European Heart Journal (2004) **25**, 459–467

Clinical research

Genotypes and haplotypes predisposing to myocardial infarction: a multilocus case-control study

Martin D. Tobin[a,*], Peter S. Braund[b], Paul R. Burton[a], John R. Thompson[a], Richard Steeds[c], Kevin Channer[c], Suzanne Cheng[d], Klaus Lindpaintner[e], Nilesh J. Samani[b]

[a] *Centre for Biostatistics and Genetic Epidemiology, Department of Health Sciences, University of Leicester, 22–28 Princess Road West, Leicester LE1 6TP, UK*
[b] *Cardiology Group, Department of Cardiovascular Sciences, University of Leicester, Clinical Sciences Wing, Glenfield Hospital, Groby Road, Leicester, UK*
[c] *Department of Cardiology, The Royal Hallamshire Hospital, Glossop Road, Sheffield S10 2JF, UK*
[d] *Roche Molecular Systems Inc, 1145 Atlantic Avenue, Alameda, CA 94501, USA*
[e] *Hoffmann-La Roche Ltd, Roche Genetics, Pharmaceuticals Division, CH-4070 Basel, Switzerland*

Received 11 July 2003; revised 31 October 2003; accepted 27 November 2003

See page 451 for the editorial comment on this article[†]

KEYWORDS
Genetics;
Myocardial infarction;
Risk factors

Aim To identify polymorphisms and haplotypes in candidate genes that predispose to myocardial infarction (MI) using a multilocus approach.
Methods and results 1052 subjects, comprising 547 acute MI cases and 505 controls were studied. The association between MI and 58 SNPs in 35 candidate genes (generating 61 016 individual genotypes), and between MI and estimated haplotypes at 14 loci encompassing 16 genes was investigated. Two individual gene variants and haplotypes at two loci showed statistical association with MI. The α-adducin 460trp variant (OR 0.73, 95% CI 0.59–0.91, $P = 0.006$) and the cholesteryl ester transfer protein −629A variant (OR 0.82, 95% CI 0.68–0.97, $P = 0.025$) were both associated with a significant protective effect on MI, as was the paraoxonase 1/paraoxonase 2 haplotype comprising met55 and gln192 in paraoxonase 1 and cys311 in paraoxonase 2 (OR 0.52, 95% CI 0.39–0.77, $P = 0.001$). The apolipoprotein C III haplotypes CCTTCG and ATCCCG at positions −641*−482*−455*1100*3175*3206 were associated with an increased risk of MI, odds ratios 1.41 (95% CI 1.06–1.76, $P = 0.023$) and 1.71 (95% CI 1.28–2.14, $P = 0.038$), respectively.
Conclusions We report associations of two polymorphisms and haplotypes at two loci with risk of MI that warrants testing in future studies. Furthermore, we demonstrate the application of a multilocus assay in the setting of a large association study and the additional benefit gained from the study of haplotypes to identify variants influencing risk of coronary heart disease.

[*] Corresponding author. Tel.: +44-0-116-252-3156; fax: +44-0-116-252-3272.
E-mail address: mt47@leicester.ac.uk (M.D. Tobin).
[†] doi:10.1016/j.ehj.2004.01.019.

0195-668X/$ - see front matter
doi:10.1016/j.ehj.2003.11.014

Introduction

Myocardial infarction is a complex disorder resulting from the combined effects of multiple environmental and genetic factors.[1] Given that myocardial infarction is a leading cause of death in the Western world, its prevention is a public health priority. One approach to preventing myocardial infarction is to use observational genetic studies to identify disease-susceptibility genes that will further our understanding of the etiological pathways and thus indicate new potential targets for intervention. Genetic linkage and association studies have already identified several candidate genes that may predispose to coronary artery narrowing, such as cholesteryl ester transfer protein (CETP),[2] and myocardial infarction, such as *SERPINE1* (*PAI 1*), connexin 37, and stromelysin-1.[3] However, the inability to replicate many of these findings has cast doubt on the implications of many of these studies. The power of most genetic association studies has been low, and few studies have been published with more than 500 cases and 500 controls.[3,4]

Haplotype-based analysis could prove a more powerful approach to dissect the genetic architecture of complex diseases.[5] The human genome is comprised of haplotype blocks, which are chromosomal segments that are preserved intact over many generations, interspersed by recombination hotspots.[6] Alleles within such ancestrally preserved haplotype blocks tend to be in strong linkage disequilibrium with one another, i.e., alleles at different sites within the haplotype block are strongly associated with one another at the level of the population. It is becoming clear that over small regions relatively few haplotypes may account for the vast majority of diversity[7] and variations in haplotypes may be powerful predictors of disease.

Recently developed assays capable of simultaneously genotyping multiple loci allow the efficient investigation of the role of many genetic polymorphisms within a single study. Furthermore, the development of methods to estimate haplotype frequencies from unrelated cases and controls now allows the study of the combined effects of a number of polymorphisms that are in linkage disequilibrium with one another.[8–12] Our aim was to identify individual polymorphisms and haplotypes that predispose to myocardial infarction using a multilocus approach[13] to interrogate 63 single nucleotide polymorphisms (SNPs) in 35 candidate genes.

Methods

We carried out a case-control association study in two UK centres. The methods have been described previously.[14] Briefly, we recruited 549 consecutive Caucasian cases <75 years who had survived admission to the coronary care units in Leicester and Sheffield. All cases satisfied the World Health Organisation criteria for myocardial infarction.[15] The 505 Caucasian controls, recruited contemporaneously, were healthy adult *visitors* to patients with noncardiovascular illnesses at the two hospitals. This provided an epidemiologically robust control population derived from the same geographic area as the cases. Control subjects with a history of MI or angina were excluded. The study had been designed to have 90% power to detect a 1.5-fold increase in the risk of myocardial infarction associated with the angiotensin converting enzyme DD genotype.[16]

Subjects completed a standard questionnaire about their personal history, which included data on age, sex, smoking status, history of hypertension and history of diabetes. We measured height and weight to calculate body mass index (BMI), and collected blood samples for the measurement of serum total cholesterol and DNA analysis. The study was approved by local research ethics committees and informed consent was obtained from participants.

