SUPPLEMENTARY MATERIAL

SUPPLEMENTARY METHODS

Atherosclerosis Risk in Communities (ARIC) Study

The traditional risk factors (TRFs) in the ARIC coronary heart disease (CHD) Risk Score are age, gender, race (not included in this analysis), smoking, diabetes, systolic blood pressure, antihypertensive medication use, total cholesterol, and high-density lipoprotein cholesterol (HDL-C) [1]. Blood pressure measurements were performed 3 times with the participant seated with a random-zero sphygmomanometer, and the last 2 measurements were averaged. Questionnaires were administered to assess current use of antihypertensive medication. Plasma total cholesterol was measured by enzymatic method [2], and HDL-C was measured by dextran-magnesium precipitation of non–high-density lipoproteins [3]. Diabetes was defined as a fasting glucose level ≥126 mg/dL, nonfasting level ≥200 mg/dL, self-reported physician diagnosis of diabetes, or diabetic medication use. Cigarette smoking status was defined as "current" or "not current."

Rotterdam Study

The Rotterdam Study is a population-based cohort study that assesses the occurrence of risk factors for chronic diseases in the elderly [4-6]. In brief, all inhabitants of Ommoord, a district of Rotterdam in the Netherlands, who were 55 years or older were invited and 7983 agreed to participate. Written informed consent was obtained from all participants, and the Medical Ethics Committee of the Erasmus Medical Center approved the study. Given that the ARIC Study is

composed of individuals younger than 65 years at baseline, we restricted the analysis in the Rotterdam study to participants who were under 65 years of age at baseline (n=2068).

To identify incident CHD, information from baseline (1990–1993) until January 1, 2007, was collected. Fatal or nonfatal myocardial infarction reported by general practitioners in the research area, letters from medical specialists, and discharge reports for hospitalized patients were the sources of information used. Incident CHD was defined as fatal or nonfatal myocardial infarction (ICD-10 code I21), coronary artery bypass grafting, or percutaneous transluminal coronary angioplasty.

The baseline examinations for the Rotterdam Study took place in 1990–1993. Participants were visited at home for an interview. Information on current health status, medical history, use of medication, and smoking status was obtained during the interview. The interview was followed by 2 visits at the research center for blood sampling and further examinations.

All participants were asked whether they had ever had a heart attack. A 12-lead electrocardiogram (ECG) was stored digitally and analyzed by using the Modular ECG Analysis System. Myocardial infarction found on ECG was based on criteria partly derived from the Minnesota code. A history of myocardial infarction was considered present in case of a self-report of myocardial infarction confirmed by ECG or additional clinical information, or the presence of an ECG characteristic of prior myocardial infarction [7,8].

For genotyping, the Illumina 550K array was used. The exclusion criteria for SNPs were minor allele frequency ≤1%, Hardy-Weinberg equilibrium $P<10^{-5}$, or SNP call rate ≤90% and resulted in data on 530,683 SNPs. Imputation was done with reference to HapMap release 22 CEU using the maximum likelihood method implemented in MACH. The final population for this analysis comprised 5974 individuals.

Assessment of clinical events started at baseline, and follow-up examinations were carried out periodically in 1995–1996, 1997–1999, and 2002–2004. Participants were continuously monitored for fatal and nonfatal cardiovascular events through automated linkage with files from general practitioners and pharmacies working in the study district of Ommoord. In addition, all medical records of the participants under the care of general practitioners outside the study area were checked annually. Two research physicians independently coded all reported events according to the *International Classification of Diseases,* 10th edition (ICD-10). Codes on which the research physicians disagreed were discussed to reach consensus. Finally, a medical expert in cardiovascular disease, whose judgment was considered final, reviewed all events. Information on vital status was obtained regularly from municipal health authorities in Rotterdam.

Framingham Offspring Study

The Framingham Offspring Study cohort included the children of the original Framingham cohort, spouses of those children, and the Third Generation cohort (the children of the Offspring cohort) [9,10]. Subjects (n=2339) were followed from study entry in 1971–1975 until the first occurrence of incident CHD, death, or 2006 if the person did not have a CHD event. CHD was defined as nonfatal myocardial infarction diagnosed by ECG or elevated enzymes, or death due to CHD.

