A two-step genetic study on quantitative precursors of coronary artery disease in a homogeneous Indian population: Case–control association discovery and validation by transmission-disequilibrium test

SANJUKTA MALLIK[1] and PARTHA P MAJUMDER[1,2,3,*]

[1]*Human Genetics Unit, Indian Statistical Institute, Kolkata, India*

[2]*TCG-ISI Centre of Population Genomics, Kolkata, India*

[3]*National Institute of Biomedical Genomics, Kalyani, India*

**Corresponding author (Email, ppm@isical.ac.in)*

In spite of its strong familiality, gene identification for coronary artery disease (CAD) has not yielded a consistent picture. One major reason for this is that families or cases and controls were not recruited from a homogeneous population. We, therefore, attempted to map genes underlying 10 quantitative traits (QTs) that are known precursors of CAD in a homogeneous population (Marwari) of India. The QTs are apolipoprotein B (ApoB), C-reactive protein (CRP), fibrinogen (FBG), homocysteine (HCY), lipoprotein (a) (LPA), cholesterol – total (CHOL-T), cholesterol – HDL (CHOL-H), cholesterol – LDL (CHOL-L), cholesterol – VLDL (CHOL-V) and triglyceride (TG). We assayed 209 SNPs in 31 genes among members of Marwari families. After log-transformation and covariate-adjustment of the QTs, a two-step analysis was performed. In Step-1, data on unrelated individuals were analysed for association with the SNPs. In Step-2, for validation of Step-1 results, a quantitative transmission-disequilibrium test on parent–offspring data was performed for each SNP found to be significantly associated with a QT in Step-1 on an independent sample set drawn from the same population. Statistically significant results found for the various QTs and SNPs were: rs3774933, rs230528, rs230521, rs1005819 and rs1609798 (intronic, NFKB1) with APOB; rs5361 (Missense, R>S, *SELE*) and rs4648004 (Intronic, *NFKB1*) with FBG; rs4220 (Missense, K>R, *FGB*) with HCY; and rs3025035 (Intronic, *VEGFA*) with CHOL-H. SNPs in *SELE*, *VEGFA*, *FGB* and *NFKB1* genes impact significantly on levels of quantitative precursors of CAD in Marwaris.

[Mallik S and Majumder PP 2011 A two-step genetic study on quantitative precursors of coronary artery disease in a homogeneous Indian population: Case–control association discovery and validation by transmission-disequilibrium test. *J. Biosci.* **36** 857–868] **DOI** 10.1007/s12038-011-9148-4

1. Introduction

The etiology of coronary artery disease (CAD) is complex. Genetic and environmental factors interact in the precipitation of the disease. It is known that multiple biochemical processes and pathways are involved in CAD (Scheuner 2001). These include pathways pertaining to lipid and apolipoprotein metabolism, inflammatory response, endothelial function, platelet function, thrombosis, fibrinolysis, homocysteine metabolism, insulin sensitivity and blood

Keywords. Candidate gene; coronary artery disease; genomic polymorphism; quantitative trait

Abbreviations used: ApoB, apolipoprotein B; CAD, coronary artery disease; CHOL-H, cholesterol – HDL; CHOL-L, cholesterol – LDL; CHOL-T, cholesterol – total; CHOL-V, cholesterol – VLDL; CRP, C-reactive protein; EGF, epidermal growth factor; ESE, exonic splicing enhancer; FBG, fibrinogen; GWA, genome-wide association; HCY, homocysteine; LPA, lipoprotein (a); MI, myocardial infarction; QT, quantitative trait; TF, transcription factor; TG, triglycerides

Supplementary materials pertaining to this article are available on the *Journal of Biosciences* Website at *http://www.ias.ac.in/jbiosci/dec2011/supp/mallik.pdf*

pressure regulation. Each of these biochemical pathways comprises enzymes, receptors and ligands, which are encoded by genes. Variations in these genes can alter the function of the constituents within a metabolic pathway. While many new genes are being identified by large genome-wide association (GWA) studies (The CAD Genetics Consortium 2011; Reilly *et al.* 2011), some other similar large GWA studies are failing to even find a single significant association (Lettre *et al.* 2011). These facts underscore the complexity of CAD and perhaps point to the need of carefully designed studies.

There have been many studies to identify genetic factors underlying CAD. However, the results have not been consistent (Hingorani 2004; Peyser 1997). One of the primary reasons for the lack of consistency of results is the lack of homogeneity, both genetic and environmental, of population groups from which patients, controls or families were recruited (Hingorani 2004). We have carried out this study in a young, isolated population group, which, we believe, is helpful in many ways. First, there is reduced environmental heterogeneity in an isolated population that enhances genetic effects on a complex phenotype, such as CAD, relative to environmental effects. Second, members of an isolated population are genetically less diverse, which implies that affected individuals in such a population are more likely to have similar susceptibility genes. Third, because younger populations are likely to have fewer meiotic events separating the individuals in them, the sizes of the chromosomal segments that are shared by individuals with the same phenotype are likely to be larger. This is helpful in gene-mapping studies because shared genomic segments are likely to harbor variants that influence the expression of the phenotype under consideration.

Heritable clinical end points, such as CAD, are usually binary (scored as affected or unaffected), but almost invariably have quantitative precursor states that are also highly heritable. Quantitative precursors, such as cholesterol levels, are known to determine the end-point risk of CAD. It is, therefore, of considerable interest to map the genes underlying such quantitative traits (QTs), because these genes are the ultimate determinants of the clinical end point, along with other precipitating environmental factors (Majumder and Ghosh 2005).

With a view to mapping genes underlying quantitative precursors of CAD, we have studied 10 QTs in a young, isolated population, the Marwaris of Kolkata (India). We chose this population because our earlier genetic epidemiological studies had shown that members of this community have high prevalence of hypertension and adverse lipid profiles (Majumder *et al.* 1996).

In this study, we chose those QTs that are known to be important precursors of CAD. Notable among these are levels of C-reactive protein (CRP), homocysteine (HCY), LDL cholesterol (Kelly *et al.* 2002; Brull *et al.* 2003; Falchi *et al.* 2004). Circulating highly sensitive CRP has emerged as one of the most powerful independent predictors of cardiovascular disease risk and cardiovascular death (Tice *et al.* 2003). CRP appears to be involved in all stages of atherogenesis, since it is present in atherosclerotic lesions but not in the normal vessel wall (Torzewski *et al.* 2000). CRP potently down-regulates endothelial nitric oxide (NO) synthase transcription and destabilizes endothelial NO synthase mRNA *in vitro*, resulting in decreased basal and lowered release of stimulated NO, a key endothelium-derived relaxing factor (Verma *et al.* 2002b). In a synchronous fashion, CRP stimulates endothelin-1 and interleukin 6 release from endothelial cells (Verma *et al.* 2002a), and decreases the production of the potent vasodilator prostacyclin (Venugopal *et al.* 2003), shifting the balance towards endothelial dysfunction. CRP plays a role in the destabilization of the atheroma's fibrous cap by stimulating matrix metalloproteinase 1 release, which degrades collagen and other matrix scaffolding proteins (Williams *et al.* 2004). Elevated level of HCY is an independent risk factor for cardiovascular disease (Wilhelmsen *et al.* 1984). HCY damages vascular endothelial cells, resulting in lipid deposition and plaque formation. An adverse lipid profile comprising low level of HDL cholesterol (CHOL-H) and high levels of LDL cholesterol (CHOL-L), VLDL cholesterol (CHOL-V) and triglyceride (TG), plays a crucial role for the development of CAD. High cholesterol level in blood leads to deposition of the molecule in arteries, which is the primary step initiating atherosclerosis.