DNA was extracted from whole blood.[14] A multilocus assay was used to genotype 63 polymorphisms within 35 candidate genes previously implicated in cardiovascular disease.[13,17] The extracted genomic DNA was amplified by two separate multiplex PCRs with biotinylated primers. The amplified PCR products were detected colorimetrically with immobilised sequence-specific oligonucleotide probes arrayed in a line array on nylon membrane strips. Correct allele identification was confirmed by direct sequencing or RFLP analysis of genomic DNA samples carrying the different genotypes for each polymorphism. We excluded data on five SNPs (apolipoprotein B (*APOB*) (arg3500gln), cholesteryl ester transfer protein [*CETP*] (asp442gly), *CETP* intron 14 G(+1)-to-A (Int14A), *CETP* intron 14 (+3)T ins (Int14T) and tumor necrosis factor (*TNF* superfamily, member 2) [*TNF*] (G(-244)A)) from our analysis, as fewer than five individuals possessed the variant allele. To check for possible laboratory errors and possible relatedness of cases and controls, we plotted the mean allele sharing identical by state across all polymorphisms and its standard deviation for all possible pairs of individuals in the study. This procedure, implemented using the GRR package (Graphical Relationship Representation),[18] demonstrated identical genotypes across all 58 polymorphisms from two supposedly unrelated cases, which were therefore excluded from the study. No other outliers indicating errors were identified.

Statistical analysis was carried out using *STATA 7.0*. In the first stage, the association between each of the 58 polymorphisms with adequate information and myocardial infarction was assessed using logistic regression to adjust for the effects of age and sex, under an additive genetic model (genetic models are explained in Table 3). Other covariates were not adjusted for at this stage so that we could assess their possible involvement in the causal pathway. Furthermore, because genotype is effectively randomly allocated at birth (a process termed 'Mendelian randomisation') the effects of genetic variants are unlikely to be confounded by lifestyle risk factors.[19] Maximum likelihood estimates of odds ratios and their 95% confidence intervals were obtained and the significance of each polymorphism was assessed using a likelihood ratio test. In the second stage, polymorphisms that showed significant association with myocardial infarction ($P < 0.05$) under an additive model were also assessed under dominant and recessive models, and the influence of other covariates was assessed. A significance level of $P < 0.05$ was adopted for this exploratory study.

A number of software programs have now been developed for estimating haplotype frequencies from genotype data on unrelated cases and controls.[8–12] We estimated haplotype frequencies using software implemented in *STATA 7.0*,[8,20] that uses a log-linear model embedded within an Expectation–Maximization (EM) algorithm.[21,22] To describe the relationship between haplotypes and disease status, we performed maximum likelihood estimation of the log-linear model using the iterative proportional fitting algorithm[23] and obtained profile likelihood estimates of the confidence intervals of haplotype odds ratios.[8,20,24] To compare the consistency of our findings, we also

estimated haplotype frequencies using *Arlequin*, a software commonly employed for this purpose that is also based on an EM algorithm.[9] We tested for evidence of hidden population admixture or stratification using the methods described by Pritchard and Rosenberg[25], utilising data from unlinked markers.

Results

Characteristics of the 1052 participants included in the study are shown in Table 1. Myocardial infarction cases were more likely to be male, older, hypertensive, diabetic, and smokers. Genotype frequencies in cases and controls are shown in Table 2. No evidence for population stratification was found for any of the polymorphisms tested. Under an additive genetic model, two polymorphisms showed statistically significant evidence of association with myocardial infarction: α-adducin [*ADD1*] (gly460trp) and cholesteryl ester transfer protein [*CETP*] (C-629A). These polymorphisms were also assessed under dominant and recessive models (Table 3). The *ADD1* 460trp variant was associated with a significant protective effect on myocardial infarction under a dominant, recessive or additive model (Table 3), which remained significant under each model after adjusting for smoking, history of hypertension or history of diabetes (in addition to age and sex). The *CETP* −629A variant was associated with a significant protective effect under a recessive or additive model (Table 3), which remained significant after adjustment for other risk factors.

Haplotypes at two loci showed statistically significant evidence of association with myocardial infarction (Table 4). The apolipoprotein C III (*APOC3*) haplotypes CCTTCG and ATCCCG at positions −641*−482*−455*-1100*3175* 3206 were associated with an increased risk of myocardial infarction: CCTTCG (OR 1.41 (95% CI 1.06−1.76), $P = 0.023$) and ATCCCG (OR 1.71 (95% CI 1.28−2.14), $P = 0.038$). The paraoxonase 1 (*PON1*)/paraoxonase 2 (*PON2*) haplotype comprising met55 and gln192 in *PON1* and cys311 in *PON2* was associated with a significant protective effect on myocardial infarction risk (OR 0.52 (95% CI 0.39−0.77), $P = 0.001$). Haplotype frequencies estimated using *Arlequin*[9] were consistent with the frequencies we obtained and presented in Table 4.

