

# Inflammatory Gene Haplotype-Interaction Networks Involved in Coronary Collateral Formation

Jian Zhang<sup>a</sup> Jakub J. Regieli<sup>b,e</sup> Maria Schipper<sup>c</sup> Mark M. Entius<sup>b</sup>  
Faming Liang<sup>f</sup> Jeroen Koerselman<sup>e</sup> Hendrik J.T. Ruven<sup>d</sup>  
Yolanda van der Graaf<sup>e</sup> Diederick E. Grobbee<sup>e</sup> Pieter A. Doevendans<sup>b</sup>

<sup>a</sup>Institute of Mathematics and Statistics, University of Kent, Canterbury, Kent, UK; <sup>b</sup>Department of Cardiology, UMC Utrecht, <sup>c</sup>Department of Biostatistics, Utrecht University, Utrecht, <sup>d</sup>Department of Clinical Chemistry St. Antonius Hospital, Nieuwegein, and <sup>e</sup>Julius Centre for Health Sciences and Primary Care, UMC Utrecht, Utrecht, The Netherlands; <sup>f</sup>Department of Statistics, Texas A&M University, College Station, Tex., USA

## Key Words

Atherosclerosis · Genetics · Inflammation · Coronary heart disease · Collateral circulation · Risk factors

## Abstract

**Objectives:** Formation of collateral circulation is an endogenous response to atherosclerosis, and is a natural escape mechanism by re-routing blood. Inflammatory response-related genes underlie the formation of coronary collaterals. We explored the genetic basis of collateral formation in man postulating interaction networks between functional Single Nucleotide Polymorphisms (SNPs) in these inflammatory gene candidates. **Methods:** The contribution of 41 genes as well as the interactions among them was examined in a cohort of 226 coronary artery disease patients, genotyped for 54 candidate SNPs. Patients were classified to the extent of collateral circulation. Stepwise logistic regression analysis and a haplotype entropy procedure were applied to search for haplotype interactions among all 54 polymorphisms. Multiple testing was addressed by using the false discovery rate (FDR) method. **Results:** The population comprised 84 patients with and 142 without visible collaterals. Among the

41 genes, 16 pairs of SNPs were implicated in the development of collaterals with the FDR of 0.19. Nine SNPs were found to potentially have main effects on collateral formation. Two sets of coupling haplotypes that predispose to collateral formation were suggested. **Conclusions:** These findings suggest that collateral formation may arise from the interactions between several SNPs in inflammatory response related genes, which may represent targets in future studies of collateral formation. This may enhance developing strategies for risk stratification and therapeutic stimulation of arteriogenesis.

Copyright © 2008 S. Karger AG, Basel

## Introduction

Formation of collateral circulation is an endogenous response to occlusive atherosclerotic arterial disease, and could be seen as a natural escape mechanism by re-routing blood to areas jeopardized by critical blood flow [1,

J. Zhang and J.J. Regieli contributed equally to this work.

## KARGER

Fax +41 61 306 12 34  
E-Mail karger@karger.ch  
www.karger.com

© 2008 S. Karger AG, Basel  
0001-5652/08/0000-0000\$24.50/0

Accessible online at:  
www.karger.com/hhe

Prof. P.A. Doevendans  
Department of Cardiology, University Medical Centre Utrecht  
Heidelberglaan 100  
NL-3584 CX, Utrecht (The Netherlands)  
Tel. +31 30 250 6176, Fax +31 30 250 5423, E-Mail p.doevendans@umcutrecht.nl

© S. Karger AG, Basel  
**PROOF Copy  
for personal  
use only**  
ANY DISTRIBUTION OF THIS  
ARTICLE WITHOUT WRITTEN  
CONSENT FROM S. KARGER  
AG, BASEL IS A VIOLATION  
OF THE COPYRIGHT.

2]. In ischemic heart disease, collaterals potentially improve prognosis [3], by preserving myocardial function and perfusion [4, 5] and reducing infarct size [6, 7]. Understanding the genomic program leading to collateral formation may be of fundamental importance to improve diagnosis, treatment and prevention of coronary artery diseases. Despite its importance, currently the determinants and mechanisms of collateral artery formation in man are incompletely understood. A major trigger in the etiology of collateral formation is chronic ischemia and the extent of atherosclerotic burden. However, in the presence of a certain fixed degree of stenosis [8] or even sudden occlusion [9], there is ample evidence to support a large variation between individuals in extent of collateral formation. This variation is only partly attributable to clinical characteristics. Seminal studies on etiology of neovascularization [10] have shown that several biochemical pathways underlie formation of collaterals. Notably the processes occurring are the sprouting of new capillaries (angiogenesis) [11] and the maturation of precursor collaterals (arteriogenesis) [12] involving remodelling by monocyte activation and inflammation. Accordingly, the use of immunosuppressants in coronary stents is associated with less functional collateral circulation [13]. A recent study seeking a genomic model of arteriogenesis using microarray data for temporal functional analyses of gene expression in mice showed that inflammatory response-related genes comprised the largest cluster among the up-regulated genes [14]. These fascinating insights in other species support a genetic component in the formation of collaterals, where the action of multiple genes in inflammatory response pathways may be required. Once a set of candidate genes has been identified, a number of potentially modifying genetic variants may exist in these genes. These genetic variants, in concert with specific gene-environment effects and gene-gene interactions, may play an important role in the pathogenesis of collateral formation. For example, the effect of an allele of a certain SNP at a locus may be mediated by specific alleles on other loci. Therefore, any single gene or locus based study may not be sufficient for identification of the determinants or modifying factors in the development of collaterals.

In the current study, we postulated that the genetic basis of collateral formation is comprised of interaction networks between functional SNPs in several inflammatory genes. We therefore analysed genotypes of patients with coronary artery disease, using a panel of polymorphisms in 41 genes that are related to various inflammatory processes. We first analyze the data with a stepwise

logistic regression [15]. Then we conduct a search for locus-locus interactions through use of the haplotype entropy procedure [16].

## Methods

### *Study Population*

The source population for the present study originated from the Second Manifestations of Arterial Disease (SMART) Study, an ongoing prospective cohort study at the University Medical Centre Utrecht designed to establish the prevalence of concomitant arterial diseases and risk factors for atherosclerosis in a high-risk population [17]. The local ethics committee approved the study, and all participants gave their written informed consent. Clinical information was obtained using a standardized health questionnaire. DNA was extracted from buffycoats using an extraction kit (QIagen biosystems) and stored at  $-80^{\circ}\text{C}$ . For the present cross-sectional study, based on a case-cohort study investigating determinants and prognostic value of coronary collateral formation, 226 consecutive patients referred for revascularization and included between January 1998 and July 2002 were enrolled. All patients had symptoms of stable angina pectoris at the time of enrollment.