In addition, fibrinogen (FBG) has been found to be a strong predictor of primary and secondary ischemic coronary events in many epidemiological studies (Haines *et al.* 1983; Wilhelmsen *et al.* 1984; Heinrich *et al.* 1994; Meade *et al.* 1993); high fibrinogen levels are associated with an increased risk of disease (Ernst and Resch 1993; Danesh *et al.* 1998; Maresca *et al.* 1999). Further, SNPs located in genes *FGB* and *TCF1* (transcription factor 1) have been found to be associated with fibrinogen levels (Robinson *et al.* 1995; Liu *et al.* 2001; Soria *et al.* 2005). Apolipoprotein B (the main component of low-density lipoproteins) level has been found to be associated with CAD and increased cholesterol level in the blood (Ichinose and Kuriyama 1995). TG level is an independent risk factor for coronary heart disease (Hokanson and Austin 1996). Numerous studies have shown that TG levels are at least partially controlled by genetic factors (Austin *et al.* 1987; Mitchell *et al.* 1996; Pérusse *et al.* 1997; Duggirala *et al.* 2000; Sonnenberg *et al.* 2004). Duggirala *et al.* (2000) reported significant linkage of TG levels to markers in the chromosomal region 15q11-q13 in Mexican American families, and to the region 7q35-q36 (Sonnenberg *et al.* 2004). Lipoprotein (a) (LPA) serum levels have been found to be associated with coronary heart disease or

myocardial infarction (MI) in several large observational studies (Suzuki *et al.* 1997; Resby and Berg 2000; Holmer *et al.* 2003). In view of these findings, in this study we chose to dissect the genetic underpinnings of these quantitative precursors (ApoB, CRP, FBG, HCY, LPA, cholesterol levels and TG) of coronary artery disease.

2. Materials and methods

2.1 *Subjects, ascertainment of families and data collected*

The study participants belonged to a single ethnic, maritally isolated population, the Marwaris settled in Kolkata. The Marwari is an endogamous community with their ancestral home in western India (Rajasthan), where they currently number about 5.6 million individuals. A small group of individuals moved to the city of Kolkata about 400 years ago in pursuit of business. The group underwent demographic expansion and currently number about 1 million in the city. Most individuals belonging to this community are vegetarians. After obtaining an ethics approval from the Committee for the Protection of Research Risks to Humans, Indian Statistical Institute, Kolkata, India, data were collected on members belonging to 144 nuclear families, each of whom provided written informed consent to participate. Each extended family, from which the nuclear families were drawn, was ascertained on the basis of presence of at least two first-degree relatives who have had one or more of the following: (a) undergone coronary bypass surgery or angioplasty, or (b) had a myocardial infarction and survived. These families were contacted by us on the basis of referrals provided by cardiologists who had treated one or more of the family members. Data pertaining to a large number of variables – including demographic, anthropometric, blood pressures and pulse rates in various positions, behavioural (lifestyle) – were collected from each individual recruited into this study (table 1). Within the set of study participants, two non-overlapping subsets were identified: Subset-1 comprised unrelated individuals (n=157) and Subset-2 comprised nuclear families (n=327 individuals in 144 nuclear families).

2.2 *Biochemical quantitative traits assayed and DNA isolation*

A 12 h fasting blood sample (10 mL) was collected by venipuncture from each participant recruited into this study. Serum and plasma samples were analysed for estimation of quantitative levels of 10 parameters that are clinically relevant to CAD. These were: ApoB, CRP, FBG, HCY, LPA, CHOL-T, CHOL-H, CHOL-L, CHOL-V and TG (the normal ranges of these parameters are given in supplementary table 1). High-molecular-weight DNA was isolated for each study subject using Qiagen columns.

2.3 *Candidate genes studied, with justifications, and SNPs genotyped*

Thirty-one genes that possibly influence the levels of the quantitative precursors of coronary artery disease were chosen as candidates for this gene-mapping study. Relevant details on each gene are provided in supplementary table 2.

Genes on the inflammatory pathway that were chosen included *CRP*, *IL10*, *IL1A*, *IL1B*, *IL8*, *NFKB1*, *TNF* and *IL6*. CRP stimulates the expression, up-regulation and secretion of many genes and their products, which consequently stimulates the total inflammatory pathway that causes plaque formation (Morrow and Ridker 2000; Ridker *et al.* 2001). IL10 has been reported to enhance oxidized LDL-induced foam cell formation (Halvorsen *et al.* 2005), which is an important event in plaque formation. Genetic variations in *IL6* have been found to be associated with susceptibility to atherosclerosis (Ridker *et al.* 2000). Moreover, *IL6* is an important regulator of *CRP* expression, which in turn up-regulates products of adhesion molecule genes, such as *ICAM1* and *SELE*. We have studied the following adhesion molecule genes: *SELP*, *SELL*, *SELE*, *VCAM1* and *ICAM1*. *FGB* and *TCF1* are associated with fibrinogen levels (Liu *et al.* 2001; Soria *et al.* 2005) and hence were included. The *F3* gene encodes coagulation factor 3 that initiates coagulation (Ardissino *et al.* 2000) and thus is an important candidate gene.

From each of these genes (including 2 kb upstream from the starting nucleotide position of the gene), the most informative subset of SNPs (at ~3 kb density; SNPs that are within 3 kb of each other show significant linkage disequilibrium in most populations, indicative of redundancy of information) has been chosen for genotyping to increase cost–time efficiency. For all SNPs, data on reference allele frequency, heterozygosity and D′ value with the preceding SNP were downloaded from the HapMap (*http://www.hapmap.org/cgiperl/gbrowse/hapmap_B36/*) and Hapmart (*http://hapmart.hapmap.org/BioMart/martview*) databases. These data were downloaded for all the four HapMap populations (CEU, CHB, JPT and YRI). Using a locally developed computer program, relevant information on SNPs from the downloaded data were extracted and the informative subset of SNPs were chosen. In addition, SNPs in these genes that result in non-conservative amino acid changes and earlier found to be significantly associated with cardiovascular diseases were included wherever possible.

A set of 209 maximally informative SNP markers in the 31 candidate genes was selected for genotyping. The list of SNPs selected from the 31 candidate genes and their characteristics are given in supplementary table 3. Genotyping was carried out using the Sequenom platform.