Table 1 Characteristics of the 1052 participants in the study

Characteristic	Controls ($n = 505$)	MI cases ($n = 547$)
Age in years (sd)	58.6 (10.7)	61.9 (9.2)
Male (%)	62.0	68.0
Body mass index (sd)	25.7 (3.6)	25.9 (3.9)
Current smoker (%)	17.0	40.2
Hypertension (%)	16.8	31.0
Diabetes (%)	2.0	8.7

Discussion

We examined the relationship between 58 polymorphisms in 35 candidate genes and myocardial infarction in 1052 subjects (generating 61 016 individual genotypes). Our study, which is one of the largest such multilocus genetic association studies of myocardial infarction involving healthy controls published to date, revealed two polymorphisms and haplotypes at two loci significantly associated with myocardial infarction risk. To our knowledge, only two other studies have examined simultaneously the association of multiple candidate gene polymorphisms with risk of coronary heart disease. Topol et al.[26] reported associations between three thrombospondin gene family variants and myocardial infarction in their study of 72 individual SNPs in 352 premature coronary artery disease cases and 418 controls recruited from a population of white Americans. Yamada et al.[3] reported associations between the C1019T polymorphism in the connexin 37 gene and the 4G-668/5G polymorphism in the plasminogen-activator inhibitor type 1 gene in their study of 112 polymorphisms in 2819 Japanese myocardial infarction cases and 2242 controls. Although large, the power of Yamada's study was modest, due to the similarity of cases and controls (controls were selected who had at least one 'risk factor', such as diabetes or hypertension), and by their adjustment for many potential intermediates on the causal pathway in all of their analyses. Notably, haplotype analysis was not undertaken in either study.[3,26]

In our study, the *ADD1* 460 trp allele was found to be protective against myocardial infarction. α-Adducin is a ubiquitously expressed cytoskeleton protein involved in the formation of actin-spectrin lattice, actin polymerisation and cell signal transduction,[27–28] including an effect on Na-KATPase. There is considerable experimental and clinical evidence that the *ADD1* (gly460trp) polymorphism is associated with functional effects related to sodium and water homeostasis. Paradoxically, it is also the 460trp allele that is associated with salt retention and tendency to hypertension,[29] effects that one might expect to increase the risk of myocardial infarction. Consistent with this, Psaty et al.[30] have reported a protective effect of the 460Trp allele against myocardial infarction and stroke in hypertensive patients taking diuretics. If further studies confirm that the 460Trp allele is also primarily protective against myocardial infarction, then this is likely to be through a cellular mechanism different from its effect leading to hypertension; alternatively, the observed association may be indicative of linkage disequilibrium to a protective gene variant.

We found that the *CETP* −629 A allele was protective against myocardial infarction. The corresponding haplotype analysis was also consistent with this finding, with a trend towards a protective effect of haplotypes containing the *CETP* −629 A allele (Table 4). CETP is a hydrophobic glycoprotein and has a key role in the metabolism of high density lipoproteins (HDL). CETP mediates the transfer of cholesterol esters from HDL to other lipoproteins and uptake of cholesterol by the liver.

Table 2 Genotype frequency in cases and in controls and association with myocardial infarction under an additive genetic model