### *Angiographic Assessment*

Coronary angiograms originated from the participants' diagnostic work-up preceding a scheduled coronary intervention at the UMC Utrecht. Two experienced observers, blinded to all patient characteristics, independently reviewed the angiograms. The Rentrop classification [18] was used to determine the extent of collateral formation (grade 0: no filling of collateral vessels; grade 1: filling of collateral vessels without any epicardial filling of the recipient artery; grade 2: partial epicardial filling by collateral vessels of the recipient artery; and grade 3: complete epicardial filling by collateral vessels of the recipient artery). The patients were assigned to either the case group (with visible collaterals, i.e., Rentrop score  $\geq 1$ ) or the control group (without visible collaterals, i.e., Rentrop score = 0). Severity of coronary artery disease was defined (single, two, or three vessel disease) as the degree of the most severe stenosis (50–90, 90–99, or 100% stenosis). A 50% diameter reducing stenosis was regarded as significant [19].

### *Genotyping*

Genes involved in various inflammatory processes were considered as candidates. A SNP assay designed by Roche Molecular Systems was used in this study. This assay for detecting bi-allelic variants contained mostly coding non-synonymous SNPs previously reported in genes involved in inflammatory processes. Each DNA sample was amplified using two multiplex polymerase chain reactions, and the alleles were genotyped simultaneously using an array of immobilized, sequence-specific oligonucleotide probes as described elsewhere [20]. The assay was composed of a panel of 51 SNPs, described previously [21, 22] and listed in table 1. One of them [GC(C35717A)] was not used in the analysis afterward because the poor quality of genotyping on this locus. In addition, four extra polymorphisms (Haptoglobin(Hp2FS>Hp1S), MCP1(A-2518G), VEGF(G405C), TLR4(A896G)) were

**Table 1.** SNPs used in the present study

Block and SNP (symbol)	Location	Gene name	rs#	OMIM #
<b>1:00</b>				
T707C [V1]	1p32-p31	VCAM1	1041163	192225
A153C [V2]	1q22-q25	SELE	5361	131210
G40A [V3]	1q21-24	SELP	6131	173610
G75271T [V4]	1q21-24	SELP	6133	
C8700A [V5]	1q31-32	IL10	1800872	124092
<b>2:00</b>				
T549C [V6]	2q12-21	IL1A	1800587	147760
C1423T [V7]	2q14	IL1B	16944	147720
C4336T [V8]	2q14	IL1B	1143634	
C875T [V9]	2q33	CTLA4	5742909	123890
A1241C [V10]	2q33	CTLA4	231775	
<b>3:00</b>				
G46295A [V11]	3p21	CCR2	1799864	601267
C320T [V12]	3p21.3	CCR3	5742906	601268
Wt/ $\Delta$ 580-611 [V13]	3p21	CCR5	333	601373
G59029A [V14]	3p21	CCR5	1799987	
G482A [V15]	3p24-26	IL5RA	2290608	147851
<b>4:00</b>				
G35706 [V16]	4q12-13	GC	7041	139200
<b>5:00</b>				
C582T [V17]	5q23-31	IL4	2243250	147780
C4045T [V18]	5q31	IL13	1295686	147683
A1633G [V19]	5q31-32	ADRB2	1042713	109690
C1666G [V20]	5q31-32	ADRB2	1042714	
C2078T [V21]	5q31-32	ADRB2	1800888	
C2232T [V22]	5q22-32	CD14	2569190	158120
C883A [V23]	5q31	TCF7	5742913	189908
A383T [V24]	5q31.1	TCF7	244656	189908
C4244T [V25]	5q31-35	IL9	2069885	146931
T2600C [V26]	5q31.1	CSF2	25882	138960
A620C [V27]	5q35	LTC4S	730012	246530
<b>6:00</b>				
G405C [V28] (*)	6p12	VEGF	192240	192240
A1069G [V29]	6p21.3	LTA	909253	153440
G3787A [V30]	6p21.3	TNF	1800629	191160
G3857A [V31]	6p21.3	TNF	361525	
<b>7:00</b>				
G589C [V32]	7p15-21	IL6	1800796	147620
G987C [V33]	7p15-21	IL6	1800795	
A498G [V34]	7q35-36	NOS3	1800779	163729
G7002T [V35]	7q35-36	NOS3	1799983	
<b>8:00</b>				
A896G [V36] (*)	9q32-33	TLR4	603030	603030
A2416G [V37]	9q32-34	C5	187611	120900
<b>9:00</b>				
G880A [V38]	10q11.1	SDF1	1801157	600835
<b>10:00</b>				
A7297G [V39]	11q13	FCERB1	569108	147138
G587A [V40]	11q11-ter	UGB	3741240	192020

**Table 1** (continued)

Block and SNP (symbol)	Location	Gene name	rs#	OMIM #
11:00				
T12022C [V41]	12q13.1	VDR	2228570	601769
G45082A [V42]	12q13.1	VDR	1544410	
12:00				
Hp2FS>Hp1S [V43]*	16q22.1	Haptoglobin (Hp)		140100
A398G [V44]	16p11.2-12.1	IL4R	1805010	147781
T1682C [V45]	16p11.2-12.1	IL4R	1805015	
A1902G [V46]	16p11.2-12.1	IL4R	1801275	
13:00				
A-2518G [V47]*	17q11.2-12	MCP1	158105	158105
G361A [V48]	17q21.1-21.2	SCYA11	3744508	601156
C231T [V49]	17q11.2-12	NOS2A	1137933	163730
G1169A [V50]	17q21.1-21.2	SCYA11	4795895	601156
14:00				
A120T [V51]	19p13.2	ICAM1	5491	147840
G657A [V52]	19p13.2	ICAM1	1799969	
C629T [V53]	19q13.1	TGFB1	1800469	190180
C364G [V54]	19p13.2-13.3	C3	2230199	120700

\* These SNPs were genotyped using Restriction Fragment Length Polymorphism assays.

genotyped separately using Restriction Fragment Length Polymorphism assays. This gives 54 polymorphisms totally. These genetic variants were numbered [V1 through V54] and grouped into blocks according to chromosomes where they were located (see table 1).

#### Statistical Analysis

##### Search for Main Factors

To detect main effects of individual genes, we conducted a logistic regression analysis on the genotype data following the stepwise approach of Cordell and Clayton [15]. We took the Rentrop score as a response variable and the polymorphisms as covariates. The covariate will take value of  $-0.5$  if the genotype at the corresponding locus is 0 (homozygous with wild type), and  $0.5$  if the genotype is 1 (homozygous with rare variant), or 2 (heterozygous), or missing. We set the threshold  $p$  value of  $0.05$  in the stepwise selection of these polymorphisms.

##### Search for Haplotype Interactions

As pointed out by Zhang et al. [16], the above logistic regression approach may be unable to detect the contributions of haplotype interactions to collateral formation. This is because the logistic regression approach allows only for testing genotype interactions; possibly miss the interactions reflected only at the haplotype level. To contend with this disadvantage, we applied the haplotype entropy procedure of Zhang et al. [16] to search for interacting pairs of polymorphisms between and within 14 blocks defined in table 1 that predispose to collateral formation. This was done in two stages. In the first stage we searched for the interactions between and within chromosome blocks in each of two

individual groups by performing a permutation procedure. The significance of the results were shown by both  $p$  values and  $Z$ -scores of Zhang et al. [16].

