nature COMMUNICATIONS

ARTICLE

https://doi.org/10.1038/s41467-018-07340-5 OPEN

GWAS and colocalization analyses implicate carotid intima-media thickness and carotid plaque loci in cardiovascular outcomes

Nora Franceschini, Claudia Giambartolomei et al.[#]

Carotid artery intima media thickness (cIMT) and carotid plaque are measures of subclinical atherosclerosis associated with ischemic stroke and coronary heart disease (CHD). Here, we undertake meta-analyses of genome-wide association studies (GWAS) in 71,128 individuals for cIMT, and 48,434 individuals for carotid plaque traits. We identify eight novel susceptibility loci for cIMT, one independent association at the previously-identified *PINX1* locus, and one novel locus for carotid plaque. Colocalization analysis with nearby vascular expression quantitative loci (cis-eQTLs) derived from arterial wall and metabolic tissues obtained from patients with CHD identifies candidate genes at two potentially additional loci, *ADAMTS9* and *LOXL4*. LD score regression reveals significant genetic correlations between cIMT and plaque traits, and both cIMT and plaque with CHD, any stroke subtype and ischemic stroke. Our study provides insights into genes and tissue-specific regulatory mechanisms linking atherosclerosis both to its functional genomic origins and its clinical consequences in humans.

Correspondence and requests for materials should be addressed to J.L.M.B. (email: johan.bjorkegren@mssm.edu) or to C.J.O. (email: Christopher.ODonnell@va.gov). #A full list of authors and their affiliations appears at the end of the paper.

Atherosclerosis is characterized by an accumulation of lipid-rich and inflammatory deposits (plaques) in the subintimal space of medium and large arteries. Plaque enlargement leads to blood flow limitation, organ ischemia, and/or tissue necrosis. Plaque rupture can lead to abrupt vascular occlusion, which underlies clinical cardiovascular events, including myocardial infarction and ischemic stroke. Coronary heart disease (CHD) accounts for one in seven deaths, and stroke accounts for one in 20 deaths in the US[1]. Because atherosclerosis has a long pre-clinical phase, early detection of atherosclerosis using non-invasive methods may help identify individuals at risk for atherosclerotic clinical events[2], and provides an opportunity for prevention. Subclinical atherosclerosis can be detected by B-mode ultrasound measurement of common carotid artery intima-media thickness (cIMT) or carotid plaques[1].

Subclinical and clinical atherosclerosis has known genetic components[3]. Genome-wide association studies (GWAS) of subclinical atherosclerosis have previously identified three loci significantly associated with cIMT at *ZHX2*, *APOC1*, and *PINX1*, and two loci associated with common carotid artery plaque at *PIK3CG* and *EDNRA*[4]. An exome-wide-association study identified significant associations of the *APOE* ε2 allele with cIMT and coronary artery calcification[5]. The *APOE* single nucleotide polymorphism (SNP) rs7412 is in linkage disequilibrium (LD) with the *APOC1* variant, thus representing the same signal. Additional GWAS-identified associations were reported for carotid plaque at the 9p21 and *SFXN2* loci[6], and for cIMT at the *CFDP1-TMEM170A* locus[7]. However, these prior studies were of limited sample size and genomic coverage, and failed to investigate the etiological role that subclinical atherosclerosis may have on atherosclerotic clinical events.

Herein, we perform a large meta-analysis of GWAS of subclinical atherosclerosis by analyzing 1000 Genomes imputed genotype data obtained from collaborations between the Cohorts for Heart and Aging Research in Genomic Epidemiology (CHARGE) consortium[8] and the University College London-Edinburgh-Bristol (UCLEB) consortium[9]. One of the greatest challenges in the translation of GWAS findings to biological understanding is related to the limited access to RNA expression data from disease-relevant tissues. Consequently, we sought to reliably identify the tissue-specific gene regulatory functions responsible for the GWAS signals by prioritizing candidate genes for established and novel loci of cIMT and carotid plaque using statistical methods for colocalization[10]. These methods integrate identified loci with expression quantitative loci (eQTLs) inferred from cardiovascular disease-relevant genetics of RNA expression, the Stockholm-Tartu Atherosclerosis Reverse Network Engineering Task (STARNET) study, where arterial wall and metabolic-related RNA samples were collected from up to 600 patients with CHD[11]. We also evaluate the relationships of cIMT and carotid plaque with clinically apparent CHD and stroke using summary data from two large consortia. In summary, our study sequentially assesses the genetic epidemiology and tissue-specific patterns of gene regulation involved in the formation of subclinical atherosclerosis traits across cardiovascular disease-related tissues.