Gene and polymorphism	Frequency in controls (%), $N = 505$			Frequency in cases (%), $N = 547$			*P*-value adjusted for age and sex
	AA	AB	BB	AA	AB	BB	
Lipoprotein Lp(a) (*LPA*) (C93T)	71.29	26.73	1.98	75.09	23.44	1.47	0.164
Lipoprotein Lp(a) (*LPA*) (G121A)	68.85	27.18	3.97	68.38	27.39	4.23	0.902
Apolipoprotein A-IV (*APOA4*) (thr347ser)	66.34	29.90	3.76	64.72	32.54	2.74	0.801
Apolipoprotein A-IV (*APOA4*) (gln360his)	81.98	17.62	0.40	83.73	15.36	0.91	0.501
Apolipoprotein B (*APOB*) (thr71ile)	44.95	44.36	10.69	44.61	42.41	12.98	0.498
Apolipoprotein C-III (*APOC3*) (C(-641)A)	41.47	46.03	12.50	37.32	52.57	10.11	0.716
Apolipoprotein C-III (*APOC3*) (C(-482)T)	56.04	38.22	5.74	53.20	42.60	4.20	0.790
Apolipoprotein C-III (*APOC3*) (T(-455)C)	42.38	45.35	12.28	38.57	51.92	9.51	0.788
Apolipoprotein C-III (*APOC3*) (C1100T)	58.61	34.06	7.33	54.48	38.21	7.31	0.262
Apolipoprotein C-III (*APOC3*) (C3175G)	83.56	15.84	0.59	83.36	15.90	0.73	0.973
Apolipoprotein C-III (*APOC3*) (T3206G)	43.96	43.76	12.28	38.94	47.90	13.16	0.141
Apolipoprotein E (*APOE*) (no. of epsilon 4 alleles)	72.82	25.79	1.39	72.06	24.82	3.13	0.495
Peroxisome proliferative activated receptor, γ (*PPARG*) (pro12ala)	75.45	23.76	0.79	79.34	18.83	1.83	0.492
Adrenergic, β-3-, receptor (*ADRB3*) (trp64arg)	85.74	13.86	0.40	87.57	12.43	0.00	0.521
Lipase, hepatic (*LIPC*) (C(-480)T)	66.14	30.89	2.97	62.16	32.72	5.12	0.149
Lipoprotein lipase (*LPL*) (T(-93)G)	97.62	2.38	0.00	97.44	2.56	0.00	0.675
Lipoprotein lipase (*LPL*) (asp9asn)	97.62	2.38	0.00	97.62	2.38	0.00	0.785
Lipoprotein lipase (*LPL*) (asn291ser)	97.03	2.97	0.00	96.34	3.66	0.00	0.434
Lipoprotein lipase (*LPL*) (ser447term)	79.60	19.60	0.79	80.44	19.01	0.55	0.614
Paraoxonase 1 (*PON1*) (met55leu)	13.07	46.53	40.40	15.72	43.88	40.40	0.372
Paraoxonase 1 (*PON1*) (gln192arg)	51.68	41.78	6.53	53.20	37.66	9.14	0.875
Paraoxonase 2 (*PON2*) (ser311cys)	56.24	37.82	5.94	58.32	36.93	4.75	0.314
Low density lipoprotein receptor (*LDLR*) (exon 18 NcoI +/−)	45.13	44.73	10.14	46.89	42.67	10.44	0.885
Cholesteryl ester transfer protein, plasma (*CETP*) (C(-631)A)	84.75	15.05	0.20	83.55	16.09	0.37	0.485
Colesteryl ester transfer protein, plasma (*CETP*) (C(-629)A)	23.96	48.32	27.72	26.14	53.56	20.29	0.025*
Colesteryl ester transfer protein, plasma (*CETP*) (ile405val)	44.36	43.37	12.28	44.06	45.34	10.60	0.707
5,10-Methylenetetrahydrofolate reductase (*MTHFR*) (C677T)	42.38	45.54	12.08	43.14	44.97	11.88	1.000
Nitric oxide synthase 3 (endothelial cell) (*NOS3*) (A(-922)G)	38.61	47.92	13.47	37.11	48.26	14.63	0.498
Nitric oxide synthase 3 (endothelial cell) (*NOS3*) (C(-690)T)	83.76	15.05	1.19	79.89	19.20	0.91	0.090
Nitric oxide synthase 3 (endothelial cell) (*NOS3*) (glu298asp)	42.38	44.75	12.87	43.38	44.30	12.32	0.758
Angiotensin I converting enzyme (peptidyl-dipeptidase A) 1 (*ACE*) (intron 16 Ins/Del)	24.16	47.92	27.92	21.02	48.99	29.98	0.217
Angiotensin II receptor, type 1 (*AGTR1*) (A1166C)	52.87	38.81	8.32	47.80	43.04	9.16	0.229
Angiotensinogen (*AGT*) (met235thr)	39.01	44.75	16.24	38.76	46.07	15.17	0.953
Natriuretic peptide precursor A (*NPPA*) (G664A)	88.51	11.49	0.00	91.22	8.59	0.18	0.207
Natriuretic peptide precursor A (*NPPA*) (T2238C)	70.63	26.59	2.78	69.06	28.91	2.03	0.908
Adducin 1 (α) (*ADD1*) (gly460trp)	60.99	33.66	5.35	67.64	29.80	2.56	0.006*
Sodium channel, nonvoltage-gated 1 α (*SCNN1A*) (trp493arg)	94.46	5.54	0.00	94.70	5.30	0.00	0.958
Sodium channel, nonvoltage-gated 1 α (*SCNN1A*) (ala663thr)	43.56	43.76	12.67	44.69	43.59	11.72	0.849
Guanine nucleotide binding protein (G protein), β polypeptide (*GNB3*) (C825T)	45.94	44.36	9.70	47.62	42.67	9.71	0.698
Adrenergic, β-2-, receptor, surface (*ADRB2*) (arg16gly)	11.88	47.52	40.59	11.58	49.82	38.60	0.793
Adrenergic, β-2-, receptor, surface (*ADRB2*) (gln27glu)	29.31	51.49	19.21	27.47	53.11	19.41	0.484
Matrix metalloproteinase 3 (stromelysin 1, progelatinase) (*MMP3*) (5A/6A)	26.93	48.32	24.75	24.68	51.19	24.13	0.723
Selectin E (endothelial adhesion molecule 1) (*SELE*) (ser128arg)	82.38	16.63	0.99	81.72	17.73	0.55	0.881
Selectin E (endothelial adhesion molecule 1) (*SELE*) (leu554phe)	92.67	6.73	0.59	89.76	10.24	0.00	0.264
Intercellular adhesion molecule 1 (CD54), human rhinovirus receptor (*ICAM1*) (gly214arg)	77.62	20.79	1.58	77.51	21.39	1.10	0.799
Tumor necrosis factor (*TNF* superfamily, member 2) (*TNF*) (G(-376)A)	97.43	2.57	0.00	96.34	3.66	0.00	0.276
Tumor necrosis factor (*TNF* superfamily, member 2) (*TNF*) (G(-308)A)	66.73	28.91	4.36	66.67	29.85	3.48	0.742
Tumor necrosis factor (*TNF* superfamily, member 2) (*TNF*) (G(-238)A)	87.13	12.28	0.59	88.12	11.88	0.00	0.504
Lymphotoxin α (*TNF* superfamily, member 1) (*LTA*) (thr26asn)	43.56	44.16	12.28	41.32	45.16	13.53	0.446

Table 2 (*continued*)

Gene and polymorphism	Frequency in controls (%), $N = 505$			Frequency in cases (%), $N = 547$			*P*-value adjusted for age and sex
	AA	AB	BB	AA	AB	BB	
Coagulation factor II (thrombin) (*F2*) (G20210A)	96.63	3.37	0.00	97.81	2.19	0.00	0.205
Coagulation factor V (proaccelerin, labile factor) (*F5*) (arg506gln)	95.45	4.55	0.00	95.25	4.75	0.00	0.983
Coagulation factor VII (serum prothrombin conversion accelerator) (*F7*) ((-323) 0/10bp ins)	77.62	21.78	0.59	76.78	22.30	0.91	0.645
Coagulation factor VII (serum prothrombin conversion accelerator) (*F7*) (arg353gln)	80.16	19.44	0.40	78.98	20.29	0.73	0.595
Serine proteinase inhibitor (plasminogen activator inhibitor type 1), member 1 (*SERPINE1*) (5G/4G)	20.99	46.93	32.08	19.74	51.19	29.07	0.662
Serine proteinase inhibitor (plasminogen activator inhibitor type 1), member 1 (*SERPINE1*) (G11053T)	20.59	45.94	33.47	18.68	50.00	31.32	0.930
Fibrinogen, B β polypeptide (*FGB*) (G(-455)A)	67.13	29.31	3.56	69.78	27.47	2.75	0.175
Integrin α-2 precursor (Platelet membrane glycoprotein Ia, GPIa) (*ITGA2*) (G873A)	32.87	49.31	17.82	36.70	46.79	16.51	0.158
Integrin, β 3 (platelet glycoprotein IIIa, GPIIIa) (*ITGB3*) (leu33pro)	65.74	31.68	2.57	69.78	27.66	2.56	0.170

A refers to baseline allele (usually the commonest allele) and B to variant allele, AA refers to baseline homozygote, AB to heterozygote and BB to variant homozygote.

