

Inheritance of coronary artery disease in men: an analysis of the role of the Y chromosome



Fadi J Charchar, Lisa D S Bloomer, Timothy A Barnes, Mark J Cowley, Christopher P Nelson, Yanzhong Wang, Matthew Denniff, Radoslaw Debiec, Paraskevi Christofidou, Scott Nankervis, Anna F Dominiczak, Ahmed Bani-Mustafa, Anthony J Balmforth, Alistair S Hall, Jeanette Erdmann, Francois Cambien, Panos Deloukas, Christian Hengstenberg, Chris Packard, Heribert Schunkert, Willem H Ouweland, Ian Ford, Alison H Goodall, Mark A Jobling, Nilesh J Samani, Maciej Tomaszewski

Summary

Background A sexual dimorphism exists in the incidence and prevalence of coronary artery disease—men are more commonly affected than are age-matched women. We explored the role of the Y chromosome in coronary artery disease in the context of this sexual inequity.

Methods We genotyped 11 markers of the male-specific region of the Y chromosome in 3233 biologically unrelated British men from three cohorts: the British Heart Foundation Family Heart Study (BHF-FHS), West of Scotland Coronary Prevention Study (WOSCOPS), and Cardiogenics Study. On the basis of this information, each Y chromosome was tracked back into one of 13 ancient lineages defined as haplogroups. We then examined associations between common Y chromosome haplogroups and the risk of coronary artery disease in cross-sectional BHF-FHS and prospective WOSCOPS. Finally, we undertook functional analysis of Y chromosome effects on monocyte and macrophage transcriptome in British men from the Cardiogenics Study.

Findings Of nine haplogroups identified, two (R1b1b2 and I) accounted for roughly 90% of the Y chromosome variants among British men. Carriers of haplogroup I had about a 50% higher age-adjusted risk of coronary artery disease than did men with other Y chromosome lineages in BHF-FHS (odds ratio 1.75, 95% CI 1.20–2.54, $p=0.004$), WOSCOPS (1.45, 1.08–1.95, $p=0.012$), and joint analysis of both populations (1.56, 1.24–1.97, $p=0.0002$). The association between haplogroup I and increased risk of coronary artery disease was independent of traditional cardiovascular and socioeconomic risk factors. Analysis of macrophage transcriptome in the Cardiogenics Study revealed that 19 molecular pathways showing strong differential expression between men with haplogroup I and other lineages of the Y chromosome were interconnected by common genes related to inflammation and immunity, and that some of them have a strong relevance to atherosclerosis.

Interpretation The human Y chromosome is associated with risk of coronary artery disease in men of European ancestry, possibly through interactions of immunity and inflammation.

Funding British Heart Foundation; UK National Institute for Health Research; LEW Carty Charitable Fund; National Health and Medical Research Council of Australia; European Union 6th Framework Programme; Wellcome Trust.

Introduction

Of all human chromosomes, the haploid Y chromosome contains the smallest number of genes. The main part of the Y chromosome (male-specific region; MSY) is transmitted intact from father to son and contains single and multicopy genes that encode about 27 distinct proteins.^{1,2} The fundamental biological role of the Y chromosome is to impart male characteristics.³ However, there are also data linking the Y chromosome to the cardiovascular system. For example, polysomy of the Y chromosome (mostly 47, XYY karyotype) was linked to increased cardiovascular mortality,⁴ and a common biallelic polymorphism of the MSY was associated with blood pressure, circulating concentrations of total cholesterol, LDL cholesterol, proatherogenic B-phenotype of LDL cholesterol molecules, and paternal history of coronary artery disease.^{5–8} Although not all studies have replicated these associations, the accumulated evidence^{9,10} lends support to the notion that genetic variation within

the MSY could play a part in determining cardiovascular risk in men. In view of the haploid nature of the MSY and its low level of recombination, the usual methods of linkage disequilibrium-based mapping applied to autosomal chromosomes cannot be used in investigation of its variation. The most appropriate strategy is the analysis of the Y chromosome phylogenetic tree. Defined by a series of biallelic single nucleotide polymorphisms (SNPs), MSY can be partitioned into 20 major haplogroups that descend from a common ancestor, Y-chromosomal Adam.²

We directly examined association between the Y chromosome and coronary artery disease. We first examined whether common Y chromosome haplogroups were associated with risk of coronary artery disease in white British men recruited into the cross-sectional British Heart Foundation Family Heart Study (BHF-FHS).¹¹ We next evaluated the association of Y chromosome lineages with prospective development of coronary artery disease

Lancet 2012; 379: 915–22

Published Online

February 9, 2012

DOI:10.1016/S0140-

6736(11)61453-0

See [Comment](#) page 873

School of Health Sciences, University of Ballarat, Ballarat, VIC, Australia (F J Charchar PhD, S Nankervis PhD, A Bani-Mustafa PhD);

Department of Cardiovascular

Sciences (L D S Bloomer MSc,

T A Barnes PhD, C P Nelson PhD,

M Denniff BSc, R Debiec MD,

P Christofidou MSc,

A H Goodall PhD, N J Samani MD,

M Tomaszewski MD) and

Department of Genetics

(M A Jobling PhD), University of

Leicester, Leicester, UK; Peter

Wills Bioinformatics Centre,

Garvan Institute for Medical

Research, Sydney, NSW,

Australia (M J Cowley PhD);

Division of Health and Social

Care Research, King's College

London, London, UK

(Y Wang PhD); Institute of

Cardiovascular and Medical

Sciences (A F Dominiczak MD)

and Robertson Centre for

Biostatistics (I Ford PhD),

College of Medicine, Veterinary

and Life Sciences, and

Department of Vascular

Biochemistry (C Packard MD),

University of Glasgow, Glasgow,

UK; Division of Epidemiology,

LIGHT, School of Medicine,

University of Leeds, Leeds, UK

(A J Balmforth MD, A S Hall MD);

Universität zu Lübeck,

Medizinische Klinik II, Lübeck,

Germany (J Erdmann PhD,

H Schunkert MD); INSERM UMRS

937, Pierre and Marie Curie

University and Medical School,

Paris, France (F Cambien PhD);

