood cotinine levels, although this was only available for Belfast AAA cohort samples. In an effort to keep statistical analysis as uniform as possible, the decision was made to use history of smoking as the covariate.

7.4.7.3.2 Population stratification

Assessment of population stratification was completed using two differing methods. A full discussion of the limitations of these methods can be found in Chapter 3. Importantly, due to lack of funds it was not possible to complete any form of analysis for population stratification on the Belfast AAA cohort. While it is possible that significant population stratification exists in this cohort, as the smallest of the three cohorts the potential presence of population stratification would have only a very small (if any) effect during meta-analysis. However, future work should include some form of population stratification assessment in this cohort.

8 Discussion

8.1 CAD as a predominantly male auto-immune disease

The involvement of monocytes and macrophages, cytokines, adhesion molecules and bacterial products in the pathogenesis of CAD has led to its widespread acceptance as an immunological disease (Chapter 1) (Libby & Theroux 2005; Watkins & Farrell 2006). Recent work has further revealed that CAD may result from a breach of self-tolerance, and it is evident that atherosclerotic plaque auto-antigens are targeted by IgM and IgG (Nilsson & Hansson 2008). The formation of foam cells (lipid-laden macrophages) is an auto-immune process whereby scavenger receptors on macrophages detect and phagocytose endogenous oxidised LDL-C (Nilsson & Hansson 2008). Foam cell formation and auto-antigen presentation further leads to an expansion of T and B cell clones specific against LDL-C (amongst other modified endogenous structures) (Palinski *et al.* 1989; Palinski & Witztum 2000). CAD risk factors such as smoking, hypertension, hyperlipidaemia and diabetes may act through auto-immune pathways that damage the arterial intima and increase oxidation of LDL-C (Ambrose *et al.* 2003).

CAD is more common in men than women, and this sexual dimorphism has been documented in many populations (further detail in Chapter 1) (Lowry *et al.* 1983; Ness & Aronow 1999; Tomaszewski *et al.* 2009; Roger *et al.* 2011). The potential causes that may underlie this apparent male disadvantage in CAD include sex differences in sexually dimorphic risk factors, relative concentrations of sex steroids and traits traditionally perceived as typical for one sex such as aggression (further details in Chapter 1, 3 and 4). However, none of these factors individually explains why CAD is significantly more common in men than age-matched women, or indeed, why a paternal history of CAD is one of the strongest risk factors for CAD (Miller 2012).

8.1.1 The Y chromosome and CAD

Seminal work conducted by Charchar *et al.* revealed that the Y chromosome is associated with risk of CAD (Charchar *et al.* 2012). Indeed, haplogroup I of the Y chromosome increased the risk of CAD by ~50%, independently of known CAD risk factors (further details can be found it Chapter 1) (Charchar *et al.* 2012). In an effort to elucidate the mechanism through which this lineage of the Y chromosome predisposes its carriers to CAD, a comprehensive investigation into the role of the Y chromosome in CAD risk factors and "male-associated" or "sex specific" traits was completed (Chapter 3 and 4). In contrast to existing data showing the association between the Y chromosome and lipid levels and BP (Ellis *et al.* 2000; Charchar *et al.* 2002; Shankar *et al.* 2007; Russo *et al.* 2006), Y chromosome haplogroup had no effect on any of the conventional CAD risk factors (further details in Chapter 3) in the general population. The Y chromosome haplogroups also had no statistically effect on levels of aggression and sex steroids (Chapter 4). These data revealed that the association between haplogroup I of the Y chromosome and CAD is not mediated by traditional cardiovascular risk factors, sex steroids or "male-associated" / "sex specific" traits. It is most likely that haplogroup I increases risk of CAD through different mechanisms than those studied before.

8.1.2 The Y chromosome and the immune system

At the time of writing, data on the role of the MSY in human health and disease were extremely limited. Due to its haploid nature (present in single copy in physiologically normal men), the Y chromosome is routinely removed from GWAS investigations. The association between MSY and male-specific cancers (Ferlin *et al.* 2007) and infertility

(Totonchi *et al.* 2012) is well-established, but there is a paucity of data on the role of human Y chromosome in processes and disorders not directly related to male reproduction. In an effort to delineate the pathways that underlie the association between the Y chromosome and CAD risk, a hypothesis-free investigation based on genome-wide transcriptome data was conducted. The results from this investigation (Chapter 5) revealed for the first time that the Y chromosome haplogroup might directly modulate the immune system. Indeed, pathways associated with inflammation and auto-immunity associated with haplogroup I are hypothesised to predispose its carriers to CAD (further details in Chapter 5). Furthermore, the decreased expression level of two MSY genes (*UTY1* and *PRKY*, further detail in Chapter 6) in carriers of haplogroup I, implies that one or more of these transcript(s) may possibly contribute to modulation of the immune system with a potential to increase the risk of CAD.

8.1.3 UTY1

UTY1 is a well-known X degenerate gene, primarily thought to be involved in self-tolerance through the production of a male-specific minor histocompatability complex (Simpson *et al.* 1997; Vogt *et al.* 2000). Data shown here provide further support for the hypothesis that a decrease in expression level of *UTY1* may be associated with general immune dysfunction in carriers of haplogroup I (as highlighted in pathway analysis Chapter 5). This could potentially lead to an increased risk of CAD through as yet unknown mechanisms.

8.1.3.1 UTY1 and the immune system

As previously discussed (Chapter 6), *UTY1* is a well-known male-specific minor histocompatibility complex protein thought to be involved in pathways such as graft rejection. Minor histocompatibility complexes are a group of cell membrane allo-antigens that trigger T lymphocyte proliferative and cytolytic activity (Perreault *et al.* 1990). Their primary role is to bind both exogenous and endogenous peptides for presentation to T cell receptors (Perreault *et al.* 1990). Unlike MHC antigens, which are recognised by both B and T lymphocytes, minor histocompatibility complexes are recognised only by T lymphocytes (Perreault *et al.* 1990). Research into the role of minor histocompatibility complexes in disease is currently restricted to graft vs. host disease (Sun *et al.* 2007) and miscarriage disorders (Christiansen *et al.* 2011). However, elements of this pathway have also been associated with CAD. A number of small studies have associated variants in MHC (which are heavily involved in both innate and adaptive immunity) with CAD (Swanberg *et al.* 2005) and a large multi-centre meta-analysis identified that one SNP (rs3869109) in a locus containing *HLA-C, HLA-B* and *HCG27* was associated with CAD (Davies *et al.* 2012). The MHC system and minor histocompatibility complexes are inextricably linked, implying that variation in minor histocombatibility complexes, such as reduced gene expression of *UTY1*, may predispose to CAD through pathways akin to those containing MHC SNPs associated with CAD. Taken together, it is tempting to speculate that the reduced gene expression level of *UTY1* in carriers of haplogroup I may translate into changes in the immune system that ultimately may predispose to CAD through the disruption of self-tolerance. This hypothesis is further supported by data showing that single X chromosome women (those with Turner syndrome) have higher rates of CAD than individuals with normal karyotype counterparts. Given that *UTX* escapes X inactivation, a decreased gene

expression level of this gene, and potentially its Y chromosome counterpart as seen in carriers of haplogroup I, may be the root cause of increased levels of CAD in those with Turner syndrome (Greenfield *et al.* 1998; Kozlowska-Wojciechowska *et al.* 2006). However, this hypothesis requires significant further research.

8.1.3.2 The origin of the haplogroup I specific variant of UTY1

While it is most likely that the disease-causing variant of *UTY1* occurred on the Y chromosome following haplogroup I divergence from the remaining Y chromosome phylogenetic tree, it is possible that the underlying sequence variation was transferred from the X chromosome during aberrant translocation during meiosis. It is widely accepted that the MSY does not recombine with the X chromosome during meiosis; however, there are reports disputing this. It has been shown that recombination can occur between the X and Y chromosome outside of the PAR regions, at recombination "hot-spots" such as the *PRKY*/*PRKX* and *VCY* loci (Trombetta *et al.* 2010; Rosser *et al.* 2009). The transfer of genetic information resulting from aberrant recombination during meiosis implies that a mutation generated on the X chromosome may be transferred to the Y chromosome (one haplogroup). Due to the lack of frequent recombination, a mutation on the Y chromosome may remain while its X chromosome counterpart is repaired. It is possible that a mutation originating from the X chromosome may underlie the differential expression of *UTY1* in carriers of haplogroup I, and ultimately the predisposition to CAD of its carriers. However, without further knowledge of the location of the mutation, the origin of such a mutation remains speculative.

8.1.4 PRKY

Protein kinases are a large group of regulatory enzymes known to alter substrate properties through the phosphorylation of serine, threonine or tyrosine residues (Hanks & Hunter 1995). Specifically, *PRKY* and its X chromosome homolog, *PRKX*, are cyclic AMP protein kinases that mediate physiological responses elicited by hormones, regulate gene expression and activate transcription factors including NFκB (Zimmermann *et al.* 1999; Pearce *et al.* 2010). Fully functioning protein kinases contain up to three regulatory domains (the activation segment, turn motif and hydrophobic motif) (Yang *et al.* 2002; Komander *et al.* 2005). However, no evidence of the turn motif has been identified in either *PRKY* or *PRKX*, nor does *PRKY* contain a hydrophobic motif (Pearce *et al.* 2010). Specifically, *PRKY* is a transcribed pseudogene with no documented function (as discussed in section 6.4.1.2.2) while *PRKX* is involved in cellular differentiation and epithelial morphogenesis.

Due to the broad action of protein kinases their involvement in pathologies is wide. To date, proteins contained in the cyclic AMP pathway have been associated with a number of diseases ranging from cancer (Carpten *et al.* 2007) to diabetes (Sakamoto *et al.* 2003). However, *PRKY* and its X chromosome homolog have only been implicated in infertility and sex reversal resulting from a genomic translocation during recombination (Schiebel *et al.* 1997; Jobling *et al.* 1998; Rosser *et al.* 2009). While a number of protein kinases have been associated with atherosclerosis (Liu *et al.* 2012) and heart failure (Chaanine & Hajjar 2011; van Heerebeek *et al.* 2012), no such association has been identified between *PRKY* or *PRKX* and CAD.

The lack of documented function of *PRKY* strongly implies that the identified association is biologically non-functional. However, it is possible that a haplogroup I-

specific mutation or genomic variant of *PRKY* does predispose carriers to CAD. As with *UTY1*, it is plausible that this mutation may have arisen on the X chromosome and transferred to *PRKY* during an aberrant crossing-over event, or a *de novo* mutation may have occurred on the Y chromosome. Either possibility may result in a haplogroup-specific inheritance pattern of the mutation. Without further detail of the location of the mutation, this remains speculative. The potential effect of a mutation is difficult to predict due to the lack of documented function of *PRKY*. However, the most plausible action would affect the likelihood of transcript degradation through nonsense-mediated decay. The mutation may also have an effect on gene expression through altered DNA methylation patterns. The relatively close proximity of *UTY1* and *PRKY* imply the presence of a communal promoter which may harbour a variant specific to haplogroup I.

8.2 The role of the Y chromosome in AAA

Following the successful identification of haplogroup I of the Y chromosome as a significant risk factor of CAD, it was hypothesised that this haplogroup may act as a risk factor for all CVD (details of this work can be found in Chapter 7). Given the extreme sexual dimorphism and major immunological pathogenesis, AAA was chosen to investigate the role of haplogroup I in other CVD pathologies. Haplogroup I of the Y chromosome was not associated with the predisposition to AAA. Furthermore, the investigation of other common Y chromosome haplogroups showed that the Y chromosome as a whole does not have an effect on the male predisposition to AAA in European ancestry populations. These data add to the growing body of work stating that the genetic background of AAA and CAD are significantly different (further detail in Chapter 7).

8.3 The role of the Y chromosome in disease

Variation in the Y chromosome has now been associated not only with CAD but also HIV infection, autism, BP and lipids (Ellis *et al.* 2000; Charchar *et al.* 2002; Russo *et al.* 2006; Arredi *et al.* 2007; Shankar *et al.* 2007; Serajee & Mahbubul Huq 2009; Sezgin *et al.* 2009). Work shown here indicates for the first time that the Y chromosome may contribute to these diseases /phenotypes through its male-specific effect on the immune system (of particular relevance in CAD and HIV infection). This finding has important consequences for research into the pathogenesis of other male predominated, immune-based diseases. To date, autoimmune diseases are widely accepted as being female dominated (Amur *et al.* 2012). However, these data suggestive of the effect of the Y chromosome on immunity including possibly self-tolerance imply that male-dominated immune diseases, such as psoriasis (McCoombe *et al.* 2009), may be associated with MSY. Furthermore, the association between the Y chromosome and the immune system exemplifies how little is known of the functional potential of this small portion of the human genome. It can be hypothesised that the Y chromosome has many other, as yet unidentified, physiological effects which may alter disease prevalence and response to treatment. While the sex chromosomes pose unique problems during genetic research, our and other research have highlighted their importance in complex disease and further research is required in order to fill this large knowledge gap.

8.4 Future work

8.4.1 Identification of the functional variant

The research discussed here represents the initial steps into elucidating how the Y chromosome affects the immune system and predisposition to CAD.

A significant advance in understanding would be made should the functional variant that underlies the association between haplogroup I and decreased expression of *UTY1* and/or *PRKY* be identified. The causal variant may reside within the sequence of one of the two genes. Given the apparent co-regulation of both *UTY1* and *PRKY*, it is possible that the functional variant may lie within a communal promoter of these genes. However, due to the genomic structure of MSY as one LD block, the functional variant could potentially reside anywhere in the MSY. In an effort to further restrict the location of the potential variant, an investigation into the expression level of each gene transcript should be completed. Potentially, one specific transcript would show differential regulation in carriers of haplogroup I, implying the variant resides in the variable sequence of that transcript. Re-sequencing of the gene transcript could then be used in an effort to identify single base mutations or small insertion/deletions.

While it is possible that only one transcript of the gene contains a functional variant, it is also possible that a higher resolution lineage of haplogroup I (for example Ia1) contains the functional variant. Though an investigation into the role of higher resolution haplogroup I lineages was not possible here, the development of a larger replication cohort may facilitate this investigation. An association between a higher resolution lineage of haplogroup I and expression of *UTY1* and *PRKY* would allow further bioinformatic investigations into differences between the haplogroup I lineages and may highlight potential functional variants *in silico*. Once the functional variant is identified, either through gene transcript analysis or through sub-haplogroup analysis,

functional experiments may be conducted to elucidate how this mutation impacts upon the immune system.

8.5 General limitations

8.5.1 Y chromosome haplogroup analysis

The premises of this research are based on the assumption that the human MSY does not recombine with the X chromosome, and as such is inherited intact between father and son. While this assumption generally holds true, there is evidence for the recombination that can occur between the X and Y chromosome outside of the PAR regions, particularly at the *PRKY* (Rosser *et al.* 2009) and *VCY* loci (Trombetta *et al.* 2010). This raises the fundamental question of whether all men in a single haplogroup are indeed identical at the Y chromosome sequence level. However, until further work proves that recombination within the MSY frequently occurs, haplogroup analysis remains the best method of Y chromosome investigation (Arredi *et al.* 2007; Serajee & Mahbubul Huq 2009; Sezgin *et al.* 2009).

8.5.2 Ancestry of populations used

Whilst the cohorts used in this study were of European ancestry, there were some differences in the specific country of origin. Specifically, YMCA cohorts as well as all renal tissues (Chapter 6) were of Polish origin while one AAA cohort was recruited in Ireland (Chapter 7). Theoretically, different population origins may contribute to an increased risk of population stratification, which could lead to false positive associations (Thomas & Witte 2002; Cardon & Palmer 2003). A well-known example of population stratification leading to a spurious association is an analysis of a HLA

haplotype and non-insulin-dependent diabetes mellitus in a cohort of European and Pima Indian individuals (Williams *et al.* 1996; Thomas *et al.* 2002; Cardon *et al.* 2003). It was found that a HLA haplotype increased risk of non-insulin-dependent diabetes mellitus, however, the association disappeared when Pima Indians were investigated in isolation. It was later found that both non-insulin-dependent diabetes mellitus and the "risk" haplotype were much more common in Pima Indian than European individuals, indicating significant population stratification within the study group (Cardon *et al.* 2003). While prevalence of Y chromosome haplogroups is known to vary widely across geographical populations, the prevalence of haplogroup I in the populations studied here were consistent across populations (ranging from 10-20% prevalence in the general population). Given the consistent prevalence of haplogroup I and the common ancestral background from which all haplogroup I carriers would have originated, the difficulties in analysis of the different ancestral populations is limited.

8.5.3 Hidden population stratification

With the exception of the Polish kidney cohorts and Belfast AAA cohort, all study cohorts underwent some form of analysis of population stratification. Two widely validated methods of population stratification were used; the benefits, drawbacks and a comparison of each method can be found in Chapter 3. No level of population stratification was identified in any of the cohorts investigated. However, the presence of differences between carriers of haplogroup I and all other haplogroups was not investigated in Leicester AAA cohort controls, Leeds AAA cohort, Belfast AAA cohort, YMCA 1 or YMCA 2 due to financial constraints. While it is not impossible that there are genetic differences between carriers of haplogroup I and all other haplogroups in

these cohorts, no evidence of population stratification have been identified in cohorts where a thorough population stratification analysis has been completed i.e. in BHF-FHS and CTS (Charchar *et al.* 2012).

While population stratification within individual cohorts was assessed where possible, there are no means of assessing the presence of population substructure between different cohorts. However, there was no population heterogeneity in several association analyses.

8.6 Other methods of Y chromosome analysis

While a significant amount of work has gone into developing and perfecting the Y chromosome phylogeneic tree, there are major limitations in using it in the search for positive genetic associations with both physiological traits and diseases (including potential recombination within the MSY and use of smaller lineages). Though limitations of haplogroup analysis do exist, there are few alternative methods of analysis. One such alternative involves the use of STRs in the generation of haplotype analysis. However, while this method can be used successfully in conjunction with haplogroup analysis, when used in isolation it generally lacks statistical power, since haplotypes based on STRs may be person- or family-specific. Similarly, the use of the Y chromosome *Hin*dIII restriction site has yielded important advancements in the knowledge of Y chromosome biology; however, this method of analysis is only a very crude assessment of variation within the MSY.

The use of mouse strains such as the *S***ry-** mice (i.e. a Y chromosome where *Sry* has been deleted) can potentially provide interesting data resulting from gene deletion. However, identification of risk variants residing on the mouse Y chromosome has the same difficulties as their identification in human tissue. Animal models such as these

are invaluable in downstream research, such as in identifying a potential role on *UTY1* in the immune system, but have provided little benefit during the initial identification of a risk genomic signature.

The development of next-generation sequencing may provide further insight into the Y chromosome. However, Wei *et al.* identified no potentially relevant sequence variations in haplogroup I (Wei *et al.* 2012). Problematically, next-generation sequencing highlights many hundred of SNPs within the MSY making haplogroup resolution difficult. Potentially, STR analysis may be the way forward in identifiying Y chromosome mutations with pathological relevance.

9 Conclusions

Previous research has shown that haplogroup I of the Y chromosome significantly increases risk of CAD in its carriers. Data shown here confirm that haplogroup I of the Y chromosome is not associated with several major sexually dimorphic CAD risk factors or traits typically described as “male-associated” or “sex-specific”. As such, the modulation of these phenotypes is extremely unlikely to underlie the association between haplogroup I and CAD.