Table 1. List of variables on which data were collected from each individual recruited into the study

Variable set	Sl. no.	Description of variable (unit of measurement/categories)	Variable code
Demographic	1	Age (In completed years)	AGE
	2	Gender (Male/Female)	SEX
	3	Marital Status (Married/Unmarried/Widowed)	MSTAT
	4	No. of Children	NCHLD
Anthropometric	1	Height (cm)	HT
	2	Weight (kg)	WT
	3	Body Mass Index=Wt/Ht^2 (kg/m^2)	BMI
	4	Waist Height (cm) In Supine Position	HT-W
	5	Hip Circumference (cm)	CIR-H
	6	Thigh Circumference (cm)	CIR-T
	7	Waist Circumference (cm)	CIR-W
	8	Biceps (mm) Skinfold Thickness	SFT-BI
	9	Triceps (mm) Skinfold Thickness	SFT-TR
	10	Subscapular (mm) Skinfold Thickness	SFT-SB
	11	Suprailiac (mm) Skinfold Thickness	STF-SP
	12	Abdomen (mm) Skinfold Thickness	SFT-AB
	13	Medial Calf (mm) Skinfold Thickness	SFT-MC
Blood pressure and pulse rate	1	Diastolic Blood Pressure (mmHg), Supine	DBP-SUP
	2	Systolic Blood Pressure (mmHg), Supine	SBP-SUP
	3	Diastolic Blood Pressure (mmHg), Supine To Standing After Five Min.	DBP-SS5
	4	Diastolic Blood Pressure (mmHg), Supine To Standing Immediate	DBP-SS0
	5	Systolic Blood Pressure (mmHg), Supine To Standing Immediate	SBP-SS0
	6	Pulse Rate, Supine To Standing Immediate (No./min)	PUL-SS0
	7	Pulse Rate, Supine (No./min)	PUL-SUP
	8	Systolic Blood Pressure (mmHg), Supine To Standing After Five Min.	SBP-SS5
	9	Pulse Rate, Supine To Standing After Five Min (No./min)	PUL-SS5
Behavioural/lifestyle/medication*	1	Alcohol consumption (Graded: 0,1,2,3)	ALC
	2	Smoking pattern (Graded: 0,1,2,3)	SMOK
	3	Tobacco chewing pattern (Graded: 0,1,2,3)	TOB
	4	Exercise pattern (Graded: 0,1,2,3,4)	EXER
	5	Anger propensity (Binary: 0,1)	ANGR
	6	Irritability propensity (Binary: 0,1)	IRRI
	7	Tension propensity (Binary: 0,1)	TENS
	8	Combined score of Anger & Irritability & Tension (Graded: 0,1,2,3)	AIT-COM
	9	Stress propensity (Binary: 0,1)	STRS
	10	Hypertension Medication (Binary: 0,1)	M-HTEN
	11	Diabetes Medication (Binary: 0,1)	M-DIAB
	12	CVD Medication (Binary: 0,1)	M-CVD
	13	Cholesterol Medication (Binary: 0,1)	M-CHL

* Detailed data on behaviour, lifestyle, current and past medications used were collected by personal interviews. The grades 0,1,2 and 3 indicate the increasing intensity of the variable, with 0 indicating absence. Binary codes 0 and 1 represent, respectively, absence and presence.

Analyses of genotype cluster plots were carried out using Typer (Sequenom) software package. In this study, only those genotype calls that belonged to the highest two quality classes – 'conservative' and 'moderate', were accepted.

2.4 *Statistical analyses*

Maximum-likelihood estimates of allele frequencies and their standard deviations were estimated using MAXLIK

(Reed and Schull 1968). Tests of equality of genotype frequencies with those expected under Hardy–Weinberg equilibrium were also carried out using MAXLIK. The genotype data were checked for Mendelian inconsistencies using PEDCHECK (*http://watson.hgen.pitt.edu/register/docs/pedcheck.html*).

The 10 QTs were log-transformed to induce normality to their frequency distributions. Using stepwise linear regression analysis, appropriate adjustments were made for significant covariates, which included demographic, anthropometric and behavioral variables, and use of anti-hypertensive, cholesterol lowering, anti-diabetic, etc., drugs (table 1). All further analyses were carried out using the covariate-adjusted residual values of each QT.

Analysis of variance (ANOVA) was used for testing the null hypothesis of equality of mean values among genotypes on unrelated individuals (Subset-1; $n=157$), using SPSS. To obtain insights into the results of association analysis, we also estimated the coefficient (D′) of linkage disequilibrium (LD) between pairs of relevant SNPs. All LD computations were carried out using HAPLOVIEW (Barrett *et al.* 2005) (*http://www.broad.mit.edu/mpg/haploview/*).

Finally, for validation of the association results, quantitative transmission-disequilibrium test (Q-TDT) was carried out for each covariate-adjusted quantitative trait using data on nuclear families (Subset-2; comprising 144 nuclear families with a total of 327 individuals; among which 241 individuals were from 111 families in each of which data and samples were available on both parents, and 86 individuals were from 33 families in which data were available on only one parent) for each SNP that turned out to be significant in Data Subset-1 using ANOVA. Q-TDT was performed using the PLINK (*http://pngu.mgh.harvard.edu/~purcell/plink/fanal.shtml#qfam*) package. PLINK was also used to test whether there was significant association between the relevant quantitative trait and haplotypes reconstructed using data on these loci.

3. Results

Demographic details of the unrelated individuals (Subset-1) are provided in table 2. The vast majority of study participants were above 45 years of age. Age- and gender-specific distribution of individuals belonging to various categories for each of the qualitative behavioural/lifestyle variables is presented in table 3. In table 4 are presented the mean values of the quantitative traits studied, by age group × gender classes. Most study participants did not exercise regularly, but were also not habitual consumers of alcohol or tobacco (table 3). Because of the mode of ascertainment of families, mean values of lipid and other variables, but not blood pressure, which are known to be risk factors for cardiovascular diseases, are outside of the normal ranges (table 4).

Table 2. Numbers (%) of individuals in various age-group and gender categories among unrelated study participants

Age group (years)	Male *n* (%)	Female *n* (%)
31–45	1 (0.6)	17 (11.0)
46–60	21 (13.4)	42 (26.8)
≥ 61	44 (28.0)	32 (20.4)
Total	66 (42.0)	91 (58.0)

Of the 209 SNPs assayed from the 31 candidate genes, 30 SNPs that belonged to 22 genes turned out to be monomorphic in this population. No polymorphic variant was observed in *CD14* and *TNF* in the study population, possibly because of founder effects. The reference allele frequencies are presented in supplementary table 4. Genotype frequencies at each of the 179 polymorphic loci were, after Bonferroni correction, in accord with Hardy–Weinberg expectations.

Age and/or gender were a significant correlate for some (FBG, HCY, CHOL-T, CHOL-H and CHOL-L) but not all QTs. Body fatness, as represented by high BMI and/or high values of skinfold thickness, was a significant correlate for most of the QTs (ApoB, CRP, CHOL-T, CHOL-L, CHOL-V and TG). Lack of exercise turned out to be a significant predictor of an elevated level of HCY. Expectedly, the effect of medication (anti-hypertensive, anti-diabetic, cholesterol-lowering, etc.) was found to be significant on 8 of the QTs (CRP, FBG, HCY, LPA, CHOL-T, CHOL-L, CHOL-V and TG), but interestingly not on quantitative levels of APOB and CHOL-H. Results of stepwise regression analysis for identifying significant covariates for each of the 10 QTs considered are provided in table 5.