In the second stage, the interactions predisposing collaterals were then found by contrasting the interaction patterns observed for cases with the interaction patterns for controls. In this stage, we set the thresholds ( $p_1$ ,  $p_2$ ) for the observed  $p$  values. If the  $p$  value of the patients without visible collaterals was larger than  $p_2$  and the  $p$  value of the patients with visible collaterals was less than or equal to  $p_1$ , we assumed an up interaction associated with collaterals. This infers that, in contrast to the patients without visible collaterals, there is a significant interaction between two chromosome blocks under consideration in the patients with visible collaterals. If the  $p$  value of the patients with visible collaterals was larger than  $p_2$  and the  $p$  value of the patients without visible collaterals was less than or equal to  $p_1$ , we assumed a down interaction associated with collaterals. This infers that, in contrast to the patients without visible collaterals, there is no significant interaction between two chromosome blocks under consideration in the patients with visible collaterals. Zhang et al. [16] did a simulation study on how to set the thresholds ( $p_1$ ,  $p_2$ ). Here, we applied two settings of thresholds, notably  $[0.02, 0.15]$  for stronger interactions and  $[0.05, 0.15]$  for detecting relatively weaker interactions. The opting for these thresholds for  $p$  values was validated by the simulation in Zhang et al. [16].

The search for interacting SNP pairs above involves several hundreds of the tests. There will be a potential impact of multiple testing on the false positive error. Since these tests are highly correlated, a conventional Bonferroni adjustment for multiple testing would be too conservative. To address this issue, we adopted FDR,

**Table 2.** Pairs of SNPs in linkage disequilibrium in both the case and control groups (with and without visible collateral respectively)

Block pair/chrom. loc. SNP pair/gene pair	Visible collaterals		Invisible collaterals	
	p value	Z score	p value	Z score
[3,6]/[3,6] [V11:V30]/[CCR2:TNF]	0.015	-0.643	0.005	-6.234
[1,1]/[1,1] [V2:V3]/[SELE:SELP]	0.000	18.538	0.000	-8.668
[2,2]/[2,2] [V6:V8]/[IL1A:IL1B]	0.000	-8.153	0.000	-15.940
[V7:V8]/[IL1B:IL1B]	0.005	-0.796	0.040	-2.208
[V9:V10]/[CTLA4:CTLA4]	0.000	-5.189	0.005	-4.106
[3,3]/[3,3] [V11:V14]/[CCR2:CCR5]	0.000	-0.725	0.040	-2.956
[V13:V14]/[CCR5:CCR5]	0.000	-0.818	0.000	-5.111
[5,5]/[5,5] [V19:V20]/[ADRB2:ADRB2]	0.000	-24.281	0.000	-12.471

a widely used significance measure for the overall error rate in multiple testing [23]. FDR is defined by the expected proportion of false positives among the tests called significant. Here we adopted a truncated Z-score mixture model based procedure (unpublished report, Zhang and Liang 2007) to identify subsets of the significant tests with the pre-specified FDR values. In the procedure the truncated Z-scores are assumed to follow a mixture model,

$$f(z) = \pi_0 f_0(z) + (1 - \pi_0) f_1(z),$$

where  $\pi_0$  is the probability that a null hypothesis is true, and  $f_0$  and  $f_1$  are fitted by the exponential power mixtures (unpublished report, Zhang and Liang 2007). Note that the device of truncation has been used to deal with the potential outliers in the test scores. The so-called Expectation-Conditional-Maximisation algorithm are then applied to calculate estimators,  $\hat{\pi}_0$  and  $\hat{F}_0$ , of  $\pi_0$  and  $F_0$ , where  $F_0$  is the distribution function of  $f_0$ . The FDR can be estimated by using the formula

$$\hat{Fdr}(z_0) = \frac{N \hat{\pi}_0 (1 - \hat{F}_0(z_0))}{\#\{z_i : z_i \geq z_0\}},$$

where  $N$  is the number of the test scores and  $\#\{z_i : z_i \geq z_0\}$  denotes the number of tests with test score larger than or equal to the threshold  $z_0$ . We select the threshold  $z_0$  so that  $\hat{Fdr}(z_0)$  reach the pre-specified FDR value.

#### Haplotyping the Interaction Networks

To explain the potential mechanism behind these interaction networks, collateral-predisposing coupling haplotypes were identified within these networks. This requires addressing potential over-fitting, since the dimension of space of the candidate haplotypes is much larger than the sample size. To tackle this problem we first expanded the above networks by including those SNPs which are in linkage disequilibrium with some SNPs in the above networks (table 2). Then we haplotyped parts of a network (usually including 3 or 4 SNPs) and merged these parts in an agglomerative way. We stopped this merging process if no of the resulting

haplotypes are approximately equally frequent in the control group and in the case group. Note that this strategy is based on the assumption that among many haplotype combinations, only a few are believed to be really coupling in both groups (these haplotype combinations are called base-line haplotype combinations). To show the strength of the evidence that a coupling haplotype combination is associated with collateral formation, we calculated the odds ratio of cases ( $\text{Rentrop} \geq 1$ ) relative to controls ( $\text{Rentrop} = 0$ ). More specifically, the odds ratio  $OR = (n_{11} * n_{22}) / (n_{21} * n_{12})$ , where  $n_{11}$  and  $n_{12}$  are the frequencies of the baseline haplotype combination in the two groups while  $n_{21}$  and  $n_{22}$  are the frequencies of a haplotype combination in the two groups. The 95% CI for the odds ratio was calculated based on Woolf's method [24].

## Results

### Angiographic Assessment

Collaterals were visible in 84 (37%) of the 226 participants included in the analyses. This prevalence of collaterals agreed with other similar populations previously described [25] in our centre. Multi-vessel disease and smoking occurred more often in the group with visible collateral than in the group without visible collateral (see table 3).

### Search for Main Risk Factors

The stepwise selection procedure on the genotype data provided only marginal suggestion for major effects of 9 polymorphisms on collateral formation (with  $0.025 < p \text{ values} < 0.05$ ). These polymorphisms were located in 9 genes listed in table 4.

**Table 3.** Main characteristics of study subjects (n = 226)\*

Characteristics	Visible collaterals (n = 84)	Invisible collaterals (n = 142)	p value
Age, years	57 ± 9.5	58 ± 9	0.5705
Male	72 (86)	113 (80)	0.3279
Clinical conditions			
Systemic hypertension	25 (30)	62 (44)	0.0531
Hyperlipidemia	71 (85)	117 (82)	0.8184
Diabetes	23 (27)	23 (16)	0.0647
Alcohol consumption	69 (82)	117 (82)	1
Smoking	32 (38)	31 (22)	0.0131
Previous MI	41 (49)	54 (39)	0.1336
Duration since MI until PTCA, years**	1 (0-19)	1 (0-8)	
Previous PTCA and/or CABG	26 (31)	47 (33)	0.8522
Multi-vessels disease	47 (56)	42 (30)	<0.0001

\* Number presented in count (%) or mean ± standard deviation. MI = myocardial infarction; PTCA = percutaneous transluminal coronary angioplasty; CABG = coronary artery bypass grafting.

