Meta-analysis of genome-wide association studies from the CHARGE consortium identifies common variants associated with carotid intima media thickness and plaque.

Joshua C. Bis, PhD; Maryam Kavousi, MD, MSc; Nora Franceschini, MD, MPH; Aaron Isaacs, PhD; Gonçalo R. Abecasis, PhD; Ulf Schminke, MD; Wendy Post, MD; Albert V. Smith, PhD; L. Adrienne Cupples, PhD; Hugh S. Markus, MD; Reinhold Schmidt, MD; Jennifer E. Huffman, MSc; Terho Lehtimäki, MD, PhD; Jens Baumert, PhD; Thomas Münzel, MD; Susan R. Heckbert, MD, PhD; Abbas Dehghan, MD, PhD; Kari North, PhD; Ben Oostra, PhD; Steve Bevan, PhD; Eva-Maria Stoegerer, MD; Caroline Hayward, PhD; Olli Raitakari, MD, PhD; Christa Meisinger, MD, MPH; Arne Schillert, PhD; Serena Sanna, PhD; Henry Völzke, MD; Yu-Ching Cheng, PhD; Bolli Thorsson, MD; Caroline S. Fox, MD, MS; Kenneth Rice, PhD; Fernando Rivadeneira, MD, PhD; Vijay Nambi, MD; Eran Halperin, PhD; Katja E. Petrovic, MSc; Leena Peltonen, MD, PhD; H. Erich Wichmann, MD, PhD; Renate B. Schnabel, MD, MSc; Marcus Dörr, MD; Afshin Parsa, MD, MPH; Thor Aspelund, PhD; Serkalem Demissie, PhD; Sekar Kathiresan, MD; Muredach P. Reilly, MBBCH, MSCE; the CARDIoGRAM Consortium; Kent Taylor, PhD; Andre Uitterlinden, PhD; David J. Couper, PhD; Matthias Sitzer, MD; Mika Kähönen, MD, PhD; Thomas Illig, PhD; Philipp S. Wild, MD; Marco Orru, MD; Jan Lüdemann, PhD; Alan R. Shuldiner, MD; Gudny Eiriksdottir, MSc; Charles C. White, MPH; Jerome I. Rotter, MD; Albert Hofman, MD, PhD; Jochen Seissler, MD; Tanja Zeller, PhD; Gianluca Usala, PhD; Florian Ernst, PhD; Lenore J. Launer, PhD; Ralph B. D'Agostino, Sr., PhD; Daniel H. O'Leary, MD; Christie Ballantyne, MD; Joachim Thiery, MD, MBA; Andreas Ziegler, Dr. rer. nat. habil.; Edward G Lakatta, MD; Ravi Kumar Chilukoti, MSc; Tamara B. Harris, MD, PhD; Philip A. Wolf, MD; Bruce M. Psaty, MD, PhD; Joseph F Polak, MD, MPH; Xia Li, MD, MPH; Wolfgang Rathmann, MD, MSPH; Manuela Uda, PhD; Eric Boerwinkle, PhD; Norman Klopp, PhD; Helena Schmidt, MD PhD; James F Wilson, DPhil; Jorma Viikari, MD, PhD; Wolfgang Koenig, MD; Stefan Blankenberg, Prof Dr med; Anne B. Newman, MD, MPH; Jacqueline Witteman, PhD; Gerardo Heiss, MD, PhD; Cornelia van Duijn, PhD; Angelo Scuteri, MD, PhD; Georg Homuth, PhD; Braxton D. Mitchell, PhD; Vilmundur Gudnason , MD, PhD; Christopher J. O'Donnell, MD, MPH

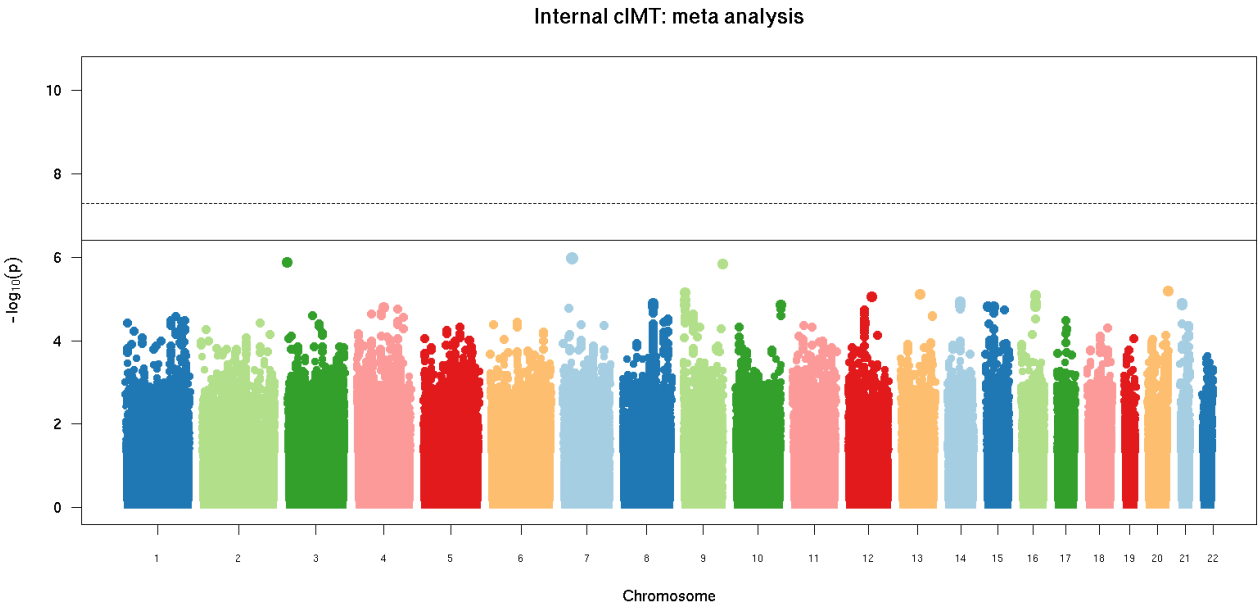
Correspondence should be addressed to J.C.B. (joshbis@uw.edu) or C.J.O. (odonnellc@nhlbi.nih.gov)

Contents

Supplementary Figures

Supplementary Figure 1: Genome-wide Manhattan plot for internal cIMT.

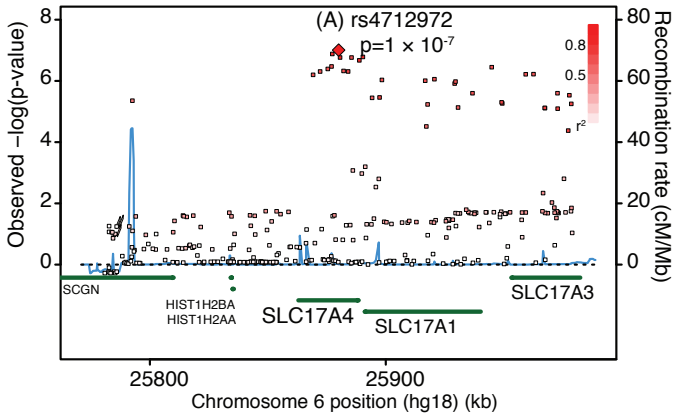
Genome-wide Manhattan plot showing the individual p-values (based on discovery meta-analysis) against their genomic position for internal carotid IMT. Within each chromosome, shown on the *x*-axis, the results are plotted left to right from the p-terminal end. The solid line indicates the threshold for follow-up, $p<4 \times 10^{-7}$, the dashed line indicates the threshold for genome-wide significance, $p<5 \times 10^{-8}$.



Supplementary Figure 2: Regional plot for suggestive common carotid IMT locus.

Plots are centered on the most significant SNP at locus along with the meta-analysis results for SNPs in the 100kb region surrounding it. All SNPs are plotted with their discovery meta-analysis p-values against their genomic position, with the most significant SNP in the region indicated as a diamond and other SNPs shaded according to their pairwise correlation (r^2) with the signal SNP. The light blue line represents the estimated recombination rates. Gene annotations are shown as dark green lines.

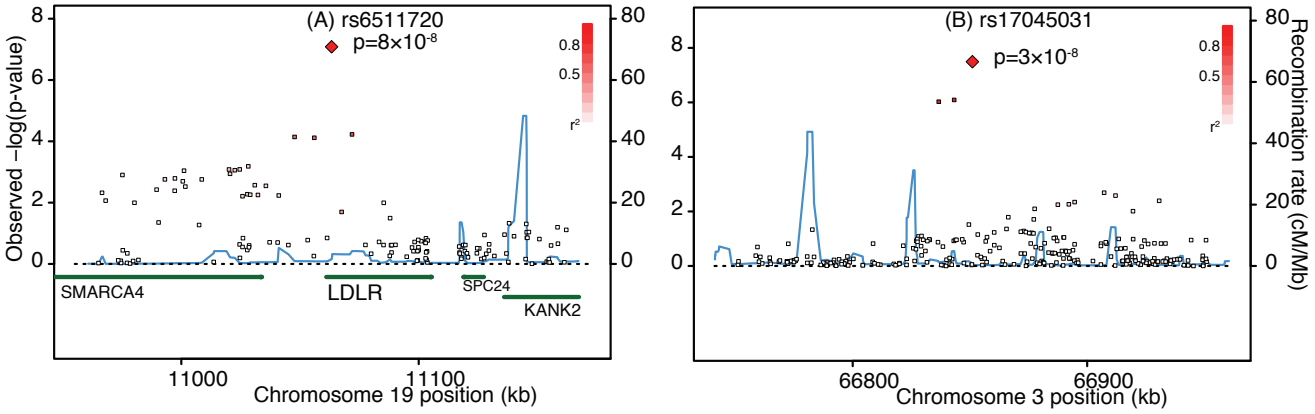
Supplementary Figure 2: Regional plot for suggestive common carotid IMT locus.



Supplementary Figure 3: Regional plots for suggestive plaque loci.

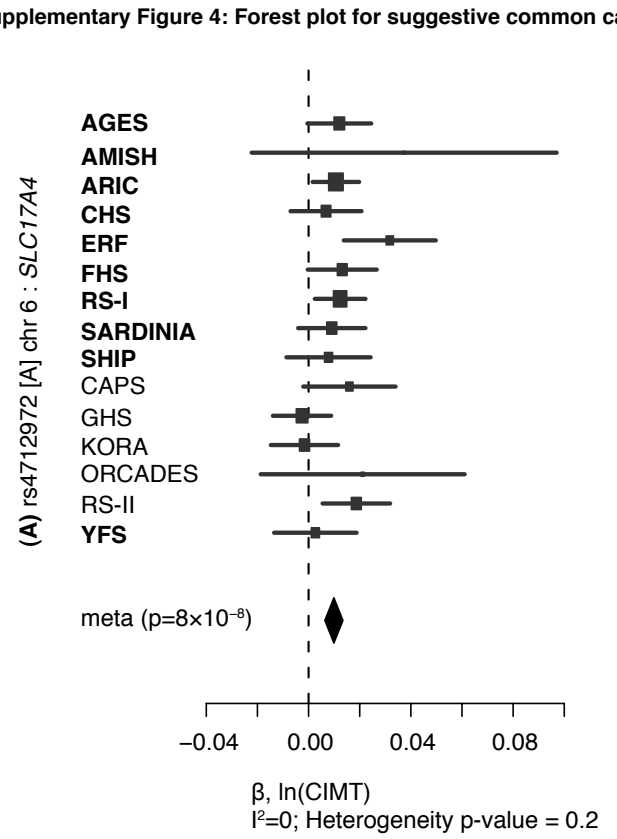
Plots are centered on the most significant SNP at each locus along with the meta-analysis results for SNPs in the 100kb region surrounding it. All SNPs are plotted with their discovery meta-analysis p-values against their genomic position, with the most significant SNP in the region indicated as a diamond and other SNPs shaded according to their pairwise correlation (r^2) with the signal SNP. The light blue line represents the estimated recombination rates. Gene annotations are shown as dark green lines.