In an effort to identify novel pathways that potentially link the Y chromosome with CAD, a hypothesis-free experiment was conducted based on transcriptome-wide analysis of human monocytes and macrophages of those with haplogroup I and all other haplogroups. Inflammatory and auto-immune pathways in macrophages show significant differences in expression level between carriers of haplogroup I and other Y chromosome haplogroups. Both inflammation and auto-immunity have been previously implicated in CAD. Interestingly, the identification of this association in macrophages and not monocytes implies that the Y chromosome may exert its effect on the predisposition of CAD at the stage of cell differentiation from monocyte to macrophage. Furthermore, both *UTY1* and *PRKY* were significantly down-regulated in macrophages from carriers of haplogroup I when compared to men with other haplogroups. They may represent potential drivers of the Y chromosome effect on the immune system and cardiovascular risk.

Finally, it was found that the CAD risk haplogroup, haplogroup I, does not increase the risk of AAA in its carriers, highlighting the differences in genetic background and pathogenesis between CAD and AAA.

This study reveals a completely new role of the Y chromosome in human biology. Research shows that the Y chromosome has functions different from sex determination and spermatogenesis, including a potential regulatory effect on the immune system and cardiovascular risk.

10 References

Ailawadi G, Eliason JL, Roelofs KJ, Sinha I, Hannawa KK, Kaldjian EP, Lu G, Henke PK, Stanley JC, Weiss SJ, Thompson RW, Upchurch GR Jr. Gender differences in experimental aortic aneurysm formation. *Arterioscler Thromb Vasc Biol* 2004; **24:**1216-1222

Aksglaede L, Skakkebaek NE, Juul A. Abnormal sex chromosome constitution and longitudinal growth: serum levels of insulin-like growth factor (IGF)-I, IGF binding protein-3, luteinizing hormone, and testosterone in 109 males with 47,XXY, 47,XYY, or sex-determining region of the Y chromosome (SRY)-positive 46,XX karyotypes. *J Clin Endocrinol Metab* 2008;**93:**169-173

Aksut SV, Aksut G, Karamehmetoglu A, Oram E.The determination of serum oestradiol, testosterone and progesterone in acute myocardial infarction. *Jpn Heart J* 1986;**27:**825–837

Aligeti VR & Horn HR. Turner's syndrome and coronary artery disease. *Am J Cardiol* 2007;**99:**741-742

Allocco DJ, Kohane IS, Butte AJ. Quantifying the relationship between co-expression, co-regulation and gene function. *BMC Bioinformatics* 2004;**5:**18

Ambrose JA, Gould RB, Kurian DC, DeVoe MC, Pearlstein NB, Coppola JT, Siegal FP. Frequency of and outcome of acute coronary syndromes in patients with human immunodeficiency virus infection. *Am J Cardiol* 2003;**92:**301-303

Amur S, Parekh A, Mummaneni P. Sex differences and genomics in autoimmune diseases. *J Autoimmune* 2012;**38:**254-265

Angelopoulou R, Lavranos G, Manolakou P. Establishing sexual dimorphism in humans. *Coll Antropol* 2006;**30:**653-658

Appleton SL, Ruffin RE, Wilson DH, Taylor AW & Adams RJ. Cardiovascular disease risk associated with asthma and respiratory morbidity might be meditated by short-acting β2-agonists. *J Allergy Clin Immunol* 2009;**123:**124-130

Arnett D. Relevance of genetics and genomics for prevention and treatment of cardiovascular disease: a scientific statement from the American Heart Association Council on epidemiology and prevention, the stroke council, and the functional genomics and translational biology interdisciplinary working group. *Circulation* 2007;**115:**2878-2901

Arredi B, Ferlin A, Speltra E, Bedin C, Zuccarello D, Ganz F, Marchina E, Stuppia L, Krausz C, Foresta C. Y-chromosome haplogroups and susceptibility to azoospermia factor c microdeletion in an Italian population. *J Med Genet* 2007;**44:**205-208

Arroyo-Espliguero R, Avanzas P, Jeffery S, Kaski JC. CD14 and toll-like receptor 4: a link between infection and acute coronary events? *Heart* 2004;**90:**983-988

Assmann G, Cullen P, Schulte H. Simple scoring scheme for calculating the risk of acute coronary events based on the 10-year follow-up of the prospective cardiovascular Münster (PROCAM) study. *Circulation* 2002;**105:**310-315

Augenlicht LH & Kobrin D. Cloning and screening of sequences expressed in a mouse colon tumor. *Cancer Res* 1982;**42:**1088-1093

Augenlicht LH, Wahrman MZ, Halsey H, Anderson L, Taylor J, Lipkin M. Expression of cloned sequences in biopsies of human colonic tissue and in colonic carcinoma cells induced to differentiate in vitro. *Cancer Res* 1987;**47:**6017-6021

Augenlicht LH, Taylor J, Anderson L, Lipkin M. Patterns of gene expression that characterize the colonic mucosa in patients at genetic risk for colonic cancer. *Proc Natl Acad Sci USA* 1991;**15:**3286-3289

Aulchenko YS, Ripatti S, Lindqvist I, Boomsma D, Heid IM, Pramstaller PP, Penninx BW, Janssens AC, Wilson JF, Spector T, Martin NG, Pedersen NL, Kyvik KO, Kaprio J, Hofman A, Freimer NB, Jarvelin MR, Gyllensten U, Campbell H, Rudan I, Johansson A, Marroni F, Hayward C, Vitart V, Jonasson I, Pattaro C, Wright A, Hastie N, Pichler I, Hicks AA, Falchi M, Willemsen G, Hottenga JJ, de Geus EJ, Montgomery GW, Whitfield J, Magnusson P, Saharinen J, Perola M, Silander K, Isaacs A, Sijbrands EJ, Uitterlinden AG, Witteman JC, Oostra BA, Elliott P, Ruokonen A, Sabatti C, Gieger C, Meitinger T, Kronenberg F, Döring A, Wichmann HE, Smit JH, McCarthy MI, van Duijn CM, Peltonen L; ENGAGE Consortium. Loci influencing lipid levels and coronary heart disease risk in 16 European population cohorts. *Nat Genet* 2009;**41**:47-55.

Austin H, Hill HA, Flanders WD, Greenberg RS. Limitations in the application of case-control methodology. *Epidemiol Rev* 1994;**16:**65-76

Balaresque P, Bowden GR, Adams SM, Leung HY, King TE, Rosser ZH, Goodwin J, Moisan JP, Richard C, Millward A, Demaine AG, Barbujani G, Previderè C, Wilson IJ, Tyler-Smith C, Jobling MA. A predominantly neolithic origin for European paternal lineages. *PLoS Biol* 2010;**8:**1000285

Barba A, Estallo L, Rodríguez L, Baquer M, Vega de Céniga M. Detection of abdominal aortic aneurysm in patients with peripheral artery disease. *Eur J Vasc Endovasc Surg* 2005;**30:**504-508

Barbouche R, Miquelis R, Jones IM, Fenouillet E. Protein-disulfide isomerase-mediated reduction of two disulfide bonds of HIV envelope glycoprotein 120 occurs post-CXCR4 binding and is required for fusion. *J Biol Chem* 2003;**278:**3131-3136

Barth JD, Jansen H, Hugenholtz PG, Birkenhager JC. Post heparin lipases, lipids and related hormones in men undergoing coronary arteriography to assess atherosclerosis. *Atherosclerosis* 1983;**48:**235–241

Barrett-Connor E. Sex Differences in Coronary Heart Disease. Why Are Women So Superior? The 1997 Ancel Keys Lecture. *Circulation.*1997;**95:**252-264

Barrett-Connor E & Khaw KT. Endogenous sex hormones and cardiovascular disease in men. A prospective population based study. *Circulation* 1988;**78:**539–545

Bass KM, Newschaffer CJ, Klag MJ, Bush TL. Plasma lipoprotein levels as predictors of cardiovascular death in women. *Arch Intern Med* 1993;**153:**2209–2216

Baumann G, Reza P, Chatterton R, Green D, Krumlovsky F. Plasma estrogens, androgens, and von Willebrand factor in men on chronic hemodialysis. *Int J Artif Organs* 1988;**11:**449–453

Bell P, Mantor C, Jacocks MA. Congenital abdominal aortic aneurysm: a case report. *J Vasc Surg* 2003;**38:**190-193

Bem SL. The measurement of psychological androgyny. *J Consult Clin Psychol* 1974;42:15-162

Bengtsson H, Bergqvist D, Ekberg O, Janzon L. A population based screening of abdominal aortic aneurysms (AAA). *Eur J Vasc Surg* 1991;**5:**53-57

Bengtsson H, Bergqvist D, Sternby NH. Increasing prevalence of abdominal aortic aneurysms.A necropsy study. *Eur J Surg* 1992;**158:**19-23

Bergqvist D, Mani K, Wanhainen A. Epidemiological aspects on abdominal aortic aneurysms (AAA). *PAAJ* 2010;**1**

Best VA, Price JF, Fowkes FG. Persistent increase in the incidence of abdominal aortic aneurysm in Scotland, 1981–2000. *Br J Surg* 2003;**90:**1510-1515

Bhagavatula MR, Fan C, Shen GQ, Cassano J, Plow EF, Topol EJ, Wang Q. Transcription factor MEF2A mutations in patients with coronary artery disease. *Hum Mol Genet* 2004;**13:**3181-3188

Bhupathy P, Haines CD, Leinwand LA. Influence of sex hormones and phytoestrogens on heart disease in men and women. *Womens Health* 2010;**6:**77-95

Bianchi NO. Y chromosome structural and functional changes in human malignant diseases. *Mutat Res* 2009;**682:**21-27

Bilginer Y, Ozaltin F, Basaran C, Duzova A, Besbas N, Topaloglu R, Ozen S &Bakkalogu A. Evaluation of intima media thickness of the common and internal carotid arteries with inflammatory markers in familial Mediterranean fever as possible predictors of atherosclerosis. *Rheumatol Int* 2008;**28:**1211-1216

Biros E, Norman PE, Walker PJ, Nataatmadja M, West M, Golledge J. A single nucleotide polymorphism in exon 3 of the kallikrein 1 gene is associated with large but not small abdominal aortic aneurysm. *Atherosclerosis* 2011;**217:**452-457

Blanchard JF, Armenian HK, Friesen PP. Risk factors for abdominal aortic aneurysm: results of a case-control study. *Am J Epidemiol* 2000;**151:**575-583

Blankenberg S, Ruppercht HJ, Bickel C, Peetz D, Hafner G, Tiret L & Meyer J. Circulating cell adhesion molecules and death in patients with coronary artery disease. *Circulation* 2001;**104:**1336-1342

Bloomer LDS, Bown MJ, Tomaszewski M. Sexual dimorphism of abdominal aortic aneurysms: A striking example of "male disadvantage" in cardiovascular disease. *Atherosclerosis* 2012;[Epub ahead of print]

Boddy AM, Lenk GM, Lillvis JH, Nischan J, Kyo Y, Kuivaniemi H. Basic research studies to understand aneurysm disease. *Drug News Perspect* 2008;**21**:142.

Bogaert V, Taes Y, Konings P, Van Steen K, De Bacquer D, Goemaere S, Zmierczak H, Crabbe P, Kaufman JM. Heritability of blood concentrations of sex-steroids in relation to body composition in young adult male siblings. *Clin Endocrinol* 2008;**69:**129-35

Boyle SH, Michalek JE, Suarez EC.Covariation of psychological attributes and incident coronary heart disease in U.S. Air Force veterans of the Vietnam war. *Psychosom Med* 2006;**68:**844-50

Bradley DT, Badger SA, Bown MJ, Sayers RD, Hughes AE. Coding polymorphisms in the genes of the alternative complement pathway and abdominal aortic aneurysm. *Int J Immunogenet* 2011;**38:**243-248

Breier C, Muhlberger V, Drexel H, Herold M, Lisch HJ, Knapp E, Braunsteiner H. Essential role of post-heparin lipoprotein lipase activity and of plasma testosterone in coronary artery disease. *Lancet* 1985;**1:**1242–1244

Bossé Y, Lemire M, Poon AH, Daley D, He JQ, Sandford A, White JH, James AL, Musk AW, Palmer LJ, Raby BA, Weiss ST, Kozyrskyj AL, Becker A, Hudson TJ, Laprise C. Asthma and genes encoding components of the vitamin D pathway. *Respir Res* 2009;**10:**98

Bown MJ, Lloyd GM, Sandford RM, Thompson JR, London NJ, Samani NJ, Sayers RD. The interleukin-10-1082 'A' allele and abdominal aortic aneurysms. *J Vasc Surg* 2007;**46:**687-693

Bown MJ, Braund PS, Thompson J, London NJ, Samani NJ, Sayers RD. Association between the coronary artery disease risk locus on chromosome 9p21.3 and abdominal aortic aneurysm. *Circ Cardiovasc Genet* 2008;**1:**39-42

Bown MJ, Jones GT, Harrison SC, Wright BJ, Bumpstead S, Baas AF, Gretarsdottir S, Badger SA, Bradley DT, Burnand K, Child AH, Clough RE, Cockerill G, Hafez H, Scott DJ, Futers S, Johnson A, Sohrabi S, Smith A, Thompson MM, van Bockxmeer FM, Waltham M, Matthiasson SE, Thorleifsson G, Thorsteinsdottir U, Blankensteijn JD, Teijink JA, Wijmenga C, de Graaf J, Kiemeney LA, Assimes TL, McPherson R; CARDIoGRAM Consortium; Global BPgen Consortium; DIAGRAM Consortium; VRCNZ Consortium, Folkersen L, Franco-Cereceda A, Palmen J, Smith AJ, Sylvius N,

Wild JB, Refstrup M, Edkins S, Gwilliam R, Hunt SE, Potter S, Lindholt JS, Frikke-Schmidt R, Tybjærg-Hansen A, Hughes AE, Golledge J, Norman PE, van Rij A, Powell JT, Eriksson P, Stefansson K, Thompson JR, Humphries SE, Sayers RD, Deloukas P, Samani NJ. Abdominal aortic aneurysm is associated with a variant in low-density lipoprotein receptor-related protein 1. *Am J Hum Genet* 2011;**89:**619-627

Braunersreuther V & Mach F. Leukocyte recruitment in atherosclerosis: potential targets for therapeutic approaches? *Cell Mol Life Sci* 2006;**63:**2079–88

Brion M, Sobrino B, Blanco-Verea A, Lareu MV, Carracedo A. Hierarchical analysis of 30 Y-chromosome SNPs in European populations. *Int J Legal Med* 2004;**119:**10-15

Broeckel U, Hengstenberg C, Mayer B, Holmer S, Martin LJ, Comuzzie AG, Blangero J, Nurnberg P, Reis A, Riegger GA, Jacob HJ, Schunkert H. A comprehensive linkage analysis for myocardial infarction and its related risk factors. *Nat Genet* 2002;**30:**210–214

Burt VL, Whelton P, Roccella EJ, Brown C, Cutler JA, Higgins M, Horan MJ, Labarthe D. Prevalence of hypertension in the US adult population. Results from the third national health and nutrition examination survey, 1988-1991. *Hypertension* 1995;**25:**305-313

Busby GB, Brisighelli F, Sánchez-Diz P, Ramos-Luis E, Martinez-Cadenas C, Thomas MG, Bradley DG, Gusmão L, Winney B, Bodmer W, Vennemann M, Coia V, Scarnicci F, Tofanelli S, Vona G, Ploski R, Vecchiotti C, Zemunik T, Rudan I, Karachanak S, Toncheva D, Anagnostou P, Ferri G, Rapone C, Hervig T, Moen T, Wilson JF, Capelli C. The peopling of Europe and the cautionary tale of Y chromosome lineage R-M269. *Proc Biol Sci* 2012;**279**:884-92.

Buss AH & Durkee A. An inventory for assessing different kinds of hostility. *J Consult Psychol* 1957;**21:**343-349

Buss AH & Perry M. The aggression questionnaire. *J Persc Soc Med* 1992;**63:**452-459

Büsst CJ, Bloomer LD, Scurrah KJ, Ellis JA, Barnes TA, Charchar FJ, Braund P, Hopkins PN, Samani NJ, Hunt SC, Tomaszewski M, Harrap SB. The epithelial sodium channel γ-subunit gene and blood pressure: family based association, renal gene expression, and physiological analyses. *Hypertension* 2011;**58:**1073-1078

Capelli C, Redhead N, Abernethy JK, Gratrix F, Wilson JF, Moen T, Hervig T, Richards M, Stumpf MP, Underhill PA, Bradshaw P, Shaha A, Thomas MG, Bradman N, Goldstein DB. A Y chromosome census of the British Isles. *Curr Biol* 2003;**13:**979-984

Cardon LR&Palmer LJ.Population stratification and spurious allelic association. *Lancet* 2003;**361:**598-604

Carpeggiani C, Coceani M, Landi P, Michelassi C, L'abbate A. ABO blood group alleles: A risk factor for coronary artery disease. An angiographic study. *Atherosclerosis* 2010;**211:**461-466

Carpten JD, Faber AL, Horn C, Donoho GP, Briggs SL, Robbins CM, Hostetter G, Boguslawski S, Moses TY, Savage S, Uhlik M, Lin A, Du J, Qian YW, Zeckner DJ, Tucker-Kellogg G, Touchman J, Patel K, Mousses S, Bittner M, Schevitz R, Lai MH, Blanchard KL, Thomas JE. A transforming mutation in the pleckstrin homology domain of AKT1 in cancer. *Nature* 2007;**448:**439-344

Case LK, Toussaint L, Moussawi M, Roberts B, Saligrama N, Brossay L, Huber SA, Teuscher C. Chromosome Y regulates survival following murine coxsackievirus B3 infection. *G3* (*Bethesda*) 2012;**2:**115–21

Castleden WM, Mercer JC. Abdominal aortic aneurysms in Western Australia: descriptive epidemiology and patterns of rupture. *Br J Surg* 1985;**72:**109-112

Cauley JA, Gutai JP, Kuller LH, Dai WS. Usefulness of sex steroid hormone levels in predicting coronary artery disease in men. *Am J Cardiol* 1987;**60:**771–777

Cavalca V, Veglia F, Squellerio I, Marenzi G, Minardi F, De Metrio M, Cighetti G, Boccotti L, Ravagnani P, Tremoli E. Glutathione, vitamin E and oxidative stress in coronary artery disease: relevance of age and gender. *Eur J Clin Invest* 2009;**39:**267-272

Cengiz K, Alvur M, Dindar U. Serum creatinine phosphokinase, lactic dehydrogenase, oestradiol, progesterone, and testosterone levels in male patients with acute myocardial infarction and unstable angina pectoris. *Mater Med Pol* 1991;**23**:195–198

Chaanine AH, Hajjar RJ. AKT signalling in the failing heart. *Eur J Heart Fail* 2011;**13:**825-829

Chalmel F, Lardenois A, Evrard B, Mathieu R, Feig C, Demougin P, Gattiker A, Schulze W, Jégou B, Kirchhoff C, Primig M. Global human tissue profiling and protein network analysis reveals distinct levels of transcriptional germline-specificity and identifies target genes for male infertility. *Hum Reprod* 2012;[Epub ahead of print]

Charchar FJ, Tomaszewski M, Padmanabhan S, Lacka B, Upton MN, Inglis GC, Anderson NH, McConnachie A, Zukowska-Szczechowska E, Grzeszczak W, Connell JM, Watt GC, Dominiczak AF. The Y chromosome effect on blood pressure in two European populations. *Hypertension* 2002;**39:**353-356

Charchar FJ, Tomaszewski M, Lacka B, Zakrzewski J, Zukowska-Szczechowska E, Grzeszczak W, Dominiczak AF. Association of the Human Y chromosome with cholesterol levels in the general population. *Hypertension* 2003;**24:**308-312