\*\* Number presented in median (10–90% percentiles). The p values were calculated based on the two-sample test for proportions or for means.

**Table 4.** The loci and alleles that have main effects on the formation of collaterals

Gene	Locus	Genotypic allele	p value
IL1A	V6 (T549C)	1 (TT)	0.0325
VDR	V42 (G45082A)	0 (AA)	0.0488
SCYA11	V48 (G361A)	1 (GG)	0.0278
CD14	V22 (C2232T)	1 (CC)	0.0351
TCF7	V23 (C883A)	0 (AA)	0.0474
SDF1	V38 (G880A)	0 (AA)	0.0443
VEGF	V28 (G405C)	1 (GG)	0.0426
LTC4S	V27 (A620C)	0 (CC)	0.0470
NOS3	V34 (A498G)	0 (GG)	0.0360

#### Search for Haplotype Interaction

To apply the haplotype entropy procedure [16] to the pairs of polymorphisms between and within 14 blocks defined in table 1, we first set the thresholds for the p values at the level of  $(p_1, p_2) = (0.02, 0.15)$ . Consequently, down interaction was observed in 9 pairs of polymorphisms [V1:V14, V4:V11, V14:V28, V16:V38, V29:V49, V30:V49, V37:V40, V18:V25, V23:V26]. Up interaction was observed in 4 pairs of polymorphisms [V10:V38, V13:V41, V16:V26, V28:V30]. Setting the thresholds at the level of  $(p_1, p_2) = (0.05, 0.15)$ , we detected 3 additional down-interacting [V22:V37, V2:V5, V48:V50] and 6 additional

up-interacting pairs [V14:V29, V15:V29, V17:V23, V17:V26, V33:V35, V43:V44]. See table 5 for more details. These interactions might influence the formation of collaterals via interaction networks as schemed in figure 2, which also displays Linkage Disequilibrium between pairs of SNPs in both the cases and controls (detected using the same procedure).

To address the issue of false discovery in the above findings and select the statistically validated interactions, we applied the FDR controlling procedure of Zhang and Liang to the Z-scores obtained in the multiple tests above. We identified two subsets of interacting pairs with two pre-specified FDR values. At the FDR value of 0.056, we identified 8 significant up- or down-interacting SNP pairs: [V1:V14, V4:V11, V10:V38, V14:V28, V37:V40, V18:V25, V28:V30, and V43:V44]. Among these pairs  $0.056 \times 8 \approx 0.44$  pairs are expected to be false positive. If we relax the FDR value to 0.19, we found 16 up- or down-interacting SNP pairs: [V1:V14, V4:V11, V10:V38, V14:V28, V14:V29, V15:V29, V22:V37, V37:V40, V2:V5, V17:V23, V17:V26, V18:V25, V23:V26, V28:V30, V33:V35, and V43:V44]. Among the above 16 pairs of interacting SNPs,  $0.19 \times 16 \approx 3$  pairs are expected to be false positive.

Table 3 indicates that the case and control groups in the study were approximately matched for all baseline features, except those related to smoking and multi-vessel disease (disease severity). To assess the possible confounding effects of these two factors on collateral forma-

**Table 5.** Haplotype interactions that may influence the formation of collaterals based on all subjects

Block pair/chrom. loc. or SNP pair/gene pair	No visible collaterals		Visible collaterals		Interaction	FDR level	
	p value	Z score	p value	Z Score		0.056	0.19
[1,3]/[1,3]	0.005	-2.106	0.110	-1.266	down	no	yes
[V1:V14]/[VCAM1:CCR5]	0.010	-2.798	0.580	0.216	down	yes	yes
[V4:V11]/[SELP:CCR2]	0.005	-3.210	0.980	-0.080	down	yes	yes
[2,9]/[2,10]	0.925	1.263	0.010	-3.073	up		
[V10:V38]/[CTLA4:SDF1]	0.430	0.052	0.000	-4.395	up	yes	yes
[3,6]/[3,6]	0.000	-3.216	0.350	-0.299	down	yes	yes
[V14:V28]/[CCR5:VEGF]	0.015	-2.541	0.960	0.837	down	yes	yes
[V14:V29]/[CCR5:LTA]	0.985	1.540	0.030	-2.156	up	no	yes
[V15:V29]/[IL5RA:LTA]	0.833	-0.149	0.030	-2.114	up	no	yes
[3,11]/[3,12]	0.655	0.552	0.045	-2.017	up	no	no
[V13:V41]/[CCR5:VDR]	0.670	0.414	0.015	-0.489	up	no	no
[4,5]/[4,5]	0.525	0.081	0.000	-2.104	up	no	no
[V16:V26]/[GC:CSF2]	0.433	-0.324	0.005	-0.886	up	no	no
[4,9]/[4,10]	0.000	-0.756	0.170	-0.299	down	no	no
[V16:V38]/[GC:SDF1]	0.020	-0.725	0.215	-0.330	down	no	no
[4,12]/[4,16]	0.965	1.897	0.020	-2.110	up	no	no
[5,8]/[5,9]	0.940	1.579	0.150	-0.962		no	no
[V22:V37]/[CD14:C5]	0.045	-1.939	0.980	1.540	down	no	yes
[6,13]/[6,17]	0.055	-1.658	0.725	0.576		no	no
[V29:V49]/[LTA:NOS2A]	0.008	-0.817	0.763	0.716	down	no	no
[V30:V49]/[TNF:NOS2A]	0.018	-0.513	0.373	-0.384	down	no	no
[8,10]/[9,11]	0.085	-1.506	0.750	0.648		no	no
[V37:V40]/[C5:UGB]	0.020	-2.722	0.480	0.114	down	yes	yes
[1,1]/[1,1]							
[V2:V5]/[SELE:IL10]	0.040	-1.950	0.465	-0.165	down	no	yes
[5,5]/[5,5]							
[V17:V23]/[IL4:TCF7]	0.995	1.453	0.045	-2.208	up	no	yes
[V17:V26]/[IL4:CSF2]	0.455	0.011	0.030	-2.124	up	no	yes
[V18:V25]/[IL13:IL9]	0.005	-3.061	0.765	0.457	down	yes	yes
[V23:V26]/[TCF7:CSF2]	0.015	-2.335	0.410	-0.778	down	no	yes
[6,6]/[6,6]							
[V28:V30]/[VEGF:TNF]	0.210	-0.763	0.000	-12.493	up	yes	yes
[7,7]/[7,7]							
[V33:V35]/[IL6:NOS3]	0.720	0.486	0.045	-2.295	up	no	yes
[12,12]/[16,16]							
[V43:V44]/[Hp:IL4R]	0.525	0.155	0.040	-2.079	up	yes	yes
[13,13]/[17,17]							
[V48:V50]/[SCYA11:SCYA11]	0.030	-0.911	0.525	-0.104	down	no	no

The SNP/gene pairs which showed significant up- or down-interactions via the contrast of the interaction patterns between the group with visible collaterals and that without visible collaterals. The significance of these interactions was measured by the haplotype-entropy based p values and Z values [16].