A total of 70 SNPs from 23 genes were found (supplementary table 5) to have significantly different mean values among genotypes in the Subset-1, comprising unrelated individuals. Several SNPs (rs3917657, rs6133, rs6128, rs2235302 and rs3917744) in *SELP* were associated with multiple QTs (ApoB, LPA, CHOL-T, CHOL-H, CHOL-L, CHOL-V and TG). Similarly, many SNPs in *NFKB1* (chromosome 4q) were found to be significantly associated with the QTs APOB (11 SNPs, spanning 111 kb), CHOL-T (13 SNPs, spanning 78 kb) and CHOL-L (9 SNPs, spanning 49.9 kb), respectively. Several SNPs in *TCF1* (chromosome 12q) were also found to be significantly associated with LPA (7 SNPs, spanning 17.5 kb). Obviously, these SNPs were not all independently associated with the traits, but showed significant association likely due to strong linkage disequilibrium among the SNPs themselves (as exemplified in supplementary figure 1).

We tested for transmission-disequilibrium of these 70 SNPs that turned out to be significant in Subset-1. The

Table 3. Age- and gender-specific percentages of individuals belonging to various categories for the qualitative behavioural/lifestyle variables

		Unrelated individuals (n=157)					
		Age group (years)					
		31–45		46–60		≥61	
Qualitative trait	Grades	Male n=1	Female n=17	Male n=21	Female n=42	Male n=44	Female n=32
		n (%)					
ALC	0		17 (100)	20 (95.2)	42 (100)	40 (90.9)	32 (100)
	1	1 (100)				1 (2.3)	
	2			1 (4.8)		1 (2.3)	
	3					2 (4.5)	
	Total	1 (100)	17 (100)	21 (100)	42 (100)	44 (100)	32 (100)
AIT-COM	0	1 (100)	4 (23.5)	4 (19)	5 (11.9)	10 (22.7)	3 (9.4)
	1		2 (11.8)	3 (14.3)	12 (28.6)	9 (20.5)	11 (34.4)
	2		6 (35.3)	3 (14.3)	17 (40.5)	14 (31.8)	11 (34.4)
	3		5 (29.4)	11 (52.4)	8 (19)	11 (25)	7 (21.9)
	Total	1 (100)	17 (100)	21 (100)	42 (100)	44 (100)	32 (100)
ANGR	0	1 (100)	9 (52.9)	14 (66.7)	15 (35.7)	22 (50)	7 (21.9)
	1		8 (47.1)	7 (33.3)	27 (64.3)	22 (50)	25 (78.1)
	Total	1 (100)	17 (100)	21 (100)	42 (100)	44 (100)	32 (100)
EXER	0		7 (41.2)	6 (28.6)	17 (40.5)	15 (34.1)	12 (37.5)
	1		6 (35.3)	10 (47.6)	16 (38.1)	23 (52.3)	9 (28.1)
	2			2 (9.5)	2 (4.8)	1 (2.3)	2 (6.3)
	3		1 (5.9)	1 (4.8)	1 (2.4)	3 (6.8)	1 (3.1)
	4	1 (100)	3 (17.6)	2 (9.5)	6 (14.3)	2 (4.5)	8 (25)
	Total	1 (100)	17 (100)	21 (100)	42 (100)	44 (100)	32 (100)
IRRI	0	1 (100)	10 (58.8)	7 (33.3)	16 (38.1)	18 (40.9)	8 (25)
	1		7 (41.2)	14 (66.7)	26 (61.9)	26 (59.1)	24 (75)
	Total	1 (100)	17 (100)	21 (100)	42 (100)	44 (100)	32 (100)
SMOK	0	1 (100)	17 (100)	19 (90.5)	42 (100)	38 (86.4)	32 (100)
	1						
	2			1 (4.8)		4 (9.1)	
	3			1 (4.8)		2 (4.5)	
	Total	1 (100)	17 (100)	21 (100)	42 (100)	44 (100)	32 (100)
STRS	0	1 (100)	3 (17.6)	6 (28.6)	8 (19)	11 (25)	7 (21.9)
	1		14 (82.4)	15 (71.4)	34 (81)	33 (75)	25 (78.1)
	Total	1 (100)	17 (100)	21 (100)	42 (100)	44 (100)	32 (100)
TOB	0	1 (100)	14 (82.4)	15 (71.4)	34 (81)	34 (77.3)	28 (87.5)
	1		1 (5.9)	1 (4.8)	2 (4.8)	2 (4.5)	
	2		2 (11.8)	1 (4.8)	6 (14.3)	6 (13.6)	4 (12.5)
	3			4 (19)		2 (4.5)	
	Total	1 (100)	17 (100)	21 (100)	42 (100)	44 (100)	32 (100)
TENS	0	1 (100)	9 (52.9)	16 (76.2)	17 (40.5)	26 (59.1)	19 (59.4)
	1		8 (47.1)	5 (23.8)	25 (59.5)	18 (40.9)	13 (40.6)
	Total	1 (100)	17 (100)	21 (100)	42 (100)	44 (100)	32 (100)
M-CHL	0	1 (100)	17 (100)	20 (95.2)	38 (90.5)	35 (79.5)	27 (84.4)
	1			1 (4.8)	4 (9.5)	9 (20.5)	5 (15.6)
	Total	1 (100)	17 (100)	21 (100)	42 (100)	44 (100)	32 (100)

Table 3. (continued)

		Unrelated individuals (n=157)					
		Age group (years)					
		31–45		46–60		≥61	
Qualitative trait	Grades	Male n=1	Female n=17	Male n=21	Female n=42	Male n=44	Female n=32
M-CVD	0	1 (100)	16 (94.1)	12 (57.1)	38 (90.5)	21 (47.7)	25 (78.1)
	1		1 (5.9)	9 (42.9)	4 (9.5)	23 (52.3)	7 (21.9)
	Total	1 (100)	17 (100)	21 (100)	42 (100)	44 (100)	32 (100)
M-DIAB	0	1 (100)	17 (100)	16 (76.2)	35 (83.3)	35 (79.5)	29 (90.6)
	1			5 (23.8)	7 (16.7)	9 (20.5)	3 (9.4)
	Total	1 (100)	17 (100)	21 (100)	42 (100)	44 (100)	32 (100)
M-HTEN	0	1 (100)	15 (88.2)	13 (61.9)	34 (81)	17 (38.6)	16 (50)
	1		2 (11.8)	8 (38.1)	8 (19)	27 (61.4)	16 (50)
	Total	1 (100)	17 (100)	21 (100)	42 (100)	44 (100)	32 (100)

transmission-disequilibrium tests were carried out on parent–offspring data (Subset-2). Nine SNPs in 4 genes were found to significantly deviate from transmission equilibrium (table 6). Thus, these 9 SNPs turned out to be 'significant' in both ANOVA and Q-TDT analyses. (Mean values of the quantitative traits among individuals belonging to different genotypes at each of these 9 SNP loci are depicted in supplementary figure 2.)

The two SNPs rs230521 and rs1005819 in *NFKB1*, located 40 kb apart, that were found to be significantly associated with ApoB levels were in strong LD (r^2=0.96). The two inferred haplotypes CC and GT, comprising SNPs rs230521 and rs1005819, were present among 69% and 31%, respectively, in the pooled set of individuals. As expected, both the haplotypes were significantly (p-value < 0.05) associated with the ApoB level, but did not show any enhancement of the strength of association over the individual SNPs.