Supplementary Figure 3: Regional plots for suggestive plaque loci.



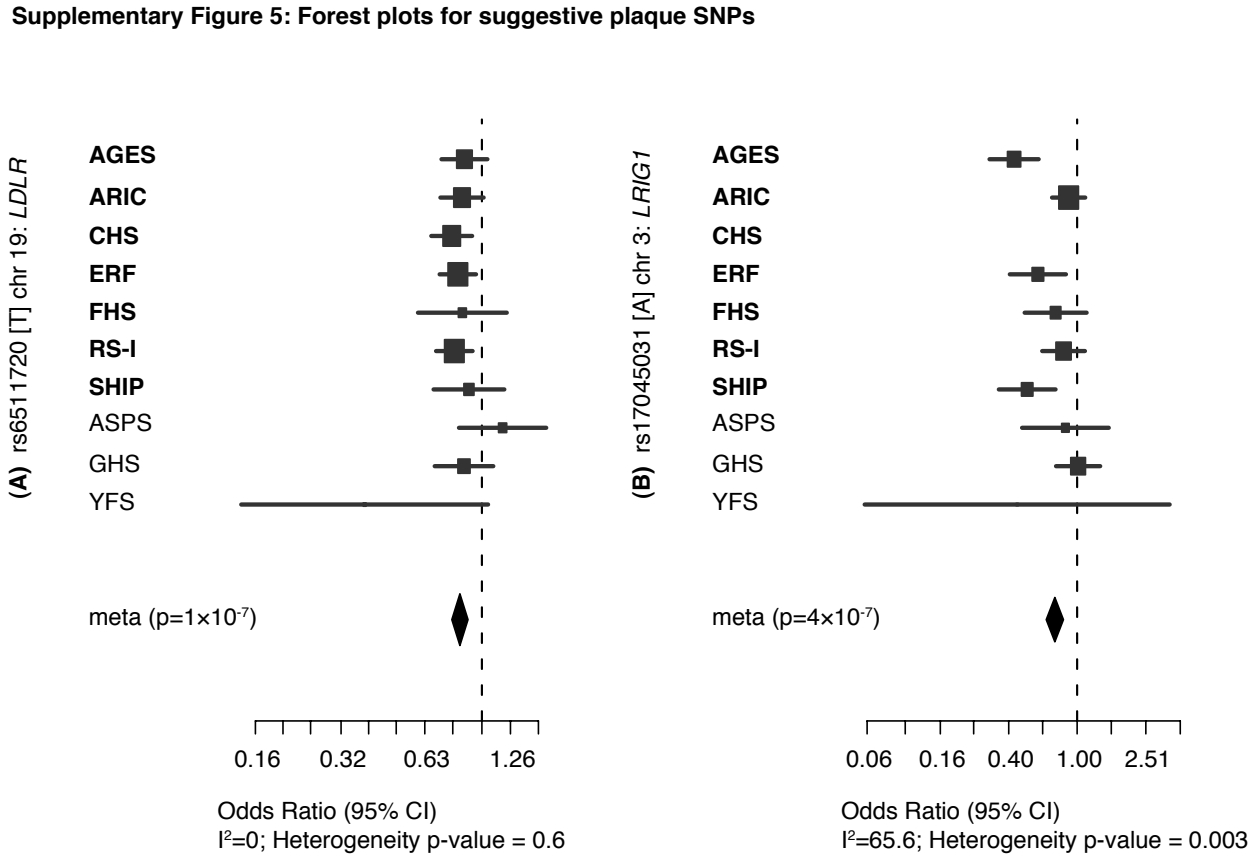
Supplementary Figure 4: Forest plot for suggestive common carotid IMT SNP association.

Plots show the study-specific association estimates (β) and 95% confidence intervals for the nine discovery and second stage studies, presented as bars. The scale is ln(cIMT). The association estimate and confidence interval for the meta-analysis combining discovery and second stage results is presented as a diamond. Blank spaces indicate occasions in which a particular study was not able to provide results for a given SNP.



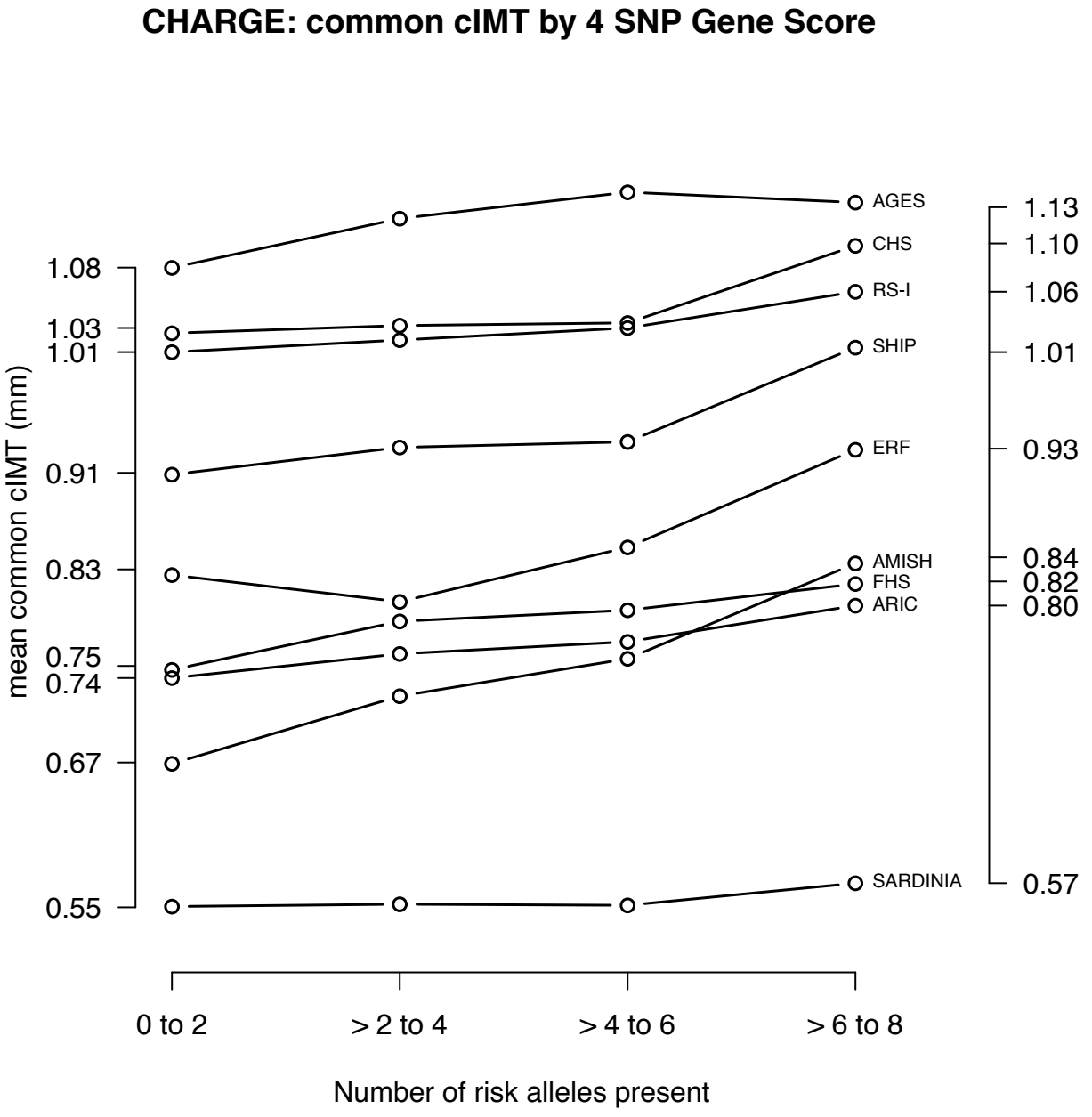
Supplementary Figure 5: Forest plots for suggestive plaque SNP associations.

Plots show the study-specific association estimates (OR) and 95% confidence intervals for the nine discovery and second stage studies, presented as bars. The association estimate and confidence interval for the meta-analysis combining discovery and second stage results is presented as a diamond. Blank spaces indicate occasions in which a particular study was not able to provide results for a given SNP.



Supplementary Figure 6: common cIMT by 4 SNP gene score.

Each of the nine discovery studies created a gene score, a sum of the number of alleles for the genomewide significant and suggestive SNPs from our analysis of common cIMT, coding the SNPs in terms of the allele associated with higher common cIMT in the meta-analysis. We present the mean common cIMT from each study according to categories of the gene score.



Supplementary Tables

Supplementary Table 1: Published evidence for heritability of carotid artery phenotypes

Study	Common cIMT		Internal cIMT		Carotid Plaque		Sample Size	Ethnicity
	Minimally adjusted (%)	Risk factor adjusted (%)	Minimally adjusted (%)	Risk factor adjusted (%)	Minimally adjusted (%)	Risk factor adjusted (%)	N persons/ N families	
European Ancestry								
Zannad 1998[1]		33					315/76	French
Jartti 2002 [2]	36						-/74 twin pairs	Finnish
Fox 2003 [3]	67	38	43	35			906/586	US
Swan 2003 [4]	31						-/264 twin pairs	British
Moskau 2005 [5]		61				47*, 26**	556/154	German
Mayosi 2005 [6].	24						854/224	British
Sayed-Tabatabaei 2005 [7]	41	34					930/1	Dutch
Ryabikov 2007 [8]		54					286/81	Russian
Rampersaud 2008 [9]	29	25					478/279	US Amish
European Ancestry & African Americans								
Lange 2002 [10]	32	41					252/122	US (T2DM)
Zhao 2008 [11].	69	59					-/98 twin pairs	US
Hispanic Ancestry								
Duggirala 1996 [12]	86	92	87	86			88/46	Mexican
Hunt 2002 [13]					28	23	750/29	Mexican Amer.
Juo 2004[14]	39		12				440/77	Carribean
Wang 2005 [15]	45	40					413/91	Mexican Amer.
Kao 2005 [16]	16						620/24	Mexican Amer.
Xiang 2002 [17]	72	64					286/54	Multiple H (HTN)
Sacco 2009 [18], Dong 2010 [19]		65				50	<=1390/100	Dominican H
Native Americans								
North 2002 [20]		21					950/32	Native American

Table modified from Manolio et al, 2004. [21]

*Heritability of plaque=carotid stenosis; ** Heritability of plaque=carotid stenosis, adjusted for common cIMT.

Abbreviations: cIMT=carotid intimal medial thickness, US=United States, T2DM=type 2 diabetes mellitus, HTN=hypertensio

Supplementary Table 2: Correlation between carotid phenotypes in ARIC, CHS, and FHS

Correlation:	Common cIMT & Internal cIMT	Common cIMT & Plaque	Internal cIMT & Plaque
ARIC	0.36	0.39	0.60
CHS	0.67	0.27	0.43
FHS	0.49	0.36	0.52

Note: Study-specific Pearson correlation coefficients, shown above, were calculated using natural-log transformed values for the continuous phenotypes of common and internal cIMT.