'yes' = Stands for passing the FDR criteria.

tion, we excluded the subjects who have the smoking habit or multi-vessel disease and performed the haplotype entropy procedure on the remaining 99 subjects (75 controls and 24 cases). The results, summarised in table 6, show a reduction in the number of significant interacting SNP pairs. Only 5 of the 16 previously identified interac-

tions remain significant with the FDR less than 0.19: [V14:V28, V15:V29, V37:V40, V17:V23, V18:V25, and V23:V26]. Note that the testing power of the haplotype entropy procedure is decreasing in the sample size. This implies that the above reduction might be caused by the small size of the remaining case group.

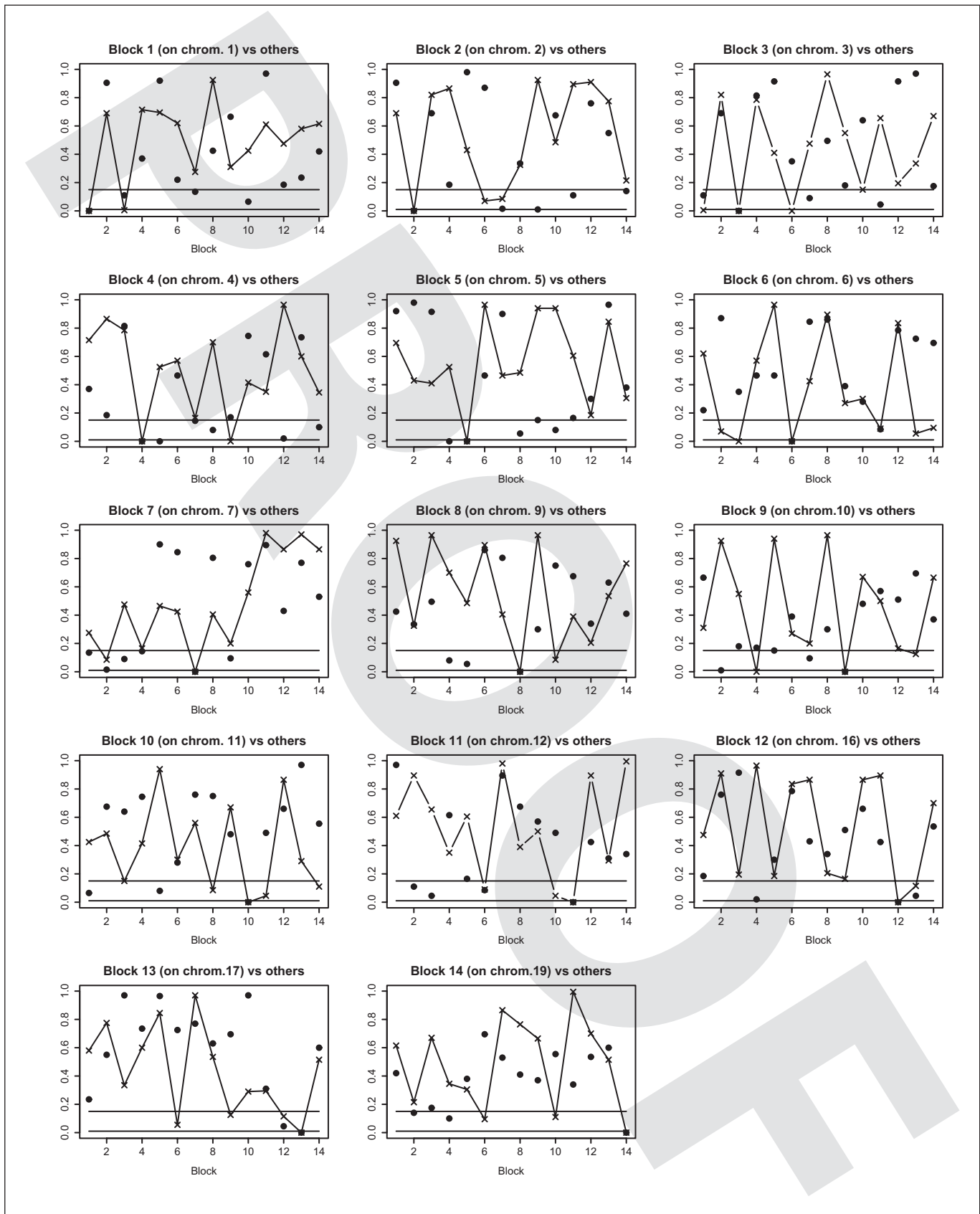
**Table 6.** Haplotype interactions that may influence the formation of collaterals based on non-smoking and non multi-vessel diseased subjects

Block pair/chrom. loc. or SNP pair/gene pair	No visible collaterals		Visible collaterals		Interaction
	p value	Z score	p value	Z score	
[1,3]/[1,3]					
[V1:V14]/[VCAM1:CCR5]	0.125	-1.211	0.495	-0.131	
[V4:V11]/[SELP:CCR2]	0.060	-2.573	1.000	0.411	
[2,9]/[2,10]					
[V10:V38]/[CTLA4:SDF1]	0.270	-0.648	0.110	-1.961	
[3,6]/[3,6]					
[V14:V28]/[CCR5:VEGF]	0.005	-2.900	0.845	0.424	down
[V14:V29]/[CCR5:LTA]	0.995	1.733	0.405	-0.140	
[V15:V29]/[IL5RA:LTA]	0.930	0.073	0.975	0.889	
[3,11]/[3,12]					
[V13:V41]/[CCR5:VDR]	0.745	0.5823	0.035	-3.613	up
[4,5]/[4,5]					
[V16:V26]/[GC:CSF2]	0.510	-0.107	0.305	-0.588	
[4,9]/[4,10]					
[V16:V38]/[GC:SDF1]	0.205	-0.409	0.410	-0.109	
[5,8]/[5,9]					
[V22:V37]/[CD14:C5]	0.605	0.536	0.455	0.127	
[6,13]/[6,17]					
[V29:V49]/[LTA:NOS2A]	0.015	-0.772	0.620	0.617	down
[V30:V49]/[TNF:NOS2A]	0.555	-0.439	0.495	-0.408	
[8,10]/[9,11]					
[V37:V40]/[C5:UGB]	0.010	-2.844	0.990	1.164	down
[1,1]/[1,1]					
[V2:V5]/[SELE:IL10]	0.710	0.369	0.135	-0.768	
[5,5]/[5,5]					
[V17:V23]/[IL4:TCF7]	0.990	1.789	0.000	-10.68	up
[V17:V26]/[IL4:CSF2]	0.390	-0.480	0.200	-0.674	
[V18:V25]/[IL13:IL9]	0.010	-3.162	0.155	-1.071	down
[V23:V26]/[TCF7:CSF2]	0.010	-2.877	0.665	-0.660	down
[6,6]/[6,6]					
[V28:V30]/[VEGF:TNF]	0.990	1.175	0.905	0.980	
[7,7]/[7,7]					
[V33:V35]/[IL6:NOS3]	0.905	1.089	0.645	0.383	
[12,12]/[16,16]					
[V43:V44]/[Hp:IL4R]	0.595	0.298	0.070	-2.346	
[13,13]/[17,17]					
[V48:V50]/[SCYA11:SCYA11]	0.115	-1.938	0.705	-0.207	

The SNP/gene pairs which showed significant up- or down-interactions via the contrast of the interaction patterns between the group with visible collaterals and that without visible collaterals. The significance of these interactions was measured by the haplotype-entropy based p values and Z values [16].