Charchar FJ, Tomaszewski M, Lacka B, Zakrzewski J, Zukowska-Szczechowska E, Grzeszczak W, Dominiczak AF. Association of the human Y chromosome with cholesterol levels in the general population. *Arterioscler Thromb Vasc Biol* 2004;**24:**308-312

Charchar F, Zimmerli L, Tomaszewski M. The pressure of finding human hypertension genes: new tools, old dilemmas. *J Hum Hypertens* 2008;**22:**821-828

Charchar FJ, Bloomer LDS, Barnes TA, Cowley MJ, Nelson CP, Wang Y, Denniff M, Debiec R, Christofidou P, Nankervis S, Dominiczak AF, Bani-Mustafa A, Balmforth AJ, Hall AS, Erdmann J, Cambien F, Deloukas P, Hengstenberg C, Packard C, Schunkert H, Ouwehand WH, Ford I, Goodall AH, Jobling MA, Samani NJ, Tomaszewski M. Inheritance of coronary artery disease in men: an analysis of the role of the Y chromosome. *Lancet* 2012;**379:**915-922

Cheanvechai V, Harthun NL, Graham LM, Freischlag JA, Gahtan V. Incidence of peripheral vascular disease in women: Is it different from that in men? *J Thorac Cardiovasc Surg* 2004;**127:**314-317

Cheng SW, Ting AC, Tsang SH. Epidemiology and outcome of aortic aneurysms in Hong Kong. *World J Surg* 2003;**27:**241-245

Choudhary SK, Bhan A, Talwar S, Goyal M, Sharma S, Venugopal P. Tubercular pseudoaneurysms of aorta. *Ann Thorac Surg* 2001;**72:**1239-1244

Christiansen OB, Steffensen R, Nielsen HS. Anti-HY responses in pregnancy disorders. *Am J Reprod Immunol* 2011;**66:**93-100

Chute CG, Baron JA, Plymate SR, Kiel DP, Pavia AT, Lozner EC, O'Keefe T, MacDonald GJ. Sex hormones and coronary artery disease. *Am J Med* 1987;**83:**853–859

Clifton MA. Familial abdominal aortic aneurysms. *Br J Surg* 1977;**64:**765-766

Cline MS, Smoot M, Cerami E, Kuchinsky A, Landys N, Workman C, Christmas R, Avila-Campilo I, Creech M, Gross B, Hanspers K, Isserlin R, Kelley R, Killcoyne S,

Lotia S, Maere S, Morris J, Ono K, Pavlovic V, Pico AR, Vailaya A, Wang PL, Adler A, Conklin BR, Hood L, Kuiper M, Sander C, Schmulevich I, Schwikowski B, Warner GJ, Ideker T, Bader GD. Integration of biological networks and gene expression data using Cytoscape. *Nat Protoc* 2007;**2:**2366-2382

Cokkinos DV, Krajcer Z, Leachman RD. Coronary artery disease in hypertrophic cardiomyopathy. *Am J Cardiol* 1985;**55:**1437-1438

Contu D, Morelli L, Zavattari P, Lampis R, Angius E, Frongia P, Murru D, Maioli M, Francalacci P, Todd JA, Cucca F. Sex-related bias and exclusion mapping of the nonrecombinant portion of chromosome Y in human type 1 diabetes in the isolated founder population of Sardinia. *Diabetes* 2002;**51:**3573-3576

Cornuz J, Sidoti Pinto C, Tevaearai H, Egger M. Risk factors for asymptomatic abdominal aortic aneurysm. *Eur J Public Health* 2004;**14:**343-349

Coronary Artery Disease (C4D) Genetics Consortium. A genome-wide association study in Europeans and South Asians identifies five new loci for coronary artery disease. *Nat Genet* 2011;**43**:339-44.

Cox D, Brennan M, Moran N. Integrins as therapeutic targets: lessons and opportunities. *Nat Rev Drug Discov* 2010;**10:**804-820

Cronbach LJ. Coefficient alpha and the internal structure of tests. *Psychometrika* 1951;**16:**297-334

Curtis RK, Oresic M, Vidal-Puig A. Pathways to the analysis of microarray data. *Trends Biotechnol* 2005;**23:**429-435

Curtiss LK. Reversing atherosclerosis. *New Engl J Med* 2009;**360:**1144-1146

Daugherty A, Cassis LA. Mouse models of abdominal aortic aneurysms. *Arterioscler Thromb Vasc Biol* 2004;**24:**429-434

Dave S & Hagen J. Myocardial infarction during intravenous immunoglobulin infusion in a 65-year-old man with common variable immunodeficiency and subsequent successful repeated administration. *Ann Allergy Asthma Immunol* 2007;**99:**567-570

Davies RW, Wells GA, Stewart AF, Erdmann J, Shah SH, Ferguson JF, Hall AS, Anand SS, Burnett MS, Epstein SE, Dandona S, Chen L, Nahrstaedt J, Loley C, König IR, Kraus WE, Granger CB, Engert JC, Hengstenberg C, Wichmann HE, Schreiber S, Tang WH, Ellis SG, Rader DJ, Hazen SL, Reilly MP, Samani NJ, Schunkert H, Roberts R, McPherson R. A genome-wide association study for coronary artery disease identifies a novel susceptibility locus in the major histocompatibility complex. *Circ Cardiocasc Genet* 2012;**5:**217-225

Day DA & Tuite MF. Post-transcriptional gene regulatory mechanisms in eukaryotes: an overview. *J Endocrinol* 1998;**157:**361-371

De Boer OJ, van der Wal AC, Becker AE. Atherosclerosis, inflammation, and infection. *J Pathol* 2000;**190:**237-243

Deeks JJ, Altman DG, Bradburn MJ. Statistical methods for examining heterogeneity and combining results from several studies in meta-analysis. In Systematic Reviews in Health Care: Meta-Analysis in Context, 2nd Edition. 2001

Delbridge ML & Graves JA. Mammalian Y chromosome evolution and the male-specific functions of Y chromosome-borne genes. *Rev Reprod* 1999;**4:**101-109

Di Angelantonio E, Chowdhury R, Sarwar N, Aspelund T, Danesh J, Gudnason V. Chronic kidney disease and risk of major cardiovascular disease and non-vascular mortality: prospective population based cohort study. *BMJ* 2010;**341:**doi: 10.1136

Dornas WC, Oliveira TT, Augusto LE, Nagem TJ. Experimental atherosclerosis in rabbits. *Arg Bras Cardiol* 2010;**95:**272-278

Du P, Kibbe WA, Lin SM. lumi: a pipeline for processing Illumina microarray. *Bioinformatics* 2008;**24:**1547-1548

Dugger KJ, Zinn KR, Weaver C, Bullard DC, Barnum SR. Effector and suppressor roles for LFA-1 during the development of experimental autoimmune encephalomyelitis. *J Neuroimmunol* 2009;**206:**22-27

Edfeldt K, Bennet AM, Eriksson P, Frostegård J, Wiman B, Hamsten A, Hansson GK, de Faire U, Yan ZQ. Association of hypo-responsive toll-like receptor 4 variants with risk of myocardial infarction. *Eu Heart J* 2004;**25:**1447-1453

Ehrt S, Schnappinger D, Bekiranov S, Drenkow J, Shi S, Gingeras TR, Gaasterland T, Schoolnik G, Nathan C. Reprogramming of the macrophage transcriptome in response to interferon-gamma and Mycobacterium tuberculosis: signalling roles of nitric oxide synthase-2 and phagocyte oxidase. *J Exp Med* 2001;**194:**1123-1140

El-Maarri O, Walier M, Behne F, van Üüm J, Singer H, Diaz-Lacava A, Nüsgen N, Niemann B, Watzka M, Reinsberg J, van der Ven H, Wienker T, Stoffel-Wagner B, Schwaab R, Oldenburg J. Methylation at global LINE-1 repeats in human blood are affected by gender but not by age or natural hormone cycles. *PLoS One* 2011;**9:**e16252

Ellis JA, Stebbing M, Harrap SB. Association of the Human Y chromosome with high blood pressure in the general population. *Hypertension* 2000;**36:**731-733

Elmore JR, Obmann MA, Kuivaniemi H, Tromp G, Gerhard GS, Franklin DP, Boddy AM, Carey DJ. Identification of a genetic variant associated with abdominal aortic aneurysms on chromosome 3p12.3 by genome-wide association. *J Vasc Surg.* 2009;**49:**1525-1531

English KM, Mandour O, Steeds RP, Jones TH, Channer KS. Men with coronary disease have lower levels of androgens than those with coronary angiograms. *Eur Heart J* 2000;**21:**890–895

Eriksson EE. Leukocyte recruitment to atherosclerotic lesions, a complex web of dynamic cellular and molecular interactions. *Curr Drug Targets Cardiovasc Haematol Disord* 2003;**3:**309-325

Eriksson EE. Mechanisms of leukocyte recruitment to atherosclerotic lesions: future prospects. *Curr Opin Lipidol* 2004;**2:**553-558

Evans M, Fryzek JP, Elinder CG, Cohen SS, McLaughlin JK, Nyrén O, Fored CM. The natural history of chronic renal failure: results from an unselected, population - based, inception cohort in Sweden. *Am J Kidney Dis* 2005;**46:**863-870

Ewens WJ & Spielman RS. The transmission/disequilibrium test: history, subdivision, and admixture. *Am J Hum Genet* 1995;**57:**455-464

Falush D, Stephens M, Pritchard JK. Inference of population structure using multilocus genotype data: linked loci and correlated allele frequencies. *Genetics* 2003;**164:**1567-1587

Falush D, Stephens M, Pritchard JK. Inference of population structure using multilocus genotype data: dominant markers and null alleles. *Mol Ecol Notes* 2007;**7:**574-578

Farrall M, Green FR, Peden JF, Olsson PG, Clarke R, Hellenius ML, Rust S, Lagercrantz J, Franzosi MG, Schulte H, Carey A, Olsson G, Assmann G, Tognoni G, Collins R, Hamsten A, Watkins H. Genome-wide mapping of susceptibility to coronary artery disease identifies a novel replicated locus on chromosome 17. *PLoS Genet* 2006;**2:**e72

Feldman HA, Johannes CB, McKinlay JB, Longcope C. Low dehydroepiandrosterone sulfate and heart disease in middle-aged men: cross-sectional results from the Massachusetts male aging study. *Ann Epidemiol* 1998;**8:**217–228

Feldman HA, Longcope C, Derby CA, Johannes CB, Araujo AB, Coviello AD, Bremner WJ, McKinlay JB. Age trends in the level of serum testosterone and other hormones in middle-aged men: longitudinal results from the Massachusetts male aging study. *J Clin Endocrinol Metab* 2002;**87:**589-598

Ferlin A, Speltra E, Garolla A, Selice R, Zuccarello D, Foresta C. Y chromosome haplogroups and susceptibility to testicular cancer. *Mol Hum Reprod* 2007;**13:**615-619

Filipovic M, Goldacre MJ, Roberts SE, Yeates D, Duncan ME, Cook-Mozaffari P. Trends in mortality and hospital admission rates for abdominal aortic aneurysm in England and Wales, 1979-1999. *Br J Surg* 2005;**92**:968-75.

Fincham AG, Bessem CC, Lau EC, Pavlova Z, Shuler C, Slavkin HC, Snead ML. Human developing enamel proteins exhibit a sex-linked dimorphism. *Calcif Tissue Int* 1991;**48:**288-290

Flaquer A, Rappold GA, Wienker TF, Fischer C. The human pseudoautosomal regions: a review for genetic epidemiologists. *Eur J Hum Genet* 2008;**16:**771-779

Folsom AR, Stevens J, Schreiner PJ, McGovern PG. Body mass index, waist/hip ratio, and coronary heart disease incidence in African Americans and whites. Atherosclerosis Risk in Communities Study Investigators. *Am J Epidemiol* 1998;**148:**1187-1194

Fowkes FG, Macintyre CC, Ruckley CV. Increasing incidence of aortic aneurysms in England and Wales. *BMJ* 1989:**298:**33-35

Francke S, Manraj M, Lacquemant C, Lecoeur C, Lepretre F, Passa P, Hebe A, Corset L, Yan SL, Lahmidi S, Jankee S, Gunness TK, Ramjuttun US, Balgobin V, Dina C, Froguel P. A genome-wide scan for coronary heart disease suggests in Indo-Mauritians a susceptibility locus on chromosome 16p13 and replicates linkage with the metabolic syndrome on 3q27. *Hum Mol Genet* 2001;**10:**2751–2765

Franzen J & Fex G. Low serum apolipoprotein A-1 in acute myocardial infarction survivors with normal HDL cholesterol. *Atherosclerosis* 1986;**59:**37–42

Frenzel H, Hampe J, Huse K, Mascheretti SM, Croucher PP, Lynch NJ, Siebert R, Platzer M, Schreiber S. Mutation detection and physical mapping of the CD11 gene cluster in association with inflammatory bowel disease. *Immunogenetics* 2002;**53:**835-842

Fuller M, Meikle PJ, Hopwood JJ. Glycosaminoglycan degradation fragments in mucopolysaccharidosis I. *Glycobiology* 2004;**14:**443-450

Galland RB, Simmons MJ, Torrie EP. Prevalence of abdominal aortic aneurysm in patients with occlusive peripheral vascular disease. *Br J Surg* 1991;**78:**1259-1260

Gallo LC & Matthews KA.Understanding the association between socioeconomic status and physical health: do negative emotions play a role? *Psychol Bull* 2003;**129:**10-51

García-León A, Reyes GA, Vila J, Pérez N, Robles H, Ramos MM. The Aggression Questionnaire: a validation study in student samples. *Span J Phsychol* 2002;**5:**45-53

Gatewood JD, Wills A, Shetty S, Xu J, Arnold AP, Burgoyne PS, Rissman EF. Sex chromosome complement and gonadal sex influence aggressive and parental behaviours in mice. *J Neurosci* 2006;**26:**2335-2342

Genest JJ, McNamara JR, Salem DN, Schaefer EJ. Prevalence of risk factors in men with premature coronary artery disease. *Am J Cardiol* 1991;**67:**1185-1189

Genovesi S, Giussani M, Pieruzzi F, Vigorita F, Arcovio C, Cavuto S, Stella A. Results of blood pressure screening in a population of school-aged children in the province of Milan: role of overweight. *J Hypertens* 2005;**23:**493–497

Giulietti A, Overbergh L, Valckx D, Decallonne B, Bouillon R, Mathieu C. An overview of real-time quantitative PCR: applications to quantify cytokine gene expression. *Methods* 2001;**25:**386-401

Giusti B, Saracini C, Bolli P, Magi A, Sestini I, Sticchi E, Pratesi G, Pulli R, Pratesi C, Abbate R. Genetic analysis of 56 polymorphisms in 17 genes with abdominal aortic aneurysm involved in methionine metabolism in patients. *J Med Genet* 2008;**45:**721-730

Gluckman PD, Hanson MA, Buklijas T, Low FM, Beedle AS. Epigenetic mechanisms that underpin metabolic and cardiovascular diseases. *Nat Rev Endocrinol* 2009;**5:**401-408

Golledge AL, Walker P, Norman PE, Golledge J. A systematic review of studies examining inflammation associated cytokines in human abdominal aortic aneurysm samples. *Dis Markers* 2009;**26:**181-188

Gomes MN, Choyke PL, Wallace RB. Infected aorticaneurysms.A changing entity. *Ann Surg* 1992;**215:**435-442

Gordon S. Alternative activation of macrophages. *Nat Rev Immunol* 2003;**3:**23-35

Götz MJ, Johnstone EC, Ratcliffe SG. Criminality and antisocial behaviour in unselected men with sex chromosome abnormalities. *Psychol Med* 1999;**29:**953-962

Goulmy E. Human minor histocompatibility antigens: new concepts for marrow transplantation and adoptive immunotherapy. *Immunol Rev* 1997;**157:**125-140

Goumans MJ, de Boer TP, Smits AM, van Laake LW, van Vliet P, Metz CH, Korfage TH, Kats KP, Hochstenbach R, Pasterkamp G, Verhaar MC, van der Heyden MA, de Kleijn D, Mummery CL, van Veen TA, Sluijter JP, Doevendans PA. TGF-beta1 induces efficient differentiation of human cardiomyocyte progenitor cells into functional cardiomyocytes in vitro. *Stem Cell Res* 2009;**1:**138-149

Graves JA. Evolution of the mammalian Y chromosome and sex-determining genes. *J Exp Zool* 1998;**281:**472-481

Greenfield A, Carrel L, Pennisi D, Philippe C, Quaderi N, Siggers P, Steiner K, Tam PP, Monaco AP, Willard HF, Koopman P. The UTX gene escapes X inactivation in mice and humans. *Hum Mol Genet* 1998;**7:**737-742

Gretarsdottir S, Baas AF, Thorleifsson G, Holm H, den Heijer M, de Vries JP, Kranendonk SE, Zeebregts CJ, van Sterkenburg SM, Geelkerken RH, van Rij AM, Williams MJ, Boll AP, Kostic JP, Jonasdottir A, Jonasdottir A, Walters GB, Masson G, Sulem P, Saemundsdottir J, Mouy M, Magnusson KP, Tromp G, Elmore JR, Sakalihasan N, Limet R, Defraigne JO, Ferrell RE, Ronkainen A, Ruigrok YM, Wijmenga C, Grobbee DE, Shah SH, Granger CB, Quyyumi AA, Vaccarino V, Patel

RS, Zafari AM, Levey AI, Austin H, Girelli D, Pignatti PF, Olivieri O, Martinelli N, Malerba G, Trabetti E, Becker LC, Becker DM, Reilly MP, Rader DJ, Mueller T, Dieplinger B, Haltmayer M, Urbonavicius S, Lindblad B, Gottsäter A, Gaetani E, Pola R, Wells P, Rodger M, Forgie M, Langlois N, Corral J, Vicente V, Fontcuberta J, España F, Grarup N, Jørgensen T, Witte DR, Hansen T, Pedersen O, Aben KK, de Graaf J, Holewijn S, Folkersen L, Franco-Cereceda A, Eriksson P, Collier DA, Stefansson H, Steinthorsdottir V, Rafnar T, Valdimarsson EM, Magnadottir HB, Sveinbjornsdottir S, Olafsson I, Magnusson MK, Palmason R, Haraldsdottir V, Andersen K, Onundarson PT, Thorgeirsson G, Kiemeney LA, Powell JT, Carey DJ, Kuivaniemi H, Lindholt JS, Jones GT, Kong A, Blankensteijn JD, Matthiasson SE, Thorsteinsdottir U, Stefansson K. Genome-wide association study identifies a sequence variant within the *DAB2IP* gene conferring susceptibility to abdominal aortic aneurysm. *Nat Genet* 2010;**42:**692-622

Gourdy P, Araujo LM, Zhu R, Garmy-Susini B, Diem S, Laurell H, Leite-de-Moraes M, Dy M, Arnal JF, Bayard F, Herbelin A. Relevance of sexual dimorphism to regulatory T cells: oestradiol promotes IFNγ production by invariant natural killer T cells. *Blood* 2005;**105:**2415-2420

Grootenboer N, Bosch JL, Hendriks JM, van Sambeek MR. Epidemiology, aetiology, risk of rupture and treatment of abdominal aortic aneurysms: does sex matter? *Eur J Vasc Endovasc Surg* 2009;**38:**278-284

González P, García-Castro M, Reguero JR, Batalla A, Ordóñez AG, Palop RL, Lozano I, Montes M, Alvarez V, Coto E. The Pro279Leu variant in the transcription factor MEF2A is associated with myocardial infarction. *J Med Genet* 2006;**43:**167-169