Results

Study description. The study design is shown in Fig. 1. We undertook meta-analysis of GWAS in individuals of European ancestry for cIMT (up to 71,128 participants from 31 studies) and carotid plaque (up to 48,434 participants from 17 studies; 21,540 with defined carotid plaque) (Supplementary Table 1). cIMT and plaque were evaluated using high-resolution B-mode ultrasonography and reading protocols as previously reported[4]. Carotid plaque was defined by atherosclerotic thickening of the common carotid artery wall or the proxy measure of luminal stenosis greater than 25% (Supplementary Table 2). Each cohort performed association analyses using standardized protocols (Methods) for variants imputed based on the 1000 Genomes Project (1000G) phase 1 v3 reference. Extensive quality control (QC) was applied to data, and there was little evidence for population stratification in any of the studies for either trait (Supplementary Table 3). The study-specific results were combined using fixed-effect meta-analyses, given the low heterogeneity across studies (0% heterogeneity)[12].

GWAS meta-analyses of cIMT and carotid plaque. For cIMT, 11 loci had at least one SNP association that reached the genome-wide association threshold ($p < 5 \times 10^{-8}$), of which eight were newly described and three have been previously reported (Table 1). The closest genes for the eight loci were: 1q32.2 intergenic (rs201648240), *ATP6AP1L* (rs224904), *AIG1* (rs6907215), *PIK3CG* (rs13225723), *MCPH1* (rs2912063), *SGK223* (rs11785239), *VTI1* (rs1196033), and *CBFA2T3* (rs844396). For three loci previously reported, the closest genes were *ZHX2* (rs148147734), *PINX1* (rs200482500), and *APOE* (rs7412).

The *PIK3CG* is a newly described locus for cIMT, but has been previously reported in a GWAS of carotid plaque[4]. The two

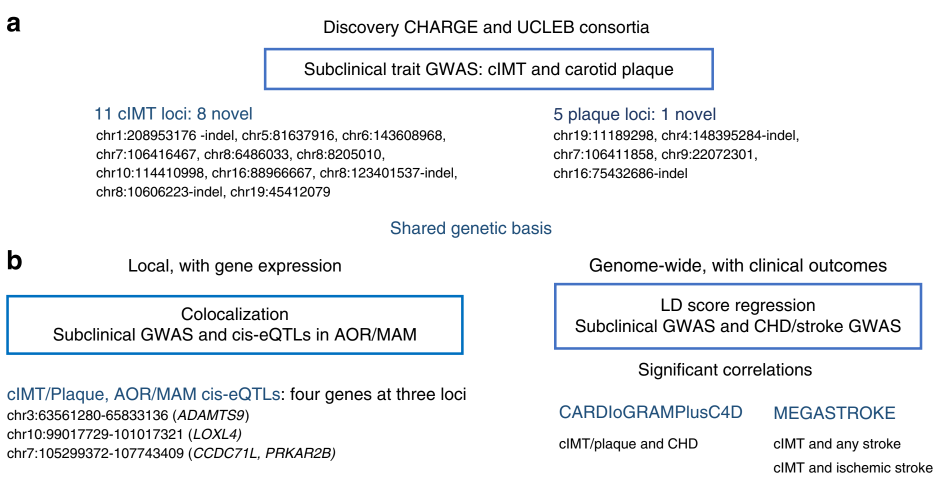


Fig. 1 Overall study design. **a** GWAS meta-analyses of cIMT and carotid plaque for gene discovery. **b** Local and genome-wide shared genetic basis using gene expression and clinical outcomes GWAS data

signals on chromosome 8 near *MCPH1* (rs2912063) and *SGK223* (rs11785239) were confirmed to be independent through conditional analysis (Supplementary Table 4). At the *PINX1* locus, the lowest association *p*-value variant (rs200482500) was not in LD with the previously reported associated variant in the region (rs6601530, $r^2 = 0.0$, Table 1), thus representing an independent signal at this locus. Two additional loci for cIMT had an SNP that reached suggestive evidence for association ($p < 1.0 \times 10^{-7}$) including an SNP nearby *APOB* (rs515135) and an intronic low frequency variant at *ATG4B* (rs139302128, minor allele frequency [MAF] = 0.03) (Supplementary Table 5).

The GWAS meta-analysis for carotid plaque identified five loci, of which one has not been previously described (nearby gene *LDLR*) (Table 1). At four known loci associated with carotid plaque (nearby genes *EDNRA*, *PIK3CG*, *CFDP1-TMEM170A*, and at the 9p21 region), the most significantly associated variants were in LD with the previously reported SNPs (Table 1)[4,6,7], indicating that these SNPs mark the same association at each locus. Two suggestive loci ($p < 10^{-7}$) were also identified nearby the genes *TMCO5B* and *STEAP2-AS1* (Supplementary Table 5). Conditional analyses confirmed the presence of a single independent signal at each locus. Manhattan and QQ plots from the meta-analysis of cIMT and carotid plaque are shown in Supplementary Figure 1 and regional plots in Supplementary Figure 2. Forest Plots for all loci are shown in Supplementary Figure 3.

Regulatory annotations of GWAS SNPs for cIMT/carotid plaque. To better define potentially causal variants within the identified genetic risk loci, we jointly analyzed the GWAS data with functional genomic information such as annotations on active transcription sites or open chromatin regions (i.e., performed a fine-mapping functional genome-wide association analysis using fGWAS[13]). Only variants in the *PINX1* region were found to have a high probability that its association with cIMT is driven by SNPs that fall within transcription sites in adipose-derived mesenchymal stem cells at a DNaseI-hypersensitive site (Supplementary Figure 4), a finding that provides a down-stream mechanistic explanation for the cIMT signal in the *PINX1* locus.