Wellcome Trust Sanger

Institute, Cambridge, UK

(P Deloukas PhD,

W H Ouweland MD); Klinik und

Poliklinik für Innere Medizin II,

Universität Regensburg,

Regensburg, Germany

(C Hengstenberg MD);

Department of Haematology, University of Cambridge and National Health Service Blood and Transplant, Cambridge, UK (W H Ouweland); and National Institute for Health Research Biomedical Research Unit in Cardiovascular Disease, Glenfield Hospital, Leicester, UK (C P Nelson, A H Goodall, N J Samani, M Tomaszewski)

Correspondence to: Dr Maciej Tomaszewski, Department of Cardiovascular Sciences, University of Leicester, Glenfield Hospital, Leicester LE3 9QP, UK mt142@le.ac.uk

See Online for webappendix

in the West of Scotland Coronary Prevention Study (WOSCOPS), a clinical trial that examined the benefits of statin treatment for primary prevention of coronary artery disease in middle-aged Scottish men.^{12,13} Finally, we explored the signatures of Y-chromosomal evolution on the human transcriptome in monocytes and macrophages from British men recruited into the Cardiogenics Study.¹⁴

Methods

Genetic association analysis

The cross-sectional case-control analysis included 811 men with coronary artery disease recruited into BHF-FHS and 633 male controls from the UK Blood Service collection of common controls, which is part of the Wellcome Trust Case Control Consortium (WTCCC). Cases in the BHF-FHS had a validated history of coronary artery disease, defined as myocardial infarction or angina (verified by exercise stress test or angiography) or coronary artery bypass surgery or percutaneous transluminal coronary angioplasty before their 66th birthday and a strong family history of coronary artery disease with at least one similarly affected sibling.¹¹ In the present analysis we used an available subsample of biologically unrelated men from the original collection. The control participants had no known history of coronary artery disease and were recruited nationally in equal proportions in each decade from 30 to 70 years. Apart from sex, age group, and geographic region, no other information was available for the controls. Both cases and controls were previously genotyped with Affymetrix GeneChip Human Mapping 500K Array (Santa Clara, USA; in the WTCCC genome-wide association studies)¹⁵ and, more recently, with Illumina HumanCVD BeadChip array (50K IBC array; San Diego, USA).¹⁶

In the original WOSCOPS, 6595 statin-naïve men aged 45–64 years (mean 55.2 years), who had total serum cholesterol between 6.5 and 7.8 mmol/L at recruitment, an LDL cholesterol greater than 4.5 mmol/L on one occasion before randomisation, and no history of myocardial infarction were assigned to receive 40 mg of pravastatin or placebo daily. During a mean follow-up of 4.9 years, 503 participants had a primary endpoint event (a composite of non-fatal myocardial infarction and death from coronary artery disease) and 77 further men had a coronary revascularisation procedure.¹³ These 580 men were regarded as cases for the purpose of nested case-control studies done here. They were each matched with two participants (controls) who remained event-free during the study on the basis of age (with 2-year age categories), duration of follow-up, and smoking status, as previously described. Altogether, 1740 men (580 cases and 1160 controls) were identified for the nested case-control analysis.¹² Leucocyte DNA, extracted from blood collected at recruitment, was available for 1542 participants (484 cases and 1058 controls). The demographic characteristics of the missing samples were similar to those studied (data not shown). Geno-

typing for the Y chromosome SNPs was successful in 1534 men (482 cases and 1052 controls) and these individuals were included in the genetic association analysis. All major risk factors were assessed during recruitment, as described previously. Information about socioeconomic factors (Carstairs deprivation index, educational attainment, employment status) was also reported before.¹⁷

Transcriptomic analysis

The Cardiogenics Study is a European collaboration on genetics of coronary artery disease.¹⁴ This initiative recruited 918 participants (459 patients with myocardial infarction and 459 normal controls) in five centres: Cambridge (UK), Leicester (UK), Lübeck (Germany), Regensburg (Germany), and Paris (France).¹⁴ 1533 samples (849 from monocytes and 684 from macrophages) were available for RNA-based analysis. The mRNA expression studies for the present project were undertaken in 255 men selected from all participants (134 patients with premature myocardial infarction and 121 normal controls). Selection was determined by male sex, origin (British), and availability of both DNA and RNA information. All participants were of white European origin.

All studies complied with the Declaration of Helsinki and were approved by their local ethics committee and participants gave written informed consent.

DNA analysis

On the basis of a study of Y chromosome diversity in the British population,¹⁸ we selected 11 biallelic SNPs (M9, M35, M45, M89, M170, M173, M201, M207, M269, M304, and SRY₁₀₈₃₁) for genotyping (webappendix p 2). Together, these markers define haplogroups that account for more than 95% of MSY lineages in the UK.¹⁹ On the basis of the hierarchical configuration of the 11 polymorphisms, each Y chromosome was assigned into one of 13 major European haplogroups named according to the Y Chromosome Consortium nomenclature.²⁰ DNA was extracted from peripheral leucocytes. Genotyping was done with TaqMan assays on an ABI PRISM 7900HT Sequence Detection System (Applied Biosystems, Warrington, UK). Discrimination of genotypes was done with Applied Biosystems software.

Monocyte and macrophage isolation and RNA analysis

Monocytes were isolated from whole blood by positive selection with CD14 magnetic beads using an AutoMACS system (Miltenyi Biotech, Bergisch Gladbach, Germany).²¹ Cell purity was confirmed by flow cytometry and in all samples more than 90% of cells were CD14-positive monocytes.²¹ Macrophages were obtained from culturing of monocytes for 7 days in macrophage-SFM medium (Gibco/Invitrogen, Grand Island, USA) with 50 ng/mL recombinant human M-CSF (R&D Systems, Minneapolis, USA).¹⁴ RNA was extracted from both monocytes and macrophages with TRIzol, followed by clean-up with

RNeasy columns (Qiagen, Venlo, Netherlands) and DNase-based treatment.²¹ Monocyte, macrophage, and RNA isolation were done separately in each centre using standardised procedures. Further microarray gene-expression profiling of all samples was done in one centre (Paris, France). None of the samples of RNA from the monocytes or macrophages was pooled. Each was run as an individual sample on the Illumina Human Ref-8 arrays (Illumina, San Diego, USA) containing 24 516 probes.¹⁴ The mRNA was amplified and labelled with the Illumina Total Prep RNA Amplification Kit (Ambion, Austin, USA).¹⁴ After hybridisation, array images were scanned with an Illumina BeadArray Reader and probe intensities were extracted with the gene expression module of Illumina Bead Studio software.²¹ Variance stabilisation transformation was applied to the raw intensities and quantile normalisation was done in the R statistical environment with the Lumi and Beadarray packages (version 1.8.3).¹⁴ After quality control, 12 145 probes with detectable levels of expression were included in the analysis.