*Statistically significant at $P < 0.05$.

Table 3 Comparison of dominant, recessive and additive genetic models for polymorphisms associated with myocardial infarction in the screening analysis

Gene and polymorphism	Dominant model[b]		Recessive model[b]		Additive model[b]	
	Odds ratio (95% CI[a])	*P*-value	Odds ratio (95% CI)	*P*-value	Odds ratio (95% CI)	*P*-value
CETP (C-629A)	0.87 (0.66—1.16)	0.35	0.67 (0.50—0.90)	0.006	0.82 (0.68—0.97)	0.025
ADD1 (gly460trp)	0.74 (0.57—0.96)	0.021	0.46 (0.23—0.89)	0.022	0.73 (0.59—0.91)	0.006

The dominant genetic model compares individuals with one or more polymorphic alleles (AB and BB genotypes combined in Table 2) with a baseline group with no polymorphic alleles (AA).

The additive genetic model assumes that there is a linear gradient in risk between the AA, AB and BB genotypes (BB genotype baseline). This is equivalent to a comparison of the B allele versus the A allele (baseline).

The recessive genetic model compares the BB genotype with the combined AB and AA genotypes (combined AA and AB genotypes form the baseline group).

[a] CI denotes confidence interval.

[b] Each model was adjusted for age and sex.

High plasma CETP is associated with reduced HDL, which is a strong risk factor for atherosclerosis.[31] Recent data strongly suggest that the C-629A polymorphism is functional, with the −629A allele associated with reduced promoter activity, decreased CETP mass and increased HDL levels.[32,33] Therefore, the finding of a protective effect of the −629A allele against the risk of myocardial infarction is consistent with its postulated functional effect. However, it should be noted that although the CETP locus has consistently been shown to be associated with reduced risk of myocardial infarction, there are discrepant findings as to which of the polymorphisms is associated or responsible for the effect[33–36] and it is possible that −629A allele is a marker for some other variant at the locus.

We also found a significant protective effect on myocardial infarction risk of the *PON1*/*PON2* haplotype comprising met55 and gln192 in *PON1* and cys311 in *PON2*. The paraoxonases are glycoproteins located on the surface of HDL. They are believed to play a pivotal role in preventing low density lipoprotein (LDL) oxidation,[37] which may be one mechanism by which HDL is protective against atherosclerosis. The paraoxonase genes are located as a tightly linked multigene family (also comprising *PON3*) at 7q21-q22.[38] There is evidence that genetic variation in the paraoxonase genes influences paraoxonase activity. Thus, compared with the 192arg and the 55leu alleles, both the *PON1* 192gln and 55met alleles have been shown to be associated with higher paraoxonase activity and to be more efficient at inhibiting the oxidation of LDL.[39,40] Correspondingly, several studies have shown these alleles to be associated with decreased risk of coronary heart disease.[41,42] However, the findings have not always been consistent[43,44] and recent studies have sought to identify haplotypes that may be more predictive.[45,46] Sanghera et al. found a significant interaction between the *PON1* gln192arg and the *PON2* cys311ser polymorphisms in predicting risk of

Table 4 Estimated haplotype frequencies and odds ratios for association with myocardial infarction

Haplotype	Controls $n = 505$	Cases $n = 547$	Odds ratio (95% CI)
NPPA G664A, T2238C			
GT	789.89	864.18	1.00
GC	162.11	180.82	1.02 (0.76–1.27)
AT	57.81	49.00	0.79 (0.59–1.19)
APOA (*LPA*): C93T, G121A			
CG	677.66	753.35	1.00
CA	177.34	196.29	1.00 (0.75–1.50)
TG	155.00	144.35	0.84 (0.63–1.26)
*ADR*β2: arg16gly, gln27glu			
arg gln	360.00	400.08	1.00
gly gln	196.00	191.16	0.88 (0.66–1.33)
gly glu	454.00	502.77	1.00 (0.75–1.50)
APOAIV: thr347ser, gln360his			
thr gln	728.00	792.01	1.00
thr his	188.99	207.99	1.01 (0.76–1.26)
ser gln	92.99	93.99	0.93 (0.70–1.40)
APOC3: C(-641)A, C(-482)T, T(-455)C, C1100T, C3175G, T3206G			
CCTCCT	460.31	480.34	1.00
CCTCCG	66.44	57.49	0.83 (0.63–1.25)
CCTTCG	98.88	144.01	1.41 (1.06–1.76)*
ACCCCT	72.15	80.18	1.07 (0.80–1.34)
ACCCCG	32.38	31.65	0.94 (0.71–1.41)
ATCCCT	106.07	111.47	1.00 (0.75–1.26)
ATCCCG	19.91	36.26	1.71 (1.28–2.14)*
ATCTCG	34.48	35.15	1.00 (0.76–1.26)
ATCTGG	80.52	89.02	1.07 (0.80–1.33)
CETP: C(-631)A, C(-629)A, ile405val			
CC ile	338.05	408.60	1.00
CC val	69.95	78.40	0.83 (0.62–1.25)
CA ile	265.55	252.28	0.77 (0.57–1.15)
CA val	258.45	262.72	0.83 (0.62–1.25)
AC ile	63.40	69.12	0.82 (0.62–1.23)
AC val	14.60	22.88	1.67 (0.84–4.60)
SELE: ser128arg, leu554phe			
ser leu	876.08	935.00	1.00
ser phe	39.92	56.00	1.31 (0.98–2.30)
arg leu	93.92	103.00	1.03 (0.77–1.28)
SCNN1A: trp493arg, ala663thr			
trpala	655.48	723.06	1.00
trp thr	326.51	341.94	0.95 (0.71–1.42)
arg thr	22.49	24.71	1.00 (0.75–1.50)
F7: arg353gln, (-323) 10bp ins/del			
arg ins	894.00	959.97	1.00
gln ins	13.06	15.03	1.08 (0.54–2.42)
gln del	102.94	116.97	1.07 (0.80–1.33)
LPL: T(-93)G, asp9asn, asn291ser, ser447term			
T asp asn ser	878.02	951.76	1.00
T asp asn term	104.98	109.00	0.96 (0.72–1.44)
T asn ser ser	13.63	19.25	1.30 (0.65–2.93)
G asn asn ser	11.35	12.25	0.99 (0.50–2.49)
NOS3: A(-922)G, C(-690)T, glu298asp			
AC glu	525.99	555.01	1.00
AC asp	106.01	114.99	1.03 (0.77–1.28)
GC glu	128.01	159.74	1.19 (0.89 –1.48)
GC asp	161.99	149.26	0.88 (0.66–1.31)
GT asp	88.00	112.35	1.22 (0.91–1.52)