4. Discussion

We have found many (70) significant SNP associations from the analyses of data on unrelated individuals. We, however, reasoned that since the literature on SNP associations with qualitative and quantitative traits is replete with false-positive results, it is crucial to test whether there is preferential transmission of alleles from parents to offspring with 'high' or 'low' values of the quantitative traits at the loci that were found to be significantly associated. This is the notable strength of our study compared to standard association studies, because the study design has a built-in mechanism to safeguard against drawing false-positive inferences. For loci depicting false-positive associations, one would not expect any preferential transmission of alleles from parents to offspring. Among the large number of loci that showed significant association with one or more quantitative traits in data of unrelated individuals, there was preferential transmission of alleles at only 9 loci. This underscores the need for built-in validation mechanisms in designing studies on complex disorders (such as Step-2 of our study).

We have found SNPs in *NFKB1*, *SELE*, *FGB* and *VEGFA* genes to significantly affect levels of ApoB, FBG, HCY and CHOL-H in the Marwari population. We note that exons 14–24 of *NFKB1* contain eleven non-synonymous SNPs (rs56280473, rs56147914, rs4648065, rs4648072, rs4648085, rs4648086, rs55661548, rs11556612, rs4648099, rs4648118 and rs41341444). Because these non-synonymous SNPs have low minor allele frequencies in the HapMap populations, these were not selected by our SNP-selection algorithm and hence were not assayed in this study. However, the associations we have found for two intronic SNPs in *NFKB1* with APOB are in LD with one or more of these non-synonymous SNPs that may be causal SNPs.

FASTSNP (function analysis and selection tool for single nucleotide polymorphisms, *http://fastsnp.ibms.sinica.edu.tw*) showed that the change in allele from A to C at rs5361 in *SELE*, that showed significant transmission-disequilibrium in this study, has damaging consequences. rs5361 is involved in changing protein structure and also with a change in the number of exonic splicing enhancer (ESE) motifs. The

Table 4. Mean values of quantitative traits among unrelated individuals belonging to the different age group×gender classes

	Mean±s.e. (unrelated individuals, n=157)					
Age group (years)	31–45		46–60		≥61	
Variable	Male n=1	Female n=17	Male n=21	Female n=42	Male n=44	Female n=32
BMI (kg/m^2)	33.5	27.9±0.93	26.7±1.20	27.8±0.60	24.3±0.46	28.1±0.90
CIR-H (cm)	116.0	104.7±1.78	100±2.36	103.0±2.27	94.4±0.99	105.9±2.12
CIR-T (cm)	65.0	53.4±0.95	48.3±1.27	51.6±1.72	45.1±0.86	49.6±1.04
CIR-W (cm)	102.0	80.2±1.81	91.6±2.48	82.3±1.19	88.5±1.22	82.9±2.05
HT (cm)	173.7	152.1±1.23	167.4±1.21	153.8±1.20	163.9±1.24	150.8±0.94
HT-W (cm)	25.0	20.4±0.68	23.3±0.81	20.3±0.41	22.3±0.37	20.7±0.43
SFT-AB (mm)	50.0	30.1±1.28	29.2±1.93	31.8±1.51	27.4±1.30	29.3±1.87
SFT-BI (mm)	14.9	15.4±0.67	8.0±0.62	11.5±0.66	7.3±0.53	10.1±0.74
SFT-MC (mm)	23.5	24.0±1.03	15.2±2.12	19.1±1.10	10.6±0.70	18.8±1.31
SFT-TR (mm)	29.7	23.3±0.58	14.8±0.89	20.3±0.66	13.5±0.61	19.3±0.81
SFT-SB (mm)	50.0	25.6±1.28	28.0±2.48	25.6±1.22	23.2±1.17	25.1±1.68
STF-SP (mm)	50.0	18.8±1.30	23.6±1.80	18.4±1.07	18.5±0.91	16.5±1.46
WT (kg)	101.0	64.4±1.92	74.4±3.08	65.7±1.60	65.4±1.44	63.9±2.08
DBP-SUP (mmHg)	82.0	75.8±2.85	82.7±2.30	80.9±1.45	83.0±1.65	79.8±1.94
DBP-SS0 (mmHg)	91.0	81.1±2.33	83.4±2.61	83.4±1.56	82.7±1.65	82.0±2.19
DBP-SS5 (mmHg)	82.0	81.9±2.72	84.9±2.13	84.5±1.65	80.4±1.54	83.1±2.36
PUL-SUP (No./min)	63.0	75.8±2.47	72.0±3.04	76.4±1.62	71.7±1.72	72.8±2.18
PUL-SS0 (No./min)	76.0	82.1±2.61	85.2±3.44	86.2±2.11	81.1±2.11	84.3±2.00
PUL-SS5 (No./min)	80.0	81.7±2.93	83.4±3.33	90.0±2.00	81.9±2.11	81.5±2.06
SBP-SUP (mmHg)	115.5	119.9±4.57	137.1±4.85	125.2±2.42	141.7±3.44	133.8±4.17
SBP-SS0 (mmHg)	116.0	119.3±4.10	130.4±5.49	121.5±2.67	134.0±3.27	129.5±3.99
SBP-SS5 (mmHg)	112.0	118.7±4.19	130.3±5.73	118.7±2.73	131.9±3.24	130.7±3.96
GLUC (mg/dL)	91.0	98.3±2.93	112.7±6.10	122.8±8.10	116.8±5.93	111.9±5.86
HbA1C (%)	5.8	5.5±0.17	6.3±0.31	6.5±0.27	6.3±0.18	6.2±0.15
INS (μU/mL)	12.5	12.0±1.88	16.3±4.63	12.5±1.52	20.3±10.58	12.0±1.43
CHOL-T (mg/dL)	193.0	185.4±6.36	166.3±11.13	179.4±4.95	163.4±4.49	191.8±6.56
CHOL-H (mg/dL)	36.0	46.6±1.98	37.5±1.72	44.6±1.74	39.9±1.54	47.2±2.05
CHOL-L (mg/dL)	138.0	116.7±5.68	104.5±9.87	108.5±4.44	100.9±4.12	117.7±4.94
CHOL-V (mg/dL)	19.0	22.2±3.12	24.2±2.28	25.4±1.60	22.6±1.77	26.9±3.32
TG (mg/dL)	93.0	118.6±15.79	123.3±11.26	128.3±7.76	119.7±8.30	141.2±15.36
LPA (mg/dL)	13.0	30.8±5.33	27.2±4.80	31.0±3.92	27.8±3.87	36.7±5.16
APOB (mg/dL)	99.1	99.6±4.68	91.5±5.90	92.4±3.37	91.9±3.00	95.0±4.61
CRP (mg/dL)	0.8	0.4±0.08	0.2±0.05	0.4±0.06	0.3±0.07	0.6±0.09
FBG (mg/L)	4592.0	3715.4±328.78	3456.6±214.79	3487.9±130.92	3700.1±159.55	4170.3±176.23
HCY (μmol/L)	27.2	15.9±1.48	22.1±1.88	18.2±1.34	17.4±1.39	18.7±1.57

C allele state of the SNP creates two ESE binding sites for SF2/ASF and SRp55 (splicing regulatory proteins). This may result in inappropriate splicing of the mRNA transcript resulting in abnormal protein product. For rs4220 in *FGB*, the presence of a G allele instead of an A allele leads to reduction of exonic splicing enhancer sites from six to three. We have found that rs3025035 in *VEGFA*, which is significantly associated with HDL-cholesterol level, is located in a transcription factor (TF) binding site. The T allele at this locus results in loss of transcription factor binding. Although, microRNAs (miRNAs) are predicted to control the activity of approximately 30% of all protein-coding genes in mammals, no known miRNA gene was found in the intronic regions where the significant SNPs were located.