Statistical and bioinformatic analysis

The initial unadjusted analysis of association between Y chromosome haplogroups (haplogroup I vs all others) and coronary artery disease as well as other qualitative traits was done with the χ^2 test. The age-adjusted and fully adjusted (for available demographic and clinical variables) analyses were done with binary logistic regression models. Crude and age-adjusted comparisons of normally distributed continuous variables with Y chromosome haplogroups were done with *t* tests (crude comparison) and linear regression models (age-adjusted). Quantitative traits with non-Gaussian distributions (triglycerides and C-reactive protein) were analysed with the Mann-Whitney test and log-transformed before multiple regression. A fixed-effect inverse-variance-weighted meta-analysis was used to combine the age-adjusted results from the BHF-FHS and WOSCOPS studies.

To detect population stratification (which can lead to spurious associations), we used data for autosomal SNPs from BHF-FHS and Cardiogenics participants, genotyped with GeneChip Human Mapping 500K Array (Affymetrix) in BHF-FHS and Human 610 Quad Custom array (Illumina) in Cardiogenics. The standard quality control filters (SNPs with minor allele frequency <1%, Hardy-Weinberg equilibrium χ^2 *p*<0.001 in controls, and call rate <95%) were applied before the analysis. Subsequently, correlated SNPs (those showing multiple correlation coefficient ≥ 0.2) were removed from the data with PLINK software (version 1.07). The pruned subset of independent SNPs was 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 of both cohorts (against three HapMap populations and stratified on the Y chromosome

	Cases	Controls	p value
BHF-FHS			
n	811	633	..
Age (years)	60.6 (7.5)	45.5 (11.7)	<0.0001
WOSCOPS			
n	482	1052	..
Age (years)	56.9 (5.1)	56.8 (5.2)	0.533
Body-mass index (kg/m ²)	26.0 (3.2)	25.6 (3.2)	0.026
Systolic blood pressure (mm Hg)	139.5 (17.2)	135.7 (17.1)	0.0001
Diastolic blood pressure (mm Hg)	85.5 (10.5)	83.7 (10.2)	0.001
HDL cholesterol (mmol/L)	1.1 (0.2)	1.1 (0.3)	<0.0001
LDL cholesterol (mmol/L)	5.0 (0.5)	4.9 (0.4)	0.004
Triglycerides (mmol/L)	2.1 (1.6 to 2.8)	1.9 (1.5 to 2.6)	0.005
Pravastatin-based treatment	194 (40%)	539 (51%)	0.0001
Glucose (mmol/L)	4.9 (0.7)	4.8 (0.7)	0.025
Diabetes	9 (2%)	13 (1%)	0.464
C-reactive protein (mg/L)	2.7 (1.2 to 4.9)	1.9 (0.9 to 3.9)	<0.0001
Smoking	258 (54%)	581 (55%)	0.571
Alcohol consumption ≥ 20 units per week	79 (16%)	186 (18%)	0.584
Carstairs deprivation index*	0.4 (-1.7 to 3.0)	0.5 (-1.8 to 3.0)	0.943
Highest educational attainment†	1.6 (0.9)	1.7 (0.9)	0.784
In employment	304 (63%)	690 (66%)	0.338

Data are n (%), mean (SD), or median (IQR). *Continuously scored measure based on four factors: lack of car ownership; Registrar General's Social Class classification of IV or V; overcrowded households; and male unemployment. †Average score based on four categories: secondary school with leaving certificate (no graduation); school leaving certificate (with graduation); further education, but no degree; and university degree or similar.

Table 1: Characteristics of men from British Heart Foundation Family Heart Study (BHF-FHS) and West of Scotland Coronary Prevention Study (WOSCOPS) stratified by case-control status

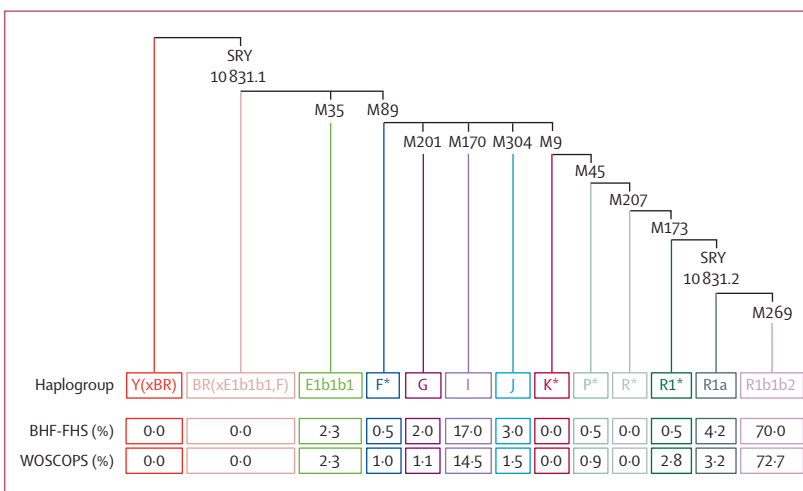


Figure 1: Phylogenetic tree of the Y chromosome and frequency of haplogroups in the British Heart Foundation Family Heart Study (BHF-FHS) and West of Scotland Coronary Prevention Study (WOSCOPS) Each vertical line represents one branch (haplogroup) of the Y chromosome phylogenetic tree. The 13 most common Y chromosome haplogroups are shown and are lined up from the phylogenetically oldest on the left (Y[xBR]) to the youngest on the right (R1b1b2). The vertical length or height of each line corresponds to the age of each haplogroup (the longest line being the oldest haplogroup). The symbols at the top of each line (SRY10831.1, M35, M89, M201, M170, M304, M9, M45, M207, M173, SRY10831.2, and M269) are the identification numbers of the single nucleotide polymorphisms that define specific haplogroups according to the Y Chromosome Consortium nomenclature. The numbers in boxes reflect prevalence (percentage) of each haplogroup in BHF-FHS (upper row) and WOSCOPS (lower row) populations.

haplogroup). We also compared the first four dimensions extracted from the multidimensional scaling analysis between men with haplogroup I and carriers of the other haplogroups using *t* tests.