To further explore the regulatory functions of variants in the identified loci for cIMT and carotid plaque, we investigated whether the identified lead SNPs were also eQTLs using vascular RNAseq data from GTEx (aorta, coronary and tibial arteries, heart atrial appendage, and heart left ventricle) and from the coronary artery disease cohort of STARNET (i.e., from the atherosclerotic-lesion-free internal mammary artery [MAM] and atherosclerotic aortic root [AOR]). Lead SNP associated with cIMT and carotid plaque (rs13225723) in the *PIK3CG* locus was found to be vascular-specific eQTLs for *CCDC71L* and *PRKAR2B* in GTEx aorta as well as in STARNET AOR and MAM tissues (Table 2, Fig. 2), suggesting that the genetic regulation of these two genes are responsible for risk variation in cIMT and carotid plaque development in this locus.

Colocalization analysis of GWAS data and STARNET eQTLs. To identify further candidate genes in tissues affected by atherosclerosis that had strong evidence of sharing the same variant for cIMT and carotid plaque as found in our GWAS, we conducted pairwise colocalization analysis of these genetic variants with *cis*-eQTLs in the STARNET study[10].

The pairwise colocalization analysis is based on coloc, a Bayesian statistical methodology that tests pairwise colocalization of SNPs in GWAS with eQTLs and, in this fashion, generates posterior probabilities for each locus weighting the evidence for competing hypothesis of either no colocalization or sharing of a distinct SNP at each locus[10]. We used summary statistics from all SNPs within a 200-kb window around each gene covered by the eQTL datasets ($N = 18{,}705$, see Methods), and analyzed each

Table 1 Loci significantly associated with cIMT and plaque GWAS

SNP	Chr:position	Nearest coding gene	Alleles (effect/ other)	Effect allele freq.	Beta (SE)	*p*	*N*
Newly identified loci for cIMT							
rs201648240	1:208953176-indel	*LINC01717*	−/AA	0.83	−0.0062 (0.0011)	4×10^{-9}	54,752
rs224904	5:81637916	*ATP6AP1L*	C/G	0.95	−0.0088 (0.0016)	5×10^{-8}	68,962
rs6907215	6:143608968	*AIG1*	T/C	0.60	−0.0040 (0.0007)	5×10^{-8}	64,586
rs13225723	7:106416467	*PIK3CG*	A/G	0.22	0.0052 (0.0009)	3×10^{-9}	68,070
rs2912063	8:6486033	*MCPH1*	A/G	0.71	0.0045 (0.0008)	9×10^{-9}	67,401
rs11785239	8:8205010	*SGK223*	T/C	0.65	−0.0043 (0.0008)	9×10^{-9}	67,107
rs11196033	10:114410998	*VTI1A*	A/C	0.48	0.0042 (0.0008)	4×10^{-8}	57,995
rs844396	16:88966667	*CBFA2T3*	T/C	0.30	−0.0051 (0.0009)	6×10^{-9}	50,377
Newly identified loci for plaque							
rs200495339	19:11189298-indel	*LDLR*	−/G	0.11	−0.1023 (0.0179)	1×10^{-8}	36,569
Known loci for cIMT							
rs148147734[a]	8:123401537-indel	*ZHX2*	−/G	0.54	0.0050 (0.0007)	3×10^{-11}	58,141
rs200482500[a]	8:10606223-indel	*PINX1*	−/GTACC	0.52	0.0056 (0.0008)	7×10^{-12}	58,141
rs7412[a]	19:45412079	*APOE*	T/C	0.08	−0.0119 (0.0015)	1×10^{-14}	44,607
Known loci for plaque							
rs11413744[b]	4:148395284-indel	*EDNRA*	−/T	0.86	−0.1586 (0.0253)	4×10^{-10}	39,577
rs17477177[b]	7:106411858	*PIK3CG*	T/C	0.79	−0.1305 (0.0197)	4×10^{-11}	47,863
rs9632884[b]	9:22072301	9p21	C/G	0.48	0.1127 (0.0163)	5×10^{-12}	45,943
rs113309773[b]	16:75432686-indel	*CFDP1- TMEM170A*	−/C	0.46	−0.1259 (0.0194)	9×10^{-11}	37,104

p = *p*-values of association from linear regression analysis, *N* = total number in meta-analyses
[a]Published cIMT SNP in LD with our most significant SNP: rs11781551 ($r^2 = 0.95$ with rs148147734), rs6601530 ($r^2 = 0$ with rs200482500), and rs445925 ($r^2 = 0.60$ with rs7412)
[b]Published plaque SNP in LD with our most significant SNP: rs1878406 ($r^2 = 0.98$ with rs11413744), rs17398575 ($r^2 = 0.8$ with rs17477177), rs9644862 ($r^2 = 0.79$ with rs9632884), and rs4888378 ($r^2 = 0.94$ with rs113309773)