Table 4 (*continued*)

Haplotype	Controls $n = 505$	Cases $n = 547$	Odds ratio (95% CI)
SERPINE1: 5G/4G, G11053T			
4G G	424.00	456.09	1.00
5G G	15.99	21.93	1.28 (0.64–2.88)
5G T	433.00	474.07	1.02 (0.76–1.28)
4G T	136.99	141.91	0.97 (0.72–1.45)
PON1 met55leu, gln192arg, PON2 ser311cys			
met gln ser	308.24	370.56	1.00
met gln cys	55.12	34.18	0.52 (0.39–0.77)*
leu gln ser	260.42	276.28	0.88 (0.66–1.32)
leu gln cys	109.22	106.98	0.81 (0.61–1.22)
leu arg ser	190.34	193.16	0.84 (0.63–1.27)
leu arg cys	83.02	105.59	1.06 (0.79–1.32)
TNF G(-376)A, G(-308)A, G(-238)A, *LTA*: thr26asn			
GGG thr	595.00	634.01	1.00
GGG asn	157.00	193.31	1.16 (0.87–1.44)
GGA thr	55.00	45.00	0.77 (0.58–1.15)
GAG asn	190.00	201.68	0.99 (0.75–1.50)
GAG thr	13.00	19.99	1.47 (0.72–3.25)

Haplotypes with frequencies of <2% not shown in table.
*Significant at $P < 0.05$.

coronary heart disease in Asian Indians.[41] Increased risk was restricted to carriers of both the *PON1* arg and the *PON2* ser alleles. Our findings are consistent with this and indicate that a haplotype that also includes the met55 allele provides the best discrimination in identifying risk associated with the PON locus.

Haplotype analysis for the polymorphisms at positions −641, −482, −455, 1100, 3175 (often referred to as the SstI S1/S2 site), and 3206 in the *APOC3* gene identified two haplotypes associated with increased risk for myocardial infarction. APOC3 is a glycoprotein synthesised in the liver and intestine that is physically associated with very low density lipoprotein (VLDL), chylomicron remnants and high density lipoprotein (HDL).[47] Previous studies have associated single polymorphisms in *APOC3* with risk of coronary heart disease[48] and altered plasma triglyceride levels.[49,50] Dammerman et al.[49] found a 3-fold increased risk of hypertriglyceridemia associated with a haplotype that included the −482T and SstI S2 variants, while Surguchov et al.[50] reported a 4-fold increased risk of hypertriglyceridemia associated with a haplotype that included −641A, −482T, −455C and SstI S2 variants. This latter haplotype shares the same variants at the first three positions as the ATCCCG haplotype that we found to be associated with myocardial infarction risk, but we did not observe association of the more common ATCTGG (i.e., also S2) haplotype with risk. Therefore, some *APOC3* haplotypes may influence myocardial infarction risk via their effect on plasma triglyceride levels while other haplotypes may play a role via other mechanisms.

There are several notable polymorphisms, which have been associated with risk of myocardial infarction in other studies that did not show any association in our study. These include the Thr26Asn polymorphism in the lymhotoxin-α gene (LTA)[51] and the 5G/4G and 5A/6A polymorphisms in the plasminogen activator inhibitor type 1 and stromelysin 1 genes, respectively.[3] Some of the differences may relate to the nature of the population studied. Specifically, allele frequencies are known to vary significantly between Japanese[3,51] and Caucasian populations. Methods of recruitment also need to be considered. All our cases were recruited after admission into hospital. If a polymorphism has an effect, not only on risk of event but also on acute survival, then this could obscure any association. Finally, although our study was adequately powered to detect a moderate effect of a gene with a minor allele frequency similar to that of the angiotensin converting enzyme polymorphism, the power is reduced with lower allele frequencies. Thus, the lack of replication of associations between less common polymorphisms and myocardial infarction that have been reported in previous studies in our study alone would not rule out the possibility of such an association.

We adopted a range of methods to ensure that the study maintained high degree of scientific rigour, including detailed genotype and relationship error checking, testing for hidden population stratification and checking the consistency of haplotype frequency estimates. Furthermore, we minimised the number of hypotheses tested using an initial screen based on an additive genetic model. However, any genetic association study including such a large number of polymorphisms should be regarded as exploratory in character and our results must be interpreted accordingly. Hence, corrections for multiple testing, which are considered appropriate for hypothesis-testing studies, but over-conservative for exploratory studies, were not applied. However, the issue of correction for multiple testing is becoming increasingly important as multilocus analyses become popular. This is illustrated by the fact that despite our study being one of the larger studies to date, if we had applied a Bonferroni correction, no individual polymorphisms would have reached statistical signifi-

cance ($P < 0.001$) and only one haplotype (paroxonase) would have reached borderline significance. Indeed, a study of over 5000 cases and 5000 controls would have been required to reach nominal significance ($P < 0.05$) after correction, to identify a polymorphism with a 50% allele frequency associated with an OR of 1.2, when 50 independent tests are being undertaken. The study sizes required increase dramatically for less common alleles. This highlights the fact that apart from much larger studies, efficient strategies for pooling of data from different studies will be required as we enter a new era of genetic association analyses for complex traits.