Supplementary Table 3: Discovery, Second stage, and Combined meta-analysis for common cIMT and plaque, suggestive loci

				Discovery GWAS (cIMT)					Second Stage Meta-analysis (cIMT)					Combined Meta-analysis (cIMT)		
SNP	Chr	Nearest gene	Alleles	AF	β	SE	N	p-value	AF	β	SE	N	p-value	β	SE	p-value
rs4712972	6	SLC17A4	A/G	0.12	0.0121	0.0023	30,565	9.8×10^{-8}	0.12	0.0056	0.0031	10,394	0.1	0.0099	0.0018	7.8×10^{-8}

				Discovery GWAS (plaque)				Second Stage Meta-analysis (plaque)				Combined Meta-analysis (plaque)	
SNP	Chr	Nearest gene	Alleles	AF	OR (95% CI)	N	p-value	AF	OR (95% CI)	N	p-value	OR (95% CI)	p-value
rs6511720	19	LDLR	T/G	0.13	0.83 (0.77 - 0.89)	18,671	8.2×10^{-8}	0.05	0.92 (0.76 - 1.12)	4,281	0.4	0.84 (0.78 - 0.89)	1.0×10^{-7}
rs17045031	3	LRIG1	A/G	0.03	0.70 (0.61 - 0.79)	22,076	3.2×10^{-8}	0.04	0.97 (0.74 - 1.25)	5,820	0.8	0.74 (0.66 - 0.83)	3.8×10^{-7}

Alleles indicates the coded (named first) & non-coded allele; AF indicates allele frequency for the coded allele, an average weighted by study size; OR indicates odds ratio, CI, confidence interval; N indicates effective sample size, calculated by taking the sum of each study's sample size multiplied by the SNP's imputation quality. When more than one SNP at a locus surpassed our p-value threshold, we presented the SNP with the lowest p-value.

Supplementary Table 4: Discovery meta-analysis for internal cIMT

SNP	Chr	Position	Closest Gene	Alleles	AF	β	SE	p-value	N
rs10276782	7	24771765	DFNA5	T/C	0.11	-0.0550	0.0113	1.0×10^{-6}	7392
rs9823028	3	1165111	CNTN6	A/G	0.01	0.2326	0.0481	1.3×10^{-6}	3602
rs2773822	9	134850854	GFI1B	T/C	0.12	-0.0507	0.0105	1.4×10^{-6}	5912
rs2275292	20	62131496	PRPF6	T/C	0.87	0.0731	0.0162	6.5×10^{-6}	2709
rs10815111	9	4901930	RCL1	T/C	0.6	-0.0245	0.0054	7.1×10^{-6}	10844
rs17090381	13	73197831	KLF12	T/C	0.98	-0.0922	0.0206	7.9×10^{-6}	9516
rs11864736	16	63818912	LOC283867	A/C	0.12	0.0365	0.0082	8.0×10^{-6}	10641
rs7195314	16	63816846	LOC283867	T/C	0.12	0.0363	0.0082	8.3×10^{-6}	10685
rs1037232	12	75866037	CSRP2	A/G	0.01	0.1657	0.0373	8.9×10^{-6}	5288

Alleles indicate coded (named first) & non-coded alleles; AF indicates allele frequency for the coded allele;
N indicates effective sample size, calculated by taking the sum of each study's sample size multiplied by the SNP's imputation quality

Supplementary Table 5: Comparison of meta-analysis results for common cIMT, plaque, and internal cIMT

(A) Top SNPs from the combined discovery and second stage meta-analysis of common cIMT

				Combined common cIMT analysis				Combined plaque analysis			Discovery internal cIMT analysis			
SNP	Chr	Alleles	Closest gene	β	SE	N	p-value	OR (95% CI)	N	p-value	β	SE	N	p-value
rs11781551	8	a/g	ZHX2	-0.0078	0.0012	41,295	2×10^{-11}	0.96 (0.93 – 1.00)	30,968	0.05	-0.0132	0.0054	10,769	0.02
rs445925	19	a/g	APOC1	-0.0156	0.0028	17,185	2×10^{-8}	0.82 (0.76 - 0.89)	14,097	4×10^{-6}	-0.0466	0.0151	2,939	0.002
rs6601530	8	g/a	PINX1	0.0078	0.0014	32,631	2×10^{-8}	1.09 (1.04 - 1.14)	22,629	0.0001	0.0130	0.0062	9,039	0.03
rs4712972	6	a/g	SLC17A4	0.0099	0.0018	40,959	8×10^{-8}	1.09 (1.02 - 1.15)	30,690	0.006	-0.0012	0.0082	10,486	0.9

(B) Top SNPs from the combined discovery and second stage meta-analysis of plaque

				Combined common cIMT analysis				Combined plaque analysis				Discovery internal cIMT analysis			
SNP	Chr	Alleles	Closest gene	β	SE	N	p-value	β	SE	N	p-value	β	SE	N	p-value
rs17398575	7	a/g	PIK3CG	0.0067	0.0014	37,533	3×10^{-6}	0.1619	0.0231	29,255	2×10^{-12}	0.0199	0.0066	10,356	0.002
rs1878406	4	t/c	EDNRA	0.0087	0.0018	38,071	7×10^{-7}	0.1993	0.0291	29,827	7×10^{-12}	0.0260	0.0083	10,201	0.002
rs6511720	19	t/g	LDLR	-0.0065	0.0022	26,649	0.003	-0.1778	0.0334	22,952	1×10^{-7}	-0.0236	0.0010	6,875	0.02
rs17045031	3	a/g	LRIG1	-0.0033	0.0036	34,090	0.4	-0.2986	0.0588	27,896	4×10^{-7}	0.0136	0.0184	8,000	0.5

Table shows results from the meta-analysis of the combined discovery and second stage for common cIMT and plaque; results for internal cIMT represent discovery findings since no internal cIMT SNPs were carried on to the second stage. Alleles indicates the coded & non-coded allele; AF indicates allele frequency for the coded allele, an average weighted by study size; N (eff) indicates effective sample size, calculated by taking the sum of each study's sample size multiplied by the SNP's imputation quality.

Supplementary Table 6: Meta-analysis results for common cIMT, plaque, and internal cIMT for 23 SNPs replicated or identified by the CARDIoGRAM Consortium

				Common cIMT		Plaque		Internal cIMT	
SNP	Band	Genes	Risk Allele	β	p-value	OR	p-value	β	p-value
rs599839	1p13.3	SORT1	A	0.0015	0.4	1.07	0.005	0.0137	0.04
rs17114036	1p32.2	PPAP2B	T	0.0061	0.02	1.05	0.2	-0.0025	0.8
rs11206510	1p32.3	PCSK9	A	0.0001	1.0	1.04	0.2	0.0100	0.2
rs17465637	1q41	MIA3	C	n/a	n/a	n/a	n/a	n/a	n/a
rs6725887	2q33.1	WDR12	C	0.0015	0.5	1	1.0	0.0001	1.0
rs2306374	3q22.3	MRAS	C	0.0012	0.5	1.04	0.2	-0.0005	0.9
rs17609940	6p21.31	ANKS1A	G	-0.0004	0.8	1	0.9	-0.0135	0.1
rs12190287	6q23.2	TCF21	C	0.0006	0.7	1.04	0.1	0.0072	0.3
rs12526453	6p24.1	PHACTR1	C	-0.0017	0.2	1.02	0.4	-0.0010	0.9
rs3798220	6q25.3	LPA	C	n/a	n/a	n/a	n/a	n/a	n/a
rs11556924	7q32.2	ZC3HC1	C	0.0003	0.8	1.06	0.01	0.0114	0.07
rs4977574	9p21.3	CDKN2A/B, ANRIL	G	-0.0006	0.7	1.06	0.003	0.0025	0.6
rs579459	9q34.2	ABO	C	-0.0006	0.8	1.01	0.7	0.0089	0.2
rs1746048	10q11.21	CXCL12	C	-0.0013	0.5	0.98	0.5	0.0117	0.2
rs12413409	10q24.32	CYP17A1, CNNM2, NT5C2	G	0.0029	0.2	1.03	0.4	0.0119	0.2
rs964184	11q23.3	ZNF259, APOA5-A4-C3-A1	G	0.0032	0.1	1.05	0.1	0.0099	0.2
rs3184504	12q24.12	SH2B3	T	-0.0008	0.6	1	0.9	0.0056	0.3
rs4773144	13q34	COL4A1, COL4A2	G	-0.0036	0.02	1.01	0.8	-0.0046	0.4
rs2895811	14q32.2	HHIPL1	C	0.0021	0.2	1.03	0.2	0.0073	0.2
rs3825807	15q25.1	ADAMTS7	A	0.0036	0.01	1.04	0.1	0.0065	0.2
rs12936587	17p11.2	RASD1, SMCR3, PEMT	G	0.0001	1.0	0.98	0.3	-0.0029	0.6
rs216172	17p13.3	SMG6, SRR	C	-0.0004	0.8	0.97	0.3	-0.0090	0.3
rs46522	17q21.32	UBE2Z, GIP, ATP5G1, SNF8	T	0.0003	0.8	0.98	0.3	-0.0026	0.6
rs1122608	19p13.2	LDLR	G	0.0049	0.003	1.09	0.0008	0.0084	0.2
rs9982601	21q22.11	MRPS6	T	-0.0038	0.07	0.96	0.2	-0.0119	0.1

Table shows results from the discovery meta-analysis for SNPs reported to be associated to coronary artery disease; effect estimates (β, or OR) from the meta-analysis are presented in terms of the literature risk allele; "n/a" indicates that rs17465637 (MIA3) and rs3798220 (LPA) were not available in our discovery meta-analysis.

Supplementary Table 7: Association of APOE e4 allele and rs445925 with common cIMT

Study	Model	SNP	β	SE	p-value
AGES	1	APOE4	−0.0018	0.0047	0.7
	2	APOE4, adjusted for rs445925	−0.0004	0.0048	0.9
		rs445925 (A), adjusted for APOE4	−0.0128	0.0060	0.03
CHS	1	APOE4	0.0015	0.0066	0.8
	2	APOE4, adjusted for rs445925	0.0032	0.0066	0.6
		rs445925 (A), adjusted for APOE4	−0.0211	0.0082	0.01
FHS	1	APOE4	−0.0002	0.0071	0.98
	2	APOE4, adjusted for rs445925	−0.0001	0.0071	0.99
		rs445925 (A), adjusted for APOE4	−0.0103	0.0135	0.4433
RS-I	1	APOE4	0.0050	0.0050	0.3
	2	APOE4, adjusted for rs445925	0.0060	0.0050	0.3
		rs445925 (A), adjusted for APOE4	−0.0230	0.0060	0.0001

Model #1 included the number of copies of the APOE4 allele as a predictor of ln(common cIMT); Model #2 included both the APOE4 allele as well as a term for the rs445925 SNP. Models were adjusted for the same parameters that were included in the common cIMT GWAS.

Supplementary Note

Section 1A: Study descriptions, discovery phase

Our analyses were performed within the Cohorts for Heart and Aging Research in Genomic Epidemiology (CHARGE) consortium,[22] which includes five large population-based prospective cohort studies: the Aging Gene-Environment Susceptibility-Reykjavik Study (AGES), the Atherosclerosis Risk in Communities (ARIC) Study, the Cardiovascular Health Study (CHS), the Framingham Heart Study (FHS), and the Rotterdam Study I (RS-I).

Four additional community-based studies – the Old Order Amish (Amish) Study, the Erasmus Ruchpen Family (ERF) Study, the SardiNIA Study, and the Study of Health in Pomerania (SHIP) – collaborated with CHARGE for these analyses. For all studies participating in the meta-analyses, each participant provided written informed consent and the Institutional Review Board at the parent institution for each respective study approved the study protocols.

The studies contributing to this meta-analysis have been described in detail elsewhere. A brief overview follows.