Table 5. Results of stepwise regression analysis for identifying significant covariates for each of the 10 quantitative traits considered

QT	List of significant predictors, and their standardized regression coefficients								Multiple corr. coeff. [r]	F-ratio (degrees of freedom)	p-value
	1	2	3	4	5	6	7	8			
APOB	DBP-SS5 0.097	CIR-W 0.199	IRRI −0.112						0.271	11.887 (3, 449)	<0.0001
CRP	BMI 0.149	DBP-SS0 0.201	M-HTEN 0.104	SEX 0.249	AGE 0.228	SFT-SU 0.184	SBP-SUP −0.248		0.522	23.851 (7, 445)	<0.0001
FBG	AGE 0.231	SEX 0.234	HT-W 0.174	M-HTEN 0.113					0.405	21.976 (4, 448)	<0.0001
HCY	HT-W 0.139	EXER −0.114	SEX −0.126	M-DIAB −0.101					0.25	7.456 (4, 448)	<0.0001
LPA	HT −0.158	M-CHL 0.144							0.212	10.577 (2, 450)	<0.0001
CHOL-T	AGE 0	SFT-AB 0	M-CHL 0.014	DBP_SS5 0	CIR-H 0	M-CVD 0.012	IRRI 0.007	M-DIAB 0.013	0.467	15.471 (8, 444)	<0.0001
CHOL-H	SEX 0.137	WT −0.298	ALC 0.155	CIR-H 0.178	PUL_SUP 0.091	SMOK −0.098			0.374	12.108 (6, 446)	<0.0001
CHOL-L	SFT-AB 0.17	M-CHL −0.19	M-DIAB −0.205	AGE 0.213	M-CVD −0.159	IRRI −0.094	SBP_SS5 0.098		0.444	15.585 (7, 445)	<0.0001
CHOL-V	CIR-W 0.146	M-DIAB 0.099	SFT-AB 0.22	SBP_SS5 0.111	STRS −0.075	SFT-SU 0.155	SFT-B −0.109	PUL-SUP 0.086	0.522	20.811 (8, 444)	<0.0001
TG	CIR-W 0.253	M-DIAB 0.084	PUL-SUP 0.084	SFT-AB 0.229	SBP_SS5 0.118				0.515	32.194 (5, 447)	<0.0001

The selectin family of adhesion molecules shares a unique mosaic structure consisting of an amino-terminal lectin-like domain followed by an epidermal growth factor (EGF)-like domain, a variable number of complement regulatory repeats, a transmembrane domain and a short cytoplasmic domain (Bevilacqua and Nelson 1993). E-selectin, one of the three members of this family, has been shown to support the rolling of leukocytes on activated endothelial cells (Bevilacqua and Nelson 1993) and may also participate in the transition to stable adhesion that precedes transmigration (Milstone *et al.* 1998). Using a recombinant chimeric-protein-based analysis (Revelle *et al.* 1996) found the R>S missense mutation in Exon 4 (that is significantly associated with fibrinogen in this study population) to change the binding specificity of E-selectin. In addition to mediating leukocyte adhesion to vascular endothelium, recent studies have indicated that E-selectin may also transmit outside-in signals in vascular endothelium during leukocyte-endothelial interactions (Yoshida *et al.* 2003). Hu *et al.* (2000) reported that E-selectin-dependent leukocyte adhesion induces activation of the MAPK pathway. Yoshida *et al.* (2003) demonstrated that a

Table 6. List of SNPs and their characteristics that were found to be statistically significant, on the basis of both ANOVA and Q-TDT analyses for various quantitative precursors of coronary artery disease

QT	rs#	Chr #	Gene name	Nucleotide position	Location of SNP	Alleles	p-value
APOB	rs3774933	4	*NFKB1*	103645369	Intron1	C/T	0.013
	rs230528	4	*NFKB1*	103676615	Intron4	C/A	0.042
	rs230521	4	*NFKB1*	103682357	Intron5	G/C	0.014
	rs1005819	4	*NFKB1*	103723343	Intron10	C/T	0.016
	rs1609798	4	*NFKB1*	103756488	Intron23	C/T	0.040
FBG	rs5361	1	*SELE*	167967684	Exon4, Missense, R>S	A/C	0.004
	rs4648004	4	*NFKB1*	103680136	Intron5	A/G	0.012
HCY	rs4220	4	*FGB*	155711209	Exon8, Missense, K>R	G/A	0.017
CHOL-H	rs3025035	6	*VEGFA*	43859337	Intron7	C/T	0.010

Serine128Arginine mutation in the EGF domain of E-selectin is a potential risk factor for genetic susceptibility to MI in the Japanese population. The Ser128Arg mutation of E-selectin significantly enhanced its adhesion to leukocytes under physiological flow conditions. It has also been reported (Revelle *et al.* 1996) that the Ser128Arg mutation confers carbohydrate specificity to E-selectin. The mechanism by which this mutation causes enhanced binding activity under flow requires additional study. Observations suggest that the polymorphisms found in cell-surface receptors lead to a constitutive activation of the receptors in the absence of their ligands (Freeburn *et al.* 1998; Fernandes *et al.* 2001), possibly through constitutive oligomerization, as has been suggested for the EGF receptor, or via induction of a conformational change, as reported with CD16. Notably, polymorphic S128R-E-selectin exhibited constitutive phosphorylation of ERK1/2 and p38 MAPK without leukocyte adhesion, which indicates that a potential conformational change of E-selectin attributable to the S128R polymorphism may influence the intracellular signalling pathway of E-selectin. The participation of the ERK1/2 signalling pathway in E-selectin–dependent PMN adhesion to vascular endothelium was identified in Hu *et al.* (2000). Jilma *et al.* (2004) reported that Ser128Arg in SELE enhances thrombin generation that leads to a pronounced procoagulant effect.

The SNP rs4220 (arginine to lysine) in *FGB* identified to be a 'significant' SNP in this study is known to increase fibrinogen level (Kathiresan *et al.* 2006). However, this SNP has not been found to be associated with fibrinogen, but has been found to be associated with HCY. Homocysteine plays an independent role in the pathogenesis of atherosclerosis; it enhances coagulation during atherosclerosis, and may be indirectly be related to increasing fibrinogen level.

The serine to arginine change in *SELE* (rs5361) is 'damaging' (SIFT score=0.01; *http://blocks.fhcrc.org/sift/SIFT.html*); it enhances leukocyte-endothelial interactions (Yoshida *et al.* 2003) and increases thrombin generation (Jilma *et al.* 2004). Thrombin is associated with the coagulation pathway, the latter playing an important role in the pathophysiology of coronary artery disease.

5. Conclusions

In conclusion, this study carried out in an isolated, homogeneous, migrant population (Marwari) of Kolkata, India, has provided valuable insights into the genomic underpinnings of important quantitative precursors of coronary artery disease, not only through an initial identification of SNP associations in candidate genes but through a powerful validation mechanism of transmission-disequilibrium tests. SNPs in *SELE*, *VEGFA*, *FGB* and *NFKB1* genes impact significantly on levels of quantitative precursors of CAD in Marwaris. Saturating the identified genomic regions in which significant associations were found with only intronic SNPs with more densely spaced SNPs may help identify more causal SNPs that modulate the levels of the quantitative precursors of CAD.