We also used data available from 1444 men in the BHF-FHS to examine whether haplogroup I and other Y chromosome lineages differ in genotype distribution of SNPs associated with coronary artery disease in previous genome-wide association studies. We first identified 34 autosomal SNPs associated with coronary artery disease in three recent large-scale genetic studies.^{14,16,22} Of those, 21 were directly genotyped and 13 were imputed with the IMPUTE algorithm, as reported before.²³ Three SNPs were excluded from further analysis because of low (<80%) number of informative genotypes. Genotypes of 31 remaining SNPs were in Hardy-Weinberg equilibrium. Differences in distribution of the genotypes stratified on the Y chromosome haplogroup (I vs others) were examined by χ^2 test with Bonferroni correction ($p=0.0016$; $0.05/31$ tests) for multiple testing. Binary logistic regression models with coronary artery disease as an independent variable and age, binarised Y haplogroup status and genotypes of each SNP were constructed individually in STATA (version 11).

To detect pathways with differentially expressed genes between men of haplogroup I versus those in other

haplogroups of the Y chromosome we used gene set enrichment analysis (GSEA)—a sensitive method for detection of both large and subtle changes in pathways avoiding the use of arbitrary cutoff values to define a list of significant genes. We converted the linear regression score for each gene into a preranked list as input to GSEA in preranked mode, using 1000 permutations to assess the false discovery rate and the Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways from the Molecular Signatures Database (version 3.0).²⁴ Since gene sets are not independent, the overlaps between significant gene sets were assessed with Enrichment Map (version 1.0)²⁵ plugin for Cytoscape (version 2.8.0),²⁶ using gene sets with false discovery rate less than 0.2 and an overlap coefficient of 0.3 or higher. The analysis was adjusted for age and coronary artery disease status and done separately in monocytes and macrophages.

Role of the funding source

The funders did not contribute to study design, data interpretation, editing of the report, or decision to submit. The corresponding author had full access to the data and final responsibility for the decision to submit for publication.

Results

Table 1 summarises the demographic and clinical characteristics of BHF-FHS and WOSCOPS men stratified by case-control status. In BHF-FHS, age (the only additional clinical phenotype available in both cases and controls) differed significantly between men with coronary artery disease and controls. There were no significant differences in age, smoking, alcohol consumption, employment status, highest education attainment, deprivation score, and prevalence of diabetes between cases and normal controls in WOSCOPS. By comparison with controls, men who developed coronary artery disease in WOSCOPS had significantly higher body-mass index (BMI), systolic and diastolic blood pressures, LDL cholesterol, triglycerides, C-reactive protein, and glucose concentrations, and lower HDL cholesterol levels (table 1).

To establish whether the Y chromosome is associated with coronary artery disease, we reconstructed the phylogenetic tree for the major Y chromosome haplogroups in both cohorts. Figure 1 shows the markers that define each haplogroup, together with the proportion of participants in each haplogroup. Of nine haplogroups identified, two (R1b1b2 and I) accounted for about 90% of variation in the Y chromosome in both studies. In BHF-FHS, haplogroup I was significantly more common in men with coronary artery disease than in controls (20% [$n=166$] vs 13% [$n=80$], $p=0.0001$). Age-adjusted analysis confirmed that, compared with other haplogroups, I was associated with increased risk of coronary artery disease (odds ratio [OR] 1.75, 95% CI 1.20–2.54, $p=0.004$). To confirm and further validate these findings using a prospective design,

	Haplogroup I	Other haplogroups	p value
BHF-FHS			
n	246	1198	..
Age (years)	55.8 (11.2)	53.6 (12.3)	0.006
WOSCOPS			
n	228	1306	..
Age (years)	57.3 (5.2)	56.7 (5.2)	0.133
Body-mass index (kg/m ²)	25.6 (3.7)	25.8 (3.1)	0.469
Systolic blood pressure (mm Hg)	138.0 (15.9)	136.7 (17.4)	0.284
Diastolic blood pressure (mm Hg)	84.2 (10.1)	84.2 (10.4)	0.987
HDL cholesterol (mmol/L)	1.1 (0.3)	1.1 (0.3)	0.310
LDL cholesterol (mmol/L)	5.0 (0.5)	5.0 (0.4)	0.514
Triglycerides (mmol/L)	1.9 (1.6 to 2.6)	1.9 (1.5 to 2.6)	0.481
Pravastatin-based treatment	110 (48%)	623 (48%)	0.937
Glucose (mmol/L)	4.9 (0.8)	4.8 (0.7)	0.243
Diabetes	6 (3%)	16 (1%)	0.179
C-reactive protein (mg/L)	2.1 (1.1 to 4.0)	2.1 (1.0 to 4.3)	0.659
Smoking	115 (50%)	724 (55%)	0.185
Alcohol consumption ≥ 20 units per week	34 (15%)	231 (18%)	0.353
Carstairs deprivation index*	0.1 (-1.9 to 2.8)	0.5 (-1.7 to 3.0)	0.087
Highest educational attainment†	1.7 (0.9)	1.6 (0.9)	0.189
In employment	156 (68%)	838 (64%)	0.214

Data are n (%), mean (SD), or median (IQR). *Continuously scored measure based on four factors: lack of car ownership; Registrar General's Social Class classification of IV or V; overcrowded households; and male unemployment. †Average score based on four categories: secondary school with leaving certificate (no graduation); school leaving certificate (with graduation); further education, but no degree; and university degree or similar.