Hägg S, Skogsberg J, Lundström J, Noori P, Nilsson R, Zhong H, Maleki S, Shang MM, Brinne B, Bradshaw M, Bajic VB, Samnegård A, Silveira A, Kaplan LM, Gigante B, Leander K, de Faire U, Rosfors S, Lockowandt U, Liska J, Konrad P, Takolander R, Franco-Cereceda A, Schadt EE, Ivert T, Hamsten A, Tegnér J, Björkegren J. Multi-organ expression profiling uncovers a gene module in coronary artery disease involving transendothelial migration of leukocytes and LIM domain binding 2: the Stockholm Atherosclerosis Gene Expression (STAGE) study. *PLoS Genet* 2009;**5:**e1000754

Hallan S, Astor B, Romundstad S, Aasarød K, Kvenild K, Coresh J. Association of kidney function and albuminuria with cardiovascular mortality in older vs younger individuals: The HUNT II Study. *Arch Intern Med* 2007;**167:**2490-2496

Hämäläinen E, Tikkanen H, Härkonen M, Näveri H, Adlercreutz H. Serum lipoproteins, sex hormones and sex hormone binding globulin in middle-aged men of different physical fitness and risk of coronary heart disease. *Atherosclerosis* 1987;**67:**155–162

Handa R. Cardiovascular co-morbidity in Asians with lupus: theoretical concern or clinical reality? *Lupus* 2010;**19:**1447-1451

Handy DE, Castro R, Loscalzo J. Epigenetic modifications. Basic mechanisms and role in Cardiovascular Disease. *Circulation* 2011;**123:**2145-2156

Hanks SK & Hunter T. Protein kinases 6. The eukaryotic protein kinase superfamily: kinase (catalytic) domain structure and classification. *FASEB J* 1995;**9:**576-596

Hanks SK. Genomic analysis of the eukaryotic protein kinase superfamily: a perspective. *Genome Biol* 2003;**4:**111

Hansson GK & Libby P. The immune response in atherosclerosis: a double-edged sword. *Nat Rev Immunol* 2006;**6:**508–19

Harman SM, Metter EJ, Tobin JD, Pearson J, Blackman MR. Longitudinal effects of aging on serum total and free testosterone levels in healthy men. *J Clin Endocrinol Metab* 2001;**86:**724–731

Harrap SB, Zammit KS, Wong ZY, Williams FM, Bahlo M, Tonkin AM, Anderson ST. Genome-wide linkage analysis analysis of the acute coronary syndrome suggests a locus on chromosome 2. *Arterioscler Thromb Vasc Biol* 2002;**22:**874–878

Harris JA. A further evaluation of the Aggression Questionnaire: issues of validity and reliability. *Behav Res Ther* 1997;**35:**1047-1053

Harrison SC, Cooper JA, Li K, Talmud PJ, Sofat R, Stephens JW, Hamsten A, HIFMECH Consortium, Sanders J, Montgomery H, Neil A**,** Simon Broome Research Consortium, Humphries SE. Association of a sequence variant in DAB2IP with coronary heart disease. *Eur Heart J* 2012;**33:**881-888

Harthun N. Current issues in the treatment of women with abdominal aortic aneurysm. *Gend Med* 2007;**5:**36-43

Hauner H, Stangl K, Burger K, Busch U, Blomer H, Pfeiffer EF. Sex hormone concentration in men with angiographically assessed coronary artery disease—relationship to obesity and body fat distribution. *Klin Wochenschr* 1991;**69:**664–668

Hauser ER, Watanabe RM, Duren WL, Bass MP, Langefeld CD, Boehnke M. Ordered subset analysis in genetic linkage mapping of complex traits. *Genet Epidemiol* 2004;**27:**53–63

Hautenen A, Manttari M, Manninen V, Tenkanen L, Huttunen JK, Frick MH, Adlercreutz H. Adrenal androgens and testosterone as coronary risk factors in the Helsinki heart study. *Atherosclerosis* 1994;**105:**191–200

Haynes SG & Feinleib M. Women, work and coronary heart disease: prospective findings from the Framingham Heart Study. *Am J Public Health*1980;**70:**133-141

Haynes SG, Levine S, Scotch N, Feinleib M, Kannel WB. The relationship of psychosocial factors to coronary heart disease in the Framingham study, 1: methods and risk factors. *Am J Epidemiol* 1978;**107:**362-38.

Helgadottir A, Manolescu A, Thorleifsson G, Gretarsdottir S, Jonsdottir H, Thorsteinsdottir U, Samani NJ, Gudmundsson G, Grant SF, Thorgeirsson G,

Sveinbjornsdottir S, Valdimarsson EM, Matthiasson SE, Johannsson H, Gudmundsdottir O, Gurney ME, Sainz J, Thorhallsdottir M, Andresdottir M, Frigge ML, Topol EJ, Kong A, Gudnason V, Hakonarson H, Gulcher JR, Stefansson K. The gene encoding 5-lipoxygenase activating protein confers risk of myocardial infarction and stroke. *Nat Genet* 2004;**36:**233-239

Helgeson VS. The role of masculinity in a prognostic predictor of heart attack severity. *Sex Roles* 1990;**22:**755-774

Helgadottir A, Manolescu A, Helgason A, Thorleifsson G, Thorsteinsdottir U, Gudbjartsson DF, Gretarsdottir S, Magnusson KP, Gudmundsson G, Hicks A, Jonsson T, Grant SF, Sainz J, O'Brien SJ, Sveinbjornsdottir S, Valdimarsson EM, Matthiasson SE, Levey AI, Abramson JL, Reilly MP, Vaccarino V, Wolfe ML, Gudnason V, Quyyumi AA, Topol EJ, Rader DJ, Thorgeirsson G, Gulcher JR, Hakonarson H, Kong A, Stefansson K. A variant of the gene encoding leukotriene A4 hydrolase confers ethnicity-specific risk of myocardial infarction. *Nat Genet* 2006;**38:**68-74

Helgadottir A, Thorleifsson G, Magnusson KP, Grétarsdottir S, Steinthorsdottir V, Manolescu A, Jones GT, Rinkel GJ, Blankensteijn JD, Ronkainen A, Jääskeläinen JE, Kyo Y, Lenk GM, Sakalihasan N, Kostulas K, Gottsäter A, Flex A, Stefansson H, Hansen T, Andersen G, Weinsheimer S, Borch-Johnsen K, Jorgensen T, Shah SH, Quyyumi AA, Granger CB, Reilly MP, Austin H, Levey AI, Vaccarino V, Palsdottir E, Walters GB, Jonsdottir T, Snorradottir S, Magnusdottir D, Gudmundsson G, Ferrell RE, Sveinbjornsdottir S, Hernesniemi J, Niemelä M, Limet R, Andersen K, Sigurdsson G, Benediktsson R, Verhoeven EL, Teijink JA, Grobbee DE, Rader DJ, Collier DA, Pedersen O, Pola R, Hillert J, Lindblad B, Valdimarsson EM, Magnadottir HB, Wijmenga C, Tromp G, Baas AF, Ruigrok YM, van Rij AM, Kuivaniemi H, Powell JT, Matthiasson SE, Gulcher JR, Thorgeirsson G, Kong A, Thorsteinsdottir U, Stefansson K. The same sequence variant on 9p21 associates with myocardial infarction, abdominal aortic aneurysm and intracranial aneurysm. *Nat Genet* 2008;**40:**217-224

Heller RF, Wheeler MJ, Micallef J, Miller NE, Lewis B. Relationship of high density lipoprotein cholesterol with total and free testosterone and sex hormone binding globulin. *Acta Endocrinol* (*Copenh*) 1983;**104:**253–256

Henriques T, Zhang X, Yiannikouris FB, Daugherty A, Cassis LA. Androgen increases AT1a receptor expression in abdominal aortas to promote angiotensin II induced AAAs in apolipoprotein E deficient mice. *Arterioscler Thromb Vasc Biol* 2008;**28:**1251-1256

Herrera VL, Tsikoudakis A, Ponce LR, Matsubara Y, Ruiz-Opazo N. Sex-specific QTLs and interacting loci underlie salt-sensitive hypertensive and target organ complications in Dahl S/jrHS hypertensive rats. *Physiol Genomics* 2006;**16:**172-179

Higgins CD, Swerdlow AJ, Schoemaker MJ, Wright AF, Jacobs PA, UK Clinical Cytogenetics Group. Mortality and cancer incidence in males with Y polysomy in Britain: a cohort study. *Hum Genet* 2007;**121:**691-696

Hinohara K, Nakajima T, Takahashi M, Hohda S, Sasaoka T, Nakahara K, Chida K, Sawabe M, Arimura T, Sato A, Lee BS, Ban JM, Yasunami M, Park JE, Izumi T, Kimura A. Replication of the association between a chromosome 9p21 polymorphism and coronary artery disease in Japanese and Korean populations. *J Hum Genet* 2008;**53:**357-359

Hinterseher I, Tromp G, Kuivaniemi H. Genes and abdominal aortic aneurysm. *Ann Vasc Surg* 2011;**25:**388-412

Hirose H&Tilson MD. Abdominal aortic aneurysm as an autoimmune disease. *Ann N Y Acad Sci* 2001;**947:**416-418

Hiura Y, Fukushima Y, Yuno M, Sawamura H, Kokubo Y, Okamura T, Tomoike H, Goto Y, Nonogi H, Takahashi R, Iwai N. Validation of the association of genetic variants on chromosome 9p21 and 1q41 with myocardial infarction in a Japanese population. *Circ J* 2008;**72:**1213-1217

Howard BV, Lee ET, Cowan LD, Fabsitz RR, Howard WJ, Oopik AJ, Robbins DC, Savage PJ, Yeh JL, Welty TK. Coronary heart disease prevalence and its relation to risk factors in American Indians.The Strong Heart Study. *Am J Epidemiol* 1995;**142:**254-268

Hromadova M, Hacik T, Riecansky I. Concentration of lipid, apoprotein-B and testosterone in patients with coronarographic findings. *Klin Wochenschr* 1985;**63:**1071–1074

Hubisz MJ, Falush D, Stephens M, Pritchard JK. Inferring weak population structure with the assistance of sample group information. *Mol Ecol Resour* 2009;**9:**1322-1332

Huxley RR & Woodward M. Cigarette smoking as a risk factor for coronary heart disease in women compared with men: a systematic review and meta-analysis of prospective cohort studies. *Lancet* 2011;**378:**1297-1305

Ideker T, Thorsson V, Ranish JA, Christmas R, Buhler J, Eng JK, Bumgarner R, Goodlett DR, Aebersold R, Hood L. Integrated genomic and proteomic analyses of a systematically perturbed metabolic network. *Science* 2001;**292:**929-934

International Human Genome Sequencing Consortium. Finishing the euchromatic sequence of the human genome. *Nature* 2004;**431:**931-945

Ioannidis JP, Lau J. Pooling research results: benefits and limitations of meta-analysis. *Jt Comm J Qual Improv* 1999;**25**:462-469

Ishii T, Sootome H, Shan L, Yamashita K. Validation of universal conditions for duplex quantitative reverse transcription polymerase chain reaction assays. *Anal Biochem* 2007;**362:**201-212

Ishikawa S, Takahashi T, Sakata K, Suzuki M, Kano M, Kawashima O, Ohshima K, Oki S, Ohtaki A, Morishita Y.Abdominal aortic dilatation in Japanese residents. *Int J Angiol* 1999;**8:**27-28

Itsui Y, Sakamoto N, Kakinuma S, Nakagawa M, Sekine-Osajima Y, Tasaka-Fujita M, Nishimura-Sakurai Y, Suda G, Karakama Y, Mishima K, Yamamoto M, Watanabe T, Ueyama M, Funaoka Y, Azuma S, Watanabe M. Antiviral effects of the interferon-induced protein guanylate binding protein 1 and its interaction with the hepatitis C virus NS5B protein. *Heptatology* 2009;**50:**1727-1737

Iwase M, Kaneko S, Kim H, Satta Y, Takahata N. Evolutionary history of sex-linked mammalian amelogenin genes. *Cells Tissues Organs* 2007;**187:**49-59

Janket S, Meurman JH, Baird AE, Qvarnstrom M, Nuutinen P, Ackerson LK, Hong J, Muthukrishnan P & Van dyke TE. Salivary immunoglobulin and prevalent coronary artery disease. *J Dent Res* 2010;**89:**389-394

Jehan Z, Vallinayagam S, Tiwari S, Pradhan S, Singh L, Suresh A, Reddy HM, Ahuja YR, Jesudasan RA. Novel noncoding RNA from human Y distal heterochromatic block (Yq12) generates testis-specific chimeric CDC2L2. *Genome Res* 2007;**17**:433-40.

Jin B, Ni HC, Shen W, Li J, Li Y. Cytochrome P450 2C19 polymorphisms in associated with poor clinical outcomes in coronary artery disease patients treated with clopidogrel. *Mol Biol Rep* 2010;**16:**1697-1702

Jobling MA, Bouzekri N, Taylor PG. Hypervariable digital DNA codes for human paternal lineages: MVR-PCR at the Y-specific minisatellite, MSY1 (DYF155S1). *Hum Mol Genet* 1998;**7:**643-653

Jobling MA. Copy number variation on the Y chromosome. *Cytogenet Genome Res* 2008;**123:**253-262

Johnson T, Gaunt TR, Newhouse SJ, Padmanabhan S, Tomaszewski M, Kumari M, Morris RW, Tzoulaki I, O'Brien ET, Poulter NR, Sever P, Shields DC, Thom S, Wannamethee SG, Whincup PH, Brown MJ, Connell JM, Dobson RJ, Howard PJ, Mein CA, Onipinla A, Shaw-Hawkins S, Zhang Y, Davey Smith G, Day IN, Lawlor DA, Goodall AH; Cardiogenics Consortium, Fowkes FG, Abecasis GR, Elliott P, Gateva V; Global BPgen Consortium, Braund PS, Burton PR, Nelson CP, Tobin MD, van der Harst P, Glorioso N, Neuvrith H, Salvi E, Staessen JA, Stucchi A, Devos N, Jeunemaitre X, Plouin PF, Tichet J, Juhanson P, Org E, Putku M, Sõber S, Veldre G, Viigimaa M, Levinsson A, Rosengren A, Thelle DS, Hastie CE, Hedner T, Lee WK, Melander O, Wahlstrand B, Hardy R, Wong A, Cooper JA, Palmen J, Chen L, Stewart AF, Wells GA, Westra HJ, Wolfs MG, Clarke R, Franzosi MG, Goel A, Hamsten A, Lathrop M, Peden JF, Seedorf U, Watkins H, Ouwehand WH, Sambrook J, Stephens J, Casas JP, Drenos F, Holmes MV, Kivimaki M, Shah S, Shah T, Talmud PJ, Whittaker J, Wallace C, Delles C, Laan M, Kuh D, Humphries SE, Nyberg F, Cusi D, Roberts R, Newton-Cheh C, Franke L, Stanton AV, Dominiczak AF, Farrall M, Hingorani AD, Samani NJ, Caulfield MJ, Munroe PB. Blood pressure loci identified with a gene-centric array. *Am J Hum Genet* 2011;**89:**688-700

Jolly C, Mitar I, Sattentau QJ. Adhesion molecule interactions facilitate human immunodeficiency virus type 1-induced virological synapse formation between T cells. *J Virol* 2007;**81:**13196-13921

Jones GT, Phillips VL, Harris EL, Rossaak JI, van Rij AM. Functional matrix metalloproteinase-9 polymorphism (C-1562T) associated with abdominal aortic aneurysm. *J Vasc Surg* 2003;**38:**1363-1367

Jones GT, Harris EL, Phillips LV, van Rij AM. The methylenetetrahydrofolate reductase C677T polymorphism does not associate with susceptibility to abdominal aortic aneurysm. *Eur J Vasc Endovasc Surg* 2005;**30:**137-142

Juvonen J, Juvonen T, Laurila A, Alakärppä H, Lounatmaa K, Surcel HM, Leinonen M, Kairaluoma MI, Saikku P. Demonstration of Chlamydia pneumoniae in the walls of abdominal aortic aneurysms. *J Vasc Surg* 1997;**25:**499-505

Kabakci G, Yildirir A, Can I, Unsal I, Erbas B. Relationship between endogenous sex hormone levels, lipoproteins and coronary atherosclerosis in men undergoing coronary angiography. *Cardiology* 1999;**92:**221–225

Kaess BM, Jozwiak J, Mastej M, Lukas W, Grzeszczak W, Windak A, Piwowarska W, Tykarski A, Konduracka E, Rygiel K, Manasar A, Samani NJ, Tomaszewski M. Association between anthropometric obesity measurements and CAD: a cross-sectional survey of 16,657 subjects from 444 Polish cities. *Heart* 2010;**96:**131-135

Karafet TM, Mendez FL, Meilerman MB, Underhill PA, Zegura SL, Hammer MF. New binary polymorphisms reshape and increase resolution of the human Y chromosomal haplogroup tree. *Genome Res* 2008;**18:**830-838

Karkos CD, Mukhopadhyay U, Papakostas I, Ghosh J, Thomson GJ, Hughes R. Abdominal aortic aneurysm: the role of clinical examination and opportunistic detection. *J Endovasc Surg* 2000;**19:**299-303

Karp I, O'Loughlin J, Paradis G, Hanley J, Difranza J. Smoking trajectories of adolescent novice smokers in a longitudinal study of tobacco use. *Ann Epidemiol* 2005;**15:**445–452

Kawaguchi H, Miyoshi N, Miura N, Fujiki M, Horiuchi M, Izumi Y, Miyajima H, Nagata R, Misumi K, Takeuchi T, Tanimoto A, Yoshida H. Microminipig, a non-rodent experimental animal optimized for life science research: novel atherosclerosis model induced by high fat and cholesterol diet. *J Pharmacol Sci* 2011;**115:**115-121

Keating BJ, Tischfield S, Murray SS, Bhangale T, Price TS, Glessner JT, Galver L, Barrett JC, Grant SF, Farlow DN, Chandrupatla HR, Hansen M, Ajmal S, Papanicolaou GJ, Guo Y, Li M, Derohannessian S, de Bakker PI, Bailey SD, Montpetit A, Edmondson AC, Taylor K, Gai X, Wang SS, Fornage M, Shaikh T, Groop L, Boehnke M, Hall AS, Hattersley AT, Frackelton E, Patterson N, Chiang CW, Kim CE, Fabsitz RR, Ouwehand W, Price AL, Munroe P, Caulfield M, Drake T, Boerwinkle E, Reich D, Whitehead AS, Cappola TP, Samani NJ, Lusis AJ, Schadt E, Wilson JG, Koenig W, McCarthy MI, Kathiresan S, Gabriel SB, Hakonarson H, Anand SS, Reilly M, Engert JC, Nickerson DA, Rader DJ, Hirschhorn JN, Fitzgerald GA. Concept, design and implementation of a cardiovascular gene-centric 50 k SNP array for large-scale genomic association studies. *PLoS One* 2008;**3:**e3583

Keil JE, Sutherland SE, Knapp RG, Lackland DT, Gazes PC, Tyroler HA. Mortality rates and risk factors for coronary disease in black as compared with white men and women.*N Engl J Med* 1993;**329:**73-78

King JK, Yeh SH, Lin MW, Liu CJ, Lai MY, Kao JH, Chen DS, Chen PJ. Genetic polymorphisms in interferon pathway and response to interferon treatment in hepatitis B patients: A pilot study. *Hepatology* 2002;**36:**1416-1424

Kittles RA, Long JC, Bergen AW, Eggert M, Virkkunen M, Linnoila M, Goldman D. Cladistic association analysis of Y chromosome effects on alcohol dependence and related personality traits. *Proc Natl Acad SciUSA* 1999;**30:**4204-4209