Acknowledgements

This study was supported by grants from the TCG-ISI Centre for Population Genomics, Indian Council of Medical Research (ICMR) and the Council of Scientific & Industrial Research (CSIR Award No. F.NO.10-2(5) 2003(I)-EU.II), India. We are thankful to the patients and their family members who have participated in the study. We are grateful to all members of the TCG-ISI Centre for Population Genomics, Kolkata, (for help with data and sample collection); the BM Birla Heart Research Centre, Kolkata, for patient referrals and evaluations; The Centre for Genomic Application, New Delhi, for help with genotyping; and Quadra Medical Services, Kolkata, for biochemical assays.

References

Ardissino D, Merlini PA, Arlens R, Coppola R, Bramucci E, Lucreziotti S, Repetto A, Fetiveau R and Mannucci PM 2000 Tissue factor in human coronary atherosclerotic plaques. *Clin. Chim. Acta* **291** 235–240

Austin MA, King MC, Bawol RD, Hulley SB and Friedman GD 1987 Risk factors for coronary heart disease in adult female twins. Genetic heritability and shared environmental influences. *Am. J. Epidemiol.* **125** 308–18

Barrett JC, Fry B, Maller J and Daly MJ 2005 Haploview, analysis and visualization of LD and haplotypemaps. *Bioinformatics* **21** 263–265

Bevilacqua MP and Nelson RM 1993 Selectins. *J. Clin. Invest.* **91** 379–387

Brull DJ, Serrano N, Zito F, Jones L, Montgomery HE, Rumley A, Sharma P, Lowe GDO, World MJ, Humphries SE and Hingorani AD 2003 Human CRP gene polymorphism influences CRP levels, implications for the prediction and pathogenesis of coronary heart disease. *Arterioscler. Thromb. Vasc. Biol.* **23** 2063–2069

Danesh J, Collins R, Appleby P and Peto R 1998 Association of fibrinogen, C-reactive protein, albumin, or leukocyte count with coronary heart disease, meta-analyses of prospective studies. *JAMA* **279** 1477–1482

Duggirala R, Blangero J, Almasy L, Dyer TD, Williams KL, Leach RJ, O'Connell P and Stern MP 2000 Amajor susceptibility locus influencing plasma triglyceride concentrations is located on chromosome 15q in Mexican Americans. *Am. J. Hum. Genet.* **66** 1237–1245

Ernst E and Resch KL 1993 Fibrinogen as a cardiovascular risk factor, a meta-analysis and review of literature. *Ann. Intern. Med.* **118** 956–963

Falchi M, Forabosco P, Mocci E, Borlino CC, Picciau A, Virdis E, Persico I, Parracciani D, Angius A and Pirastu M 2004 A genomewide search using an original pairwise sampling approach for large genealogies identifies a new locus for total and low-density lipoprotein cholesterol in two genetically differentiated isolates of Sardinia. *Am. J. Hum. Genet.* **75** 1015–1031

Fernandes H, Cohen S and Bishayee S 2001 Glycosylation-induced conformational modification positively regulates receptor-receptor association, a study with an aberrant epidermal growth factor receptor (EGFRvIII/ DeltaEGFR) expressed in cancer cells. *J. Biol. Chem.* **276** 5375–5383

Freeburn RW, Gale RE and Linch DC 1998 Activating point mutations in the C chain of the GM-CSF, IL-3 and IL-5 receptors are not a major contributory factor in the pathogenesis of acute myeloid leukaemia. *Br. J. Haematol.* **103** 66–71

Haines AP, Howarth D, North WR, Goldenberg E, Stirling Y, Meade TW, Raftery EB and Millar Craig MW 1983 Haemostatic variables and the outcome of myocardial infarction. *Thromb. Haemost.* **50** 800–803

Halvorsen B, Wæhre T, Scholz H, Clausen OP, Thüsen JHVD, Müller F, Heimli H, Tonstad S, Hall C, Frøland SS, Biessen EA, Damås JK and Aukrust P 2005 Interleukin-10 enhances the oxidized LDL-induced foam cell formation of macrophages by antiapoptotic mechanisms. *J. Lipid Res.* **46** 211–219

Heinrich J, Balleisen L, Schulte H, Assmann G and Loo VDJ 1994 Fibrinogen and factor VII in the prediction of coronary risk, results from the PROCAM study in healthy men. *Arterioscler. Thromb.* **14** 54–59

Hingorani A 2004 Resolving inconsistency in the results of genetic association studies of cardiovascular disease. *Clin. Sci.* **107** 251–253

Hokanson JE and Austin MA 1996 Plasma triglyceride level is a risk factor for cardiovascular disease independent of high-density lipoprotein cholesterol level, a meta-analysis of population-based prospective studies. *J. Cardiovasc. Risk* **3** 213–219

Holmer SR, Hengstenberg C, Kraft HG, Maye B, Pöll M, Kürzinger S, Marcus Fischer M, Löwel H, Klein G, Günter AJ and Riegger Schunkert H 2003 Association of polymorphisms of the Apolipoprotein(a) gene with Lipoprotein(a) levels and myocardial infarction. *Circulation* **107** 696–701

Hu Y, Kiely JM, Szente BE, Rosenzweig A and Gimbrone MA Jr 2000 E-Selectin dependent signaling via the mitogen-activated protein kinase pathway in vascular endothelial cells. *J. Immunol.* **165** 2142–2148

Ichinose A and Kuriyama M 1995 Detection of polymorphisms in the 5-prime-flanking region of the gene for apolipoprotein(a). *Biochem. Biophys. Res. Commun.* **209** 372–378

Jilma B, Marsik C, Kovar F, Wagner OF, Jilma-Stohlawetz P and Endler G 2004 The single nucleotide polymorphism Ser128Arg in the E-selectin gene is associated with enhanced coagulation during human endotoxemia. *Blood* **105** 2380–2383

Kathiresan S, Yang Q, Larson MG, Camargo AL, Tofler GH, Hirschhorn JN, Gabriel SB and O'Donnell CJ 2006 Common genetic variation in five thrombosis genes and relations to plasma hemostatic protein level and cardiovascular disease risk. *Arterioscler. Thromb. Vasc. Biol.* **26** 1405–1412

Kelly PJ, Rosand J, Kistler JP, Shih VE, Silveira S, Plomaritoglou A and Furie KL 2002 Homocysteine, MTHFR 677 C→T polymorphism, and risk of ischemic stroke, results of a meta-analysis. *Neurology* **59** 529–536

Lettre G, Palmer CD, Young T, Ejebe KG, Allayee H, Benjamin EJ, Bennett F, Bowden DW, *et al.* 2011 Genome-wide association study of coronary heart disease and its risk factors in 8,090 African Americans: The NHLBI CARe Project. *PLoS Genet.* **7** e1001300

Liu Y, Saha N, Heng CK, Hong S and Low PS 2001 Fibrinogen genotypes (alpha and beta) are associated with plasma fibrinogen levels in Chinese. *J. Med. Genet.* **38** E31