Table 2: Characteristics of men from British Heart Foundation Family Heart Study (BHF-FHS) and West of Scotland Coronary Prevention Study (WOSCOPS) stratified by Y chromosome haplogroup status

which is less subject to confounding than is the cross-sectional design of BHF-FHS, we investigated participants recruited into WOSCOPS. In this cohort, haplogroup I was again significantly more common in men who developed coronary artery disease during a period of 4.9 years than in matched controls who remained free of the disease (18% [n=88] vs 13% [n=140], $p=0.014$). Age-adjusted analysis confirmed the association between haplogroup I and increased risk of coronary artery disease (OR 1.45, 95% CI 1.08–1.95, $p=0.012$) in this study. The WOSCOPS participants were extensively characterised for cardiovascular risk factors and traits (systolic and diastolic blood pressure, lipids, BMI, diabetes, C-reactive protein, glucose, alcohol consumption, smoking), pravastatin-based treatment, and socioeconomic factors at baseline. None of these phenotypes was significantly associated with haplogroup I in WOSCOPS (table 2). The binary logistic analysis showed that adjustment for these variables did not attenuate the association of haplogroup I with coronary artery disease (OR 1.60, 95% CI 1.16–2.19, $p=0.004$) and that after HDL cholesterol and lipid-lowering treatment, haplogroup I was the most significant predictor of coronary artery disease (webappendix p 3). A combined analysis of the age-adjusted effects from the BHF-FHS study and WOSCOPS cohort showed that on average haplogroup I increased the risk of coronary artery disease by about 50% (OR 1.56, 95% CI 1.24–1.97, $p=0.0002$).

To exclude the presence of hidden admixture and population stratification as an explanation for the association of haplogroup I and coronary artery disease noted in BHF-FHS, we took advantage of genome-wide SNP data available for both cases and controls recruited into this study.¹⁵ Using non-metric multidimensional scaling of autosomal SNPs, we showed that all cases and controls in BHF-FHS showed evidence of European ancestry and there were no major differences between men of haplogroup I and carriers of other Y chromosome haplogroups (webappendix pp 4, 14).

We also examined whether alleles of 31 autosomal SNPs associated with coronary artery disease genotyped with 500K Affymetrix array and 50K IBC array showed different frequencies in carriers of haplogroup I and other MSY lineages. None of the SNPs showed differential distribution of genotypes after correction for multiple testing (webappendix p 5). We also confirmed that haplogroup I increased the risk of coronary artery disease, independent of the autosomal SNPs identified in previous genome-wide association studies; introduction of each SNP (as an additional dependent variable) into the regression model did not attenuate the significance of association between haplogroup I and coronary artery disease (webappendix p 6).

In 255 Cardiogenics men included in the transcriptomic analysis, those with coronary artery disease (cases) were significantly older and had higher BMI than did healthy controls (webappendix p 7). However, there were no differences in either age or BMI between carriers of

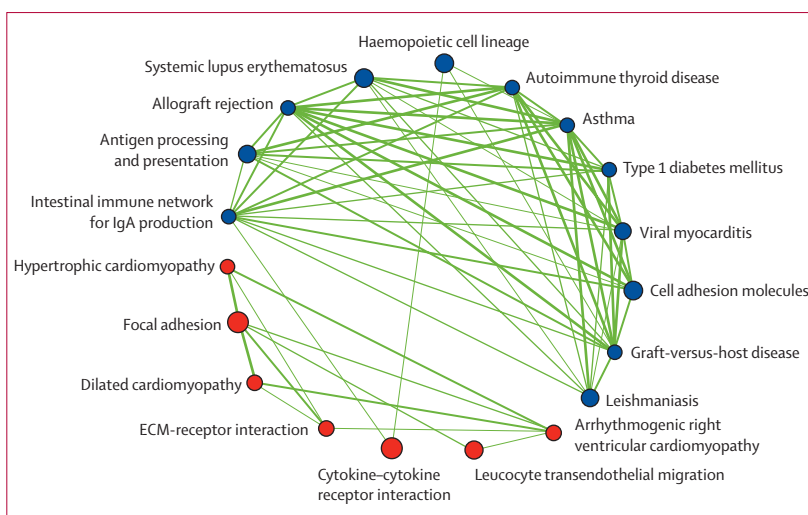


Figure 2: Immune pathways showing significant (false discovery rate <20%) differential expression in macrophages from men with haplogroup I compared with carriers of all other haplogroups
Red nodes show upregulated pathways and blue nodes show downregulated pathways in men with haplogroup I from the Kyoto Encyclopedia of Genes and Genomes. Node size is a function of the number of genes in each pathway (the larger the node, the more genes in the pathway) and the colour intensity of each node reflects the level of significance. Lines represent a network of biological links between the pathways; the thickness of each line is proportional to number of genes shared by pathways connected by the line. The high interconnection between these pathways suggests that they are not independent and represent different, yet overlapping, facets of immunity (webappendix pp 9–13). The image was generated with the Enrichment Map plugin for Cytoscape. ECM=extracellular matrix.

haplogroup I and other Y chromosome lineages in this population (webappendix p 8). To identify potential functional effects of haplogroup I in cells of relevance to coronary artery disease we examined associations between Y chromosome haplogroups and both monocyte and macrophage transcriptome profiles. We excluded the presence of population stratification and admixture in these participants, using the same approach as described earlier for the BHF-FHS cohort (webappendix pp 4, 14). The GSEA showed no differences in monocyte transcriptome between men with haplogroup I and carriers of other MSY lineages after correction for multiple testing. By contrast, analysis of macrophage transcriptome identified 30 KEGG pathways with differential expression between haplogroup I and other lineages of the Y chromosome (false discovery rate <20%; webappendix pp 9–13). Of these, 19 pathways were interconnected by common genes related to inflammation and immunity; seven pathways were upregulated and 12 were downregulated in men with haplogroup I compared with carriers of other haplogroups (figure 2).

Discussion

Our study is the first to evaluate associations between main European Y chromosome lineages and coronary artery disease as well as its underlying risk factors. The most important finding from our analysis is that haplogroup I is associated with significantly increased risk of coronary artery disease compared with other ancient lineages of the Y chromosome and that this

association, although independent of major cardiovascular risk factors and socioeconomic status, might be mediated through a genetically programmed profile of immunity and response to inflammation (panel).