The Framingham Heart Study was established in 1948 with the systematic recruitment of 5209 men and women from two-thirds of the households in the town of Framingham, Massachusetts. The methods for recruitment and clinical covariate collection have been described previously for the original Framingham Heart Study [11], the Framingham Offspring

Study cohort (5124 children of the original cohort and spouses of those children) [9], and the Third Generation cohort (4095 children of the Offspring cohort, recruited beginning in 2002) [10]. Briefly, study participants attended periodic exams in which a medical history was taken, questionnaire data were collected on environmental and behavioral risk factors, blood was drawn for laboratory measures, and a physical examination was conducted. Subjects included here were 2339 participants of the Offspring cohort who were free of CHD at baseline and had genetic data. Subjects were followed from study entry at exam 1 (1971–1975) for the first occurrence of CHD (incident CHD), or to death, or to 2006 if the person did not develop CHD. Of the 2339 individuals, 221 developed incident CHD during the follow-up period. An endpoint committee consisting of 3 physicians reviewed all suspected coronary disease events to confirm occurrence of the event, using previously agreed-upon criteria.

The TRFs for the Framingham Risk Score were measured at exam 1 and used to generate the Framingham Risk Score. Standard methods were used to obtain these measures. Smoking (yes/no) was derived from the number of cigarettes smoked per day, with those reporting none coded as zero. Blood pressure measurements were performed 2 times with the participant seated with a random-zero sphygmomanometer and were averaged. Questionnaires were administered to assess current use of antihypertensive medication, though very little medication was used in 1971–1975. Plasma total cholesterol and HDL-C were measured using standard LRC protocols. Diabetes was defined as a fasting glucose level ≥126 mg/dL or diabetic medication use.

Genotyping was conducted for the SNP Health Association Resource (SHARe) project (http://www.ncbi.nlm.nih.gov/projects/gap/cgi-bin/study.cgi?study_id=phs000007.v10.p5) using the Affymetrix 500K mapping array (250K Nsp and 250K Sty arrays), and the Affymetrix 50K supplemental gene focused array on a total of 9274 individuals from all 3 Framingham cohorts.

Genotyping resulted in 503,551 SNPs with successful call rate >95% and HWE $P>1.0\times10^{-6}$ in 8,481 individuals with call rate >97%. Imputation of 2,543,887 autosomal SNPs in HapMap release 22, build 36, CEU sample was conducted using the algorithm implemented in MACH (version 1.0.15). From a total of 534,982 genotyped autosomal SNPs in Framingham, 378,163 SNPs were used in imputation after filtering out 15,586 SNPs (Hardy-Weinberg equilibrium $P<1.0\times10^{-6}$), 64,511 SNPs (missingness >0.03), 45,361 SNPs (mishap $P<1.0\times10^{-9}$), 4857 SNPs (>100 Mendelian errors), 67,269 SNPs (frequency <0.01), 2 SNPs (due to strand issues upon merging data with HapMap), and a further 13,394 SNPs that were not present on HapMap. We used 200 biologically unrelated participants to estimate the parameters of the imputation model and subsequently applied the estimated parameters to obtain imputed SNPs for all 8481 participants. For this study, the best-guess genotype was used if the SNP was imputed. The Framingham Heart Study was approved by the institutional review boards of Boston University and the National Institutes of Health. All participants provided written informed consent.