The Aging Gene-Environment Susceptibility-Reykjavik Study (AGES):

The Reykjavik Study cohort originally comprised a random sample of 30,795 men and women born in 1907–1935 and living in Reykjavik in 1967. [23] A total of 19381 attended, resulting in 71% recruitment rate. The study sample was divided into six groups by birth year and birth date within month. One group was designated for longitudinal follow-up and was examined in all stages. One group was designated a control group and was not included in examinations until 1991. Other groups were invited to participate in specific stages of the study. Between 2002 and 2006, the AGES-Reykjavik study re-examined 5764 survivors of the original cohort who had participated before in the Reykjavik Study.

The AGES Reykjavik Study GWAS was approved by the National Bioethics Committee (00-063-V8+1) and the Data Protection Authority. DNA was genotyped using the Illumina 370CNV BeadChip array on 3,664 participants. Samples were excluded from the dataset based on sample failure, genotype mismatch with reference panel, and sex mismatch, resulting in clean genotype data on 3,219 individuals. Standard protocols for working with Illumina data were followed, with clustering score greater than 0.4. Prior to genotype imputation, SNPs were excluded using filters based on call rate (<97%), Hardy-Weinberg Equilibrium ($< 1\times10^{-6}$), mishap ($<1\times10^{-9}$), and mismatched positions between Illumina, dbSNP and/or HapMap resulting in 325,094 SNPs passing all QC (of 353,202 prior to cleaning steps). Imputation was done using MaCH against all the HapMap CEPH haplotypes (release 22/NCBI build 36) resulting in 2,533,153 total SNPs for analysis. Association analysis was conducted against all genotypes, and includes the most likely imputed gentoypes.

The Old Order Amish Study (Amish):

The Old Order Amish individuals included in this study were participants of several ongoing studies of cardiovascular health carried out at the University of Maryland. [24] Participants were relatively healthy volunteers from the Old Order Amish community of Lancaster County, PA. and their family members. Examinations were conducted at the Amish Research Clinic in Strasburg, PA. All protocols were approved by the Institutional Review Board at the University of Maryland and informed consent was obtained, including permission to use their DNA for genetic studies. Genotyping was performed using the Affymetrix GeneChip® Human Mapping 500K Array set. Genotype calls were made using the BRLMM genotype calling algorithm. A total of 364,336

informative autosomal SNPs that passed our quality-control were included in the analysis. MaCH was used for imputation after applying the following filters: 1) not in HapMap; 2) frequency < 0.01; 3) Hardy-Weinberg p-value < 1×10^{-6}; 4) missingness > 0.05.

The Atherosclerosis Risk in Communities Study (ARIC):

The ARIC study is a multi-center prospective investigation of atherosclerotic disease in a predominantly bi-racial population[25]. Men and women aged 45-64 years at baseline were recruited from 4 communities: Forsyth County, North Carolina; Jackson, Mississippi; suburban areas of Minneapolis, Minnesota; and Washington County, Maryland. A total of 15,792 individuals participated in the baseline examination in 1987-1989, with follow-up examinations in approximate 3-year intervals, during 1990-1992, 1993-1995, and 1996-1998.

ARIC Study samples were genotyped using the Affymetrix Genome-Wide Human SNP Array 6.0 (Santa Clara, California); for the current analysis only white participants were analyzed. Sample exclusion criteria included discordant with previous genotype data (n=83), genotypic and phenotypic sex mismatch (n=32), suspected first-degree relative of an included individual based on genotype data (n=297), genetic outlier as assessed by Identity by State (IBS) using PLINK [26] and >8 SD along any of the first 10 principal components in EIGENSTRAT [27] with 5 iterations (n=322). Autosomal SNPs were used for imputation after exclusion of SNPs with HWE deviation < 5×10^{-5}, call rate < 95%, or MAF< 1%.

The Cardiovascular Health Study (CHS):

The CHS is a population-based cohort study of risk factors for CHD and stroke in adults ≥65 years conducted across four field centers in the United States. [28] The original predominantly Caucasian cohort of 5201 persons was recruited in 1989-1990 from a random sample of people on Medicare eligibility lists and an additional 687 African-Americans were enrolled subsequently for a total sample of 5888.

DNA was extracted from blood samples drawn on all participants who consented to genetic testing at their baseline examination in 1989-90. In 2007-2008, genotyping was performed at the General Clinical Research Center's Phenotyping/Genotyping Laboratory at Cedars-Sinai using the Illumina 370CNV Duo® BeadChip system on the 3980 CHS participants who were free of CVD at baseline. Genotyping was successful in 3869 of 3980 persons. QC criteria used to define unsuccessful genotyping were a call rate < 95%, sex mismatch, or other sample failures. From a total of 335,887 genotyped autosomal SNPs, 306,655 SNPs that were present on HapMap and that passed quality control measures (HWE p>10-5, callrate > 0.97, < 3 replicate errors or Mendelian inconsistencies among reference CEPH trios) were used for imputation with BIMBAM version 0.99. [29] Because the other cohorts were predominantly of European ancestry, the African-American participants were excluded from this analysis to limit the potential for false positive associations due to population stratification, leaving 3,261 participants with an IMT measurement for analysis.

The Erasmus Rucphen Family Study (ERF):

The Erasmus Rucphen Family (ERF) Study is comprised of a family-based cohort embedded in the Genetic Research in Isolated Populations (GRIP) program in the southwest of the Netherlands. [30] The aim of this program is to identify genetic risk factors for the development of complex disorders.

In ERF, twenty-two families that had a minimum of five children baptized in the community church between 1850 and 1900 were identified with the help of detailed genealogical records. All living descendants of these couples, and their spouses, were invited to take part in the study. Comprehensive interviews, questionnaires, and examinations were completed at a research center in the area; approximately 3,200 individuals participated. The examination included the

determination of carotid intima media thickness and plaque scores via ultrasonography. Data collection started in June 2002 and was completed in February 2005.

In the current analyses, 2155 participants for whom phenotypic, genotypic and genealogical information was available were studied.

The Framingham Heart Study (FHS):

The methods of recruitment and data collection have been described previously for the original Framingham Heart Study cohort (5,209 participants ascertained systematically from two-thirds of the households in the town of Framingham, MA, beginning in 1948),[31] the Framingham Heart Study Offspring cohort (5,124 children of the original cohort, and spouses of those children, beginning in 1972[32] and the Third Generation cohort (4,095 children of the Offspring cohort, beginning in 2002).[33] The current study was conducted in 3,022 participants of the Offspring cohort participating in examination 6 from 1995 to 1998, who underwent contemporaneous carotid ultrasonography examination.

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 9,274 individuals from all three cohorts. To evaluate population substratification, we conducted principal component analyses using EIGENSTRAT [27]on the genotypes from 882 unrelated participants. We estimated the first 10 principal components and applied the loadings of these components to all genotyped participants. Finally, we evaluated whether any of these principal components were associated with carotid artery phenotypes. When one of the ten principal components was association with a trait, we adjust for that principal component in linear regression models. Genotyping resulted in 503,551 SNPs with successful call rate >95% and HWE $P>1\times10^{-6}$ 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 (HWE $P<1\times10^{-6}$), 64,511 SNPs (missingness >0.03), 45,361 SNPs (mishap $P<1\times10^{-9}$), 4,857 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 8,481 participants. 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.

Rotterdam Study I and Rotterdam Study II (RS I and RS II):

The Rotterdam Study is a prospective population-based cohort study to investigate the determinants of chronic diseases among participants aged 55 years and older.[34] Briefly, residents of Ommoord, a district of Rotterdam, in the Netherlands, 55 years of age or older, were asked to participate, of whom 7,983 participated (RS I). The baseline examination was conducted in 1990 - 1993 and consisted of a home interview and research center visit for blood samples. In 1999, inhabitants who turned 55 years of age or moved into the study district since the start of the study were invited of whom 3011 participated (RS II). The Medical Ethics Committee of Erasmus MC approved the study, and all participants gave informed consent. .

We used the Illumina 550K array to conduct the genotyping. Individuals were excluded for the following reasons: (1) Excess autosomal heterozygosity > 0.336 (~FDR<0.1%), (2) Mismatch between called and phenotypic gender, (3) If there were outliers identified by the IBS clustering analysis with >3 standard deviations from population mean, or IBS probabilities>97%.The exclusion criteria for SNPs were minor

allele frequency ≤1%, Hardy-Weinberg equilibrium (HWE) $p<10^{-5}$, or SNP call rate ≤90% and resulted in data on 530,683 SNPs in RS I and 495,478 SNPs in RS II. To obtain imputed data more restrictive SNP filters including a minor allele frequency >0.01, SNP Call Rate >0.98, and HWE P>1.0x10-6 were applied. Imputation was done with reference to HapMap release 22 CEU using the maximum likelihood method implemented in MaCH(version 1.0.15). The final population for the analysis comprised 5,974 individuals in RS I and 2157 individuals in RS II. Of these, 4699 subjects from RS I and 1980 subjects from RS II had IMT measurements available. The analysis was preformed using ProbABEL (version 1.1) (http://mga.bionet.nsc.ru/~yurii/ABEL/).

The SardiNIA Study (SardiNIA):

The SardiNIA Study was conceived as a study of a Sardinian founder population investigating the genetics of complex traits/phenotypes, including CV risk factors and arterial properties.[35] Over a 3 year period, 6,148 subjects were enrolled, comprising over 60% of those aged 14-102 in a cluster of 4 towns. Each subject came to the clinic before breakfast, signed consent forms, and gave a sample of fasting blood so that all tests would be uninfluenced by meals at different times. Each subject underwent a detailed medical history and full medical examination, blood pressure and anthropometric measurements, a 12-lead resting EKG, measurements of arterial structure and function, and personality testing.

A total of 4,305 related individuals participating the Study were genotyped. Genotyped individuals had four Sardinian grandparents and were selected for genotyping without regard to their phenotypic values. Relationships between genotyped individuals were verified using RELPAIR. [36] Among the individuals examined, 1,412 were genotyped with the Affymetrix Mapping 500K Array Set allowing measurement and not imputation of alleles at the level of each examined SNP. We used the SNPs that passed quality control filters evaluating data completeness (>90%), Mendelian transmission (<3 inconsistencies), and Hardy-Weinberg equilibrium ($p > 1\times10^{-6}$) to estimate genotypes for all the polymorphic SNPs genotyped by the HapMap consortium. The remaining individuals (among the set of 4,305 analyzed individuals) were related to this core set of 1,412 individuals and were genotyped with the Affymetrix 10K genotyping arrays. Genotypes for these individuals were imputed using knowledge of IBD sharing with close relatives who were more densely genotyped. [37]

The Study of Health in Pomerania (SHIP):

The Study of Health in Pomerania (SHIP) is a cross-sectional survey in West Pomerania, the north-east area of Germany. [38] A sample from the population aged 20 to 79 years was drawn from population registries. First, the three cities of the region (with 17,076 to 65,977 inhabitants) and the 12 towns (with 1,516 to 3,044 inhabitants) were selected, and then 17 out of 97 smaller towns (with less than 1,500 inhabitants), were drawn at random. Second, from each of the selected communities, subjects were drawn at random, proportional to the population size of each community and stratified by age and gender. Only individuals with German citizenship and main residency in the study area were included. Finally, 7,008 subjects were sampled, with 292 persons of each gender in each of the twelve five-year age strata. In order to minimize drop-outs by migration or death, subjects were selected in two waves. The net sample (without migrated or deceased persons) comprised 6,267 eligible subjects. Selected persons received a maximum of three written invitations. In case of non-response, letters were followed by a phone call or by home visits if contact by phone was not possible. The SHIP population finally comprised 4,310 participants (corresponding to a final response of 68.8%). The carotid arteries were assessed with ultrasonography in participants at age 45 or older. Data on IMT and carotid plaques are available in 2,438 participants, of which 2,321 consented to take part in genome-wide association studies.