Our transcriptomic analysis showed that several of 19 pathways interconnected by inflammation and immunity genes, and showing differential expression between men with haplogroup I and other lineages, might be relevant to atherosclerosis. Specifically, trafficking of leucocytes through the endothelial barrier is a well recognised process in both early²⁷ and late²⁸ stages of atherosclerosis, and leucocyte transendothelial migration (hsa04670) was identified as the most significantly upregulated pathway in gene expression profiling of atherosclerotic arterial wall.²⁸ Focal adhesions (hsa04510) control cytoskeletal or adhesion dynamics and thus affect both leucocyte motility within intima and interactions between platelets and endothelium, all of which play a part in the pathogenesis of coronary artery disease.^{29–31} The balance between proinflammatory and anti-inflammatory cytokines (cytokine-cytokine receptor interaction pathway, hsa04060) is also a well recognised mechanism underlying atherosclerotic plaque development.³² Interestingly, four of seven upregulated pathways (focal adhesion, cytokine-cytokine receptor interaction, hypertrophic cardiomyopathy [hsa05410], and extracellular

matrix-receptor interaction [hsa04512]) show the strongest enrichment for genes with previous evidence of association with coronary artery disease among more than 100 KEGG pathways.³³ Pharmacological inhibitions of these pathways were suggested as treatments with a potential to reduce or even reverse the burden of atherosclerosis in coronary circulation.^{27,30,34} These molecular networks therefore represent possible mechanisms explaining the association of haplogroup I with coronary artery disease. Furthermore, upregulation of these pathways in macrophages (but not in monocytes) suggests that differentiation (activation) of monocytes to macrophages, one of the key steps in the pathogenesis of atherosclerosis,³⁵ might be the stage at which haplogroup I exerts its molecular effects on coronary artery disease.

Despite this attractive possibility, a statistically more striking finding was the association of haplogroup I with downregulation of several pathways of the immune system. Indeed, some of the identified pathways are activated in exposure to pathogens (leishmaniasis [hsa05140], viral myocarditis [hsa05416]). Others directly represent immune response and processing (antigen processing and presentation [hsa04612] and intestinal immune network for IgA production [hsa04672]). Furthermore, some of the most significant pathways are in fact autoimmune disorders, including autoimmune thyroid disease (hsa05320) and type 1 diabetes (hsa09490). Most of the downregulated pathways interact together in regulation of adaptive immunity and operate mainly in lymphocytes or antigen-presenting cells, or both, possibly mostly through MHC class II cell surface receptors (HLA-DP, HLA-DQ, HLA-DR). Although not fully understood, dysfunction of immune response is a well established contributor to atherosclerosis and coronary artery disease.³² Previous studies suggested that the Y chromosome could play a part in regulation of the immune system—ie, men with haplogroup I of the Y chromosome were particularly vulnerable to HIV infection.³⁶ In fact, carriers of haplogroup I on retroviral therapy took a longer time to HIV suppression and had more accelerated progression to AIDS compared with carriers of other MSY lineages.³⁶ Mortality from AIDS in men of haplogroup I was also significantly higher than that in men with other haplogroups.³⁶

Taken together, these findings suggest that downregulation of adaptive immunity in carriers of haplogroup I is accompanied by upregulation of pathways underlying inflammatory response. This conclusion implies that haplogroup I carriers might have chronic derangements in homeostatic mechanisms of adaptive immunity, possibly with heightened inflammation affecting the cardiovascular system. A similar mechanism has been well documented in other complex disorders—ie, in inflammatory bowel disease, in which deficiencies in immunity status can lead to increased systemic inflammation.³⁷ Interestingly, the association between haplogroup I and increased susceptibility to coronary

Panel: Research in context

Systematic review

We searched original English-language publications in PubMed to May, 2011, using “Y chromosome”, “haplogroup”, “haplotype”, and “coronary artery disease” as keywords and found no studies that investigated the association between paternal lineages of the Y chromosome (haplogroups) and predisposition to coronary artery disease.

Interpretation

Our study is the first robust analysis of association between different lineages of the Y chromosome and coronary artery disease. We show that men who inherit haplogroup I (one of the most common Y chromosome types in Europe) from their male ancestors have a roughly 50% higher risk of coronary artery disease than do men with other types of the Y chromosome. The effect of haplogroup I on coronary artery disease is not mediated by traditional cardiovascular risk factors such as age, body-mass index, blood pressure, lipids, diabetes, smoking, alcohol consumption, socioeconomic status, or circulating concentrations of C-reactive protein. However, we noted that men with haplogroup I showed downregulation of adaptive immunity as well as upregulation of inflammatory response pathways in their macrophages compared with carriers of other Y chromosome types. These data show that predisposition to coronary artery disease in men might, at least in part, be determined by the paternal lineage of their Y chromosome and that this effect on risk of coronary artery disease is most likely mediated through immune response.

artery disease was independent of C-reactive protein. This marker of inflammation is more closely related to innate than to adaptive defence mechanisms.³⁸ Since the most distinct signature of haplogroup I in our study was identified on adaptive immunity pathways, we expect that studies with more robust markers of specific immune system will be necessary to explain the association between the Y chromosome and coronary artery disease. Such investigations should also take us closer to dissection of the mechanisms underlying the sex differences in regulation of human immunity. Indeed, there is a well recognised sexual dimorphism in response to immunisation³⁹ as well as susceptibility to autoimmune disorders.⁴⁰ The biological background of these differences is most likely multifactorial. Whether genes unique to the X chromosome, polymorphic variants of the MSY, their interactions with autosomal loci and environment, or all the above contribute to sex differences in regulation of the immune system will need further studies to elucidate.