SUPPLEMENTARY REFERENCES

1 Chambless LE, Folsom AR, Sharrett AR, et al. Coronary heart disease risk prediction in the Atherosclerosis Risk in Communities (ARIC) study. *J Clin Epidemiol* 2003;56:880-890.
2 Siedel J, Hagele EO, Ziegenhorn J, Wahlefeld AW. Reagent for the enzymatic determination of serum total cholesterol with improved lipolytic efficiency. *Clin Chem* 1983;29:1075-1080.
3 Warnick GR, Benderson J, Albers JJ. Dextran sulfate-Mg2+ precipitation procedure for quantitation of high-density-lipoprotein cholesterol. *Clin Chem* 1982;28:1379-1388.
4 Hofman A, Grobbee DE, de Jong PT, van den Ouweland FA. Determinants of disease and disability in the elderly: the Rotterdam Elderly Study. *Eur J Epidemiol* 1991;7:403-422.
5 Hofman A, Breteler MM, van Duijn CM, et al. The Rotterdam Study: objectives and design update. *Eur J Epidemiol* 2007;22:819-829.
6 Hofman A, Breteler MM, van Duijn CM, et al. The Rotterdam Study: 2010 objectives and design update. *Eur J Epidemiol* 2009;24:553-572.
7 van der Meer IM, Bots ML, Hofman A, del Sol AI, van der Kuip DA, Witteman JC. Predictive value of noninvasive measures of atherosclerosis for incident myocardial infarction: the Rotterdam Study. *Circulation* 2004;109:1089-1094.
8 de Bruyne MC, Mosterd A, Hoes AW, et al. Prevalence, determinants, and misclassification of myocardial infarction in the elderly. *Epidemiology* 1997;8:495-500.
9 Kannel WB, Feinleib M, McNamara PM, Garrison RJ, Castelli WP. An investigation of coronary heart disease in families: the Framingham offspring study. *Am J Epidemiol* 1979;110:281-290.
10 Splansky GL, Corey D, Yang Q, et al. The Third Generation Cohort of the National Heart, Lung, and Blood Institute's Framingham Heart Study: design, recruitment, and initial examination. *Am J Epidemiol* 2007;165:1328-1335.
11 Dawber TR, Meadors GF, Moore FE, Jr. Epidemiological approaches to heart disease: the Framingham Study. *Am J Public Health* 1951;41:279-281.
12 Erdmann J, Grosshennig A, Braund PS, et al. New susceptibility locus for coronary artery disease on chromosome 3q22.3. *Nat Genet* 2009;41:280-282.
13 Kathiresan S, Voight BF, Purcell S, et al. Genome-wide association of early-onset myocardial infarction with single nucleotide polymorphisms and copy number variants. *Nat Genet* 2009;41:334-341.
14 Samani NJ, Erdmann J, Hall AS, et al. Genomewide association analysis of coronary artery disease. *N Engl J Med* 2007;357:443-453.
15 Wellcome Trust Case Control Consortium. Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls. *Nature* 2007;447:661-678.
16 Helgadottir A, Thorleifsson G, Manolescu A, et al. A common variant on chromosome 9p21 affects the risk of myocardial infarction. *Science* 2007;316:1491-1493.
17 McPherson R, Pertsemlidis A, Kavaslar N, et al. A common allele on chromosome 9 associated with coronary heart disease. *Science* 2007;316:1488-1491.
18 Tregouet DA, Konig IR, Erdmann J, et al. Genome-wide haplotype association study identifies the SLC22A3-LPAL2-LPA gene cluster as a risk locus for coronary artery disease. *Nat Genet* 2009;41:283-285.
19 Bare LA, Morrison AC, Rowland CM, et al. Five common gene variants identify elevated genetic risk for coronary heart disease. *Genet Med* 2007;9:682-689.

20 Shiffman D, O'Meara ES, Bare LA, et al. Association of gene variants with incident myocardial infarction in the Cardiovascular Health Study. *Arterioscler Thromb Vasc Biol* 2008;28:173-179.
21 Shiffman D, Rowland CM, Louie JZ, et al. Gene variants of VAMP8 and HNRPUL1 are associated with early-onset myocardial infarction. *Arterioscler Thromb Vasc Biol* 2006;26:1613-1618.
22 Iakoubova OA, Tong CH, Rowland CM, et al. Association of the Trp719Arg polymorphism in kinesin-like protein 6 with myocardial infarction and coronary heart disease in 2 prospective trials: the CARE and WOSCOPS trials. *J Am Coll Cardiol* 2008;51:435-443.
23 Morrison AC, Bare LA, Chambless LE, et al. Prediction of coronary heart disease risk using a genetic risk score: the Atherosclerosis Risk in Communities study. *Am J Epidemiol* 2007;166:28-35.