Knowler WC, Williams RC, Pettitt DJ, Steinberg AG. Gm3;5,13,14 and type 2 diabetes mellitus: an association in American Indians with genetic admixture. *Am J Hum Genet* 1988;**43:**520-526

Koch S, Kucharzik T, Heidemann J, Nusrat A, Luegering A. Investigating the role of proinflammatory CD16+ monocytes in the pathogenesis of inflammatory bowel disease. *Clin Exp Immunol* 2010;**161:**332-341

Komander D, Garg R, Wan PT, Ridley AJ, Barford D. Mechanism of multi-site phosphorylation from a ROCK-I:RhoE complex structure. *EMBO J* 2008;**27:**3175-3185

Kop WJ. Chronic and acute psychological risk factors for clinical manifestations of coronary artery disease. *Psychosom Med* 1999;**61:**476-87

Kozlowska-Wojciechowska M, Jez W, Zdrojewski T, Chwojnicki K. Are young women with Turner syndrome at greater risk of coronary artery disease? *Eur J Cardiovasc Prev Rehabil* 2006;**13:**467-469

Kraemer R. Regulation of cell migration in atherosclerosis. *Curr Atheroscler Rep* 2000;**2:**445-452

Kral BG & Becker DM. Familial occurrence of abnormalities of high-density lipoprotein cholesterol. *J Clin Lipidol* 2007;**1**:31-40

Krantz DS & McCeney MK. Effects of psychological and social factors on organic disease: a critical assessment of research on coronary heart disease. *Annu Rev Psycol* 2002;**53:**341-369

Krausz C, Quintana-Murci L, Rajpert-De Meyts E, Jørgensen N, Jobling MA, Rosser ZH, Skakkebaek NE, McElreavey K. Identification of a Y chromosome haplogroup associated with reduced sperm counts. *Hum Mol Genet* 2001;**10:**1873-1877

Krausz C, Quintana-Murci L, Forti G. Y chromosome polymorphisms in medicine. *Ann Med* 2004;**36:**573-583

Kronzon I. Diagnosis and treatment of iatrogenic femoral artery pseudoaneurysm: a review. *J Am Soc Echocardiogr* 1997;**10:**236-245

Kuivaniemi H, Shibamura H, Arthur C, Berguer R, Cole CW, Juvonen T, Kline RA, Limet R, Mackean G, Norrgård O, Pals G, Powell JT, Rainio P, Sakalihasan N, van Vlijmen-van Keulen C, Verloes A, Tromp G. Familial Abdominal aortic aneurysms: Collection of 233 multiplex families. *J Vasc Surg* 2003;**37:**340-345

Kuivaniemi H, Kyo Y, Lenk G, Tromp G. Genome-wide approach to finding abdominal aortic aneurysm susceptibility genes in Humans. *Ann N Y Acad Sci*2006;**1085:**270-281

Kuo WP, Jenssen TK, Butte AJ, Ohno-Machado L, Kohane IS. Analysis of matched mRNA measurements from two different microarray technologies. *Bioinformatics* 2002;**18:**405-412

Laaser I, Theis FJ, de Angelis MH, Kolb HJ, Adamski J. Huge splicing frequency in human Y chromosomal UTY gene. *OMICS* 2011;**15:**141-154

Labropoulos B, Velonakis E, Oekonomakos P, Laskaris J, Katsimades D. Serum sex hormones in patients with coronary disease and their relationship to known factors causing atherosclerosis. *Cardiology* 1982;**69:**98–103

Lander ES & Schork NJ. Genetic dissection of complex traits. *Science* 1994;**265:**2037-2048

Lau YC & Zhang J. Expression analysis of thirty one Y chromosome genes in human prostate cancer. *Mol Carcin* 2000;**27:**308-321

Lederle FA, Johnson GR, Wilson SE, Chute EP, Littooy FN, Bandyk D, Krupski WC, Barone GW, Acher CW, Ballard DJ. Prevalence and associations of abdominal aortic aneurysm detected through screening. *Ann Intern Med* 1997;**126:**441-449

Lee WH, Kim SH, Lee Y, Lee BB, Kwon B, Song H, Kwon BS, Park JE. Tumor necrosis factor receptor superfamily 14 is involved in atherogenesis by inducing proinflammatory cytokines and matrix metalloproteinases. *Arterioscler Thromb Vasc Biol* 2001;**21:**2004-2010

Le Goff C, Mahaut C, Abhyankar A, Le Goff W, Serre V, Afenjar A, Destrée A, di Rocco M, Héron D, Jacquemont S, Marlin S, Simon M, Tolmie J, Verloes A, Casanova JL, Munnich A, Cormier-Daire V. Mutations at a single codon in Mad homology 2 domain of SMAD4 cause Myhre syndrome. *Nat Genet* 2011;**44:**85-88

Lemos B, Araripe LO, Hartl DL. Polymorphic Y chromosomes harbor cryptic variation with manifold functional consequences. *Science* 2008;**319:**91-93

Lemos B, Branco AT, Hartl DL. Epigenetic effects of polymorphic Y chromosomes modulate chromatin components, immune response, and sexual conflict. *Proc Natl Acad Sci USA* 2010; **107:**15826–31.

Lenk GM, Tromp G,Weinsheimer S, Gatalica Z, Berguer R, Kuivaniemi H. Whole genome expression profiling reveals a significant role for immune function in human abdominal aortic aneurysms. *BMC Genomics* 2007;**8:**237

Lerner DJ&Kannel WB.Patterns of coronary heart disease morbidity and mortality in the sexes: A 26-year follow-up of the Framingham population. *Am Heart J* 1986;**111:**383-390

Le Roy I, Mortaud S, Tordjman S, Donsez-Darcel E, Carlier M, Degrelle H, Roubertoux PL. Genetic correlation between steroid sulfatase concentration and initiation of attack behavior in mice. *Behav Genet* 1999;**29:**131-136

Lewallen S & Courtright P. Epidemiology in practice: case-control studies. *Community Eye Health* 1998;**11:**57-58

Lewin B. Units of transcription and translation: the relationship between heterogeneous nuclear RNA and messenger RNA. *Cell* 1975;**4:**11-20

Ley K, Miller YI, Hedrick CC. Monocyte and macrophage dynamics during atherogenesis. *Arterioscler Thromb Vasc Biol* 2011;**31:**1506-1516

Li CM & Chiu RC. The importance and limitations of prospective randomized studies for new, evolving surgical procedures: lessons from the dynamic cardiomyoplasty trial. *Pacing Clin Electrophysiol* 1996;**19:**2035-2039

Li F, Jiang L, Willis-Owen SA, Zhang Y, Gao J. Vitamin D binding protein variants associate with asthma susceptibility in the Chinese Han population. *BMC Med Genet* 2011;**12:**103

Libby P& Theroux P. Pathophysiology of coronary artery disease. *Circulation* 2005;**111:**3481-3488

Lichtenstein MJ, Yamell JW, Elwood PC, Beswick AD, Sweetnam PM, Marks V, Teale D, Riad-Fahmy D. Sex hormones, insulin, lipids, and prevalent ischemic heart disease. *Am J Epidemiol* 1987;**126:**647–657

Lin R, Dai S, Irwin RD, Heinloth AN, Boorman GA, Li L. Gene set enrichment analysis for non-monotone association and multiple experimental categories. *BMC Bioinformatics* 2008;**9:**481

Lindström S, Adami HO, Adolfsson J, Wiklund F. Y chromosome haplotypes and prostate cancer in Sweden. *Clin Cancer Res* 2008;**14**6712-6716

Ling JC, Freeman AF, Gharib AM, Arai AE, Lederman RJ, Rosing DR & Hollan SM. Coronary artery aneurysms in patients with hyper IgE recurrent infection syndrome. *Clin Immunol* 2007;**122:**255-258

Liu H, Liu W, Liao Y, Cheng L, Liu Q, Ren X, Shi L, Tu X, Wang QK, Guo AY. CADgene: a comprehensive database for coronary artery disease genes. *Nucleic Acids Res.*2011;**39:**991-996

Liu J, Mercher T, Scholl C, Brumme K, Gilliland DG, Zhu N. A functional role for the histone demethylase UTX in normal and malignant hematopoietic cells. *Exp Hematol* 2012;**40:**487-498

Lloyd-Jones DM, Nam BH, D'Agostino RB Sr, Levy D, Murabito JM, Wang TJ, Wilson PW, O'Donnell CJ. Parental cardiovascular disease as a risk factor for cardiovascular disease in middle-aged adults: a prospective study of parents and offspring. *JAMA* 2004;**291:**2204-2211

López-Pedrera C, Aguirre MÁ, Barbarroja N, Cuadrado MJ.Accelerated atherosclerosis in systemic lupus erythematosus: role of proinflammatory cytokines and therapeutic approaches. *J Biomed Biotechnol* 2010.pii: 607084

Lovell-Badge R. Aggressive behaviour: contributions from genes on the Y chromosome. *Novartis Found Symp* 2005;**268:**20-33

Lowry PJ, Mace PJ, Glover DR, Littler WA. The pattern and severity of coronary artery disease in Asians and whites living in Birmingham. *Postgrad Med J* 1983;**59:**634-635

Luft FC, Schuster H, Bilginturan N, Wienker T. 'Treasure your exceptions': what we can learn from autosomal-dominant inherited forms of hypertension. *J Hypertens* 1995;**13:**1535-1538

Luoma PV. Cytochrome P450 and gene activation--from pharmacology to cholesterol elimination and regression of atherosclerosis. *Eur J Clin Pharmacol* 2008;**64:**841-850

Luria MH, Johnson MW, Pego R, Seuc CA, Manubens SJ, Wieland MR, Wieland RG. Relationship between sex hormones, myocardial infarction, and occlusive coronary disease. *Arch Intern Med* 1982;**142:**42–44

Madaric J, Vulev I, Bartunek J, Mistrik A, Verhamme K, De Bruyne B, Riecansky I. Frequency of abdominal aortic aneurysm in patients >60 years of age with coronary artery disease. *Am J Cardiol* 2005;**96:**1214-1216

Mahmoodi BK, Nijsten M, Wijsman J, Matthews AG, van der Laan L. ABO-blood groups and risk of abdominal aortic aneurysm and peripheral obstructive artery disease: two sides of the same coin. *Thromb Res* 2012;**129:**89-90

Makomaski Illing EM, Kaiserman MJ. Mortality attributable to tobacco use in Canada and its regions. *Can J Public Health* 2004;**95:**38–44

Makos M, Nelkin BD, Lerman MI, Latif F, Zbar B, Baylin SB. Distinct hypermethylation patterns occur at altered chromosome loci in human lung and colon cancer. *Proc Natl Acad Sci USA* 1992;**89:**1929-1933

Manning G, Whyte DB, Martinez R, Hunter T, Sudarsanam S. The protein kinase complement of the human genome. *Science* 2002;**298:**1912-1934

Manolio TA. Study designs to enhance identification of genetic factors in healthy aging. *Nutr Rev* 2007;**65:**228-233

Mathias N, Bayés M, Tyler-Smith C. Highly informative compound haplotypes for the human Y chromosome. *Hum Mol Genet* 1994;**3**:115-23.

Matter WF, Estridge T, Zhang C, Belagaje R, Stancato L, Dixon J, Johnson B, Bloem L, Pickard T, Donaghue M, Acton S, Jeyaseelan R, Kadambi V, Vlahos CJ. Role of PRL-3, a human muscle-specific tyrosine phosphatase, in angiotensin-II signalling. *Biochem Biophys Res Commun* 2001;**283:**1061-1068

Marenberg ME, Risch N, Berkman LF, Floderus B, de Faire U. Genetic susceptibility to death from coronary heart disease in a study of twins. *New Engl J Med* 1994;**330:**1041-1046

Maric C. Sex, diabetes and the kidney. *Am J Physiol Renal Physiol* 2009;**296:**680-688

Marmot MG, Syme SL, Kagan A, Kato H, Cohen JB, Belsky J. Epidemiologic studies of coronary heart disease and stroke in Japanese men living in Japan, Hawaii and California: prevalence of coronary and hypertensive heart disease and associated risk factors. *Am J Epidemiol* 1975;**102:**514-525

Marquez-Vidal P, Sie P, Cambou J-P, Chap H, Perret B. Relationships of plasminogen activator inhibitor activity and lipoprotein (a) with insulin, testosterone, 17 beta-oestradiol, and testosterone binding globulin in myocardial infarction patients and healthy controls. *J Clin Endocrinol Metab* 1995;**80:**1794–1798

McCombe PA, Greer JM, Mackay IR. Sexual dimorphism in autoimmune disease. *Curr Mol Med* 2009;**9:**1058-1079

McGorrian C, Yusuf S, Islam S, Jung H, Rangarajan S, Avezum A, Prabhakaran D, Almahmeed W, Rumboldt Z, Budaj A, Dans AL, Gerstein HC, Teo K, Anand SS**,** INTERHEART Investigators. Estimating modifiable coronary heart disease risk in multiple regions of the world: the INTERHEART Modifiable Risk Score. *Eur Heart J* 2011;**32:**581-589

McKee M, Bobak M, Rose R, Shkolnikov V, Chenet L, Leon D. Patterns of smoking in Russia. *Tob Control* 1998;**7:**22-26

McLaren JE & Ramji DP. Interferon gamma: a master regulator of atherosclerosis. *Cytokine Growth Factor Res* 2006;**20:**125-135

McPherson R, Pertsemlidis A, Kavaslar N, Stewart A, Roberts R, Cox DR, Hinds DA, Pennacchio LA, Tybjaerg-Hansen A, Folsom AR, Boerwinkle E, Hobbs HH, Cohen JC.

A common allele on chromosome 9 associated with coronary heart disease. *Science* 2007;**316:**488-491

McQueen MJ, Hawken S, Wang X, Ounpuu S, Sniderman A, Probstfield J, Steyn K, Sanderson JE, Hasani M, Volkova E, Kazmi K, Yusuf S, INTERHEART study investigators. Lipids, lipoproteins, and apolipoproteins as risk markers of myocardial infarction in 52 countries (the INTERHEART study): a case-control study. *Lancet* 2008;**372:**224-233

Meesters C, Muris P, Bosma H, Schouten E, Beuving S. Psychometric evaluation of the Dutch version of the Aggression Questionnaire. *Behav Res Ther* 1996;**34:**839-843

Mehall JR, Saltzman DA, Chandler JC, Kidd JN, Wells T, Smith SD. Congenital abdominal aortic aneurysm in the infant: case report and review of the literature. *J Pediatr Surg* 2001;**36:**657-658

Mehta M, Sternbach Y, Taggert JB, Kreienberg PB, Roddy SP, Paty PS, Ozsvath KJ, Darling RC 3rd. Long-term outcomes of secondary procedures after endovascular aneurysm repair. *J Vasc Surg* 2010;**52:**1442-1449

Melton LJ 3rd, Bickerstaff LK, Hollier LH, Van Peenen HJ, Lie JT, Pairolero PC, Cherry KJ, O'Fallon WM. Changing incidence of abdominal aortic aneurysms: a population-based study. *Am J Epidemiol* 1984;**120**:379-86.

Mendoza SG, Zerpa A, Carrasco H, Colmenares G, Rangel A, Gartside PS, Kashyap ML. Oestradiol, testosterone, apolipoproteins, lipoprotein cholesterol and lipolytic enzymes in men with premature myocardial infarction and angiographically assessed coronary occlusion. *Artrey* 1983;**12:**1–23

Mercher T, Scholl C, Brumme K, Gilliland DG, Zhu N. A functional role for the histone demethylase UTX in normal and malignant hematopoietic cells. *Exp Hematol* 2012;**40:**487-498

Merico D, Isserlin R, Stueker O, Emili A, Bader GD. Enrichment map: a network-based method for gene-set enrichment visualization and interpretation. *PLoS One* 2010;**5**:e13984

Miles DR & Carey G. Genetic and environmental architecture of human aggression. *J Pers Soc Psychol* 1997;**72:**207-217

Miller DV, Oderich GS, Aubry MC, Panneton JM, Edwards WD. Surgical pathology of infected aneurysms of the descending thoracic and abdominal aorta: clinicopathologic correlations in 29 cases (1976 to 1999). *Hum Pathol* 2004;**35:**1112-1120

Miller VM. Family matters: sexual dimorphism in cardiovascular disease. *Lancet* 2012;**379:**873-875

Mitchell LE, Sprecher DL, Borecki IB, Rice T, Laskarzewski PM. Evidence for an association between dehydroepiandrosterone sulfate and non-fatal, premature myocardial infarction in males. *Circulation* 1994;**89:**89–93

Mitchell RJ, Kreskas M, Baxter E, Buffalino L, Van Oorschot RA. An investigation of sequence deletions of amelogenin (AMELY), a Y-chromosome locus commonly used for gender determination. *Ann Hum Biol* 2006;**33:**227-240

Mitra AK, Del Core MG, Agrawal DK. Cells, cytokines and cellular immunity in the pathogenesis of fibroproliferative vasculopathies. *Can J Physiol Pharmacol* 2005;**83:**701-715

Mittendorfer B. Sexual dimorphism in Human lipid metabolism. *J Nutr* 2005;**135:**681-686

Mizuguchi T, Collod-Beroud G, Akiyama T, Abifadel M, Harada N, Morisaki T, Allard D, Varret M, Claustres M, Morisaki H, Ihara M, Kinoshita A, Yoshiura K, Junien C, Kajii T, Jondeau G, Ohta T, Kishino T, Furukawa Y, Nakamura Y, Niikawa N, Boileau C, Matsumoto N. Heterozygous *TGFBR2* mutations in Marfan syndrome. *Nat Genet* 2004;**36:**855-860

Modrek B & Lee C. A genomic view of alternative splicing. *Nat Genet* 2002;**30:**13-19

Moore LT, McEvoy B, Cape E, Simms K, Bradley DG. A Y-chromosome signature of hegemony in Gaelic Ireland. *Am J Hum Genet* 2006;**78:**334-338

Morris GE, Hubbard CS, Quick CR. An abdominal aortic aneurysm screening programme for all males over the age of 50 years. *Eur J Vasc Surg* 1994;**8:**156-160

Mozaffarian D, Kamineni A, Carnethon M, Djoussé L, Mukamal KJ, Siscovick D. Lifestyle risk factors and new-onset diabetes mellitus in older adults: the cardiovascular health study. *Arch Intern Med* 2009;**169**:798-807

Mudge GH Jr, Mills RM Jr, Taegtmeyer H, Gorlin R, Lesch M. Alterations of myocardial amino acid metabolism in chronic ischemic heart disease. *J Clin Invest* 1976;**58:**1185-1192

Mühl H & Pfeilschifter J. Anti-inflammatory properties of pro-inflammatory interferon-gamma. *Int Immunopharmacol* 2003;**3:**1247-1255

Myocardial Infarction Genetics Consortium, Kathiresan S, Voight BF, Purcell S, Musunuru K, Ardissino D, Mannucci PM, Anand S, Engert JC, Samani NJ, Schunkert H, Erdmann J, Reilly MP, Rader DJ, Morgan T, Spertus JA, Stoll M, Girelli D, McKeown PP, Patterson CC, Siscovick DS, O'Donnell CJ, Elosua R, Peltonen L, Salomaa V, Schwartz SM, Melander O, Altshuler D, Ardissino D, Merlini PA, Berzuini C, Bernardinelli L, Peyvandi F, Tubaro M, Celli P, Ferrario M, Fetiveau R, Marziliano N, Casari G, Galli M, Ribichini F, Rossi M, Bernardi F, Zonzin P, Piazza A, Mannucci PM, Schwartz SM, Siscovick DS, Yee J, Friedlander Y, Elosua R, Marrugat J, Lucas G, Subirana I, Sala J, Ramos R, Kathiresan S, Meigs JB, Williams G, Nathan DM, MacRae CA, O'Donnell CJ, Salomaa V, Havulinna AS, Peltonen L, Melander O, Berglund G,

Voight BF, Kathiresan S, Hirschhorn JN, Asselta R, Duga S, Spreafico M, Musunuru K, Daly MJ, Purcell S, Voight BF, Purcell S, Nemesh J, Korn JM, McCarroll SA, Schwartz SM, Yee J, Kathiresan S, Lucas G, Subirana I, Elosua R, Surti A, Guiducci C, Gianniny L, Mirel D, Parkin M, Burtt N, Gabriel SB, Samani NJ, Thompson JR, Braund PS, Wright BJ, Balmforth AJ, Ball SG, Hall A; Wellcome Trust Case Control Consortium, Schunkert H, Erdmann J, Linsel-Nitschke P, Lieb W, Ziegler A, König I, Hengstenberg C, Fischer M, Stark K, Grosshennig A, Preuss M, Wichmann HE, Schreiber S, Schunkert H, Samani NJ, Erdmann J, Ouwehand W, Hengstenberg C, Deloukas P, Scholz M, Cambien F, Reilly MP, Li M, Chen Z, Wilensky R, Matthai W, Qasim A, Hakonarson HH, Devaney J, Burnett MS, Pichard AD, Kent KM, Satler L, Lindsay JM, Waksman R, Knouff CW, Waterworth DM, Walker MC, Mooser V, Epstein SE, Rader DJ, Scheffold T, Berger K, Stoll M, Huge A, Girelli D, Martinelli N, Olivieri O, Corrocher R, Morgan T, Spertus JA, McKeown P, Patterson CC, Schunkert H, Erdmann E, Linsel-Nitschke P, Lieb W, Ziegler A, König IR, Hengstenberg C, Fischer M, Stark K, Grosshennig A, Preuss M, Wichmann HE, Schreiber S, Hólm H, Thorleifsson G, Thorsteinsdottir U, Stefansson K, Engert JC, Do R, Xie C, Anand S, Kathiresan S, Ardissino D, Mannucci PM, Siscovick D, O'Donnell CJ, Samani NJ, Melander O, Elosua R, Peltonen L, Salomaa V, Schwartz SM, Altshuler D. Genome-wide association of early-onset myocardial infarction with single nucleotide polymorphisms and copy number variants. *Nat Genet* 2009;**41**:334-41.