Evolutionarily, the distinguishing feature of haplogroup I is its almost complete absence in indigenous populations outside Europe.² The carriers of this haplogroup probably arrived from the Middle East as hunter-gatherers during the Paleolithic era roughly 25 000 years ago and it has been suggested that they spread throughout Europe together with diffusion of the Gravettian archaeological culture.⁴¹ The gradual Neolithic expansion of farmers 10 000 years ago led to the present dominance of R1b1b2 (over I and other MSY haplogroups) in most parts of northern and western Europe.^{2,42} The present geographical distribution of one of the major lineages of haplogroup I (I1) in western Europe correlates with the well established north-south gradient in prevalence of coronary artery disease. Indeed, in northern populations (Scandinavia, Germany, Netherlands) in which the prevalence of haplogroup I ranges between 15% and 40%, mortality from coronary artery disease is significantly higher than in southern Europe (France, Apennine peninsula, Spain, and Switzerland), where haplogroup I is less prevalent (3–15%).^{43,44} The north-south gradient in coronary artery disease morbidity was also reported in the UK (2011). Unavailability of information about birthplace of participants and their male ancestors and lack of samples from different parts of the UK prevented us from examining whether men with haplogroup I are over-represented in regions known for particularly high prevalence of coronary artery disease and its modifiable lifestyle risk factors. Further studies should address the question of whether distribution of Y chromosome lineages could account (at least in part) for the geographic differences in predisposition to coronary artery disease in men.

We should acknowledge that similar to other common genetic variants associated with increased risk of cardiovascular disease, haplogroup I of the Y chromosome on its own is unlikely to offer sufficiently high positive predictive

value of coronary artery disease. Indeed, single risk factors (both genetic and non-genetic) that are fairly common in both cases and controls and coupled with a 1·5-increase in the OR of a disease are not specific enough as individual risk predictors.^{45,46} Nevertheless, the relative estimates of coronary artery disease risk in carriers of haplogroup I are not trivial from the point of view of genetic association analysis. Indeed, they are larger than that of many common autosomal alleles identified in recent genome-wide association studies. Further replication of our results in large, prospective population-based studies is necessary to provide precise estimates of coronary artery disease risk attributable to variation in the Y chromosome. Such studies could have important public health implications in view of the significant lifetime risk of coronary artery disease in men⁴⁷ and still imperfect risk stratification on the basis of traditional cardiovascular risk factors.⁴⁶

A gap remains in the understanding of disease differences both within and between populations. Here, we have identified an association that supports the hypothesis that the Y chromosome determines interindividual differences in susceptibility to coronary artery disease among British men. We also showed that this association could be mediated through immunity and inflammation-related networks. Our study revealed that the Y chromosome might have a magnified effect on men beyond sex determination despite the small number of genes it harbours in the human genome. Future resequencing efforts and functional experiments will be needed to identify the causative variants underlying the increased susceptibility to coronary artery disease in carriers of haplogroup I and to decipher complex interplay between human Y chromosome, immunity, and cardiovascular disease.

Contributions

FJC, MAJ, AFD, NJS, and MT conceived and designed the study. FJC and MT oversaw laboratory analyses and MT provided the overall supervision of the study. LDSB, TAB, MJC, CPN, YW, MD, RD, PC, SN, and AB-M did the laboratory experiments or contributed to the statistical analysis, or both. AJB, ASH, JE, FC, PD, CH, CP, HS, WHO, IF, and AHG contributed to clinical data collections. FJC, NJS, and MT drafted the report. All authors contributed to critical revision of the report.

Conflicts of interest

We declare that we have no conflicts of interest.

Acknowledgments

This study was funded by the British Heart Foundation project grant (PG/06/097/21331) to MT. NJS holds a personal chair supported by the British Heart Foundation and is a UK National Institute for Health Research (NIHR) senior investigator. FJC is supported by research grants from the LEW Carty Charitable Fund and National Health and Medical Research Council of Australia. The Cardiogenics project was supported by the European Union 6th Framework Programme (LSHM-CT-2006-037593). Contributors to Cardiogenics Transcriptome Study are listed in the webappendix p 16. AFD held a personal chair supported by the British Heart Foundation between 2006 and 2010. MAJ is a Wellcome Trust senior fellow in biomedical science (087576). Collection of healthy participants for this study was supported by the Cambridge NIHR Biomedical Research Centre and by an NIHR programme grant to WHO (RP-PG-0310-1002) and a project grant from the Wellcome Trust to WHO (WT-0841383/2/07/2). Work described in this paper is part of the research portfolio supported by the Leicester NIHR Biomedical Research Unit in Cardiovascular Disease.