Table 2 Gene expression results for significant SNPs in GTEx and STARNET tissues

SNP	eQTL[a] (Gene, *p*) GTEx		eQTL[a] (Gene, *p*) STARNET tissues	
	AOR[b]	HEART (ATR/VEN)[c]	AOR	MAM
rs201648240	*CAMK1G*, 0.0094 *AL031316.1*, 0.0040			*CD34*,0.00532 *TRAF3IP3*, 0.0097
rs6907215		*AL023584.1*, 0.005384704 (VEN)	*ENSG00000217648*, 0.00046	*ENSG00000217648*, 0.8×10^{-5}
rs13225723	*AC005050.1*, 1×10^{-10} *ENSG00000177820.5*, 7.0×10^{-5} *CCDC71L*, 5×10^{-6} *PRKAR2B*, 4×10^{-8} *PIK3CG*, 10×10^{-3}		*CCDC71L*, 6×10^{-36} *PRKAR2B*, 7×10^{-7} *SYPL1*, 0.0043	*CCDC71L*, 3×10^{-33} *PRKAR2B*, 6×10^{-8} *NAMPT*, 6×10^{-6}
rs2912063	*MCPH1*, 0.0041	*ENSG00000271743.1*, 0.0093 (VEN)	*MCPH1-AS1*, 0.0020	
rs11785239		*AC022784.1*, 0.0078 (VEN)	*ERI1*, 0.0069	*PPP1R3B*, 0.0036
rs844396	*ENSG00000141012.8*, 0.003 AC092384.2, 0.001 *CBFA2T3*, 1×10^{-7}	*ZNF469*, 0.004 (ATR) *AC092384.3*, 5×10^{-6} (ATR) *AC092384.1*, 0.002 (ATR) *CBFA2T3*, 0.0004 (ATR) *ZNF469*, 0.002 (VEN) *AC138028.4*, 0.001 (VEN) *ENSG00000224888.3*, 0.009 (VEN) *PIEZO1*, 0.0004 (VEN) *GALNS*, 0.004 (VEN)	*RPL13*, 0.0024 *ZNF276*, 0.0070 *TRAPPC2L*, 0.0091	*TRAPPC2L*, 0.0040 *ZNF276*, 0.0059
rs200495339		*ENSG00000267105.1*, 0.0005 (VEN)		
rs148147734	*DERL1*, 0.0082			
rs200482500	*AF131215.6*, 0.005 *AF131215.5*, 0.001	*AF131215.5*, 0.002 (ATR) *AF131215.6*, 0.003 (VEN) *AF131215.5*, 0.004 (VEN)		
rs7412	ENSG00000267163.1, 0.007			
rs11413744	*PRMT9*, 0.004			
rs17477177	*ENSG00000267052.1*, 6×10^{-11} *ENSG00000177820.5*, 5×10^{-6} *CCDC71L*, 4×10^{-7} *PRKAR2B*, 2×10^{-8}	*BCAP29*, 0.002 (ATR)	*CCDC71L*, 2×10^{-37} *PRKAR2B*, 6×10^{-7} *SYPL1*, 0.0091	*CCDC71L*,1×10^{-33} *PRKAR2B*, 2×10^{-8} *NAMPT*, 1×10^{-5}
rs9632884		*DMRTA1*, 0.007 (ATR)	*CDKN2B*, 2×10^{-3}	*CDKN2B*, 2×10^{-3}
rs113309773	*BCAR1*, 6×10^{-11} *ENSG00000261783.1*, 2×10^{-16} *GABARAPL2*, 0.004	*ENSG00000261783.1*, 1×10^{-5} (ATR) *ENSG00000166822.8*, 0.005 (ATR) *ENSG00000261783.1*, 0.0003 (VEN)	*ZFP1*, 4×10^{-4} AC009078.2, 0.002 *BCAR1*, 3×10^{-12} *CFDP1*, 0.002 *TMEM170A*, 0.009	

p = *p*-values of association from linear regression analysis
[a]The lead SNP from GWAS is considered an eQTL if the *cis*-association has a nominal *p*-value of association <0.01. Multiple but not all lead SNPs reach genome-wide significance ($p < 10^{-4}$).
[b]This includes aorta (AOR)
[c]This includes heart atrial (ATR) and heart left ventricle (VEN)

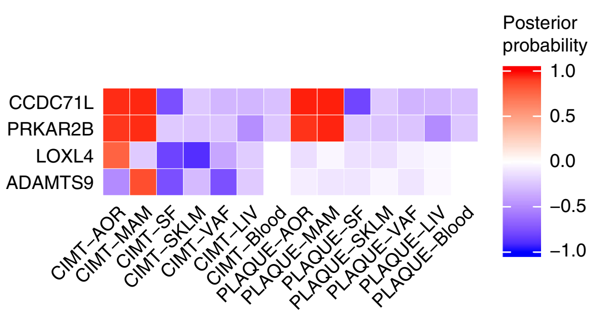


Fig. 2 Pairwise colocalization results for genes identified for cIMT and carotid plaque GWAS meta-analysis with STARNET expression datasets. Red indicates a high posterior probability of colocalization and blue a high probability of no colocalization of the same SNP with tissue eQTLs