We also conducted the search for block-level interactions among these 14 unlinked blocks, using the above entropy method. The search was performed on the case and control groups separately. The p values for these two groups were compared by plotting them in graphs, as shown in figure 1. These block pairs were selected by use of the thresholds for the p values at the level of ( $p_1, p_2$ ) =

**Fig. 1.** The p values of testing the interactions of blocks 1–14 with the other blocks, for the subjects with and without visible collaterals, respectively. The dots are for the subjects with visible collaterals and the lines with crosses are for the subjects without visible collaterals. The two straight lines are for the upper and lower thresholds (that is, 0.15 and 0.01) of p values respectively.



(0.01, 0.15). We obtained 4 interacting block pairs on the chromosome pairs [2, 10], [3, 6], [4, 5], and [4, 10]. See table 5. Among these selected pairs of blocks, [3, 6] and [4, 10] are down-interacting. The up-interaction pairs were on the chromosome pairs [2, 10] and [4, 5]. These interacting pairs have passed the FDR criteria (i.e., FDR being less than 0.19).

Finally, we adopted the strategy mentioned in the last section to find the coupling haplotypes in the networks as follows. The first set of coupling haplotypes, related to the gene/SNP combination VCAM1-SELP-CCR2-CCR5-IL5RA-VEGF-LTA-TNF-NOS2A/(V1-V4-V11-V14-V15-V28-V29-V30-V49) include: TGGGGGAGC, CGGAGCAGC, TGGGGCAGC, TGGGGGAGT, TGGAAGAGC.

**Table 7.** The coupling haplotypes in the SNP interaction network

SNP-combination/ coupling-haplotype	EPF	OR	95% CI
V1-V4-V11-V14-V15-V28-V29-V30-V49			
TGGGGGAGC	0.174	1.711	[0.686, 4.270]
CGGAGCAGC	0.068	4.857	[0.216, 109]
TGGGGCAGC	0.139	1.619	[0.611, 4.291]
TGGGGGAGT	0.075	0.405	[0.122, 1.339]
TGGAAGAGC	0.068	0.270	[0.072, 1.014]
V16-V17-V23-V26			
TCCT	0.438	2.163	[1.284, 3.643]
GCAC	0.031	5.505	[0.664, 45.655]
GCAT	0.128	1.835	[0.566, 2.995]
TTCT	0.050	2.359	[0.619, 8.997]

EPF, OR and 95% CI stand for the estimated population frequency, odds ratio, and confidence interval at the level of 95% respectively.

The second set of coupling haplotypes, related to the gene/SNP combination GC-IL4-TCF7-CSF2/(V16-V17-V23-V26), include: TCCT, GCAC, GCAT, TTCT. See table 7 for the details.

## Discussion

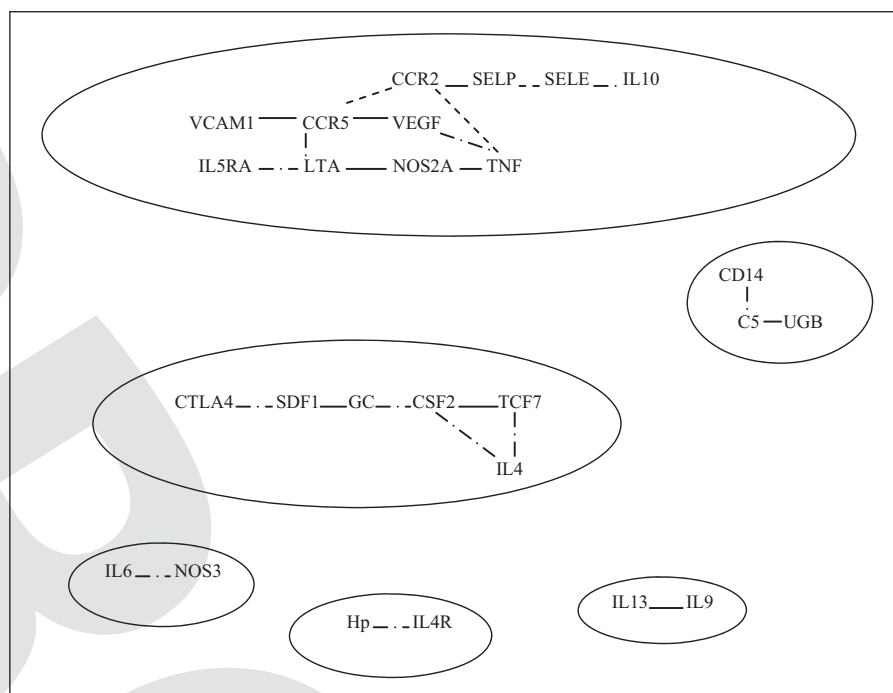
### Implications of the Results

In the present study, 54 genetic variants in candidate genes and their interactions were studied as determinants of collateral formation in the context of patients with coronary artery disease. In the first stage, the effects of 9 variants were detected using a conventional approach by running a stepwise logistic regression. In the second stage, a search for haplotype interactions using the haplotype entropy procedure put forward a set of 28 genes involved in collateral formation, including a network of 22 SNP-SNP hypothesized interactions. Twelve of the hypothesized interactions are down-interactions and the remaining 10 are up-interactions. The up-interactions would suggest that these interactions lead to a susceptibility to collateral formation, whereas the down-interactions could imply that the interactions may reduce the development of collaterals. Two sets of coupling haplotypes have been suggested. The overall multiple testing error rate has been addressed by use of the FDR controlling procedure of Zhang and Liang. In particular, 16 of these 22 interactions, where 24 genes were involved, have been found to have the estimated FDR value of 0.19. They are [VCAM1, CCR5, SELP, CCR2, CTLA4, SDF1, VEGF, LTA, IL5RA, CD14, C5, UGB, SELE, IL10, IL4, TCF7, CSF2, IL13, IL9, TNF, IL6, NOS3, Hp, and IL4R]. That is, among these 16 interactions, 3 would be expected to be significant by chance alone. These 16 interactions intro-

**Table 8.** KEGG pathway information for the 20 genes which have been identified to potentially relate to the formation of collaterals

Environmental information processing: Signalling molecules and interaction	
Cytokine-cytokine receptor interaction	CCR2, CCR5, IL6, IL4, IL4R, IL13, CSF2, IL5RA, IL9, VEGF, IL10, TNF, LTA
Signal transduction: Wnt signalling pathway	TCF7
Cell adhesion molecules for immune system	CTLA4, SELE, SELP, VCAM1
Cellular processes: Immune system	
Hematopoietic cell lineage	IL6, IL4, TNF, CD14
Complement and coagulation cascades	C5
Toll-like receptor signalling pathway	TNF, IL6
Natural killer cell mediated cytotoxicity	TNF
T cell receptor signalling pathway	CTLA4, IL4, IL10, TNF
Fc epsilon RI signalling pathway	IL4, IL13, TNF