For more on coronary artery disease morbidity in the UK see www.heartstats.org

References

- 1 Skaletsky H, Kuroda-Kawaguchi T, Minx PJ, et al. The male-specific region of the human Y chromosome is a mosaic of discrete sequence classes. *Nature* 2003; **423**: 825–37.
- 2 Jobling MA, Tyler-Smith C. The human Y chromosome: an evolutionary marker comes of age. *Nat Rev Genet* 2003; **4**: 598–612.
- 3 Graves JA, Koina E, Sankovic N. How the gene content of human sex chromosomes evolved. *Curr Opin Genet Dev* 2006; **16**: 219–24.
- 4 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**: 691–96.
- 5 Shankar RR, Charchar FJ, Eckert GJ, et al. Studies of an association in boys of blood pressure and the Y chromosome. *Am J Hypertens* 2007; **20**: 27–31.
- 6 Charchar FJ, Tomaszewski M, Lacka B, et al. Association of the human Y chromosome with cholesterol levels in the general population. *Arterioscler Thromb Vasc Biol* 2004; **24**: 308–12.
- 7 Charchar FJ, Tomaszewski M, Padmanabhan S, et al. The Y chromosome effect on blood pressure in two European populations. *Hypertension* 2002; **39**: 353–56.
- 8 Ellis JA, Stebbing M, Harrap SB. Association of the human Y chromosome with high blood pressure in the general population. *Hypertension* 2000; **36**: 731–33.
- 9 Charchar FJ, Tomaszewski M, Strahorn P, Champagne B, Dominiczak AF. Y is there a risk to being male? *Trends Endocrinol Metab* 2003; **14**: 163–68.
- 10 Ely D, Underwood A, Dunphy G, Boehme S, Turner M, Milsted A. Review of the Y chromosome, Sry and hypertension. *Steroids* 2010; **75**: 747–53.
- 11 Samani NJ, Burton P, Mangino M, et al. A genome-wide 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**: 1011–20.
- 12 Packard CJ, O'Reilly DS, Caslake MJ, et al. Lipoprotein-associated phospholipase A2 as an independent predictor of coronary heart disease. West of Scotland Coronary Prevention Study Group. *N Engl J Med* 2000; **343**: 1148–55.
- 13 Shepherd J, Cobbe SM, Ford I, et al. Prevention of coronary heart disease with pravastatin in men with hypercholesterolemia. *N Engl J Med* 1995; **333**: 1301–08.
- 14 Schunkert H, König IR, Kathiresan S, et al. Large-scale association analysis identifies 13 new susceptibility loci for coronary artery disease. *Nat Genet* 2011; **43**: 333–38.
- 15 Samani NJ, Erdmann J, Hall AS, et al. Genome-wide association analysis of coronary artery disease. *N Engl J Med* 2007; **357**: 443–53.
- 16 IBC 50K CAD Consortium. Large-scale gene-centric analysis identifies novel variants for coronary artery disease. *PLoS Genet* 2011; **7**: e1002260.
- 17 Batty GD, Wang Y, Brouillette SW, et al. Socioeconomic status and telomere length: the West of Scotland Coronary Prevention Study. *J Epidemiol Community Health* 2009; **63**: 839–41.
- 18 Capelli C, Redhead N, Abernethy JK, et al. A Y chromosome census of the British Isles. *Curr Biol* 2003; **13**: 979–84.
- 19 King TE, Jobling MA. Founders, drift, and infidelity: the relationship between Y chromosome diversity and patrilineal surnames. *Mol Biol Evol* 2009; **26**: 1093–102.
- 20 Y Chromosome Consortium. A nomenclature system for the tree of human Y-chromosomal binary haplogroups. *Genome Res* 2002; **12**: 339–48.
- 21 Heinig M, Petretto E, Wallace C, et al. A trans-acting locus regulates an anti-viral expression network and type 1 diabetes risk. *Nature* 2010; **467**: 460–64.
- 22 Coronary Artery Disease (CAD) 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.
- 23 Preuss M, König IR, Thompson JR, et al. Design of the Coronary Artery Disease Genome-Wide Replication and Meta-Analysis (CARDIoGRAM) Study—a genome-wide association meta-analysis involving more than 22,000 cases and 60,000 controls. *Circ Cardiovasc Genet* 2010; **3**: 475–83.
- 24 Subramanian A, Tamayo P, Mootha VK, et al. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc Natl Acad Sci USA* 2005; **102**: 15545–50.
- 25 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.
- 26 Cline MS, Smoot M, Cerami E, et al. Integration of biological networks and gene expression data using Cytoscape. *Nat Protoc* 2007; **2**: 2366–82.
- 27 Eriksson EE. Mechanisms of leukocyte recruitment to atherosclerotic lesions: future prospects. *Curr Opin Lipidol* 2004; **15**: 553–58.
- 28 Hägg S, Skogsberg J, Lundström J, et al. 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**: 1615–27.
- 29 Mitra SK, Hanson DA, Schlaepfer DD. Focal adhesion kinase: in command and control of cell motility. *Nat Rev Mol Cell Biol* 2005; **6**: 56–68.
- 30 Curtiss LK. Reversing atherosclerosis? *N Engl J Med* 2009; **360**: 1144–46.
- 31 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.
- 32 Hansson GK, Libby P. The immune response in atherosclerosis: a double-edged sword. *Nat Rev Immunol* 2006; **6**: 508–19.
- 33 Liu H, Liu W, Liao Y, et al. CADgene: a comprehensive database for coronary artery disease genes. *Nucleic Acids Res* 2011; **39**: D991–96.
- 34 Braunersreuther V, Mach F. Leukocyte recruitment in atherosclerosis: potential targets for therapeutic approaches? *Cell Mol Life Sci* 2006; **63**: 2079–88.
- 35 Ross R. The pathogenesis of atherosclerosis: a perspective for the 1990s. *Nature* 1993; **362**: 801–09.
- 36 Sezgin E, Lind JM, Shrestha S, et al. Association of Y chromosome haplogroup I with HIV progression, and HAART outcome. *Hum Genet* 2009; **125**: 281–94.
- 37 Marks DJB, Harbord MWN, MacAllister R, et al. Defective acute inflammation in Crohn's disease: a clinical investigation. *Lancet* 2006; **367**: 668–78.
- 38 Peisajovich A, Marnell L, Mold C, Du Clos TW. C-reactive protein at the interface between innate immunity and inflammation. *Expert Rev Clin Immunol* 2008; **4**: 379–90.
- 39 Klein SL, Jedlicka A, Pekosz A. The Xs and Y of immune responses to viral vaccines. *Lancet Infect Dis* 2010; **10**: 338–49.
- 40 McCarthy M. The “gender gap” in autoimmune disease. *Lancet* 2000; **356**: 1088.
- 41 Semino O, Passarino G, Oefner PJ, et al. The genetic legacy of Paleolithic *Homo sapiens sapiens* in extant Europeans: a Y chromosome perspective. *Science* 2000; **290**: 1155–59.
- 42 Balaesque P, Bowden GR, Adams SM, et al. A predominantly neolithic origin for European paternal lineages. *PLoS Biol* 2010; **8**: e1000285.
- 43 Rootsi S, Magri C, Kivisild T, et al. Phylogeography of Y-chromosome haplogroup I reveals distinct domains of prehistoric gene flow in Europe. *Am J Hum Genet* 2004; **75**: 128–37.
- 44 Vollset SE. Does the southern European cardiovascular mortality advantage extend to total mortality? 50-year trends in death risks between 40 and 70 years of age in Western European men and women. *Scand J Public Health* 2010; **38**: 127–34.
- 45 Wald NJ, Hackshaw AK, Frost CD. When can a risk factor be used as a worthwhile screening test? *BMJ* 1999; **319**: 1562–65.
- 46 Holmes MV, Harrison S, Talmud PJ, Hingorani AD, Humphries SE. Utility of genetic determinants of lipids and cardiovascular events in assessing risk. *Nat Rev Cardiol* 2011; **8**: 207–21.
- 47 Ulrich S, Hingorani AD, Martin J, Vallance P. Lifetime risk of developing coronary heart disease. *Lancet* 1999; **353**: 925.