eQTL-GWAS dataset pair (Supplementary Table 6). A posterior probability of ≥75% was considered strong evidence of the tissue-specific eQTL-GWAS pair influencing both the expression and GWAS trait at a particular region. Results for this analysis are shown in Table 3 and Supplementary Figure 5. The strongest evidence for an effect on gene expression within the regions identified in our standard GWAS meta-analysis was for the *CCDC71L* and *PRKAR2B* genes at the previously described chromosome 7 cIMT locus (*PIK3CG* in Table 2, Fig. 2). These genes showed evidence of colocalization for both cIMT and carotid plaque in AOR and MAM tissues (Table 3, Fig. 3). *CCDC71L* had the highest probability (>95%) for colocalization for cIMT, and MAM and AOR tissue eQTLs, and for carotid plaque, and MAM and AOR tissue eQTLs. We found a low probability of colocalization of the SNP with the *PIK3CG* gene expression (<1%).

Table 3 Colocalization of cIMT and plaque with eQTLs in tissues from patients with CHD in STARNET tissues for genes/tissues combinations that have more than 75% probability to share the same associated variant

Region (chr:start-stop)	Trait	Gene	SNP with best joint probability	*p*, BETA (SE), Tissue posterior probability (PPA)[a]			Direction of effect GWAS/eQTL
				cIMT /plaque GWAS	AOR eQTL	MAM eQTL	
chr3:63561280-65833136	cIMT	*ADAMTS9*	rs17676309 (T/C)	2×10^{-6}, -0.0035 (0.0007)	2×10^{-25}, −0.65 (0.06) PPA=0.93	1×10^{-23}, −0.61 (0.06) PPA=0.89	−/−
chr10:99017729-101017321	cIMT	*LOXL4*	rs55917128 (T/C)	5×10^{-7}, 0.0037 (0.0007)	6×10^{-8}, 0.33 (0.06) PPA=0.79		+/+
chr7:105299372-107743409	cIMT	*CCDC71L* *PRKAR2B*	rs12705390 (A/G)	5×10^{-9}, 0.0049 (0.0008)	2×10^{-37}, 0.81 (0.06) PPA=0.97 6×10^{-7}, 0.34 (0.07) PPA=0.93	1×10^{-33}, 0.755 (0.06) PPA=0.97 2×10^{-8}, 0.368 (0.06) PPA=0.96	+/+ +/+
	Plaque	*CCDC71L* *PRKAR2B*	rs12705390 (A/G)	4×10^{-8}, 0.12 (0.022)	2×10^{-37}, 0.80 (0.06) PPA=0.97 6×10^{-7}, 0.33 (0.07) PPA=0.93	1×10^{-33}, 0.75 (0.06) PPA=0.97 2×10^{-8}, 0.37 (0.06) PPA=0.96	+/+ +/+

PPA posterior probability of sharing same SNP higher than 75%, *cIMT* common carotid artery intima-media thickness, *AOR* aorta, *MAM* mammary artery
[a]This signal reaches genome-wide significance in cIMT/plaque, and reaches a high probability of being mediated by the genes in AOR and MAM

The eQTL associations at two additional loci (*ADAMTS9*, *LOXL4*) in MAM or AOR showed evidence of colocalization with cIMT or carotid plaque, although GWAS association *p*-values at these loci did not meet the genome-wide significance threshold (Table 3, Supplementary Figure 5). Albeit with weaker magnitudes, the expression of these two genes were also associated with the top colocalizing SNPs as detected in RNAseq data in GTEx aorta (rs17676309, chr3:64730121, *ADAMTS9*, $p = 0.0003$ and rs55917128, chr10:100023359, *LOXL4*, $p = 0.0005$).

Colocalization of CHD and stroke GWAS and STARNET eQTLs. We next assessed if the four genes (*CCDC71L, PRKAR2B, ADAMTS9, LOXL4)* identified through colocalization of cIMT/carotid plaque with tissue-specific eQTLs also showed evidence for colocalization with CHD and stroke traits (Supplementary Data 1 and Supplementary Figure 6). We used GWAS summary data for CHD (CARDIoGRAMPlusC4D), and stroke subtypes (MEGASTROKE) and AOR and MAM STARNET tissue eQTLs for these analyses. *CCDC71L* and *PRKAR2B* had suggestive evidence of sharing the same variant with large vessel disease stroke in both AOR and MAM tissues (probability of colocalization ≥20%, Supplementary Data 1). In contrast, there was strong evidence (≥75%) to reject a shared variant for CHD and eQTLs at this locus, thus suggesting there is atherosclerotic outcome specificity at vascular level for this locus (Supplementary Figure 5). Three of these genes, *CCDC71L*, *PRKAR2B*, and *ADAMTS9*, showed evidence for shared genetic influences of cIMT or carotid plaque on CHD/stroke outcomes when testing the joint association using moloc, a multiple-trait extension of coloc[14] (Supplementary Table 7). We also highlight the expression of *KIAA1462* gene in MAM, carotid plaque/cIMT, and CHD, which were positively correlated (Supplementary Figure 7). This gene has suggestive evidence of pairwise colocalization with carotid plaque (67% of probability of shared variant between carotid plaque and eQTL in MAM), as well as a high probability of shared variant between MAM eQTL expression of this gene, GWAS carotid plaque or cIMT, and CHD traits (Supplementary Table 7). We note, however, that the GWAS signal for outcomes across the datasets did not reach genome-wide significance and larger sample sizes may be needed to strengthen the evidence for involvement in disease outcomes.