**Fig. 2.** The interaction networks that may influence the formation of collaterals. Each connecting line indicates a relation between 2 polymorphisms: A dashed line between two genes shows the linkage disequilibrium across two groups of patients; the solid line indicates a down-interaction; the long dashed line between two genes indicates an up-interaction.



duce an interaction network as depicted in figure 2. Interestingly, based on the KEGG and GO pathway database (<http://www.genome.ad.jp/kegg/pathway> and <http://www.ebi.ac.uk/ego>), 21 of these 24 genes were found to be involved in two biological processes, namely, the environmental information processing and cell processes. In particular, [CCR2, CCR5, IL6, IL4, IL4R, IL13, CSF2, IL5RA, IL9, VEGF, IL10, TNF, and LTA] participated in the cytokine-cytokine receptor interaction. See table 8 for details. However, after correction for the disbalance of smoking and multi-vessel disease, only 5 of the above 16 interactions remain significant, where 9 genes were involved: [CCR5, VEGF, C5, UGB, IL4, TCF7, IL13, IL9, and CSF2]. This may have resulted from limited sample size. Nevertheless, our results are consistent with previous findings that collateral formation is mediated by TNF, IL6, VEGF, and CSF2 reported previously by Schultz et al. [26], Rakhit et al. [27], and Hossmann et al. [28]. This is also in agreement with a recently reported activation of the expression of ICAM1, another TNF/NFkB-induced gene in patients with coronary collaterals [26]. See table 9 for more information on the medical implications of these genes. The further biological mechanisms, by which these genes or SNPs might confer a better capacity of forming collaterals, remain to be elucidated by more experiments.

**Table 9.** Reported relevant medical implications of the candidate genes in this study

Reported relevant implications	Genes
Atherosclerosis	VCAM1 CD14 VEGF IL10 [34]
Stroke	CSF2 TGFB1 [35]
Coronary artery disease	SELE IL1A IL1B [37]
Myocardial infarction	SELP CCR2 LTA TNF IL6 SDF1 [41]
Diabetes	CTLA4 TCF7 VDR [44]
Aortic aneurysm	CCR5 [45]
Arteriogenesis	CSF2 VEGF [23]
Cardiovascular disease in diabetics	Hp [46]
Vascular wall injury	NOS2A [47]

Similar to collateral formation, the pathogenesis of atherosclerosis is also known to contain an important inflammatory component, involving the recruitment and adhesion of circulating leukocytes, particular monocytes, to injured or otherwise stimulated vascular endothelium [29, 30]. Thus, the findings in the present study imply that important genes or gene families may be shared by collateral formation and inflammatory disorders such as diabetes, cardiovascular disease, and atherosclerosis [29, 30].

### Study Limitations and strengths

To appreciate these findings, some aspects of this study merit consideration. First, the presence of collateral circulation was defined by using the dichotomized form of the Rentrop classification. This semi-quantitative method assesses spontaneously visible collateral circulation, and may not identify vessels <100  $\mu$ m in diameter or those recruitable upon coronary occlusion. More invasive methods to quantify recruitable collateral flow have been developed. However, spontaneously visible collateral arteries may prevent ischemia in 90% of patients upon subsequent coronary occlusion [31]. The functional importance of large collateral vessels is underscored by the notion that collateral flow is exponentially related to vessel radius [31]. Indeed, prognostic significance of collaterals has been demonstrated using the semi-quantitative method in patients with established coronary artery disease [31].

A second issue is the fact that, due to the small sample size, our selection of the combination of main effects of individual genes from the stepwise logistic regression is only marginally significant and might not be unique. There are other possible combinations of individual SNPs that can explain the variation between individuals. For example, we run the logistic regression on all SNPs and select these alleles and SNPs with the coefficient significance levels less than 0.05. Then we did identify a slightly different set of SNPs (IL1A, CCR2, IL5RA, GC, CD14, TCF7, LTC4S, VEGF, C5, SDF1, VDR, MCP1, SCYA11, NOS2A, and ICAM1) that significantly account for the variation in the data.

Despite these limitations, we believe our study for the first time provides a view on polymorphism-interacting networks involved in the complex processes contributing to collateral formation. Our study confirms that collateral formation may share a common set of disease-susceptibility genes with other inflammatory response related disorders. These findings offer a rationale for a large scale association study on collateral formation.

### Acknowledgements

We gratefully acknowledge Koos (J.) Plomp, cardiologist, Department of Cardiology, Heart Lung Centre Utrecht, UMC Utrecht, The Netherlands, for scoring the 100 coronary angiograms to assess interobserver variability. Moreover, Wim Gerritsen, Hatice Alpar and Anna Waldstrom are gratefully acknowledged for performing the assays. We thank Michael Edlinger for handling of the SMART data. We thank Roche Molecular Systems

(Alameda, Calif.) for providing the multiplex genotyping assays, under a research collaboration. We also thank two referees for their very constructive comments which led to a significant improvement of the presentation of our paper.

### Competing Interests

All ten authors have no conflicts of interest to disclose.

### Funding

This work was supported by grants from the Netherlands Organization for Health Research and Development (ZonMW): AGIKO grant 920-03-367 and program grant project number 904-65-095. These funding sources had no involvement in the writing of this paper or its submission.

### Appendix 1

The SMART Study Group consists of A. Algra, MD, PhD; Y. van der Graaf, MD, PhD; D.E. Grobbee, MD, PhD; G.E.H.M. Rutten, MD, PhD, Julius Center for Health Sciences and Primary Care; J.D. Banga, MD, PhD; F.L.J. Visseren, MD, PhD, Department of Internal Medicine; B.C. Eikelboom, MD, PhD; F.L. Moll, MD, PhD, Department of Vascular Surgery; L.J. Kappelle, MD, PhD, Department of Neurology; H.A. Koomans, MD, PhD, Department of Nephrology; W.P.Th.M. Mali, MD, PhD, Department of Radiology, and P.P. Th. De Jaegere, MD, PhD, Department of Cardiology, Utrecht University Medical Centre, Utrecht, The Netherlands.

### Web Resources

<http://www.ncbi.nlm.nih.gov/Omim/>  
<http://www.ncbi.nlm.nih.gov/SNP/>  
<http://www.genome.ad.jp/kegg/pathway/>  
<http://www.ebi.ac.uk/ego>

### References

- 1 Hansen JF: Coronary collateral circulation: clinical significance and influence on survival in patients with coronary artery occlusion. *Am Heart J* 1989;117:290-295.
- 2 Koerselman J, Van der Graaf Y, de Jaegere PP, Grobbee DE: Coronary collaterals: An important and underexposed aspect of coronary artery disease. *Circulation* 2003;107:2507-2511.
- 3 Seiler C: The human coronary collateral circulation. *Heart* 2003;89:1352-1357.
- 4 Khaja F, Sabbah HG, Brymer JF, Stein PD: Influence of coronary collaterals on left ventricular function in patients undergoing coronary angioplasty. *Am Heart J* 1988;117:1174-1180.