In conclusion, high-throughput genotyping assays will be in common use in association studies in the near future. We have shown here that such studies are feasible and can reveal genetic polymorphisms that warrant testing in further studies. Furthermore, we have demonstrated the feasibility and utility of using a number of polymorphisms in linkage disequilibrium to assess the association between specific haplotypes and myocardial infarction in a study of unrelated individuals. This may increase the power to identify disease-associated variants, a crucial step towards the eventual clinical application of new predictive markers.

Acknowledgments

We thank all participants in the study. We also thank Michael Grow and Arkadiy Silbergleit for their efforts in developing and optimising the genotyping assays. NJS and RS act as guarantors for the Leicester and Sheffield data, respectively. NJS is supported by the British Foundation.

References

1. Marenberg ME, Risch N, Berkman LF et al. Genetic susceptibility to death from coronary heart disease in a study of twins. *N Engl J Med* 1994;**330**:1041–6.
2. Kuivenhoven JA, Jukema JW, Zwinderman AH et al. The role of a common variant of the cholesteryl ester transfer protein gene in the progression of coronary atherosclerosis. The regression growth evaluation statin study group. *N Engl J Med* 1998;**338**:86–93.
3. Yamada Y, Izawa H, Ichihara S et al. Prediction of the risk of myocardial infarction from polymorphisms in candidate genes. *N Engl J Med* 2002;**347**:1916–23.
4. Keavney B, McKenzie C, Parish S et al. Large-scale test of hypothesised associations between the angiotensin-converting-enzyme insertion/deletion polymorphism and myocardial infarction in about 5000 cases and 6000 controls. International studies of infarct survival (ISIS) collaborators. *Lancet* 2000;**355**:434–42.
5. Cardon LR, Bell JI. Association study designs for complex diseases. *Nat Rev Genet* 2001;**2**:91–9.
6. Phillips MS, Lawrence R, Sachidanandam R et al. Chromosome-wide distribution of haplotype blocks and the role of recombination hot spots. *Nat Genet* 2003;**33**:382–7.
7. Gabriel SB, Schaffner SF, Nguyen H et al. The structure of haplotype blocks in the human genome. *Science* 2002;**296**:2225–9.
8. Mander AP. Haplotype analysis in population-based association studies. *Stata J* 2001;**1**:58–75.
9. Schneider S, Roessli D, Excoffier L. Arlequin: A software for population genetics data analysis. In: 2.000 ed. Geneva, Switzerland: Genetics and Biometry Laboratory, University of Geneva; 2000.
10. Clayton D. SNPHAP: A program for estimating frequencies of large haplotypes of SNPs. In. 1.0 ed. Cambridge: Department of Medical Genetics, Cambridge Institute for Medical Research; 2002.
11. Stephens M, Smith NJ, Donnelly P. A new statistical method for haplotype reconstruction from population data. *Am J Hum Genet* 2001;**68**:978–89.
12. Schaid DJ, Rowland CM, Tines DE et al. Score tests for association between traits and haplotypes when linkage phase is ambiguous. *Am J Hum Genet* 2002;**70**:425–34.
13. Cheng S, Grow MA, Pallaud C et al. A multilocus genotyping assay for candidate markers of cardiovascular disease risk. *Genome Res* 1999; **9**:936–49.
14. Steeds R, Adams M, Smith P et al. Distribution of tissue plasminogen activator insertion/deletion polymorphism in myocardial infarction and control subjects. *Thromb Haemost* 1998;**79**:980–4.
15. WHO. Nomenclature and criteria for diagnosis of ischemic heart disease. *Circulation* 1979;**59**:607–9 [Report of the Joint International Society and Federation of Cardiology/World Health Organization task force on standardization of clinical nomenclature].
16. Samani NJ, Martin DS, Brack M et al. Apolipoprotein E polymorphism does not predict risk of restenosis after coronary angioplasty. *Atherosclerosis* 1996;**125**:209–16.
17. Hoppe C, Cheng S, Grow M et al. A novel multilocus genotyping assay to identify genetic predictors of stroke in sickle cell anaemia. *Br J Haematol* 2001;**114**:718–20.
18. Abecasis GR, Cherny SS, Cookson WO et al. GRR: graphical representation of relationship errors. *Bioinformatics* 2001;**17**:742–3.
19. Davey Smith G, Ebrahim S. Mendelian randomization: can genetic epidemiology contribute to understanding environmental determinants of disease? *Int J Epidemiol* 2003;**32**:1–22.
20. Bitti PP, Murgia BS, Ticca A et al. Association between the ancestral haplotype HLA A30B18DR3 and multiple sclerosis in central Sardinia. *Genet Epidemiol* 2001;**20**:271–83.
21. Slatkin M, Excoffier L. Testing for linkage disequilibrium in genotypic data using the expectation–maximization algorithm. *Heredity* 1996; **76**(Pt 4):377–83.
22. Excoffier L, Slatkin M. Maximum-likelihood estimation of molecular haplotype frequencies in a diploid population. *Mol Biol Evol* 1995;**12**: 921–7.
23. Agresti A. Modelling patterns of agreement and disagreement. *Stat Methods Med Res* 1992;**1**:201–18.
24. McCullagh P, Nelder J. Generalized linear models. 2nd ed. London: Chapman and Hall; 1989.
25. Pritchard JK, Rosenberg NA. Use of unlinked genetic markers to detect population stratification in association studies. *Am J Hum Genet* 1999;**65**:220–8.
26. Topol EJ, McCarthy J, Gabriel S, et al. Single nucleotide polymorphisms in multiple novel thrombospondin genes may be associated with familial premature myocardial infarction. *Circulation* 2001;**104**: 2641–4.
27. Joshi R, Gilligan DM, Otto E, et al. Primary structure and domain organization of human alpha and beta adducin. *J Cell Biol* 1991;**115**: 665–75.
28. Matsuoka Y, Hughes CA, Bennett V. Adducin regulation. Definition of the calmodulin-binding domain and sites of phosphorylation by protein kinases A and C. *J Biol Chem* 1996;**271**:25157–66.
29. Manunta P, Cusi D, Barlassina C, et al. Alpha-adducin polymorphisms and renal sodium handling in essential hypertensive patients. *Kidney Int* 1998;**53**:1471–8.
30. Psaty BM, Smith NL, Heckbert SR, et al. Diuretic therapy, the alpha-adducin gene variant, and the risk of myocardial infarction or stroke in persons with treated hypertension. *JAMA* 2002;**287**: 1680–9.
31. Hannuksela ML, Liinamaa MJ, Kesaniemi YA, et al. Relation of polymorphisms in the cholesteryl ester transfer protein gene to transfer protein activity and plasma lipoprotein levels in alcohol drinkers. *Atherosclerosis* 1994;**110**:35–44.
32. Dachet C, Poirier O, Cambien F, et al. New functional promoter polymorphism, CETP/-629, in cholesteryl ester transfer protein (CETP) gene related to CETP mass and high density lipoprotein cholesterol levels: role of Sp1/Sp3 in transcriptional regulation. *Arterioscler Thromb Vasc Biol* 2000;**20**:507–15.
33. Klerkx AH, Tanck MW, Kastelein JJ, et al. Haplotype analysis of the CETP gene: not TaqIB, but the closely linked -629C–>A polymor-