The SHIP samples were genotyped using the Affymetrix Human SNP Array 6.0. Hybridisation of genomic DNA was done in accordance with the manufacturer's standard recommendations. The genetic data analysis workflow was created using the Software InforSense. Genetic data were stored using the database Caché (InterSystems). Genotypes were determined using the Birdseed2

clustering algorithm. For quality control purposes, several control samples where added. On the chip level, only subjects with a genotyping rate on QC probesets (QC callrate) of at least 86% were included. The overall genotyping efficiency of the GWA was 98.55 % with a minimum sample callrate of 92%. Imputation of genotypes in SHIP was performed with the software IMPUTE v0.5.0 based on HapMap II CEU dataset.

Section 1B: Study descriptions, second stage

Austrian Stroke Prevention Study (ASPS):

The ASPS study is a single center prospective follow-up study on the effects of vascular risk factors on brain structure and function in the normal elderly population of the city of Graz, Austria. The procedure of recruitment and diagnostic work-up of study participants has been described previously. [39] A total of 2007 participants were randomly selected from the official community register stratified by gender and 5 year age groups. All participants were Caucasian and had no history of neuropsychiatric disease, including previous stroke, transient ischemic attacks, and dementia. All had a normal neurologic examination. During 2 study periods between September 1991 and March 1994 and between January 1999 and December 2003 an extended diagnostic work-up including Duplex scanning of extracranial arteries, brain MRI and neuropsychological testing was done in 1076 individuals aged 45 to 85 years randomly selected from the entire cohort. In all participants who gave informed consent blood was drawn for DNA extraction since 1992. Genotyping was performed in 996 participants at the Human Genotyping Facility, Genetic Laboratory Department of Internal Medicine, Erasmus MC, Rotterdam, The Netherlands using the Illumina Human610-Quad BeadChip® system. QC criteria used to define unsuccessful genotyping were a call rate < 97%, sex mismatch, or other sample failures. SNPs that were present on HapMap and that passed quality control measures (HWE $p>1\times10^{-5}$, callrate > 0.97, < 3 replicate errors or Mendelian inconsistencies among reference CEPH trios) were used for imputation to plus strand of NCBI build 36, HapMap release #22.

Carotid Atherosclerosis Progression Study (CAPS):

The Carotid atherosclerosis progression study (CAPS) is a community-based study from Germany. Details of the study have been published before. [40] In brief, members of a German primary health care service population (n=32708) were invited to participate. Within a predefined time limit 6962 (21.3%) agreed to participate. Of these, 5,056 were invited to follow-up examination after three years and 3383 (67%) participated. 1,000 individuals in whom carotid IMT measurements were performed, and in whom there was sufficient DNA for investigation, were genotyped and data on these individuals contributed to this study. Informed written consent was obtained from all participants, and the study protocol was approved by the local ethical committee.

Genotyping was performed at Helmholtz Zentrum Munich, Germany in 2009 using the Affymetrix 6.0 chip. From a total of 906,716 genotyped autosomal SNPs, 756,948 passed the quality control measures (no-call rate < 10%, MAF > 0.01). We flagged SNPs for HWE but we have not removed them from the analysis. From the total of 993 samples, 967 samples passed the quality control measurements (no-call rate > 10%, cryptic relatedness and outliers in the multidimensional samples were removed). We used the software BEAGLE [41] for the imputation of ungenotyped SNPs. We removed imputed SNPs with MAF < 0.01 or with posterior probability < 0.9, resuling in 1,458,429 additional SNPs that were used in the analysis. Linear regression was performed on the resulting data; the inflation factor was 1.

Gutenberg Heart Study (GHS):

The GHS is designed as a population-based, prospective, observational single-center cohort study in the Rhein-Main-Region in western Mid-Germany. The primary aim of GHS is to evaluate and improve cardiovascular risk stratification.

The sample was drawn randomly from the governmental local registry offices in the city of Mainz and the district of Mainz-Bingen (In Germany, there is a legal obligation to register; therefore these registry offices contain all citizens of the defined area.) The sample was stratified 1:1 for gender and residence (urban and rural), and in equal strata for decades of age. Individuals between 35 and 74

years of age were enrolled and a written informed consent was obtained from all participants. Exclusion criteria were insufficient knowledge of the German language, and physical or psychological inability to participate in the examinations at the study center. The study protocol and the sampling design were approved by the local ethics committee, and by the local and federal data safety commissioners.

DNA was examined from blood samples of the first 3500 participants who consented to genetic testing at their baseline examination in 2007-08. Genomic DNA was isolated from buffy-coats of EDTA-Plasma samples using the method of Miller et al., (1988). For GWA, genotyping was conducted on the Affymetrix Genome-Wide Human SNP 6.0 Array. Processing of DNA samples using the Affymetrix Genome-Wide Human SNP Nsp/Sty Assay 5.0 and hybridization was done in accordance with the manufactures' standard recommendations. Genotyping was successful in 3196 persons.

Genotyping was performed using Birdseed as implemented in the Affymetrix Power Tools v1.08. Samples with a call rate < 97%, deviations in expected heterozygosity and doubtful identical by state patterns were excluded from the analysis. On the marker level, SNPs were removed with a call rate < 2%, a minor allele frequency < 1%, or a p-value < 0.0001 in the asymptotic test for deviation from Hardy-Weinberg-Equilibrium. Imputation was carried out with IMPUTE v0.5.

MONICA/KORA Augsburg Study (KORA):

The MONICA/KORA Augsburg Study consisted of a series of independent population-based epidemiological surveys of participants living in the region of Augsburg, Southern Germany. All survey participants are residents of German nationality identified through the registration office. The presented data were derived from the fourth population-based Monitoring of Trends and Determinants in Cardiovascular Disease (MONICA)/ Cooperative Health Research in the Region of Augsburg (KORA) survey S4. This cross-sectional survey covering the city of Augsburg (Germany) and two adjacent counties was conducted in 1999/2001 with 4,261 individuals aged 25 to 74 years. In a follow-up examination conducted in 2006/08 (MONICA/KORA F4), a number of 3,080 subjects participated. All participants underwent standardized examinations including blood withdrawals for plasma and DNA. For the MONICA/KORA genome-wide association study, a number of 1,814 subjects were selected from F4 samples. After excluding subjects with no IMT measurements available, the final population for the MONICA/KORA data comprised 1,593 subjects.

Genotyping for F4 was performed using Affymetrix Human SNP Array 6.0. Hybridisation of genomic DNA was done in accordance with the manufacturer's standard recommendations. Genotypes were determined using Birdseed2 clustering algorithm. For quality control purposes, we applied a positive control and a negative control DNA every 96 samples. On chip level only subjects with overall genotyping efficiencies of at least 93% were included. In addition the called gender had to agree with the gender in the MONICA/KORA study database. Imputation of genotypes was performed using maximum likelihood method with the software MaCH v1.0.9.

The Orkney Complex Disease Study (ORCADES):

The Orkney Complex Disease Study (ORCADES) is an ongoing family-based, cross-sectional study in the isolated Scottish archipelago of Orkney. Genetic diversity in this population is decreased compared to Mainland Scotland, consistent with the high levels of endogamy historically. Data for participants from a subgroup of ten islands were used for this analysis. Fasting blood samples were collected and over 200 health-related phenotypes and environmental exposures were measured in each individual. All participants gave informed consent and the study was approved by Research Ethics Committees in Orkney and Aberdeen.

We genotyped 318,237 SNPs for each individual using the Illumina HumanHap300 beadchip. Alleles were called in BeadStudio using Illumina cluster files. Subjects were excluded if they fulfilled any of the following criteria: genotypic call rate <98%, mismatch between reported and

genotypic sex, unexpectedly low genomic sharing with first degree relatives, excess autosomal heterozygosity, or outliers identified by IBS clustering analysis. We excluded SNPs on the basis of minor allele frequency (<0.02), HWE ($P<1\times10^{-10}$), call rate (<98%). Pregnant women were excluded from the study. MaCH v1.0.15 was used to impute over 2 million autosomal SNPs from HapMap build 36. Analysis was performed using ProbABEL to account for the probabilistic nature of imputed genotype calls and kinship accounted for using a mixed polygenic model (mmscore or FASTA routine).

The Cardiovascular Risk In Young Finns (YFS) study:

The Cardiovascular Risk in Young Finns (YFS) is a population-based 27 year follow up-study (http://med.utu.fi/cardio/youngfinnsstudy/). The first cross-sectional survey was conducted in 1980, when 3,596 Caucasian subjects aged 3-18 years participated. In adulthood, the latest 27-year follow-up study was conducted in 2007 (ages 30-45 years) with 2,204 participants. The study cohort for the present analysis comprised subjects who had participated in the ultrasound study in 2007 and had other risk factor data. [42]

DNA was extracted from blood samples drawn on all participants in 2001 and 2007. In 2009 genotyping was performed at the Sanger institute (UK) using the custom-built Illumina BeadChip Human670K from 2442 YFS participants (1123 males, 1319 females) including 546677 SNPs. Genotypes were called using Illumina's clustering algorithm. (Teo et al, A genotype calling algorithm for the Illumina BeadArray platform, Bioinformatics. 2007 October 15; 23(20): 2741–2746) In the start of QC protocol we had 2556 samples in YF intensity file, after initial clustering we removed 2 subjects (CR<0.90), thus the main clustering include 2554 subjects, from these 54 samples failed QC. Thus genotyping pipeline contained 2500 subject s from these 54 were removed due to Sanger genotyping pipeline QC criteria (i.e., duplicated samples, heterozygosity, low call rate, or Sequenom fingerprint discrepancy). After genotyping pipeline QC the following filters were applied to the remaining data: MAF 0.01, GENO 0.05, MIND 0.05, and HWE 1×10^{-6}. 3 of 2500 individuals were removed for low genotyping (MIND > 0.05), 11766 markers were excluded based on HWE test ($p \leq 1\times10^{-6}$), 7,746 SNPs failed missingness test (GENO > 0.05), 34596 SNPs failed frequency test (MAF < 0.01) and one individual failed gender check. None were removed by subsequent heterozygosity check. New binary files were created after removing the individual which failed the gender-check and identity-by-descent (IBD) matrix was subsequently calculated in PLINK [26]. There were 546,770 SNPs and 2,496 individuals at this point which were utilized to generate the genome file. There were 51 pairs of individuals with pi-hat greater than 0.2 thus these individuals removed due to possible relatedness. One of the pair was removed using greater missingness as criteria. After final frequency and genotyping running, there was 546,677 SNPs available from sample of 2,442 YFS subjects. From these a total of 546,677 genotyped autosomal SNPs, those SNPs that were present on HapMap and that passed quality control measures were used for imputation with MaCH version 1.0 (http://www.sph.umich.edu/csg/abecasis/MaCH/).

Section 1C: Associations with coronary artery disease, CARDIoGRAM Consortium

The CARDIoGRAM Consortium represents a GWAS meta-analysis of coronary artery disease (CAD) comprising a discovery set of 22,233 individuals with CAD (cases) and 64,762 controls and a replication stage of 56,682 additional individuals. The methods of consortium organization, genotyping, imputation, and analysis have been described previously. [43] In brief, the consortium used log-additive models adjusting for age (at first event onset for cases or recruitment for controls) and gender and taking into account the uncertainty of possibly imputed genotypes using logistic regression. The primary results and replication are described in recent Nature Genetics manuscript [44]. For the lookup of SNPs from our analyses of common cIMT and plaque, we examined results from the consortium's default meta-analysis, a fixed effect model with inverse variance weighting no indication for heterogeneity for a SNP (P for Q > 0.01).