- 5 Rogers WJ, Hood WP Jr, Mantle JA, Baxley WA, Kirklin JK, Zorn GL, Nath HP: Return of left ventricular function after reperfusion in patients with myocardial infarction: importance of subtotal stenoses or intact collaterals. *Circulation* 1984;338–349.
- 6 Charney R, Cohen M: The role of the coronary collateral circulation in limiting myocardial ischemia and infarct size. *Am Heart J* 1993;937–945.
- 7 Habib GB, Heibig J, Forman SA, Brown BG, Roberts R, Terrin ML, Bolli R: Influence of coronary collateral vessels on myocardial infarct size in humans. Results of phase I thrombolysis in myocardial infarction (TIMI) trial. The TIMI Investigators. *Circulation* 1991;83:739–746.
- 8 Rentrop KP, Cohen M, Blanke H, Phillips RA: Changes in collateral channel filling immediately after controlled coronary artery occlusion by an angioplasty balloon in human subjects. *J Am Coll Cardiol* 1985;5:587–592.
- 9 Wustmann K, Zbinden S, Windecker S, Meier B, Seiler C: Is there functional collateral flow during vascular occlusion in angiographically normal coronary arteries? *Circulation* 2003;107:2213–2220.
- 10 Simons M: Angiogenesis: Where do we stand now? *Circulation* 2005;111:1556–1566.
- 11 Carmeliet P: Angiogenesis in health and disease. *Nat Med* 2003;653–660.
- 12 Schaper W, Scholz D: Factors regulating arteriogenesis. *Arterioscler Thromb Vasc Biol* 2003;23:1143–1151.
- 13 Meier P, Zbinden R, Togni M, Wenaweser P, Windecker S, Meier B, Seiler C: Coronary collateral function long after drug-eluting stent implantation. *J Am Coll Cardiol* 2007;49:15–20.
- 14 Lee CW, Stabile E, Kinnaird T, Shou M, Devaney JM, Epstein SE, Burnett MS: Temporal patterns of gene expression after acute hindlimb ischemia in mice: insights into the genomic program for collateral vessel development. *J Am Coll Cardiol* 2004;43:474–482.
- 15 Cordell HJ, Clayton DG: A unified stepwise regression procedure for evaluating the relative effects of polymorphisms within a gene using case/control or family data: application to HLA in type 1 diabetes. *Am J Hum Genet* 2002;124–141.
- 16 Zhang J, Liang F, Dassen WR, Doevendans PA, De Gunst M: Search for haplotype interactions that influence susceptibility to type 1 diabetes, through use of unphased genotype data. *Am J Hum Genet* 2003;1385–1401.
- 17 Simons PC, Algra A, van de Laak MF, Grobbee DE, van der GY: Second manifestations of ARterial disease (SMART) study: Rationale and design. *Eur J Epidemiol* 1999;15:773–781.
- 18 Cohen M, Rentrop KP: Limitation of myocardial ischemia by collateral circulation during sudden controlled coronary artery occlusion in human subjects: A prospective study. *Circulation* 1986;74:469–476.
- 19 Scanlon PJ, Faxon DP, Audet AM, Carabello B, Dehmer GJ, Eagle KA, Legako RD, Leon DF, Murray JA, Nissen SE, Pepine CJ, Watson RM, Ritchie JL, Gibbons RJ, Cheitlin MD, Gardner TJ, Garson A Jr, Russell RO Jr, Ryan TJ, Smith SC Jr: ACC/AHA guidelines for coronary angiography: Executive summary and recommendations. A report of the American College of Cardiology/American Heart Association Task Force on Practice Guidelines (Committee on Coronary Angiography) developed in collaboration with the Society for Cardiac Angiography and Interventions. *Circulation* 1999;2345–2357.
- 20 Cheng S, Grow MA, Pallaud C, Klitz W, Erlich HA, Visvikis S, Chen JJ, Pullinger CR, Malloy MJ, Siest G, Kane JP: A multilocus genotyping assay for candidate markers of cardiovascular disease risk. *Genome Res* 1999;936–949.
- 21 Barcellos LF, Begovich AB, Reynolds RL, Caillier SJ, Brassat D, Schmidt S, Grams SE, Walker K, Steiner LL, Cree BA, Stillman A, Lincoln RR, Pericak-Vance MA, Haines JL, Erlich HA, Hauser SL, Oksenberg JR: Linkage and association with the NOS2A locus on chromosome 17q11 in multiple sclerosis. *Ann Neurol* 2004;55:793–800.
- 22 Monraats PS, Pires NM, Agema WR, Zwinderman AH, Schepers A, de Maat MP, Doevendans PA, de Winter RJ, Tio RA, Waltenberger J, Frants RR, Quax PH, van Vlijmen BJ, Atsma DE, van der LA, van der Wall EE, Jukema JW: Genetic inflammatory factors predict restenosis after percutaneous coronary interventions. *Circulation* 2005;112:2417–2425.
- 23 Storey JD, Tibshirani R: Statistical significance for genomewide studies. *Proc Natl Acad Sci USA* 2003;100:9440–9445.
- 24 Agresti A: On logit confidence intervals for the odds ratio with small samples. *Biometrics* 1999;55:597–602.
- 25 Nathoe HM, Koerselman J, Buskens E, van DD, Stella PR, Plokker TH, Doevendans PA, Grobbee DE, de Jaegere PP: Determinants and prognostic significance of collaterals in patients undergoing coronary revascularization. *Am J Cardiol* 2006;98:31–35.
- 26 Schultz A, Lavie L, Hochberg I, Beyar R, Stone T, Skorecki K, Lavie P, Roguin A, Levy AP: Interindividual heterogeneity in the hypoxic regulation of VEGF: significance for the development of the coronary artery collateral circulation. *Circulation* 1999;547–552.
- 27 Rakhit RD, Seiler C, Wustmann K, Zbinden S, Windecker S, Meier B, Eberli FR: Tumour necrosis factor-alpha and interleukin-6 release during primary percutaneous coronary intervention for acute myocardial infarction is related to coronary collateral flow. *Coron Artery Dis* 2005;147–152.
- 28 Hossmann KA, Buschmann IR: Granulocyte-macrophage colony-stimulating factor as an arteriogenic factor in the treatment of ischaemic stroke. *Expert Opin Biol Ther* 2005;1547–1556.
- 29 Hansson GK: Inflammation, atherosclerosis, and coronary artery disease. *N Engl J Med* 2005;1685–1695.
- 30 Ross R: Atherosclerosis – an inflammatory disease. *N Engl J Med* 1999;115–126.
- 31 Werner GS, Ferrari M, Heinke S, Kuethle F, Surber R, Richartz BM, Figulla HR: Angiographic assessment of collateral connections in comparison with invasively determined collateral function in chronic coronary occlusions. *Circulation* 2003;107:1972–1977.