Myres NM, Rootsi S, Lin AA, Järve M, King RJ, Kutuev I, Cabrera VM, Khusnutdinova EK, Pshenichnov A, Yunusbayev B, Balanovsky O, Balanovska E, Rudan P, Baldovic M, Herrera RJ, Chiaroni J, Di Cristofaro J, Villems R, Kivisild T, Underhill PA. A major Y-chromosome haplogroup R1b Holocene era founder effect in Central and Western Europe. *Eur J Hum Genet* 2011;**19:**95-101

Nakano K. Psychometric evaluation on the Japanese adaptation of the Aggression Questionnaire. *Behav Res Ther* 2001;**39:**853-858

Naydeck BL, Sutton-Tyrrell K, Schiller KD, Newman AB, Kuller LH. Prevalence and risk factors for abdominal aortic aneurysms in older adults with and without isolated systolic hypertension. *Am J Cardiol* 1999;**83:**759-764

Naylor AR, Webb J, Fowkes FG, Ruckley CV. Trends in abdominal aortic aneurysm surgery in Scotland (1971-1984). *Eur J Vasc Surg* 1988;**2:**217-221

Ness J & Aronow WS. Prevalence of coronary artery disease, ischemic stroke, peripheral arterial disease, and coronary revascularization in older African-Americans, Asians, Hispanics, whites, men, and women. *Am J Cardiol* 1999;**84:**932-933

Neve BP, Fruchart JC, Staels B. Role of the peroxisome proliferator-activated receptors (PPAR) in atherosclerosis. *Biochem Pharmacol* 2000;**60:**1245-1250

Nickkholgh B, Noordam MJ, Hovingh SE, van Pelt AM, van der Veen F, Repping S. Y chromosome TSPY copy numbers and semen quality. *Fertil Steril* 2010;**95:**1744-1747

Nilsson J & Hansson GK. Autoimmunity in atherosclerosis: a protective response losing control? *J Intern Med* 2008;**263:**464-478

Noordam MJ & Repping S. The human Y chromosome: a masculine chromosome. *Curr Opin Genet Dev* 2006;**16:**225-232

Norman PE, Castleden WM, Hockey RL.Prevalence of abdominal aortic aneurysm in Western Australia. *Br J Surg* 1991;**78:**1118-1121

Nuwaysir EF, Huang W, Albert TJ, Singh J, Nuwaysir K, Pitas A, Richmond T, Gorski T, Berg JP, Ballin J, McCormick M, Norton J, Pollock T, Sumwalt T, Butcher L, Porter D, Molla M, Hall C, Blattner F, Sussman MR, Wallace RL, Cerrina F, Green RD. Gene expression analysis using oligonucleotide arrays produced by maskless photolithography. *Genome Res* 2002;**12:**1749-1755

Ober C, Loisel DA, Gilad Y. Sex-specific genetic architecture of human disease. *Nat Rev Genet* 2008;**9:**911-922

Obukofe B, Sayers RD, Thompson J, Sandford RM, London NJ, Samani NJ, Bown MJ. Is there an association between angiotensin converting enzyme (ACE) genotypes and abdominal aortic aneurysm? *Eur J Vasc Endovasc Surg* 2010;**40:**457-460

Ocaña E, Pérez-Requena J, Bohórquez JC, Brieva JA, Rodríguez C. Chemokine receptor expression on infiltrating lymphocytes from abdominal aortic aneurysms: role of CXCR4-CXCL12 in lymphoid recruitment. *Atherosclerosis* 2008;**200:**264-270

Ogata T, Shibamura H, Tromp G, Sinha M, Goddard KA, Sakalihasan N, Limet R, MacKean GL, Arthur C, Sueda T, Land S, Kuivaniemi H. Genetic analysis of polymorphisms in biologically relevant candidate genes in patients with abdominal aortic aneurysms. *J Vasc Surg* 2005;**41:**1036-1042

Ogeng'o JA, Olabu BO, Kilonzi JP. Pattern of aortic aneurysms in an African country. *J Thorac Cardiovasc Surg* 2010;**4:**797-800

Orwoll E, Lambert LC, Marshall LM, Blank J, Barrett-Connor E, Cauley J, Ensrud K, Cummings SR, Osteoporotic Fractures in Men Study Group. Endogenous testosterone levels, physical performance, and fall risk in older men. *Arch Intern Med* 2006;**166:**2124-2131

Owens AP 3rd, Rateri DL, Howatt DA, Moore KJ, Tobias PS, Curtiss LK, Lu H, Cassis LA, Daugherty A. MyD88 deficiency attenuates angiotensin II-induced abdominal aortic aneurysm formation independent of signalling through Toll-like receptors 2 and 4. *Arterioscler Thromb Vasc Biol* 2011;**31:**2813-2819

Pajukanta P, Cargill M, Viitanen L, Nuotio I, Kareinen A, Perola M, Terwilliger JD, Kempas E, Daly M, Lilja H, Rioux JD, Brettin T, Viikari JS, Ro¨nnemaa T, Laakso M, Lander ES, Peltonen L. Two loci on chromosomes 2 and X for premature coronary heart disease identified in early- and late-settlement populations of Finland. *Am J Hum Genet* 2000;**67:**1481–1493

Palinski W, Rosenfeld ME, Ylä-Herttuala S, Gurtner GC, Socher SS, Butler SW, Parthasarathy S, Carew TE, Steinberg D, Witztum JL.Low density lipoprotein undergoes oxidative modification in vivo. *Proc Natl Acad Sci USA* 1989;**86:**1372-1376

Palinski W & Witztum JL. Immune responses to oxidative neoepitopes on LDL and phospholipids modulate the development of atherosclerosis. *J Intern Med* 2000;**247:**371-380

Pan L, Ober C, Abney M. Heritability estimation of sex-specific effects on human quantitative traits. *Genet Epidemiol* 2007;**31:**338–347

Pandey LK, Pandey S, Gupta J, Saxena AK. Loss of the AZFc region due to a human Y-chromosome microdeletion in infertile male patients. *Genet Mol Res* 2010;**29:**1267-1273

Paradis G, Lambert M, O'Loughlin J, Lavallée C, Aubin J, Delvin E, Lévy E, Hanley JA. Blood pressure and adiposity in children and adolescents. *Circulation* 2004;**110:**1832–1838

Parkhurst GF & Decker JP.Bacterial aortitis and mycotis cnaeurysm of the aorta. *Am J Pathol* 1955;**31:**821-835

Patel RS & Ye S. Genetic determinants of coronary heart disease: new discoveries and insights from genome-wide association studies. *Heart* 2011;**97:**1463-1473

Pearce G, Audzevich T, Jessberger R. SYK regulates B-cell migration by phosphorylation of the F-actin interacting protein SWAP-70. *Blood* 2011;**117:**1574-1584

Pease AC, Solas D, Sullivan EJ, Cronin MT, Holmes CP, Fodor SP. Light-generated oligonucleotide arrays for rapid DNA sequence analysis. *Proc Natl Acad Sci USA* 1994;**91:**5022-5026

Perreault C, Décary F, Brochu S, Gyger M, Bélanger R, Roy D. Minor histocompatibility antigens. *Blood* 1990;**76:**1269-1280

Phillips GB, Yano K, Stemmerman GN. Serum sex hormone levels and myocardial infarction in the Honolulu Heart Program. Pitfalls in prospective studies on sex hormones. *J Clin Epidemiol* 1988;**41:**1151–1156

Phillips GB, Castelli WP, Abbott RD, McNamara PM. Association of hyperestrogenemia and coronary heart disease in men in the Framingham cohort. *Am J Med* 1983;**74:**863–869

Phillips GB, Pinkernell BH, Jing TY. The association of hypotestosteronemia with coronary artery disease in men. *Arterioscler Thromb Vasc Biol* 1994;**14:**701–706

Piedrahita JA, Zhang SH, Hagaman JR, Oliver PM, Maeda N. Generation of mice carrying a mutant apolipoprotein E gene inactivated by gene targeting in embryonic stem cells. *Proc Natl Acad Sci USA* 1992;**89:**4471-4475

Pietrobelli A, Faith MS, Allison DB, Gallagher D, Chiumello G, Heymsfield SB. Body mass index as a measure of adiposity among children and adolescents: a validation study. *J Pediatr* 1998;**132:**204–210

Pleumeekers HJ, Hoes AW, van der Does E, van Urk H, Hofman A, de Jong PT, Grobbee DE. Aneurysms of the abdominal aorta in olderadults. The Rotterdam Study. *Am J Epidemiol* 1995;**142:**1291-1299

Pola R, Gaetani E, Santoliquido A, Gerardino L, Cattani P, Serricchio M, Tondi P, Flore R, Grande M, Carbonin P, Fadda G, Pola P. Abdominal aortic aneurysm in normotensive patients: association with angiotensin-converting enzyme gene polymorphism. *Eur J Vasc Endovasc Surg* 2001;**21:**445-449

Popovici V, Goldstein DR, Antonov J, Jaggi R, Delorenzi M, Wirapati P. Selecting control genes for RT-QPCR using public microarray data. *BMC Bioinformatics* 2009;**10:**42

Powell JT, Sweeting MJ, Brown LC, Gotensparre SM, Fowkes FG, Thompson SG. Systematic review and meta-analysis of growth rates of small abdominalaortic aneurysms. *Br J Surg* 2011;**98:**609-618

Prescott E, Hippe M, Schnohr P, Hein HO, Vestbo J. Smoking and risk of myocardial infarction in women and men: longitudinal population study. *BMJ* 1998;**316:**1043-1047

Pritchard JK, Stephens M, Donnelly P. Inference of population structure using multilocus genotype data.*Genetics* 2000;**155:**945-959

Puech-Leão P, Molnar LJ, Oliveira IR, Cerri GG. Prevalence of abdominal aortic aneurysms – a screening program in São Paulo, Brazil. *Sao Paulo Med J* 2004;**122:**158-160

Railo M, Isoluoma M, Keto P, Salo JA. Surgical treatment of inflammatory abdominal aortic aneurysms: a long-term follow-up of 19 patients. *Ann Vasc Surg* 2005;**19:**361-366

R Development Core Team (2010). R: A language and environment for statistical computing.

Rahimi K, Maerz HK, Zotz RJ, Tárnok A. Pre-procedural expression of Mac-1 and LFA-1 on leukocytes for prediction of late restenosis and their possible correlation with advanced coronary artery disease. *Cytometry B Clin Cytom* 2003;**53:**63-69

Rasmussen TE, Hallett JW Jr. New insights into inflammatory abdominal aortic aneurysms. *Eur J Vasc Endovasc Surg* 1997;**14:**329-332

Rasmussen TE, Hallett JW Jr, Schulte S, Harmsen WS, O'Fallon WM, Weyand CM. Genetic similarity in inflammatory and degenerative abdominal aortic aneurysms: a study of human leukocyte antigen class II disease risk genes. *J Vasc Surg* 2001;**34:**84-89

Reiser S. Effect of dietary sugar on metabolic risk factors associated with heart disease. *Nutr Health* 1985;**3:**203-216

Reitsma JB, Pleumeekers HJ, Hoes AW, Kleijnen J, de Groot RM, Jacobs MJ, Grobbee DE, Tijssen JG. Increasing Incidence of aneurysms of the abdominal aorta in the Netherlands. *J Vasc Endovasc Surg* 1996;**12:**446-451

Rhee SH & Waldman ID. Genetic and environmental influences on antisocial behaviour: a meta-analysis of twin and adoption studies. *Psychol Bull* 2002;**128:**490-529

Rice T, Sprecher DL, Borecki IB, Mitchell LE, Laskarzewski PM, Rao DC. Cincinnati myocardial infarction and hormone family study: family resemblance for testosterone in random and MI families. *Am J Med Genet* 1993;**47:**542–549

Ridker PM. High-sensitivity C-reactive protein, inflammation, and cardiovascular risk: from concept to clinical practice to clinical benefit. *AM Heart J* 2004;**148;**19-26

Roger VL, Go AS, Lloyd-Jones DM, Adams RJ, Berry JD, Brown TM, Carnethon MR, Dai S, de Simone G, Ford ES, Fox CS, Fullerton HJ, Gillespie C, Greenlund KJ, Hailpern SM, Heit JA, Ho PM, Howard VJ, Kissela BM, Kittner SJ, Lackland DT, Lichtman JH, Lisabeth LD, Makuc DM, Marcus GM, Marelli A, Matchar DB, McDermott MM, Meigs JB, Moy CS, Mozaffarian D, Mussolino ME, Nichol G, Paynter NP, Rosamond WD, Sorlie PD, Stafford RS, Turan TN, Turner MB, Wong ND, Wylie-Rosett J. Heart disease and stroke statistics--2012 update: a report from the American Heart Association. *Circulation* 2012;**125:**e2-e220

Rootsi S, Magri C, Kivisild T, Benuzzi G, Help H, Bermisheva M, Kutuev I, Barać L, Pericić M, Balanovsky O, Pshenichnov A, Dion D, Grobei M, Zhivotovsky LA, Battaglia V, Achilli A, Al-Zahery N, Parik J, King R, Cinnioğlu C, Khusnutdinova E, Rudan P, Balanovska E, Scheffrahn W, Simonescu M, Brehm A, Goncalves R, Rosa A, Moisan JP, Chaventre A, Ferak V, Füredi S, Oefner PJ, Shen P, Beckman L, Mikerezi I, Terzić R, Primorac D, Cambon-Thomsen A, Krumina A, Torroni A, Underhill PA, Santachiara-Benerecetti AS, Villems R, Semino O. Phylogeography of Y-chromosome haplogroup I reveals distinct domains of prehistoric gene flow in europe. *Am J Hum Genet* 2004;**75:**128-137

Rosser ZH, Balaresque P, Jobling MA. Gene conversion between the X chromosome and the male-specific region of the Y chromosome at a translocation hotspot. *Am J Hum Genet* 2009;**85:**130-134

Rotival M, Zeller T, Wild PS, Maouche S, Szymczak S, Schillert A, Castagné R, Deiseroth A, Proust C, Brocheton J, Godefroy T, Perret C, Germain M, Eleftheriadis M, Sinning CR, Schnabel RB, Lubos E, Lackner KJ, Rossmann H, Münzel T, Rendon A, Cardiogenics Consortium, Erdmann J, Deloukas P, Hengstenberg C, Diemert P, Montalescot G, Ouwehand WH, Samani NJ, Schunkert H, Tregouet DA, Ziegler A, Goodall AH, Cambien F, Tiret L, Blankenberg S. Integrating genome-wide genetic variations and monocyte expression data reveals trans-regulated gene modules in humans. *PLoS Genet* 2011;**7:**e1002367

Rozanski A, Blumenthal JA, Kaplan J. Impact of psychological factors on the pathogenesis of cardiovascular disease and implications for therapy. *Circulation* 1999;**99:**2192-2217

Rozen S, Marszalek JD, Alagappan RK, Skaletsky H, Page DC. Remarkably little variation in proteins encoded by the Y chromosome's single-copy genes, implying effective purifying selection. *Am J Hum Genet* 2009;**85:**923-928

Russo P, Venezia A, Lauria F, Strazzullo P, Cappuccio FP, Iacoviello L, Barba G, Siani A. *Hin*dIII polymorphism of the Y chromosome, blood pressure, and serum lipids: no evidence of association in three white populations. *Epidemiology* 2006;**19:**331-338

Russo P, Siani A, Miller MA, Karanam S, Esposito T, Gianfrancesco F, Barba G, Lauria F, Strazzullo P, Cappuccio FP. Genetic variants of Y chromosome are associated with a protective lipid profile in blackmen. *Arterioscler Thromb Vasc Biol* 2008;**28:**1569-1574

Ryter SW, Lee SJ, Smith A, Choi AM. Autophagy in vascular disease. *Proc Am Thorac Soc* 2010;**7:**40-47

Sakamoto H & Kurabayashi M. Cardiovascular effects of an immunosuppressive agent cyclosporin A. *Int J Immunopathol Pharmacol* 2003;**15:**75-79

Salem MK, Rayt HS, Hussey G, Rafelt S, Nelson CP, Sayers RD, Naylor AR, Nasim A. Should Asian men be included in abdominal aortic aneurysm screening programmes? *Eur J Vasc Endovasc Surg* 2009;**38:**748-749

Salido EC, Yen PH, Koprivnikar K, Yu LC, Shapiro LJ. The human enamel protein gene amelogenin is expressed from both the X and the Y chromosomes. *Am J Hum Genet* 1992;**50:**303-316

Samani NJ, Burton P, Mangino M, Ball SG, Balmforth AJ, Barrett J, Bishop T, Hall A, BHF Family Heart Study Research Group. A genomewide linkage study of 1,933 families affected by premature coronary artery disease: The British Heart Foundation (BHF) Family Heart Study. *Am J Hum Genet* 2005;**77:**1011-1020

Samani NJ, Erdmann J, Hall AS, Hengstenberg C, Mangino M, Mayer B, Dixon RJ, Meitinger T, Braund P, Wichmann HE, Barrett JH, König IR, Stevens SE, Szymczak S, Tregouet DA, Iles MM, Pahlke F, Pollard H, Lieb W, Cambien F, Fischer M, Ouwehand W, Blankenberg S, Balmforth AJ, Baessler A, Ball SG, Strom TM, Braenne I, Gieger C, Deloukas P, Tobin MD, Ziegler A, Thompson JR, Schunkert H, WTCCC and the Cardiogenics Consortium. Genome-wide association analysis of coronary artery disease. *N Engl J Med* 2007;**357:**443-453

Samani W & Shively V. Abdominal aortic aneurysm as a complex multifactorial disease interactions of polymorphisms of inflammatory genes, features of autoimmunity, and current status of MMPs. *Ann N Y Acad Sci* 2006;**1085:**117-132

Scarborough P, Wickramasinghe K, Bhatnagar P, Rayner M. Trends in coronary heart disease 1961-2011. British Heart Foundation: London 2011.