Section 2A: Participant characteristics, discovery phase

Characteristic	AGES N=3073	Amish N=1,054	ARIC N= 7,767	CHS N= 3,261	ERF N= 1,809	FHS N=3,004	RS I N= 4,699	SardiNIA N=4,235	SHIP N= 2,309
Age, years	76.4 (5.4)	48.1 (15.9)	54.3 (5.7)	72.3 (5.4)	48.5 (14.5)	58.5 (9.7)	68.9 (8.70)	43.5 (17.5)	61.8 (9.5)
Women, %	57.7%	49.4%	53%	61%	56.5	53.3%	59.3%	56.2%	48.6%
Hypertension, %	80.6%	9.3%	27%	51%	51.4%	40.5%	55.9%	29.1%	72.4%
Diabetes, %	11.6%	2.1%	8%	12%	6.1%	8.6%	10%	4.8%	10.1%
Current smoker, %	12.6%	9.4% **	25%	11%	39.4%	15.6%	23.4%	20.2%	19.2%
Total cholesterol, mg/dL	217.9 (44.5)	211.3 (48.1)	214.7 (40.5)	213.0 (38.9)	214.4 (42.6)	205.9 (39.7)	256.0 (46.8)	208.6 (42.1)	234.3 (47.9)
HDL cholesterol, mg/dL	61.0 (17.1)	55.7 (14.8)	50.7 (16.8)	55.3 (15.8)	49.5 (14.1)	51.1 (16.1)	51.8 (13.9)	64.4 (14.9)	55.3 (17.8)
Triglyceride, mg/dL	107.0 (59.0)	74.9 (47.1)	136.0 (89.5)	140.4 (76.4)	118.6 (68.1)	142.3 (138.6)	N/A	87.2 (61.4)	177.6 (134.8)
BMI, kg/m2	27.1 (4.5)	26.9 (4.7)	26.9 (4.7)	26.3 (4.5)	26.80 (4.7)	27.9 (5.1)	26.3 (3.7)	25.3 (4.7)	28.5 (4.6)
Prevalent CVD, %	21.9%	6.9%	5%	0%*	3.1% ***	10.4%	30.8%	1.7%	8.4%
IMT common carotid	0.97 (0.1)	0.74 (0.2)	0.77 (0.2)	1.03 (0.2)	0.82 (0.2)	0.74 (0.2)	1.02 (0.2)	0.54 (0.1)	0.93 (0.2)
IMT internal carotid	n/a	n/a	0.86 (0.48) N=4906	1.23 (0.53) N=3260	n/a	0.79 (0.52) N=2796	n/a	n/a	n/a
Plaque/stenosis present, %	66.9% N=3,073	n/a	18% N=7,957	66.0% N=3,138	66.4% N=1,945	17.7% N=3,019	61% N=3,752	n/a	70.1% N=2,321

Numbers in table are Mean (SD) or percentage. N in the column headers indicates the number of participants with common carotid IMT available. BMI=body mass index; HDL=high density lipoprotein. Diabetes was defined as fasting blood glucose > 125 mg/dL, a random blood glucose of > 200 mg/dL, or use of insulin or oral hypoglycemic agents; hypertension was defined as blood pressure $\geq$140/90 mmHg or on anti-hypertensive medication; current cigarette smoking was defined as self-reported cigarette smoking of at least 1 cigarette per day for a year at any attended exam; cardiovascular disease was defined as coronary heart disease, stroke or transient ischemic attack, or congestive heart failure.

Section 2B: Participant characteristics, second stage

Characteristic	ASPS N=837	ORCADES N=687	CAPS N= 967	GHS N=3,194	KORA N=1,593	RS II N=1,980	YFS N=1,982
Age, years	65.2 (8.0)	53.6 (15.3)	48.7 (13.2)	55.9 (10.9)	60.4 (9.0)	64.7 (7.9)	37.7 (5.0)
Women, %	57%	53%	50.5%	48.8%	51%	54.4%	55%
Hypertension, %	72.5%	21%	21.7%	53.0%	45%	60.3%	6%
Diabetes, %	10%	2%	3.1%	7.4%	8%	10.6%	2.2%
Current smoker, %	11.4%	8%	15.3%	18.3%	14%	19.6%	18%
Total cholesterol, mg/dL	262.5 (45.9)	224.7 (45.2)	215.0 (39.9)	224.2 (41.4)	222.1 (39.0)	223.6 (38.3)	194.3 (34.8)
HDL cholesterol, mg/dL	56.2 (17.2)	65.0 (15.5)	n/a	56.9 (16.1)	56.8 (14.6)	53.1 (14.2)	51.8 (12.8)
Triglyceride, mg/dL	133.1 (74.4)	117.8 (58.5)	n/a	126.1 (74.9)	132.5 (94.5)	141.4 (77.7)	122.6 (80.5)
BMI, kg/m2	26.7 (6.1)	27.69 (4.8)	26.74 (4.1)	27.2 (4.8)	28.2 (4.8)	27.3 (4.2)	26.0 (4.8)
Prevalent CVD, %	0%	8%	n/a	7.2%	3%	10.6%	0%
IMT common carotid	n/a	0.57 (0.21)	0.726 (0.187)	0.7 (0.1)	0.88 (0.13)	0.99 (0.17)	0.66 (0.10)
Plaque/stenosis present, %	60.5% N=837	n/a	n/a	34.8% N=3,161	n/a	n/a	2.4% N=2,015

Numbers in table are Mean (SD) or percentage. N in the column headers indicates the number of participants with common carotid IMT available. BMI=body mass index; HDL=high density lipoprotein. Diabetes was defined as fasting blood glucose > 125 mg/dL, a random blood glucose of > 200 mg/dL, or use of insulin or oral hypoglycemic agents; hypertension was defined as blood pressure $\geq$140/90 mmHg or on anti-hypertensive medication; current cigarette smoking was defined as self-reported cigarette smoking of at least 1 cigarette per day for a year at any attended exam; cardiovascular disease was defined as coronary heart disease, stroke or transient ischemic attack, or congestive heart failure.

Section 3A: Phenotype information, discovery phase

	AGES	Amish	ARIC	CHS	ERF	FHS	RS I	SardiNIA	SHIP
Years measured	2002-2006	2003-2008	1987-89	1989-90 (Baseline)	2002 - 2005	1996-1998 (Cycle 6)	1990-1993	2001-2004 (Baseline)	1997-2001 (Baseline)
Site measured (common carotid)	Near & far wall Left & right	Far wall Left & right	Far wall Left & right	Near & far wall Left & right	Near & far wall Left & right	Near & far wall Left & right	Near & far wall Left & right	Far wall Right	Far wall Left & right
Device	Acuson Sequoia C256, system 5,07 (AS) and GE Vivid 7 version 5.1.1 (GE).	ATL Phillps HDI 5000 Ultrasound machine.	Duplex scanner (Biosound 2000 II SA) high resolution 7,5-MHz transducer	B-mode ultrasonography (model SSA-270A; Toshiba). The ultrasound probe was a 5.0 MHz probe with imaging characteristics of a probe with a frequency of 6.7 MHz. It was used for both the CCA and the ICA. Images were not gated to either systole or diastole	Duplex scanner (ATL UltraMark IV, Advanced Technology Laboratories)	Duplex scanner (Toshiba SSH-140A with high-resolution, 7,5-MHz transducer for CCA and a 5,0-MHz transducer for the ICA)	Duplex scanner (ATL UltraMark IV, Advanced Technology Laboratories)	High-resolution B-mode carotid ultrasonography was performed by use of a linear-array 5- to 7.5-MHz transducer (HDI 3500- ATL Ultramark Inc)	B-Mode ultrasonography was performed using a 5 MHz linear array transducer with an axial resolution of less than 0.5mm and a high resolution instrument (Diasonics VST Gateway, Santa Clara, California, USA).

Section 3B: Phenotype information, second stage

	ASPS	CAPS	GHS	KORA	ORCADES	RS II	YFS
YEARS measured	1991-1994 and 1999-2003 (Baseline)	Baseline	Baseline	Baseline (2006-08)	2005-2007	2000-2001	2007
Site measured (for common carotid IMT)	n/a	Far wall Left & right	Far wall Left & right	Far wall Left & right	Far wall Right	Near & far wall Left & right	Far wall Left
Device	Color-coded B-mode Duplex-scanning (Diasonics, VingMed CFM 750)	7.5-10.0 MHz linear array transducer (P700SE, Phillips Medical System)	ie33 Ultrasound System (Philips, The Netherlands) 11 to 3 MHz linear array transducer B-Mode Ultrasonography	B-mode ultrasound sonography (Sonoline G, Siemens Medical Solutions, Munich, Germany)	Sonosite Titan B mode ultra-sonography	Duplex scanner (ATL UltraMark IV, Advanced Technology Laboratories)	Sequoia 512 ultrasound mainframe (Acuson, CA, USA), 13.0 MHz linear array transducer

Section 4A: Methods for characterization of IMT and plaque, discovery phase

AGES:

Carotid artery intima-media thickness was measured with standardized longitudinal B-mode images from the near and far wall of the common carotid artery for a predefined 10mm segment at defined interrogation angles. Standard images were obtained from 4 angles at each site. The average of all these CIMT values comprises the CIMT outcome parameter.

The presence of atherosclerotic lesions was quantified during the ultrasound examination of the left and right carotid bifurcation and internal carotid artery. The most severe lesion per segment was assessed in a semi-quantitative manner as none, minimal, moderate or severe lesion. Also, an image was stored of the observed plaques.

Amish:

A longitudinal view of the common carotid artery was obtained and the far wall IMT was analyzed using a semi-automated edge detector system over a 1 cm distance. The maximum IMT was determined. The mean of four measurements was averaged (two from the left common carotid artery and two from the right common carotid artery) from different still frames gated to the ECG. [9]

ARIC:

Carotid artery intima-media thickness measured (bilaterally) with B-mode real-time ultrasound in the common (1 cm proximal to the dilatation of the carotid bulb), bifurcation (1 cm proximal to the flow divider), and the internal carotid artery (1 cm distal to the flow divider). Standardized interrogation angles were used. Data was acquired by trained sonographers at four study sites and all images were read centrally. For each arterial segment, 11 IMT measurements (at 1mm increments) on the near and far wall were attempted, on average 5 were obtained. [45] For analyses, the mean of the maximal far wall IMT were used. [46]

Plaque was defined as the presence of a lesion defined by abnormal arterial wall thickness, shape, or texture. Acoustic shadowing was defined as a reduction in amplitude of echoes caused by intervening structures with high attenuation. [47]

CHS:

IMT was characterized with one longitudinal image of the common carotid artery and three longitudinal images of the internal carotid artery. The maximal IMT of the common carotid artery and of the internal carotid artery was defined as the mean of the maximal IMT of the near and far walls on both the left and right sides. [48]

Plaque was defined by the presence of the largest focal lesion classified by surface characteristics, echogenicity, and texture. [48]

ERF:

IMT was defined using ultrasonography of the common carotid artery, carotid bifurcation, and internal carotid artery of the left and right carotid arteries performed with a 7.5-MHz linear-array transducer (ATL UltraMark IV). When an optimal longitudinal image was obtained, it was frozen on the R wave of the ECG and stored on videotape. This procedure was repeated three times for both sides. The actual measurements of intima-media thickness were performed off-line. The interfaces of the distal common carotid artery were marked across a length of 10 mm. The beginning of the dilatation of the distal common carotid artery served as a reference point for the start of the measurement. The average of the intima-media thickness of each of the three frozen images was calculated. For each individual, the common carotid intima-media thickness was determined as the average of near- and far-wall measurements of both the left and right arteries.