Genetic correlations of cIMT/carotid plaque and clinical outcomes. To provide etiological insights into the role of measures of subclinical atherosclerosis and major atherosclerotic disease outcomes such as CHD and ischemic stroke, we quantified the genetic correlation using cross-trait LD score regression, a method that estimates genetic correlation across different traits using summary level data[15]. We used summary statistics between cIMT/carotid plaque with CHD and stroke meta-analysis of GWAS. Both cIMT and carotid plaque had positive significant genetic correlations with CHD (all $p < 0.05$ after adjusting for multiple testing), though the magnitude of the correlation was twice as strong for carotid plaque (0.52) as for cIMT (0.20) (Table 4). There was also evidence for genetic correlations between cIMT with any stroke and ischemic stroke subtype.

Pathway analysis and druggability. Gene Ontology (GO) analyses of genes identified in the loci for cIMT and carotid plaque according to our meta-analysis of GWAS (Table 1 and Supplementary Table 5) and in the colocalization analyses (Table 3, Supplementary Table 7) showed that cIMT genes are enriched in lipoprotein-related terms and cholesterol efflux, whereas carotid plaque genes are enriched in terms associated with fibroblast apoptosis (Supplementary Figure 8). Analysis of the cIMT genes using a GO Slim additionally identified several of the genes that were associated with terms describing cardiovascular development, cell adhesion, and immune processes, processes already considered relevant to atherosclerosis. Specifically, there is corroborating evidence from GO that *CCDC71L, PRKAR2B,* and *TWIST1* are associated with cIMT/carotid plaque as they are involved in lipid metabolism, with similar support that *ADAMTS9, CDH13,* and *KIAA1462* are associated with cIMT or carotid plaque risk as they are all involved in cell adhesion and, together with *TWIST1*, in cardiovascular system development (Supplementary Data 2).

From the loci associated with cIMT and carotid plaque, we identified seven genes (*ATG4B, ALPL, LDLR, APOB, EDNRA, APOE*, and *ADAMTS9*) whose encoded proteins are targets at various stages of the drug development process (Supplementary Tables 8 and 9). *ADAMTS9* gene encodes a protein likely to be druggable[16]. *ATG4B, ALPL,* and *LDLR* are proteins being targeted by compounds in pre-clinical phase (tier 2), while *APOB* and *EDNRA* are proteins targeted by drugs in clinical phase or licensed (tier 1). *APOB* is the target of an approved FDA drug for treatment of familial hypercholesterolemia. *EDNRA* gene encodes for endothelin A receptor, against which several antagonists have been developed for the treatment of pulmonary arterial