Schena M, Shalon D, Davis RW, Brown PO. Quantitative monitoring of gene expression patterns with a complementary DNA microarray. *Science* 1995;**270:**467-470

Schiebel K, Winkelmann M, Mertz A, Xu X, Page DC, Weil D, Petit C, Rappold GA. Abnormal XY interchange between a novel isolated protein kinase gene, PRKY, and its homologue, PRKX, accounts for one third of all (Y+)XX males and (Y-)XY females. *Hum Mol Genet* 1997;**6:**1985-1989

Schuler-Lüttmann S,Mönnig G, Enbergs A, Schulte H, Breithardt G, Assmann G, Kerber S, von Eckardstein A. Insulin-like growth factor binding protein-3 is associated with the presence and extent of coronary arteriosclerosis. *Arterioscler Thromb Vasc Biol* 2000;**20:**e10–e15

Schunkert H, Götz A, Braund P, McGinnis R, Tregouet DA, Mangino M, Linsel-Nitschke P, Cambien F, Hengstenberg C, Stark K, Blankenberg S, Tiret L, Ducimetiere P, Keniry A, Ghori MJ, Schreiber S, El Mokhtari NE, Hall AS, Dixon RJ, Goodall AH, Liptau H, Pollard H, Schwarz DF, Hothorn LA, Wichmann HE, König IR, Fischer M, Meisinger C, Ouwehand W, Deloukas P, Thompson JR, Erdmann J, Ziegler A, Samani NJ; Cardiogenics Consortium. Repeated replication and a prospective meta-analysis of the association between chromosome 9p21.3 and coronary artery disease. *Circulation* 2008;**117:**1675-1684.

Schunkert H, König IR, Kathiresan S, Reilly MP, Assimes TL, Holm H, Preuss M, Stewart AF, Barbalic M, Gieger C, Absher D, Aherrahrou Z, Allayee H, Altshuler D, Anand SS, Andersen K, Anderson JL, Ardissino D, Ball SG, Balmforth AJ, Barnes TA, Becker DM, Becker LC, Berger K, Bis JC, Boekholdt SM, Boerwinkle E, Braund PS, Brown MJ, Burnett MS, Buysschaert I; Cardiogenics, Carlquist JF, Chen L, Cichon S, Codd V, Davies RW, Dedoussis G, Dehghan A, Demissie S, Devaney JM, Diemert P, Do R, Doering A, Eifert S, Mokhtari NE, Ellis SG, Elosua R, Engert JC, Epstein SE, de Faire U, Fischer M, Folsom AR, Freyer J, Gigante B, Girelli D, Gretarsdottir S, Gudnason V, Gulcher JR, Halperin E, Hammond N, Hazen SL, Hofman A, Horne BD, Illig T, Iribarren C, Jones GT, Jukema JW, Kaiser MA, Kaplan LM, Kastelein JJ, Khaw KT, Knowles JW, Kolovou G, Kong A, Laaksonen R, Lambrechts D, Leander K, Lettre G, Li M, Lieb W, Loley C, Lotery AJ, Mannucci PM, Maouche S, Martinelli N, McKeown PP, Meisinger C, Meitinger T, Melander O, Merlini PA, Mooser V, Morgan T, Mühleisen TW, Muhlestein JB, Münzel T, Musunuru K, Nahrstaedt J, Nelson CP, Nöthen MM, Olivieri O, Patel RS, Patterson CC, Peters A, Peyvandi F, Qu L, Quyyumi AA, Rader DJ, Rallidis LS, Rice C, Rosendaal FR, Rubin D, Salomaa V, Sampietro ML, Sandhu MS, Schadt E, Schäfer A, Schillert A, Schreiber S, Schrezenmeir J, Schwartz SM, Siscovick DS, Sivananthan M, Sivapalaratnam S, Smith A, Smith TB, Snoep JD, Soranzo N, Spertus JA, Stark K, Stirrups K, Stoll M, Tang WH, Tennstedt S, Thorgeirsson G, Thorleifsson G, Tomaszewski M, Uitterlinden AG, van Rij AM, Voight BF, Wareham NJ, Wells GA, Wichmann HE, Wild PS, Willenborg C, Witteman JC, Wright BJ, Ye S, Zeller T, Ziegler A, Cambien F, Goodall AH, Cupples LA, Quertermous T, März W, Hengstenberg C, Blankenberg S, Ouwehand WH, Hall AS, Deloukas P, Thompson JR, Stefansson K, Roberts R, Thorsteinsdottir U, O'Donnell CJ, McPherson R, Erdmann J; CARDIoGRAM Consortium, Samani NJ. Large-scale

association analysis identifies 13 new susceptibility loci for coronary artery disease. *Nat Genet* 2011;**43:**333-338

Scott RA, Wilson NM, Ashton HA, Kay DN. Influence of screening on the incidence of ruptured abdominal aortic aneurysm: 5-year results of a randomized controlled study. *Br J Surg* 1995;**82:**1066-70.

Semino O, Passarino G, Oefner PJ, Lin AA, Arbuzova S, Beckman LE, De Benedictis G, Francalacci P, Kouvatsi A, Limborska S, Marcikiae M, Mika A, Mika B, Primorac D, Santachiara-Benerecetti AS, Cavalli-Sforza LL, Underhill PA. The genetic legacy of Paleolithic Homo sapiens sapiens in extant Europeans: a Y chromosome perspective. *Science* 2000;**290:**1155-1159

Serajee FJ & Mahbubul Huq AH. Association of Y chromosome haplotypes with autism. *J Child Neurol* 2009;**24:**1258-1261

Sesso HD, Paffenbarger RS Jr, Lee IM. Physical activity and coronary heart disease in men: The Harvard Alumni Health Study. *Circulation* 2000;**102:**975-980

Sewdarsen M, Jialal I, Vythilingum S, Desai R. Sex hormone levels in young Indian patients with myocardial infarction. *Arteriosclerosis* 1986;**6:**418–421

Sewdarsen M, Jialal I, Naidu RK. The low plasma testosterone levels of young Indian infarct survivors are not due to a primary testicular defect. *Postgrad Med J* 1988;**64:**264–266

Sewdarsen M, Vythilingum S, Jialal I, Desai RK, Becker P. Abnormalities in sex hormones are a risk factor for premature manifestation of coronary artery disease in South African Indian men. *Atherosclerosis* 1990;**83:**111–117

Sezgin E, Lind JM, Shrestha S, Hendrickson S, Goedert JJ, Donfield S, Kirk GD, Phair JP, Troyer JL, O'Brien SJ, Smith MW. Association of Y chromosome haplogroup I with HIV progression, and HAART outcome. *Hum Genet* 2009;**125:**281-294

Shadwick R. Mechanical design in arteries. *J Exp Bio* 1999;**202:**3305-3313

Shah SS, Ayub Q, Firasat S, Kaiser F, Mehdi SQ. Y haplogroups and aggressive behavior in a Pakistani ethnic group. *Aggress Behav* 2009;**35:**68-74

Shankar RR, Charchar FJ, Eckert GJ, Saha C, Tu W, Dominiczak AF, Pratt JH. Studies of an association in boys of blood pressure and the Y chromosome. *Am J Hypertens* 2007;**20:**27-31

Shapira OM, Pasik S, Wassermann JP, Barzilai N, Mashiah A. Ultrasound screening for abdominal aortic aneurysms in patients with atherosclerotic peripheral vascular disease. *J Cardiovasc Surg* (*Torino*)1990;**31:**170-172

Sharrett AR, Ballantyne CM, Coady SA, Heiss G, Sorlie PD, Catellier D, Patsch W, Atherosclerosis Risk in Communities Study Group. Coronary heart disease prediction from lipoprotein cholesterol levels, triglycerides, lipoprotein(a), apolipoproteins A-I and

B, and HDL density subfractions: The Atherosclerosis Risk in Communities (ARIC) Study. *Circulation* 2001;**104:**1108-1113

Shen GQ, Li L, Rao S, Abdullah KG, Ban JM, Lee BS, Park JE, Wang QK. Four SNPs on chromosome 9p21 in a South Korean population implicate a genetic locus that confers high cross-race risk for development of coronary artery disease. *Arterioscler Thromb Vasc Biol* 2008;**28:**360-365

Sillanpää MJ. Overview of techniques to account for confounding due to population stratification and cryptic relatedness in genomic data association analyses. *Heredity* 2011;**106:**511-519

Simoni G, Pastorino C, Perrone R, Ardia A, Gianrossi R, Decian F, Cittadini G Jr, Baiardi A, Bachi V. Screening for abdominal aortic aneurysms and associated risk factors in a general population. *Eur J Vasc Endovasc Surg* 1995;**10**:207-10.

Simpson E, Scott D, Chandler P. The male-specific histocompatibility antigen, H-Y: a history of transplantation, immune response genes, sex determination and expression cloning. *Annu Rev Immunol* 1997;**15:**39-61

Sims LM, Garvey D, Ballantyne J. Improved resolution haplogroup G phylogeny in the Y chromosome, revealed by a set of newly characterized SNPs. *PLoS One* 2009;**4:**5729-5792

Singh K, Bønaa KH, Jacobsen BK, Bjørk L, Solberg S. Prevalence of and risk factors for abdominal aortic aneurysms in a population-based study. *Am J Epidemiol* 2001;**154:**236-244

Sinha I, Cho BS, Roelofs KJ, Stanley JC, Henke PK, Upchurch GR Jr. Female gender attenuates cytokine and chemokine expression and leukocyte recruitment in experimental rodent abdominal aortic aneurysms. *Ann N Y Acad Sci* 2006;**1085:**367-379

Sjödin P & François O. Wave-of-advance models of the diffusion of the Y chromosome haplogroup R1b1b2 in Europe. *PLoS One* 2011;**6:**e21512

Skaletsky H, Kuroda-Kawaguchi T, Minx PJ, Cordum HS, Hillier L, Brown LG, Repping S, Pyntikova T, Ali J, Bieri T, Chinwalla A, Delehaunty A, Delehaunty K, Du H, Fewell G, Fulton L, Fulton R, Graves T, Hou SF, Latrielle P, Leonard S, Mardis E, Maupin R, McPherson J, Miner T, Nash W, Nguyen C, Ozersky P, Pepin K, Rock S, Rohlfing T, Scott K, Schultz B, Strong C, Tin-Wollam A, Yang SP, Waterston RH, Wilson RK, Rozen S, Page DC. The male-specific region of the human Y chromosome is a mosaic of discrete sequence classes. *Nature* 2003;**19:**825-837

Slowinska-Srzednicka J, Zgliczynski S, Ciswicka-Sznajderman M, Scredicki M, Soszynski P, Biernacka M, Woroszylska M, Ruzyllo W, Sadowski Z. Decreased plasma dehydroepiandrosterone sulfate and dihydrotestosterone concentrations in young men after myocardial infarction. *Atherosclerosis* 1989;**79:**197–203

Small M, Lowe GDO, Beastall GH, Beattie JM, McEachern M, Hutton I, Lorimar AR, Forbes CD. Serum oestradiol and ischaemic heart disease relationship with myocardial infarction but not coronary atheroma or haemostasis. *Q J Med* 1985;**57:**775–782

Smith TW & Ruiz JM. Psychosocial influences on the development and course of coronary heart disease: current status and implications for research and practice. *J Consult Clin Psychol* 2002;**70:**548-68

Soler EP & Ruiz VS. Epidemiology and risk factors of cerebral ischemia and ischemic heart diseases: Similarities and differences. *Curr Cardiol Rev* 2010;**6:**138-149

Spach KM, Blake M, Bunn JY, McElvany B, Noubade R, Blankenhorn EP, Teuscher C. Cutting edge: the Y chromosome controls the age-dependent experimental allergic encephalomyelitis sexual dimorphism in SJL/J mice. *J Immunol*. 2009;**182:**1789–93

Spark JI, Baker JL, Vowden P, Wilkinson D. Epidemiology of abdominal aortic aneurysms in the Asian community. *Br J Surg* 2001;**88:**382-384

Sprague AH & Khalil RA. Inflammatory cytokines in vascular dysfunction and vascular disease. *Biochem Pharmacol* 2009;**78:**539-552

Stemme S, Faber B, Holm J, Wiklund O, Witztum JL, Hansson GK. T lymphocytes from human atherosclerotic plaques recognize oxidized low density lipoprotein. *Proc Natl Acad Sci USA* 1995;**92:**3893-3897

Stouffs K, Lissens W, Tournaye H, Haentjens P. What about gr/gr deletions and male infertility? Systematic review and meta-analysis. *Hum Reprod Update* 2011;**17:**197-209

Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette MA, Paulovich A, Pomeroy SL, Golub TR, Lander ES, Mesirov JP. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc Natl Acad Sci USA* 2005;**102**:15545-15550

Sugimoto T, Sada M, Miyamoto T, Yao H. Genetic analysis on HLA loci in Japanese patients with abdominal aortic aneurysm. *Eur J Vasc Endovasc Surg* 2003;**26:**215-218

Suls J & Bunde J. Anger, anxiety, and depression as risk factors for cardiovascular disease: the problems and implications of overlapping affective dispositions. *Psychol Bull* 2005;**131:**260-300

Sun Y, Tawara I, Toubai T, Reddy P. Pathophysiology of acute graft-versus-host disease: recent advances. *Transl Res* 2007;**150:**197-214

Swanberg M, Lidman O, Padyukov L, Eriksson P, Akesson E, Jagodic M, Lobell A, Khademi M, Börjesson O, Lindgren CM, Lundman P, Brookes AJ, Kere J, Luthman H, Alfredsson L, Hillert J, Klareskog L, Hamsten A, Piehl F, Olsson T. MHC2TA is associated with differential MHC molecule expression and susceptibility to rheumatoid arthritis, multiple sclerosis and myocardial infarction. *Nat Genet* 2005;**37:**486-494

Swartz CM & Young MA. Low serum testosterone and myocardial infarction in geriatric male inpatients. *J Am Geriatr Soc* 1987;**35:**39–44

Szegedi K, Sonkoly E, Nagy N, Németh IB, Bata-Csörgo Z, Kemény L, Dobozy A, Széll M. The anti-apoptotic protein G1P3 is overexpressed in psoriasis and regulated by the non-coding RNA, PRINS. *Exp Dermatol* 2010;**19:**269-278

Tahara E Jr, Tahara H, Kanno M, Naka K, Takeda Y, Matsuzaki T, Yamazaki R, Ishihara H, Yasui W, Barrett JC, Ide T, Tahara E. G1P3, an interferon inducible gene 6-16, is expressed in gastric cancers and inhibits mitochondrial-mediated apoptosis in gastric cancer cell line TMK-1 cell. *Cancer Immunol Immunother* 2005;**54:**729-740

Takagi H, Manabe H, Kawai N, Goto SN, Umemoto T. Serum high-density and low-density lipoprotein cholesterol is associated with abdominal aortic aneurysm presence: a systematic review and meta-analysis. *Int Angiol* 2010;**29:**371-375

Tang H, Zhong Y, Zhu Y, Zhao F, Cui X, Wang Z. Low responder T cell susceptibility to the suppressive function of regulatory T cells in patients with dilated cardiomyopathy. *Heart* 2010;**96:**765-771

Thomas DC, Witte JS. Point: population stratification: a problem for case-control studies of candidate-gene associations? *Cancer Epidemiol Biomarkers Prev* 2002;**11:**505-512

TaqMan® Chemistry vs. SYBR® Chemistry for Real-Time PCR

Taylor D, Nagle N, Ballantyne KN, van Oorschot RA, Wilcox S, Henry J, Turakulov R, Mitchell RJ. An investigation of admixture in an Australian Aboriginal Y-chromosome STR database. *Forensic Sci Int Genet* 2012;**6:**532-538

Tedgui A & Mallat Z. Cytokines in atherosclerosis: pathogenic and regulatory pathways. *Physiol Rev* 2006;**86:**515-581

Tentolouris C, Tousoulis D, Antoniades C, Bosinakou E, Kotsopoulou M, Trikas A, Toutouzas P, Stefanadis C. Endothelial function and proinflammatory cytokines in patients with ischemic heart disease and dilated cardiomyopathy. *Int J Cardiol* 2004;**94:**301-305

Teo KK, Ounpuu S, Hawken S, Pandey MR, Valentin V, Hunt D, Diaz R, Rashed W, Freeman R, Jiang L, Zhang X, Yusuf S, INTERHEART Study Investigators. Tobacco use and risk of myocardial infarction in 52 countries in the INTERHEART study: a case-control study. *Lancet* 2006;**368**:647-658

Teo YY, Inouye M, Small KS, Gwilliam R, Deloukas P, Kwiatkowski DP, Clark TG. A genotype calling algorithm for the Illumina BeadArray platform. *Bioinformatics* 2007;**23:**2741-2746

Teupser D, Baber R, Ceglarek U, Scholz M, Illig T, Gieger C, Holdt LM, Leichtle A, Greiser KH, Huster D, Linsel-Nitschke P, Schäfer A, Braund PS, Tiret L, Stark K, Raaz-Schrauder D, Fiedler GM, Wilfert W, Beutner F, Gielen S, Grosshennig A, König IR, Lichtner P, Heid IM, Kluttig A, El Mokhtari NE, Rubin D, Ekici AB, Reis A, Garlichs CD, Hall AS, Matthes G, Wittekind C, Hengstenberg C, Cambien F, Schreiber S, Werdan K, Meitinger T, Loeffler M, Samani NJ, Erdmann J, Wichmann HE, Schunkert H, Thiery J. Genetic regulation of serum phytosterol levels and risk of coronary artery disease. *Circ Cardiovasc Genet* 2010;**3**:331-9.