The common carotid artery and the carotid bifurcation were evaluated for the presence (yes/no) of atherosclerotic lesions on both the near and far walls of the carotid arteries. Plaques were defined as a focal widening relative to adjacent segments, with protrusion into the lumen composed either

of only calcified deposits or a combination of calcification and noncalcified material. The size or extent of the lesions was not quantified. [49]

FHS:

IMT was characterized with 1 longitudinal images of the distal CCA taken at end diastole and 2 longitudinal views of the ICA at end diastole. Measurement of the peak systolic velocity in the ICA was obtained with color Doppler imaging and duplex ultrasound. [50,51]

Plaque was defined by carotid stenosis of 25% or greater. [51,52]

Rotterdam (RS I and RS II):

IMT for the common carotid artery was characterized using an average of 3 measurements across a length of 10 mm with the beginning of the dilatation of the distal CCA as the reference point. Average of near- and far- wall measurements of both the left and right arteries were calculated. Both mean and the max were computed. Bifurcation was defined the part of the artery between the common carotid artery and the tip of the flow divider. Internal Carotid Artery was defined as the part f the artery after the tip of the flow divider. The max IMT at bifurcation and ICA were used. [49]

Plaques were defined as a focal widening relative to adjacent segments, with protrusion into the lumen composed either of only calcified deposits or a combination of calcification and noncalcified material. [53]

SardiNIA:

IMT was characterized while the subject lay in the supine position in a dark, quiet room. The stabilized BP after 15 minutes from the onset of testing was used for subsequent analyses. The right CCA was examined with the head tilted slightly upward in the midline position. The transducer was manipulated so that the near and far walls of the CCA were parallel to the transducer footprint and the lumen diameter was maximized in the longitudinal plane. A region 1.5 cm proximal to the carotid bifurcation was identified, and the IMT of the far wall was evaluated as the distance between the luminal-intimal interface and the medial-adventitial interface. IMT was measured on the frozen frame of a suitable longitudinal image with the image magnified to achieve a higher resolution of detail. The IMT measurement was obtained from 5 contiguous sites at 1-mm intervals, and the average of the 5 measurements was used for analyses. [54,55]

Common carotid artery plaque was defined as focal encroachment of the arterial wall.

SHIP:

For the characterization of IMT, both carotid arteries were assessed in B-mode. The maximum far-wall IMT of the distal straight portion of the right and left common carotid arteries (CCA) proximally from the bifurcation was assessed by selecting the highest IMT value of in total 10 consecutive measurement points (in 1-mm steps) from the bulb. The 'mean of the maximum IMT' was calculated by averaging the maximum IMT of both, the right and left CCA far wall. IMT measurement was performed manually on digitally stored images by trained and certified readers.[56]

Atherosclerotic plaques were defined as a focal thickening of the vessel wall with protrusion into the vessel lumen relative to adjacent segments or as a localized roughness with increased echogenicity.

Section 4B: Methods for characterization of IMT and plaque, second stage

ASPS:

For plaque, the largest focal lesion was classified by surface characteristics, echogenicity, and texture. [57]

CAPS:

For IMT, one longitudinal image of the common carotid artery was measured using a standard protocol as previously published. [40] Images were digitally captured during the systole of a single heartbeat and mean IMT determined using semi-automated software with high inter and intra-observer reproducibility. The mean of left and right CCA-IMT was used in all analyses. ICA and bifurcation IMT was also measured, but only CCA IMT was used in this analysis

GHS:

IMT was characterized using one image loop from a longitudinal sectional view of the common carotid artery. Evaluation was performed by a semi-automatic computerized system (Qlab quantification software, IMT; Philips, NL); triggered according to the Q wave of the ECG. IMT was recorded over a region of 10mm at a position 1cm proximal to the carotid bulb and without plaque. The IMT of the common carotid artery was defined as the mean of the means of all automatized measurements of the far walls on both, the left and right side.

Plaques were defined as thickening of the IMT of at least 1,5mm in any of the arteries (ACC, ACI and ACE). The number of plaques from both sides was recorded and subjects being classified as plaque positive when at least one plaque was measured on either side.

KORA:

The mean far-wall IMT was determined using the average of the measurements of frozen images from both the left and right CCA to calculate artery thickness of the CCA, 1 cm below the bifurcation ((mean left + mean right) /2).

ORCADES:

For IMT, one longitudinal image of the common carotid artery centered ~2 cm below the bifurcation was measured three times using the Sonosite tool. The mean of the three IMT readings from the far wall on the right side was taken for analysis.

YFS:

For IMT, the left carotid artery was scanned. The image was focused on the posterior (far) wall, and images were recorded from the angle showing the greatest distance between the lumen-intima interface and the media-adventitia interface. A scan including the beginning of the carotid bifurcation and the common carotid artery was recorded and stored in digital format on optical discs for subsequent off-line analysis. A minimum of four measurements of the common carotid far wall were taken 10 mm proximal to the bifurcation to derive mean of the maximal carotid IMT.

The far and near walls of the left common carotid artery and carotid bulb area were scanned for the presence of atherosclerotic plaque, defined as a distinct area of the vessel wall protruding into the lumen >50% of the adjacent intima–media layer.

Section 5A: Analysis logistics, discovery phase

	AGES	Amish	ARIC	CHS	ERF	FHS	RS I	SardiNIA	SHIP
Adjustments	Age and sex	age, age^2, sex, age×sex, & age^2×sex	age and field center within each sex	age, sex, clinic: as adjustment in genotype-phenotype model	Age (sex adjusted in strata, as above)	Age and principal components 1-10 (sex-specific, unstandardized residuals)	Age, Sex (Sex specific unstandardized residuals)	Age, Sex	Age, sex
Analysis method	Linear regression models for quantitative traits and logistic regression for qualitative traits	Mixed model with relationship matrix to account for family structure	Linear regression models for quantitative traits and logistic regression for qualitative traits	Linear regression models for quantitative traits and logistic regression for qualitative traits	Linear mixed models	Linear mixed effect models to account for family structure	Linear regression models for quantitative traits and logistic regression for qualitative traits	Variance Component Based Analysis to Account for Relatedness[37]	Linear regression models for quantitative traits and logistic regression for qualitative traits
Analysis software	R using ProbABEL software	Mixed Model Analysis for Pedigrees (MAPP),	R using ProbABEL software	R version 2.7	R, GenABEL, ProbABEL	R version 2.6.1, using lmekin	R using ProbABEL software (version 1.1)	Merlin	SNPTEST v1.1.5; QUICKTEST v0.94

Section 5B: Analysis logistics, second stage

	ASPS	ORCADES	CAPS	GHS	KORA	RS II	YFS
Adjustments	Age, sex	Age, sex, first three principal components	Age, sex	Age, sex	Age ,sex	Age, Sex (Sex specific unstandardized residuals)	Age, sex
Analysis method	Logistic regression	Linear regression with mmscore (FASTA) to account for relatedness	Linear regression	Linear regression models for quantitative traits and logistic regression for qualitative traits	Linear regression	Linear regression models for quantitative traits	Linear regression models for quantitative traits and logistic regression for qualitative traits
Analysis software	R version 2.10	ProbABEL	Plink	SNPTEST	MaCH, ProbABEL	R using ProbABEL software (version 1.1)	ProbABEL

MAPP, developed by Dr. J. R. O'Connell at UMB (joconnel@medicine.umaryland.edu)

Plink: http://pngu.mgh.harvard.edu/~purcell/plink/

ProbABEL [58]: http://mga.bionet.nsc.ru/~yurii/ABEL/

QUICKTEST: http://toby.freeshell.org/software/quicktest.shtml

R: http://www.r-project.org

SAS: Cary, North Carolina, http://www.sas.com/presscenter/guidelines.html

R: http://www.r-project.org

SAS: Cary, North Carolina, http://www.sas.com/presscenter/guidelines.html

SNPTEST: http://www.stats.ox.ac.uk/~marchini/software/gwas/snptest.html

Section 6: Genotyping & quality control, discovery phase

	AGES	Amish	ARIC	ERF	CHS	FHS	RS I	SardiNIA	SHIP
Genotyping platforms, SNP panel	Illumina HumanHap 370CNV	Affymetrix 500K	Affymetrix Genome-Wide Human SNP Array 6.0	Version 3 Illumina Infinium II HumanHap550	Illumina HumanHap 370CNV	Affymetrix 500K (250K Nsp & 250K Sty, MIPS 50K	Version 3 Illumina Infinium II HumanHap550	Affymetrix 500K	Affymetrix Genome-Wide Human SNP Array 6.0
Gentoyping center	NIA/NIH	UMB Genomic Core Facility	Broad Institute	Erasmus MC Rotterdam	Cedars-Sinai, Rotter & Taylor	Affymetrix	Erasmus MC Rotterdam	Lanusei, Sardinia, Italy	Affymetrix, University of Greifswald
Genotyping calling algorithm	Illumina BeadStudio	BRLMM	BRLMM	Illumina BeadStudio	Illumina BeadStudio	BRLMM	Illumina BeadStudio	BRLMM	Birdseed V2
Call rate threshold	> 95%	95%	> 95%	≥ 98%	> 95%	> 97%	> 97.5%	95%	> 92%
Other exclusions	sample failures, genotyped sex different from recorded sex, discordance with prior genotyping		Discordant with previous genotyping, genotypic sex mismatch, suspected 1st degree relative of included individual by genotype, genetic outlier as assessed by Plink IBS or EIGENSTRAT		sample failures, genotyped sex different from recorded sex, discordance with prior genotyping	>1000 Mendelian errors Heterozygosity 5 SD from mean (<25.758% or >29.958%)		Markers with an excess of Mendelian inconsistencies (>2 in 67 trios; 500K or >4 in 1266 trios; 10K) removed	15 individuals were excluded due to duplicate samples (by IBS) or reported vs. genotyped gender mismatch

Section 6 (continued): Genotyping & quality control, discovery phase

	AGES	Amish	ARIC	ERF	CHS	FHS	RS I	SardiNIA	SHIP
Imputation software	Mach (version 1.0.16)	Mach (version 1.0.16)	Mach (version 1.0.16)	Mach (v1.0.1.5)	BIMBAM (v.0.99)	Mach (v1.0.1.5)	Mach (v1.0.1.5)	Mach (v1.0.1.16)	IMPUTE (v0.5.0)
Imputation: reference panel	HapMap CEU, build 36	HapMap CEU, build 36	BRLMM to Hapmap-V1 CEU backbone (build 35)	HapMap CEU trios (build 36, release 22)	HapMap CEU, Build 36	HapMap CEU, Build 36, Release 22	HapMap CEU trios (build 36, release 22)	HapMap CEU trios (build 36, release 22)	HapMap CEU trios (build 36, release 22)
Pre-imputation MAF filter	< 0.01	< 0.01	< 0.01	≤ 0.01	none	< 0.01	≤ 0.01	< 0.05	none
Pre-imputation HWE filter	$< 1x10^{-6}$	$< 1x10^{-6}$	$< 1x10^{-5}$	$< 1x10^{-6}$	$< 1x10^{-5}$	$< 1x10^{-6}$	$< 1x10^{-6}$	$< 1x10^{-6}$	none
Pre-imputation call frequency filter	< 90%	< 95%	< 95%	≤ 90%	97%	< 97%	≤ 90%	< 90%	< 98.6%

Other pre-imputation SNP filters	SNPs with mismatched position, and/or mismatched alleles to dbSNP; missing by haplotype	SNPs not in the HapMap were excluded from being used for imputation	SNPs without chromosomal location; monomorphic SNPs; and SNPs whose genotype frequencies between differed by $p<10^{-6}$		SNPs with no observed heterozygotes; SNPs with variance dosage < 0.01 excluded (~MAF 0.05); or SNPs not present in HapMap reference panel			SNPs without position assigned; poorly imputed SNPs ($r^2< 0.3$)	None
N SNPs used for imputation	329,804	338,598	704,588	460,584	306,665	378,163	491,875	356,359	869,224
Imputation quality metrics	r^2hat>0.3	None	none		variance dosage > 0.01	Ratio of dosage variance to expected >0.3		$R^2 > 0.3$	none
N imputed SNPs for analysis	2,408,991	2,543,013	2,516,204	2,543,888	2,334,835	2,543,887	2,586,725	2,252,229	2,748,910