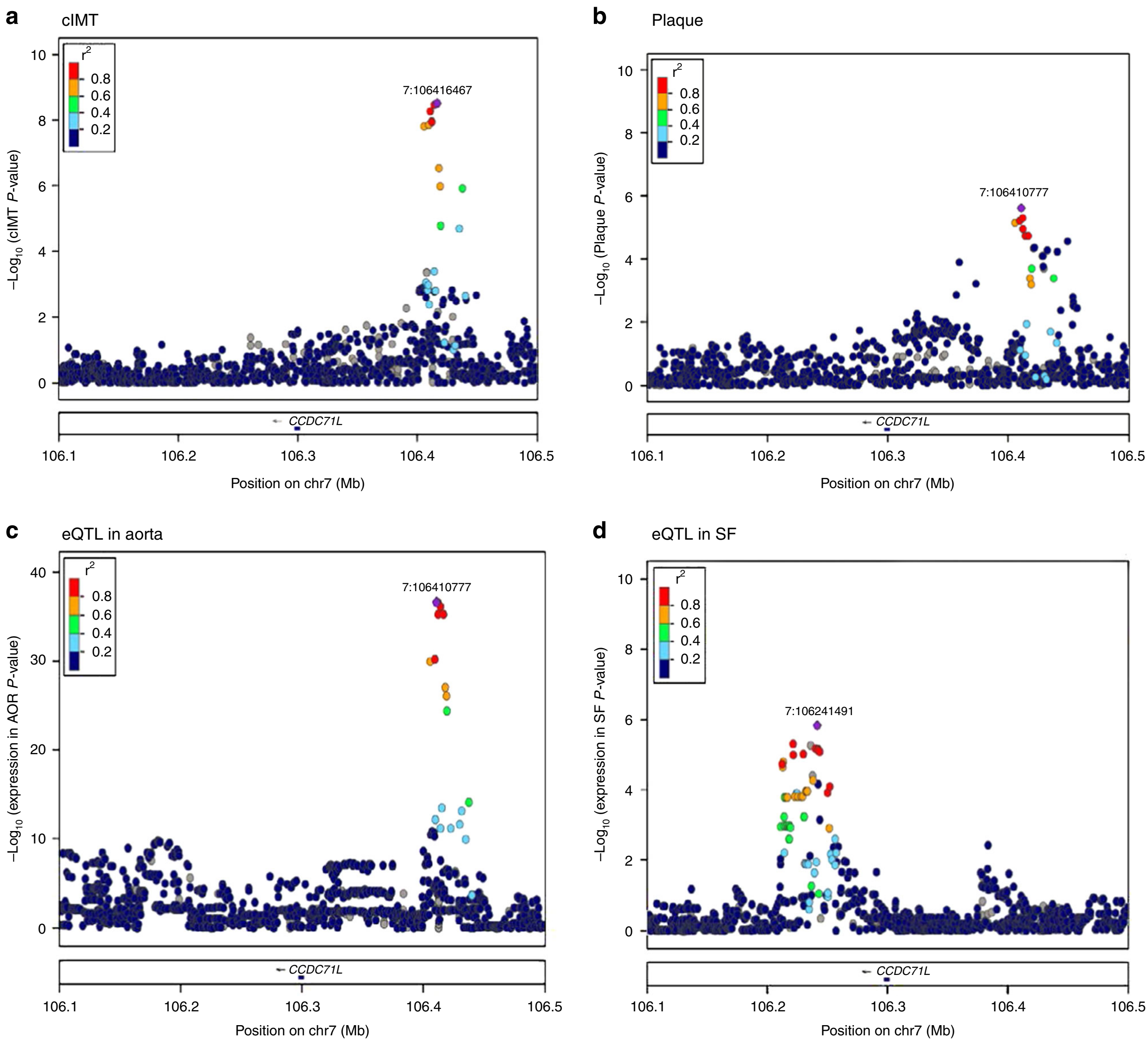


Fig. 3 Association results at the *CCDC71L* locus (chromosome 7), showing a high posterior probability of a shared variant for cIMT and carotid plaque in AOR and MAM eQTLs. −log10(*p*) SNP association *p*-values for cIMT (plot A) and carotid plaque (plot B), and eQTL in AOR (plot C) and eQTL in SF (plot D). Association results in SF tissue have a low probability of a shared signal with cIMT and carotid plaque, possibly indicating a different mechanism in this tissue. eQTLs in MAM are identical to AOR and not shown. The *p*-values were calculated by fitting a linear regression model with cIMT or plaque as dependent variable and imputed SNPs as independent variables. Each dot is an SNP and the color indicates linkage disequilibrium (r^2) with the best hit (in purple)

hypertension or which are in advanced clinical phase development for non-small cell lung cancer and diabetic nephropathy.

Discussion

We provide results of a large meta-analysis of GWAS of subclinical atherosclerosis and we integrate our results with tissue-specific gene expression data using eQTLs from both the early (MAM) and late advanced (AOR) atherosclerotic arterial wall from the STARNET study to enable reliable discovery of genes with biological evidence of an increased probability for conferring inherited risk of atherosclerosis development. Our discovery approach using GWAS meta-analyses identified 16 loci significantly associated with either cIMT or carotid plaque, of which nine are novel.

The integration of GWAS and tissue-specific *cis*-eQTLs for the joint analyses of tissue-specific eQTLs from CHD patients identified two potentially additional loci colocalizing with cIMT or carotid plaque: chr3:63561280-65833136 (*ADAMTS9*), chr10:99017729-101017321 (*LOXL4*). *ADAMTS9* is a metalloproteinase involved in thrombosis and angiogenesis and has been associated with cardiometabolic traits (waist-to-hip ratio, waist circumference, and type 2 diabetes) in GWAS, and with coronary artery calcification in a gene-by-smoking interaction GWAS[17,18]. *LOXL4* encodes a lysyl oxidase involved in crosslinks of collagen and elastin in the extracellular matrix. This family of proteins are involved in the development of elastic vessels and mechanical strength of the vessel wall, and their inhibition was associated with the development of abdominal aortic aneurysms and more severe atherosclerosis in experimental models[19].

Table 4 Genetic correlation between CHD and stroke traits with cIMT and plaque, and cIMT with plaque using LD score and meta-GWAS

Cardiovascular disease trait	Subclinical atherosclerosis trait	Genetic correlation	SE	z	p
CHD[a]	cIMT	0.20	0.05	4.1114	4×10^{-5}
Any stroke	cIMT	0.30	0.07	4.2301	2.3×10^{-5}
Ischemic stroke[b]	cIMT	0.31	0.07	4.646	3.4×10^{-6}
Cardio-embolic stroke[b]	cIMT	0.10	0.09	1.0729	0.28
Small vessel disease stroke[b]	cIMT	0.33	0.18	1.8728	0.06
CHD[a]	Carotid plaque	0.52	0.08	6.4263	1.3×10^{-10}
Any stroke[b]	Carotid plaque	0.28	0.10	2.7097	0.007
Ischemic stroke[b]	Carotid plaque	0.27	0.10	2.6578	0.008
Cardio-embolic stroke[b]	Carotid plaque	0.06	0.14	0.4684	0.64
Small vessel disease stroke[b]	Carotid plaque	−0.03	0.24	−0.1344	0.89
Plaque	cIMT	0.40	0.10	3.9667	7.3×10^{-5}