phism and a novel promoter variant are independently associated with CETP concentration. *Hum Mol Genet* 2003;**12**:111–23.
34. Eiriksdottir G, Bolla MK, Thorsson B, et al. The −629C>A polymorphism in the CETP gene does not explain the association of TaqIB polymorphism with risk and age of myocardial infarction in Icelandic men. *Atherosclerosis* 2001;**159**:187–92.
35. Corbex M, Poirier O, Fumeron F, et al. Extensive association analysis between the CETP gene and coronary heart disease phenotypes reveals several putative functional polymorphisms and gene-environment interaction. *Genet Epidemiol* 2000;**19**:64–80.
36. Freeman DJ, Samani NJ, Wilson V, et al. A polymorphism of the cholesteryl ester transfer protein gene predicts cardiovascular events in non-smokers in the west of Scotland coronary prevention study. *Eur Heart J* 2003;**24**:1833–42.
37. Mackness MI, Durrington PN. HDL, its enzymes and its potential to influence lipid peroxidation. *Atherosclerosis* 1995;**115**:243–53.
38. Humbert R, Adler DA, Disteche CM, et al. The molecular basis of the human serum paraoxonase activity polymorphism. *Nat Genet* 1993;**3**:73–6.
39. Mackness B, Mackness MI, Arrol S, et al. Effect of the human serum paraoxonase 55 and 192 genetic polymorphisms on the protection by high density lipoprotein against low density lipoprotein oxidative modification. *FEBS Lett* 1998;**423**:57–60.
40. Leviev I, Negro F, James RW. Two alleles of the human paraoxonase gene produce different amounts of mRNA. An explanation for differences in serum concentrations of paraoxonase associated with the (Leu–Met54) polymorphism. *Arterioscler Thromb Vasc Biol* 1997;**17**:2935–9.
41. Sanghera DK, Saha N, Aston CE, et al. Genetic polymorphism of paraoxonase and the risk of coronary heart disease. *Arterioscler Thromb Vasc Biol* 1997;**17**:1067–73.
42. Ruiz J, Blanche H, James RW, et al. Gln–Arg192 polymorphism of paraoxonase and coronary heart disease in type 2 diabetes. *Lancet* 1995;**346**:869–72.
43. Turban S, Fuentes F, Ferlic L, et al. A prospective study of paraoxonase gene Q/R192 polymorphism and severity, progression and regression of coronary atherosclerosis, plasma lipid levels, clinical events and response to fluvastatin. *Atherosclerosis* 2001;**154**:633–40.
44. Gardemann A, Philipp M, Hess K, et al. The paraoxonase Leu–Met54 and Gln–Arg191 gene polymorphisms are not associated with the risk of coronary heart disease. *Atherosclerosis* 2000;**152**:421–31.
45. >Sanghera DK, Aston CE, Saha N, et al. DNA polymorphisms in two paraoxonase genes (PON1 and PON2) are associated with the risk of coronary heart disease. *Am J Hum Genet* 1998;**62**:36–44.
46. Wang X, Fan Z, Huang J, et al. Extensive association analysis between polymorphisms of PON gene cluster with coronary heart disease in Chinese Han population. *Arterioscler Thromb Vasc Biol* 2003;**23**:328–34.
47. Breslow J. Familial disorders of high density lipoprotein metabolism. In: Scriver C, Beaudet A, Sly W, Valle D, editors. The metabolic basis of inherited disease. New York: McGraw-Hill; 1995. p. 1251–66.
48. Peacock RE, Hamsten A, Johansson J, et al. Associations of genotypes at the apolipoprotein AI-CIII-AIV, apolipoprotein B and lipoprotein lipase gene loci with coronary atherosclerosis and high density lipoprotein subclasses. *Clin Genet* 1994;**46**:273–82.
49. Dammerman M, Sandkuijl LA, Halaas JL, et al. An apolipoprotein CIII haplotype protective against hypertriglyceridemia is specified by promoter and 3′ untranslated region polymorphisms. *Proc Natl Acad Sci USA* 1993;**90**:4562–6.
50. Surguchov AP, Page GP, Smith L, et al. Polymorphic markers in apolipoprotein C-III gene flanking regions and hypertriglyceridemia. *Arterioscler Thromb Vasc Biol* 1996;**16**:941–7.
51. Ozaki K, Ohnishi Y, Iida A, et al. Functional SNPs in the lymphotoxin-alpha gene that are associated with susceptibility to myocardial infarction. *Nat Genet* 2002;**32**:650–4.