Teuscher C, Noubade R, Spach K, McElvany B, Bunn JY, Fillmore PD, Fillmore PD, Zachary JF, Blankenhorn EP. Evidence that the Y chromosome influences autoimmune disease in male and female mice. *Proc Natl Acad Sci USA*. 2006; **103:**8024–9

Thomas DC & Witte JS. Point: population stratification: a problem for case-control studies of candidate-gene associations? *Cancer Epidemiol Biomarkers Prev* 2002;**11:**505-512

Thompson AR, Rodway A, Mitchell A, Hafez H. Screening and ultrasound surveillance of large abdominal aortic aneurysms does not improve suitability for endovascular repair. *J Vasc Surg* 2006;**43:**265-269

Thompson AR, Drenos F, Hafez H, Humphries SE. Candidate gene association studies in abdominal aortic aneurysm disease: A review and meta-analysis. *Eur J Vasc Endovasc Surg* 2009;**35:**19-30

Thuny F, Textoris J, Amara AB, Filali AE, Capo C, Habib G, Raoult D, Mege JL. The gene expression analysis of blood reveals S100A11 and AQP9 as potential biomarkers of infective endocarditis. *PLoS One* 2012;**7:**e31490

Tivesten A, Hulthe J, Wallenfeldt K, Wikstrand J, Ohlsson C, Fagerberg B. Circulating oestradiol is an independent predictor of progression of carotid artery intima-media thickness in middle-aged men. *J Clin Endocrinol Metab* 2006;**91:**4433-4437

Tobin MD, Tomaszewski M, Braund PS, Hajat C, Raleigh SM, Palmer TM, Caulfield M, Burton PR, Samani NJ. Common variants in genes underlying monogenic hypertension and hypotension and blood pressure in the general population. *Hypertension* 2008;**51:**1658-1664

Tolbert T & Oparil S. Cardiovascular effects of estrogen. *Am J Hypertens* 2001;**14:**186-193

Tollerud DJ, Ildstad ST, Brown LM, Clark JW, Blattner WA, Mann DL, Neuland CY, Pankiw-Trost L, Hoover RN. T-cell subsets in healthy teenagers: transition to the adult phenotype. *Clin Immunol Immunopathol* 1990; **56:**88-96

Tomaszewski M, Zimmerli L, Charchar FJ, Dominiczak AF. Genetic information in the diagnosis and treatment of hypertension. *Curr Hypertens Rep* 2006;**8:**309-316

Tomaszewski M, Charchar FJ, Lynch MD, Padmanabhan S, Wang WY, Miller WH, Grzeszczak W, Maric C, Zukowska-Szczechowska E, Dominiczak AF. Fibroblast growth factor 1 gene and hypertension: from the quantitative trait locus to positional analysis. *Circulation* 2007;**116:**1915-1924

Tomaszewski M, Charchar FJ, Maric C, McClure J, Crawford L, Grzeszczak W, Sattar N, Zukowska-Szczechowska E, Dominiczak AF. Glomerular hyperfiltration: a new marker of metabolic risk. *Kidney Int* 2007;**71:**816-821

Tomaszewski M, Charchar FJ, Maric C, Kuzniewicz R, Gola M, Grzeszczak W, Samani NJ, Zukowska-Szczechowska E. Association between lipid profile and circulating concentrations of estrogens in young men. *Atherosclerosis* 2008;**203:**257-262

Tomaszewski M, Charchar FJ, Barnes T, Gawron-Kiszka M, Sedkowska A, Podolecka E, Kowalczyk J, Rathbone W, Kalarus Z, Grzeszczak W, Goodall AH, Samani NJ, Zukowska-Szczechowska E. A common variant in low-density lipoprotein receptor-related protein 6 gene (LRP6) is associated with LDL-cholesterol. *Arterioscler Thromb Vasc Biol* 2009;**29:**1316-1321

Tomaszewski M, Charchar FJ, Maric C, Kuzniewicz R, Gola M, Grzeszczak W, Samani NJ, Zukowska-Szczechowska E. Inverse associations between androgens and renal function: the Young Men Cardiovascular Association (YMCA) study. *Am J Hypertens* 2009;**22:**100-105

Tomaszewski M, Debiec R, Braund PS, Nelson CP, Hardwick R, Christofidou P, Denniff M, Codd V, Rafelt S, van der Harst P, Waterworth D, Song K, Vollenweider P, Waeber G, Zukowska-Szczechowska E, Burton PR, Mooser V, Charchar FJ, Thompson JR, Tobin MD, Samani NJ. Genetic architecture of ambulatory blood pressure in the general population: insights from cardiovascular gene-centric array. *Hypertension* 2010;**56:**1069-1076

Tomaszewski M, Charchar FJ, Nelson CP, Barnes T, Denniff M, Kaiser M, Debiec R, Christofidou P, Rafelt S, van der Harst P, Wang WY, Maric C, Zukowska-Szczechowska E, Samani NJ. Pathway analysis shows association between FGFBP1 and hypertension. *JASN* 2011;**22:**947-955

Toot J, Dunphy G, Turner M, Ely D. The SHR Y-chromosome increases testosterone and aggression, but decreases serotonin as compared to the WKY Y-chromosome in the rat model. *Behav Genet* 2004;**34:**515-524

Totonchi M, Mohseni Meybodi A, Borjian Boroujeni P, Sedighi Gilani M, Almadani N, Gourabi H. Clinical data for 185 infertile Iranian men with Y-chromosome microdeletion. *J Assist Reprod Genet* 2012;**29:**847-853.

Trigona B, Aggoun Y, Maggio A, Martic XE, Marchand LM, Beghetti M, Farpour-Lambert NJ. Preclinical non-invasive markers of atherosclerosis in children and adolescents with type 1 diabetes are influenced by physical activity. *J Pediatr* 2010;**157:**533-539

Tsorbatzoudis H. Psychometric evaluation of the Greek version of the Aggression Questionnaire. *Percept Mot Skills* 2006;**102:**703-712

UK small aneurysm trial. The U.K. Small Aneurysm Trial: design, methods and progress. The UK Small Aneurysm Trial participants. *Eur J Vasc Endovasc Surg* 1995;**9:**42-48

Underhill PA&Kivisild T. Use of Y chromosome and mitochondrial DNA population structure in tracing Human migrations. *Annu Rev Genet*. 2007;**41:**539-564

Underhill PA, Myres NM, Rootsi S, Metspalu M, Zhivotovsky LA, King RJ, Lin AA, Chow CE, Semino O, Battaglia V, Kutuev I, Järve M, Chaubey G, Ayub Q, Mohyuddin A, Mehdi SQ, Sengupta S, Rogaev EI, Khusnutdinova EK, Pshenichnov A, Balanovsky O, Balanovska E, Jeran N, Augustin DH, Baldovic M, Herrera RJ, Thangaraj K, Singh V, Singh L, Majumder P, Rudan P, Primorac D, Villems R, Kivisild T. Separating the post-glacial coancestry of European and Asian Y chromosomes within haplogroup R1a. *Eur J Hum Genet* 2010;**18:**479-484

Van Gelder RN, von Zastrow ME, Yool A, Dement WC, Barchas JD, Eberwine JH. Amplified RNA synthesized from limited quantities of heterogeneous cDNA. *Proc Natl Acad Sci USA* 1990;**87:**1663-1667

Van Gils JM, Zwaginga JJ, Hordijk PL. Molecular and functional interactions among monocytes, platelets, and endothelial cells and their relevance for cardiovascular diseases. *J Leukoc Biol* 2009;**85:**195–204

Van Heerebeek L, Hamdani N, Falcão-Pires I, Leite-Moreira AF, Begieneman MP, Bronzwaer JG, van der Velden J, Stienen GJ, Laarman GJ, Somsen A, Verheugt FW, Niessen HW, Paulus WJ. Low myocardial protein kinase g activity in heart failure with preserved ejection fraction. *Circulation* 2012;**126:**830-839

Van Vlijmen-van Keulen CJ, Pals G, Rauwerda JA. Familial abdominal aortic aneurysm: a systematic review of a genetic background. *J Endovasc Surg* 2002;**24:**105-116

Van Vlijmen-Van Keulen CJ, Rauwerda JA, Pals G. Genome-wide linkage in three Dutch families maps a locus for abdominal aortic aneurysms to chromosome 19q13.3. *Eur J Vasc Endovasc Surg* 2005;**30:**29-35

Vanbillemont G, Bogaert V, De Bacquer D, Lapauw B, Goemaere S, Toye K, Van Steen K, Taes Y, Kaufman JM. Polymorphisms of the SHBG gene contribute to the

inter-individual variation of sex steroid hormone blood levels in young, middle-aged and elderly men. *Clin Endocrinol* 2009;**70:**300-310

Vartanian K, Slottke R, Johnstone T, Casale A, Planck SR, Choi D, Smith JR, Rosenbaum JT, Harrington CA. Gene expression profiling of whole blood: comparison of target preparation methods for accurate and reproducible microarray analysis. *BCM Genomics* 2009;**10:**2

Velasco G, Savarese V, Sandorfi N, Jimenez SA, Jabbour S. 46, XX SRY-positive male syndrome presenting with primary hypogonadism in the setting of scleroderma. *Endocr Pract* 2011;**17:**95-98

Vernon SW & Roberts RE. Measuring nonspecific psychological distress and other dimensions of psychopathology. Further observations on the problem. *Arch Gen Psychiatry* 1981;**38:**1239-1247

Vogt MH, Goulmy E, Kloosterboer FM, Blokland E, de Paus RA, Willemze R, Falkenburg JH. UTY gene codes for an HLA-B60-restricted human male-specific minor histocompatibility antigen involved in stem cell graft rejection: characterization of the critical polymorphic amino acid residues for T-cell recognition. *Blood* 2000;**96:**3126-3132

Vogt PH. AZF deletions and Y chromosomal haplogroups: history and update based on sequence. *Hum Reprod Update* 2005;**11:**319-336

Voineagu I, Wang X, Johnston P, Lowe JK, Tian Y, Horvath S, Mill J, Cantor RM, Blencowe BJ, Geschwind DH. Transcriptomic analysis of autistic brain reveals convergent molecular pathology. *Nature* 2011;**474:**380-384

Von Boehmer H & Hafen K. Minor but not major histocompatibility antigens of thymus epithelium tolerize precursors of cytolytic T cells. *Nature* 1986;**320:**626-628

Wagenseil J. & Mecham R, Vascular extracellular matrix and arterial mechanisms. *Physiol Rev* 2009;**89:**957-989

Waldo SW, Li Y, Buono C, Zhao B, Billings EM, Chang J, Kruth HS. Heterogeneity of human macrophages in culture and in atherosclerotic plaques. *Am J Pathol* 2008;**172:**1112-1126

Walker DI, Bloor K, Williams G, Gillie I. Inflammatory aneurysms of the abdominal aorta. *Br J Surg* 1972;**59:**609-614

Wallace C, Rotival M, Cooper JD, Rice CM, Yang JH, McNeill M, Smyth DJ, Niblett D, Cambien F, Cardiogenics Consortium, Tiret L, Todd JA, Clayton DG, Blankenberg S. Statistical colocalization of monocyte gene expression and genetic risk variants for type 1 diabetes. *Hum Mol Genet* 2012;**21:**2815-282

Wang D, Yang H, Quiñones MJ, Bulnes-Enriquez I, Jimenez X, De La Rosa R, Modilevsky T, Yu K, Li Y, Taylor KD, Hsueh WA, Hodis HN, Rotter JI. A genome-wide

scan for carotid artery intima-media thickness: the Mexican-American Coronary Artery Disease family study. *Stroke* 2005;**36:**540-545

Wang K, Li M, Hakonarson H. Analysing biological pathways in genome-wide association studies. *Nat Rev Genet* 2010;**11:**843-854

Wanhainen A, Bergqvist D, Boman K, Nilsson TK, Rutegård J, Björck M. Risk factors associated with abdominal aortic aneurysm: A population based study with historical and current data. *J Vasc Surg* 2005;**41:**390-396

Wanhainen A. How to define an abdominal aortic aneurysm – influence on epidemiology and clinical practice. *Scand J Surg* 2008;**97:**105-109

Warren EH, Gavin MA, Simpson E, Chandler P, Page DC, Disteche C, Stankey KA, Greenberg PD, Riddell SR. The human UTY gene encodes a novel HLA-B8-restricted H-Y antigen. *J Immunol* 2000;**164:**2807-2814

Waters PD, Wallis MC, Marshall Graves JA. Mammalian sex-origin and evolution of the Y chromosome and *SRY*. *Semin Cell Dev Biol* 2007;**18:**389-400

Watkins H & Farrall M. Genetic susceptibility to coronary artery disease: from promise to progress. *Nat Rev Genet* 2006;**7:**163-173

Webb CM & Collins P. Testosterone and coronary artery disease in men. *Maturitas* 2010;**67:**15-19

Wei M, Mitchell BD, Haffner SM, Stern MP. Effects of cigarette smoking, diabetes, high cholesterol, and hypertension on all-cause mortality and cardiovascular disease mortality in Mexican Americans.The San Antonio Heart Study. *Am J Epidemiol* 1996;**144:**1058-1065

Wei W, Luo HB, Yan J, Hou YP. Exploring of new Y-chromosome SNP loci using Pyrosequencing and the SNaPshot methods. *Int J Legal Med* 2012;**126**:825-33.

Wells RS, Yuldasheva N, Ruzibakiev R, Underhill PA, Evseeva I, Blue-Smith J, Jin L, Su B, Pitchappan R, Shanmugalakshmi S, Balakrishnan K, Read M, Pearson NM, Zerjal T, Webster MT, Zholoshvili I, Jamarjashvili E, Gambarov S, Nikbin B, Dostiev A, Aknazarov O, Zalloua P, Tsoy I, Kitaev M, Mirrakhimov M, Chariev A, Bodmer WF. The Eurasian heartland: a continental perspective on Y-chromosome diversity. *Proc Natl Acad Sci USA* 2001;**98:**10244-10249

Whitcombe D, Brownie J, Gillard HL, McKechnie D, Theaker J, Newton CR, Little S. A homogeneous fluorescence assay for PCR amplicons: its application to real-time, single-tube genotyping. *Clin Chem* 1998;**44:**918-923

Whitworth JA, World Health Organization, International Society of Hypertension Writing Group. 2003 World Health Organization (WHO)/International Society of

Hypertension (ISH) statement on management of hypertension. *J Hypertens* 2003;**21:**1983-1992

Wilhelm D, Palmer S, Koopman P. Sex determination and gonadal development in mammals. *Physiol Rev* 2007;**87:**1-28

Willer CJ, Sanna S, Jackson AU, Scuteri A, Bonnycastle LL, Clarke R, Heath SC, Timpson NJ, Najjar SS, Stringham HM, Strait J, Duren WL, Maschio A, Busonero F, Mulas A, Albai G, Swift AJ, Morken MA, Narisu N, Bennett D, Parish S, Shen H, Galan P, Meneton P, Hercberg S, Zelenika D, Chen WM, Li Y, Scott LJ, Scheet PA, Sundvall J, Watanabe RM, Nagaraja R, Ebrahim S, Lawlor DA, Ben-Shlomo Y, Davey-Smith G, Shuldiner AR, Collins R, Bergman RN, Uda M, Tuomilehto J, Cao A, Collins FS, Lakatta E, Lathrop GM, Boehnke M, Schlessinger D, Mohlke KL, Abecasis GR. Newly identified loci that influence lipid concentrations and risk of coronary artery disease. *Nat Genet* 2008;**40**:161-9.

Williams RC, Hanson RL, Pettitt DJ, Sievers ML, Nelson RG, Knowler WC. HLA*A2 confers mortality risk for cardiovascular disease in Pimans. *Tissue Antigens* 1996;**47:**188-93

Wilmink AB, Quick CR. Epidemiology and potential for prevention of abdominal aortic aneurysm. *Br J Surg* 1998;**85:**155-162

Wilson SG, Adam G, Langdown M, Reneland R, Braun A, Andrew T, Surdulescu GL, Norberg M, Dudbridge F, Reed PW, Sambrook PN, Kleyn PW, Spector TD. Linkage and potential association of obesity-related phenotypes with two genes on chromosome 12q24 in a female dizygous twin cohort. *Eur J Hum Genet* 2006;**14:**340-348

Wong DR, Willett WC, Rimm EB. Smoking, hypertension, alcohol consumption, and risk of abdominal aortic aneurysm in men. *Am J Epidemiol* 2007:**165:**838-845

Woodward M, Lam TH, Barzi F, Patel A, Gu D, Rodgers A, Suh I, Asia Pacific Cohort Studies Collaboration. Smoking, quitting, and the risk of cardiovascular disease among women and men in the Asia-Pacific region. *Int J Epidemiol* 2005;**34:**1036-1045

Wu FC & von Eckardstein A. Androgens and coronary artery disease. *Endocr Rev* 2003;**24:**183-217

Wu G. Amino acids: metabolism, functions and nutrition. *Amino acids* 2009;**37:**1-17

Wu XF, Zhang J, Paskauskas S, Xin SJ, Duan ZQ. The role of estrogen in the formation of experimental abdominal aortic aneurysm. *Am J Surg* 2009;**197:**49-54

Xu S, Liu AC, Gotlieb AI. Common pathogenic features of atherosclerosis and calcific aortic stenosis: role of transforming growth factor-beta. *Cardiovasc Pathol* 2010;**19:**236-247

Yan ZQ & Hansson GK.Innate immunity, macrophage activation, and atherosclerosis. *Immunol Rev* 2007;**219:**187-203

Yang IA, Holloway JW, Ye S; Southampton Atherosclerosis Study (SAS) Group. TLR4 Asp299Gly polymorphism is not associated with coronary artery stenosis. *Atherosclerosis* 2003;**170:**187-190

Yang J, Cron P, Good VM, Thompson V, Hemmings BA, Barford D. Crystal structure of an activated Akt/protein kinase B ternary complex with GSK3-peptide and AMP-PNP. *Nat Struct Biol* 2002;**9:**940-944

Yeap BB, Hyde Z, Norman PE, Chubb SA, Golledge J. Associations of total testosterone, sex hormone-binding globulin, calculated free testosterone, and luteinizing hormone with prevalence of abdominal aortic aneurysm in older men. *J Clin Endocrinol Metab* 2010;**95:**1123-1130

Yii MK. Epidemiology of abdominal aortic aneurysm in an Asian population. *ANZ J Surg* 2003;**73:**393-3954

Yuan G, Wang J, Hegele RA. Heterozygous familial hypercholesterolemia: an under recognized cause of early cardiovascular disease. *CMAJ* 2006;**174:**1124-1129

Yusuf S, Hawken S, Ounpuu S, Dans T, Avezum A, Lanas F, McQueen M, Budaj A, Pais P, Varigos J, Lisheng L, INTERHEART Study Investigators. Effect of potentially modifiable risk factors associated with myocardial infarction in 52 countries (the INTERHEART study): case-control study. *Lancet* 2004; **364:** 937-952

Zandbergen F & Plutzky J. PPARalpha in atherosclerosis and inflammation. *Biochem Biophys Acta* 2007;**1771:**972-982

Zdravkovic S, Wienke A, Pedersen NL, Marenberg ME, Yashin AI, De Faire U. Heritability of death from coronary heart disease: a 36-year follow-up of 20 966 Swedish twins. *J Intern Med* 2002;**252:**247-254

Zhang JK, Zhao LF, Cheng J, Guo J, Lun YZ, Hong Y. Screening of genes for proteins interacting with the PS1TP5 protein of hepatitis B virus: probing a human leukocyte cDNA library using the yeast two-hybrid system. *Chin Med J* (*Engl*)2006;**119:**1884-1891

Zhao SP & Li XP. The association of low plasma testosterone level with coronary artery disease in Chinese men. *Int J Cardiol* 1998;**63:**161–164

Zhou L, Xi B, Wei Y, Shen W, Li Y. Meta-analysis of the association between the insertion/deletion polymorphism in ACE gene and coronary heart disease among the Chinese population. *J Renin Angiotensin Aldosterone Syst* 2012;**13:**296-304

Zhou X, Nicoletti A, Elhage R, Hansson GK. Transfer of CD4(+) T cells aggravates atherosclerosis in immunodeficient apolipoprotein E knockout mice. *Circulation* 2000;**102:**2919-2922

Zimmermann B, Chiorini JA, Ma Y, Kotin RM, Herberg FW. PrKX is a novel catalytic subunit of the cAMP-dependent protein kinase regulated by the regulatory subunit type I. *J Biol Chem* 1999;**274:**5370-5378

Zintzaras E, Raman G, Kitsios G, Lau J. Angiotensin-converting enzyme insertion/deletion gene polymorphic variant as a marker of coronary artery disease: a meta-analysis. *Arch Intern Med* 2008;**168:**1077-1089

Zumoff B, Troxler RG, O'Connor J, Rosenfeld RS, Kream J, Levin J, Hickman JR, Sloan AM, Walker W, Cook RL, Fukushima DK. Abnormal hormone levels in men with coronary artery disease. *Arteriosclerosis* 1982;**2:**58–67