Majumder PP and Ghosh S 2005 Mapping quantitative trait loci in humans, achievements and limitations. *J. Clin. Invest.* **115** 1419–1424

Majumder PP, Nayak S, Das RN and Bhattacharya SK 1996 Genetic and cultural determinants of high-density lipoprotein cholesterol and serum triglycerides among Marwaris of Calcutta. *Indian J. Med. Res.* **103** 112–119

Maresca G, Blasio AD, Marchioli R and Minno GD 1999 Measuring plasma fibrinogen to predict stroke and myocardial infarction, an update. *Arterioscler. Thromb. Vasc. Biol.* **19** 1368–1377

Meade TW, Ruddock V, Stirling Y, Chakrabarti R and Miller GJ 1993 Fibrinolytic activity, clotting factors, and long-term incidence of ischaemic heart disease in the Northwick Park Heart Study. *Lancet* **342** 1076–1079

Milstone DS, Fukumura D, Padgett RC, O'Donnell PE, Davis VM, Benavidez OJ, Monsky WL, Melder RJ, Jain RK and Gimbrone MA Jr 1998 Mice lacking E-selectin show normal numbers of rolling leukocytes but reduced leukocyte stable arrest on cytokine-activated microvascular endothelium. *Microcirculation* **5** 153–171

Mitchell BD, Kammerer CM, Blangero J, Mahaney MC, Rainwater DL, Dyke B, Hixson JE, Henkel RD, *et al.* 1996 Genetic and environmental contributions to cardiovascular risk factors in Mexican Americans. The San Antonio Family Heart Study. *Circulation* **94** 2159–2170

Morrow DA and Ridker PM 2000 C-Reactive protein, inflammation, and coronary risk. *Med. Clin. North Am.* **84** 149–161

Pérusse L, Rice T, Després JP, Rao DC and Bouchard C 1997 Cross-trait familial resemblance for body fat and blood lipids, familial correlations in the Quebec Family Study. *Arterioscler. Thromb. Vasc. Biol.* **17** 3270–3277

Peyser PA 1997 Genetic epidemiology of coronary artery disease. *Epidemiol. Rev.* **19** 80–90

Reed TE and Schull WJ 1968 A general maximum likelihood estimation program. *Am. J. Hum. Genet.* **20** 579–580

Reilly MP and members of the Myocardial Infarction Genetics Consortium and the Wellcome Trust Case Control Consortium 2011 Identification of *ADAMTS7* as a novel locus for coronary atherosclerosis and association of *ABO* with myocardial infarction in the presence of coronary atherosclerosis: two genome-wide association studies. *Lancet* **377** 383–392

Resby O and Berg K 2000 LPA gene, interaction between the apolipoprotein(a) size ('kringle IV' repeat) polymorphism and a pentanucleotide repeat polymorphism influences Lp(a) lipoprotein level. *J. Internal Med.* **247** 139–152

Revelle BM, Scott D and Beck PJ 1996 Single amino acid residues in the E- and P-selectin epidermal growth factor domains can determine carbohydrate binding specificity. *J. Biol. Chem.* **271** 16160–16170

Ridker PM, Rifai N, Clearfield M, Downs JR, Weis SE, Miles JS and Gotto AM 2001 Measurement of C-reactive protein for the targeting of statin therapy in the primary prevention of acute coronary events. *New Engl. J. Med.* **344** 1959–1965

Ridker PM, Rifai N, Stampfer MJ and Hennekens CH 2000 Plasma concentration of interleukin-6 and the risk of future myocardial infarction among apparently healthy men. *Circulation* **101** 1767 –1772

Robinson K, Mayer EL, Dave P, Miller MS, Green R, van Lente F, Gupta A, Marchan KK, *et al.* 1995 Hyperhomocysteinemia and low pyridoxal phosphate common and independent reversible risk factors for coronary artery disease. *Circulation* **92** 2825–2830

Scheuner MT 2001 Genetic predisposition to coronary artery disease. *Curr. Opin. Cardiol.* **16** 25–60

Sonnenberg GE, Krakower GR, Martin LJ, Olivier M, Kwitek AE, Comuzzie AG, Blangero J and Kissebah AH 2004 Genetic determinants of obesity-related lipid traits. *J. Lipid Res.* **45** 610–615

Soria JM, Almasy L, Souto JC, Buil A, Lathrop M, Blangero J and Fontcuberta J 2005 A genome search for genetic determinants that influence plasma fibrinogen levels. *Arterioscler. Thromb. Vasc. Biol.* **25** 1287–1292

Suzuki K, Kuriyama M, Saito T and Ichinose A 1997. Plasma lipoprotein(a) levels and expression of the apolipoprotein(a) gene are dependent on the nucleotide polymorphisms in its 5'-flanking region. *J. Clin. Invest.* **99** 1361–1366

The Coronary Artery Disease (CAD) Genetics Consortium 2011 A genome-wide association study in Europeans and South Asians identifies five new loci for coronary artery disease. *Nat. Genet.* **43** 339–344

Tice JA, Browner W, Tracy RP and Cummings SR 2003 The relation of C-reactive protein levels to total and cardiovascular mortality in older U.S. women. *Am. J. Med.* **114** 199–205

Torzewski M, Rist C, Mortensen RF, Zwaka TP, Bienek M, Waltenberger J, Koenig W, Schmitz G, Hombach V and Torzewski J 2000 C-reactive protein in the arterial intima, role of C-reactive protein receptor-dependent monocyte recruitment in atherosclerosis. *Arterioscler. Thromb. Vasc. Biol.* **20** 2094–2099

Venugopal SK, Devaraj S and Jialal I 2003 C-reactive protein decreases prostacyclin release from human aortic endothelial cells. *Circulation* **108** 1676–1678

Verma S, Li SH, Badiwala MV, Weisel RD, Fedak PW, Li RK, Dhillon B and Mickle DA 2002a Endothelin antagonism and interleukin-6 inhibition attenuate the proatherogenic effects of C-reactive protein. *Circulation* **105** 1890–1896

Verma S, Wang CH, Li SH, Dumont AS, Fedak PW, Badiwala MV, Dhillon B, Weisel RD, Li RK, Mickle DA and Stewart DJ 2002b A self-fulfilling prophecy, C-reactive protein attenuates nitric oxide production and inhibits angiogenesis. *Circulation* **106** 913–919

Wilhelmsen L, Svardsudd K, Korsan-Bengtsen K, Larsson B, Welin L and Tibblin G 1984 Fibrinogen as a risk factor for stroke and myocardial infarction. *New Engl. J. Med.* **311** 501–505

Williams TN, Zhang, CX, Game BA, He L and Huang Y 2004 C-reactive protein stimulates MMP-1 expression in U937 histiocytes through FcγRII and extracellular signal-regulated kinase pathway, an implication of CRP involvement in plaque destabilization. *Arterioscler. Thromb. Vasc. Biol.* **24** 61–66

Yoshida M, Takano Y, Sasaoka T, Izumi T and Kimura A 2003 E-Selectin polymorphism associated with myocardial infarction causes enhanced leukocyte endothelial interactions under flow conditions. *Arterioscler. Thromb. Vasc. Biol.* **23** 783–788

MS received 10 May 2011; accepted 12 August 2011

*e*Publication: 18 October 2011

Corresponding editor: SARAH H ELSEA