BIMBAM: http://stephenslab.uchicago.edu/software.html
MaCH: http://www.sph.umich.edu/csg/abecasis/MaCH/index.html

References

1. Zannad, F. *et al.* Genetics strongly determines the wall thickness of the left and right carotid arteries. *Hum Genet* **103**, 183-8 (1998).
2. Jartti, L. *et al.* Population-based twin study of the effects of migration from Finland to Sweden on endothelial function and intima-media thickness. *Arterioscler Thromb Vasc Biol* **22**, 832-7 (2002).
3. Fox, C.S. *et al.* Genetic and environmental contributions to atherosclerosis phenotypes in men and women: heritability of carotid intima-media thickness in the Framingham Heart Study. *Stroke* **34**, 397-401 (2003).
4. Swan, L., Birnie, D.H., Inglis, G., Connell, J.M. & Hillis, W.S. The determination of carotid intima medial thickness in adults--a population-based twin study. *Atherosclerosis* **166**, 137-41 (2003).
5. Moskau, S. *et al.* Heritability of carotid artery atherosclerotic lesions: an ultrasound study in 154 families. *Stroke* **36**, 5-8 (2005).
6. Mayosi, B.M. *et al.* Genotype at the -174G/C polymorphism of the interleukin-6 gene is associated with common carotid artery intimal-medial thickness: family study and meta-analysis. *Stroke* **36**, 2215-9 (2005).
7. Sayed-Tabatabaei, F.A. *et al.* Heritability of the function and structure of the arterial wall: findings of the Erasmus Rucphen Family (ERF) study. *Stroke* **36**, 2351-6 (2005).
8. Ryabikov, A. *et al.* Intrafamilial correlations of carotid intima-media thickness and flow-mediated dilation in a Siberian population. *Am J Hypertens* **20**, 248-54 (2007).
9. Rampersaud, E. *et al.* The association of coronary artery calcification and carotid artery intima-media thickness with distinct, traditional coronary artery disease risk factors in asymptomatic adults. *Am J Epidemiol* **168**, 1016-23 (2008).
10. Lange, L.A. *et al.* Heritability of carotid artery intima-medial thickness in type 2 diabetes. *Stroke* **33**, 1876-81 (2002).
11. Zhao, J. *et al.* Heritability of carotid intima-media thickness: a twin study. *Atherosclerosis* **197**, 814-20 (2008).
12. Duggirala, R., Gonzalez Villalpando, C., O'Leary, D.H., Stern, M.P. & Blangero, J. Genetic basis of variation in carotid artery wall thickness. *Stroke* **27**, 833-7 (1996).
13. Hunt, K.J. *et al.* Genetic basis of variation in carotid artery plaque in the San Antonio Family Heart Study. *Stroke* **33**, 2775-80 (2002).
14. Juo, S.H. *et al.* Genetic and environmental contributions to carotid intima-media thickness and obesity phenotypes in the Northern Manhattan Family Study. *Stroke* **35**, 2243-7 (2004).
15. Wang, D. *et al.* A genome-wide scan for carotid artery intima-media thickness: the Mexican-American Coronary Artery Disease family study. *Stroke* **36**, 540-5 (2005).
16. Kao, W.H. *et al.* Family history of type 2 diabetes is associated with increased carotid artery intimal-medial thickness in Mexican Americans. *Diabetes Care* **28**, 1882-9 (2005).
17. Xiang, A.H. *et al.* Heritability of subclinical atherosclerosis in Latino families ascertained through a hypertensive parent. *Arterioscler Thromb Vasc Biol* **22**, 843-8 (2002).
18. Sacco, R.L. *et al.* Heritability and linkage analysis for carotid intima-media thickness: the family study of stroke risk and carotid atherosclerosis. *Stroke* **40**, 2307-12 (2009).
19. Dong, C. *et al.* Genomewide linkage and peakwide association analyses of carotid plaque in Caribbean Hispanics. *Stroke* **41**, 2750-6 (2010).
20. North, K.E. *et al.* Heritability of carotid artery structure and function: the Strong Heart Family Study. *Arterioscler Thromb Vasc Biol* **22**, 1698-703 (2002).
21. Manolio, T.A., Boerwinkle, E., O'Donnell, C.J. & Wilson, A.F. Genetics of ultrasonographic carotid atherosclerosis. *Arterioscler Thromb Vasc Biol* **24**, 1567-77 (2004).
22. Psaty, B.M. *et al.* Cohorts for Heart and Aging Research in Genomic Epidemiology (CHARGE) Consortium: Design of Prospective Meta-Analyses of Genome-Wide Association Studies From 5 Cohorts. *Circ Cardiovasc Genet* **2**, 73-80 (2009).
23. Harris, T.B. *et al.* Age, Gene/Environment Susceptibility-Reykjavik Study: multidisciplinary applied phenomics. *Am J Epidemiol* **165**, 1076-87 (2007).
24. Post, W. *et al.* Determinants of coronary artery and aortic calcification in the Old Order Amish. *Circulation* **115**, 717-24 (2007).
25. The Atherosclerosis Risk in Communities (ARIC) Study: design and objectives. The ARIC investigators. *Am J Epidemiol* **129**, 687-702 (1989).

26. Purcell, S. *et al.* PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am J Hum Genet* **81**, 559-75 (2007).
27. Price, A.L. *et al.* Principal components analysis corrects for stratification in genome-wide association studies. *Nat Genet* **38**, 904-9 (2006).
28. Fried, L.P. *et al.* The Cardiovascular Health Study: design and rationale. *Ann Epidemiol* **1**, 263-76 (1991).
29. Servin, B. & Stephens, M. Imputation-based analysis of association studies: candidate regions and quantitative traits. *PLoS Genet* **3**, e114 (2007).
30. Aulchenko, Y.S. *et al.* Linkage disequilibrium in young genetically isolated Dutch population. *Eur J Hum Genet* **12**, 527-34 (2004).
31. Dawber, T.R. & Kannel, W.B. The Framingham study. An epidemiological approach to coronary heart disease. *Circulation* **34**, 553-5 (1966).
32. Kannel, W.B., Feinleib, M., McNamara, P.M., Garrison, R.J. & Castelli, W.P. An investigation of coronary heart disease in families. The Framingham offspring study. *Am J Epidemiol* **110**, 281-90 (1979).
33. Splansky, G.L. *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* **165**, 1328-35 (2007).
34. Hofman, A. *et al.* The Rotterdam Study: 2010 objectives and design update. *Eur J Epidemiol* **24**, 553-72 (2009).
35. Pilia, G. *et al.* Heritability of cardiovascular and personality traits in 6,148 Sardinians. *PLoS Genet* **2**, e132 (2006).
36. Epstein, M.P., Duren, W.L. & Boehnke, M. Improved inference of relationship for pairs of individuals. *Am J Hum Genet* **67**, 1219-31 (2000).
37. Chen, W.M. & Abecasis, G.R. Family-based association tests for genomewide association scans. *Am J Hum Genet* **81**, 913-26 (2007).
38. John, U. *et al.* Study of Health In Pomerania (SHIP): a health examination survey in an east German region: objectives and design. *Soz Praventivmed* **46**, 186-94 (2001).
39. Schmidt, R. *et al.* Assessment of cerebrovascular risk profiles in healthy persons: definition of research goals and the Austrian Stroke Prevention Study (ASPS). *Neuroepidemiology* **13**, 308-13 (1994).
40. Sitzer, M. *et al.* C-reactive protein and carotid intimal medial thickness in a community population. *J Cardiovasc Risk* **9**, 97-103 (2002).
41. Browning, B.L. & Browning, S.R. A unified approach to genotype imputation and haplotype-phase inference for large data sets of trios and unrelated individuals. *Am J Hum Genet* **84**, 210-23 (2009).
42. Raitakari, O.T. *et al.* Cohort profile: the cardiovascular risk in Young Finns Study. *Int J Epidemiol* **37**, 1220-6 (2008).
43. Preuss, M. *et al.* Design of the Coronary ARtery DIsease Genome-Wide Replication And Meta-Analysis (CARDIoGRAM) Study: A Genome-wide association meta-analysis involving more than 22 000 cases and 60 000 controls. *Circ Cardiovasc Genet* **3**, 475-83 (2010).
44. Schunkert, H. *et al.* Large-scale association analysis identifies 13 new susceptibility loci for coronary artery disease. *Nat Genet* **43**, 333-8 (2011).
45. Heiss, G. *et al.* Carotid atherosclerosis measured by B-mode ultrasound in populations: associations with cardiovascular risk factors in the ARIC study. *Am J Epidemiol* **134**, 250-6 (1991).
46. Espeland, M.A. *et al.* Analysis strategies for serial multivariate ultrasonographic data that are incomplete. *Stat Med* **11**, 1041-56 (1992).
47. Hunt, K.J. *et al.* Acoustic shadowing on B-mode ultrasound of the carotid artery predicts ischemic stroke: the Atherosclerosis Risk in Communities (ARIC) study. *Stroke* **32**, 1120-6 (2001).
48. O'Leary, D.H. *et al.* Use of sonography to evaluate carotid atherosclerosis in the elderly. The Cardiovascular Health Study. CHS Collaborative Research Group. *Stroke* **22**, 1155-63 (1991).
49. Bots, M.L., Hoes, A.W., Koudstaal, P.J., Hofman, A. & Grobbee, D.E. Common carotid intima-media thickness and risk of stroke and myocardial infarction: the Rotterdam Study. *Circulation* **96**, 1432-7 (1997).
50. Fox, C.S. *et al.* Genomewide linkage analysis for internal carotid artery intimal medial thickness: evidence for linkage to chromosome 12. *Am J Hum Genet* **74**, 253-61 (2004).
51. O'Leary, D.H. *et al.* Distribution and correlates of sonographically detected carotid artery disease in the Cardiovascular Health Study. The CHS Collaborative Research Group. *Stroke* **23**, 1752-60 (1992).

52. Elosua, R. *et al.* Association of APOE genotype with carotid atherosclerosis in men and women: the Framingham Heart Study. *J Lipid Res* **45**, 1868-75 (2004).
53. Hollander, M. *et al.* Carotid plaques increase the risk of stroke and subtypes of cerebral infarction in asymptomatic elderly: the Rotterdam study. *Circulation* **105**, 2872-7 (2002).
54. Scuteri, A. *et al.* Age- and gender-specific awareness, treatment, and control of cardiovascular risk factors and subclinical vascular lesions in a founder population: The SardiNIA Study. *Nutr Metab Cardiovasc Dis* (2009).
55. Scuteri, A. *et al.* The central arterial burden of the metabolic syndrome is similar in men and women: the SardiNIA Study. *Eur Heart J* (2009).
56. Ludemann, J. *et al.* [Methods for quality assurance of medical examination in epidemiological field studies: the "Study of Health in Pomerania" (SHIP)]. *Gesundheitswesen* **62**, 234-43 (2000).
57. Schmidt, R. *et al.* C-reactive protein, carotid atherosclerosis, and cerebral small-vessel disease: results of the Austrian Stroke Prevention Study. *Stroke* **37**, 2910-6 (2006).
58. Aulchenko, Y.S., Ripke, S., Isaacs, A. & van Duijn, C.M. GenABEL: an R library for genome-wide association analysis. *Bioinformatics* **23**, 1294-6 (2007).