SUPPLEMENTARY TABLE 1: Process for selecting SNPs for study inclusion

Phase 1: GWAS Approach	
	SNPs
28 SNPs initially identified using the National Human Genome Research Institute database (updated as of December 2009)	rs9818870, rs2259816 [12], rs646776, rs9982601, rs12526453, rs1122608, rs1746048, rs11206510, rs6725887, rs4977574 [13], rs599838, rs6922269, rs501120, rs17228212, rs2943634, rs17465637, rs1333049(9p21) [14], rs17672135, rs688034, rs8055236, rs1333049 [15],rs10757278 [16,17], rs2048327, rs3127599, rs7767084, rs10755578 [18], rs2259816, rs9818870 [12]
Reason for exclusion (n=number excluded)	
Association with an intermediate phenotype of CHD (n=12)	
LDL-C	rs17465637, rs1122608, rs646776, rs11206510, rs599838
Non-HDL-C	rs17228212
HDL-C	rs2943634
Lp(a)	rs2048327, rs3127599, rs7767084, rs10755578
CRP	rs2259816
Lack of association with CHD (n=4)	rs17465637, rs17672135, rs688034, rs8055236
Linkage disequilibrium with rs10757274, which resides in the 9p21 region (n=3)	rs10757278, rs1333049, rs4977574
9 SNPs included in genetic risk score	rs9818870, rs2259816, rs9982601, rs12526453, rs1746048, rs6725887, rs6922269, rs501120, rs10757274
Phase 2: Candidate Genes–Literature Review	
Reason for exclusion	**Number excluded**
SNPs were found to be related to an intermediate phenotype associated with CHD	107
The study population was not Caucasian	9
Failure to replicate the findings in other studies	3
Negative study findings or the study outcome focused on phenotypes other than clinical CHD	49
5 SNPs included in genetic risk score	rs3900940, rs20455, rs1010, rs7439293, rs2298566 [19-23]

SUPPLEMENTARY TABLE 2: Published and calculated risk estimates with CHD in ARIC and in published papers for SNPs included in the genetic risk score

SNP	Gene Region Name or Locus	HR in ARIC [95% CI]	Published Risk Estimates	Reference
rs9818870	*MRAS*	1.062 [0.949, 1.190]	OR 1.15 [1.11–1.19]	Erdmann J, et al. *Nat Gene* 2009;41:280–2.
rs2259816	*HNF1A*	1.019 [0.934, 1.113]	OR 1.08 [1.05–1.11]	Erdmann J, et al. *Nat Gene* 2009;41:280–2.
rs9982601	*SLC5A3, MRPS6, KCNE2*	1.171 [1.044, 1.314]	OR 1.2 [1.14–1.27]	Kathiresan S, et al. *Nat Gene* 2009;41:334–41.
rs12526453	*PHACTR1*	1.141 [1.043, 1.250]	OR 1.12 [1.08–1.17]	Kathiresan S, et al. *Nat Gene* 2009;41:334–41.
rs1746048	*CXCL12*	1.229 [1.077, 1.401]	OR 1.17 [1.11–1.24]	Kathiresan S, et al. *Nat Gene* 2009;41:334–41.
rs6725887	*WDR12*	1.073 [0.950, 1.213]	OR 1.17 [1.11–1.23]	Kathiresan S, et al. *Nat Gene* 2009;41:334–41.
rs6922269	*MTHFD1L*	1.010 [0.920, 1.110]	OR 1.23 [1.15–1.33]	Samani NJ, et al. *N Engl J Med* 2007;357:443–53.
rs501120	*CXCL12*	1.214 [1.066, 1.383]	OR 1.33 [1.20–1.48	Samani NJ, et al. *N Engl J Med* 2007;357:443–53.
rs3900940	*MYH15*	1.127 [1.032, 1.230]	HR 1.17 [1.07–1.28]	Bare LA, et al. *Genet Med* 2007;9:682–9.
rs1010	*VAMP8*	1.042 [0.959, 1.133]	HR 1.2 [P < 0.019]	Shiffman D, et al. *Arterioscler Thromb Vasc Biol* 2006;26:1613–8.
rs7439293	*PALLD*	1.104 [1.012, 1.205]	HR 1.11 [1.02–1.22]	Bare LA, et al. *Genet Med* 2007;9:682–9.
rs2298566	*SNX19*	1.131 [1.027, 1.247]	HR 1.12 [1.01–1.24]	Bare LA, et al. *Genet Med* 2007;9:682–9.
rs10757274	*9p21*	1.214 [1.117, 1.319]	OR 1.21 [1.04–1.40]	Newton-Cheh C, et al. *Circulation* 2009;120:2062–8.

HR indicates hazard ratio; OR, odds ratio

SUPPLEMENTARY TABLE 3: Baseline characteristics in the ARIC, Framingham Offspring, and Rotterdam Studies

Characteristic	ARIC (n=8542)	Framingham (n=2339)	Rotterdam (n=2068)
Age, years (mean± SD)	54.1±5.7	40.4±7.2	60.3±2.8
Male, %	45.1%	48.2%	43.6%
Systolic blood pressure, mm Hg (mean± SD)	118.1±16.7	122.3±15.3	132.5±20.7
Total cholesterol, mg/dL (mean± SD)	214.0±40.6	201.4±37.5	259.8±45.5
HDL-C, mg/dL (mean± SD)	51.0±16.7	51.5±15.3	52.8±14.5
LDL-C, mg/dL (mean± SD)	136.8±37.8	128.5±34.1	unavailable
Diabetes, %	7.8%	1.3%	5.3%
Current tobacco use, %	24.5%	39.4%	29.7%

HDL-C indicates high-density lipoprotein cholesterol; LDL-C, low-density lipoprotein cholesterol.

SUPPLEMENTARY TABLE 4: Analysis of maximum likelihood estimates in the Rotterdam Study

Variable	Hazard Ratio	95% Confidence Limits	
Sex	0.38	0.29	0.49
Age	1.10	1.05	1.15
Systolic blood pressure	1.003	0.9974	1.0089
Hypertension medication	1.78	1.37	2.30
Current smoker	1.41	1.10	1.81
Diabetes	1.87	1.26	2.77
HDL-C in 2nd quintile*	1.97	1.27	3.05
HDL-C in 3rd quintile*	1.79	1.19	2.67
HDL-C in 4th quintile*	1.19	0.74	1.91
HDL-C in 5th quintile*	1.34	0.91	1.98
Total cholesterol in 2nd quartile†	1.89	0.97	3.65
Total cholesterol in 3rd quartile†	2.06	1.07	3.98
Total cholesterol in 4th quartile†	2.88	1.49	5.56
GRS‡	1.09	1.03	1.14

*Compared with HDL-C in first (lowest) quintile
†Compared with total cholesterol in first (lowest) quartile
‡Modeled with age and sex only
HDL-C indicates high-density lipoprotein cholesterol; GRS, genetic risk score.

SUPPLEMENTARY TABLE 5: Analysis of maximum likelihood estimates in the Framingham Offspring Study

Variable	Hazard Ratio	95% Confidence Limits	
Sex	0.311	0.215	0.45
Age	1.036	1.013	1.06
Systolic blood pressure	1.008	0.996	1.021
Hypertension medication	1.231	0.803	1.889
Current smoker	1.954	1.507	2.535
Total cholesterol	1.009	1.005	1.013
GRS*	1.069	0.997	1.133

*Modeled with age and sex only

GRS indicates genetic risk score.

SUPPLEMENTARY TABLE 6: Analysis of maximum likelihood estimates in the ARIC Study

Variable	Hazard Ratio	95% Confidence Limits	
Sex	2.25	1.96	2.58
Age	1.4	1.25	1.56
Age square	0.77	0.63	0.95
Systolic blood pressure	1.011	1.008	1.015
Hypertension medication	1.18	1.02	1.36
Current smoker	1.55	1.36	1.77
Diabetes	2.17	1.85	2.55
HDL-C in 2nd quintile*	3.289	2.563	4.22
HDL-C in 3rd quintile*	2.577	2.037	3.261
HDL-C in 4th quintile*	2.388	1.84	3.099
HDL-C in 5th quintile*	1.86	1.45	2.4
Total cholesterol in 2nd quartile†	1.55	1.34	1.79
Total cholesterol in 3rd quartile†	2.01	1.7	2.37
Total cholesterol in 4th quartile†	2.01	1.61	2.6
GRS‡	1.092	1.063	1.12

*Compared with HDL-C in first (lowest) quintile
†Compared with total cholesterol in first (lowest) quartile
‡Modeled with age and sex only
HDL-C indicates high-density lipoprotein cholesterol; GRS, genetic risk score.

SUPPLEMENTARY FIGURE 1: Distribution of the unweighted genetic risk score in the Framingham Offspring Study (top panel), Rotterdam Study (center panel), and ARIC Study (bottom panel)