[a]CARDIoGRAMPlusC4D
[b]MEGASTROKE consortium. Unable to estimate the genetic correlations with large vessel disease

Some loci identified in our meta-analysis of GWAS include genes in known pathways for atherosclerosis, including *LDLR*, which is related to lipid pathways and CHD, and identified for associations with carotid plaque in our study. For most of the loci, however, the underlying gene implicated in signals are unknown. Our colocalization approach found both *CCDC71L* and *PRKAR2B* as the most likely genes at the chromosome 7 locus, where *PIK3CG* was previously the suggested gene. This finding is in agreement with a targeted sequencing study of subclinical atherosclerosis[15]. An additional SNP (rs342286) at this locus has been associated with platelets volume and reactivity, and cardiovascular traits. However, rs342286 is not in LD with our most significant SNP and it is not associated with cIMT or carotid plaque in our studies ($p = 0.49$ and 0.01, respectively). Of interest, the variant we identified in this study showed evidence for colocalization with cIMT/carotid plaque and large vessel disease stroke but not CHD, therefore showing tissue and outcome-specificity. *CCDC71L* has unknown function. *PRKAR2B* codes for one of the several regulatory subunits of cAMP-dependent protein kinase and its expression is ubiquitous. In vitro studies have shown that adenosine-induced apoptosis of arterial smooth muscle cells involves a cAMP-dependent pathway[20].

Measures of cIMT and carotid plaque reflect vascular pathophysiologic and atherosclerosis processes, respectively, with carotid plaque more strongly reflecting atherosclerotic clinical events. An important contribution of this study is the supporting evidence for overall genetic correlations of CHD and stroke (any cause and ischemic stroke) with subclinical atherosclerosis traits, estimated using LD score methods. Further highlighting the potential biological relevance of our findings, the genetic correlations estimates for CHD were stronger for carotid plaque than for cIMT. However, cIMT and carotid plaque GWAS were correlated, and the genetic correlations estimates with stroke were similar for cIMT or carotid plaque, and not significant for carotid plaque. The colocalization analyses provided additional insights in the relationships between subclinical atherosclerosis, clinical outcomes, and tissue-specific regulation at specific genomic regions. For example, our suggestive top gene association in multi-trait colocalization for *KIAA1462* included MAM eQTLs, carotid plaque, and CHD, supporting the shared genetic effects at this locus of atherosclerosis in carotid and coronary arteries. *KIAA1462* has been previously reported in the same locus identified by GWAS for CHD[21]. This gene encodes a protein involved in cell–cell junctions in endothelial cells[22], which was recently shown to be involved in pathologic angiogenic process in in vitro and in vivo experimental models[23]. These findings suggest that there may be important differences in vascular bed regulation at distinctive regions for atherosclerotic cardiovascular and stroke outcomes that may help to identify genes and specific targets for CHD or stroke prevention and treatment.

Additional studies in diverse and large samples across the multiple datasets are needed to explore these results further. As more summary statistics become available for other clinical endpoints beyond stroke and CHD (both in terms of larger sample size and richer genome coverage), and as further refinements in clinical phenotypes emerge (e.g. from CHD to acute coronary syndrome sub-components), strategies to integrate this knowledge using methods such as *moloc*[10] and *eCAVIAR*[24] will continue to be essential for harnessing genome-wide findings in the drug-discovery process.

In summary, our study is a large GWAS meta-analysis of cIMT and carotid plaque. Through a sequential approach of discovery and colocalization studies, we provide deeper insights into disease causal genes of subclinical cIMT and carotid plaque formation. We confirmed three loci and identified nine novel loci in the meta-analyses of cIMT and carotid plaque. Additionally, we provide strong evidence for the role of three novel genes from our integrative analysis of GWAS and eQTL data. Moreover, the identified correlations with CHD and stroke highlight novel biological pathways that merit further assessments as novel targets for drug development.

Methods

Ethics statement. All human research was approved by the relevant institutional review boards for each study, and conducted according to the Declaration of Helsinki. All participants provided written informed consent.

Populations and phenotypes. The discovery GWAS in this study consists of a collaboration between the CHARGE [8] and the UCLEB consortia[9], for genetic studies of cIMT and carotid plaque among individuals of European ancestry (Supplementary Note 1). All studies followed standardized protocols for phenotype ascertainment and statistical analyses. The descriptive characteristics of participating studies are shown in Supplementary Table 1.

cIMT and carotid plaque measures were evaluated using high-resolution B-mode ultrasonography and reading protocols as previously reported[4]. We used data from the baseline examination or the first examination in which carotid ultrasonography was obtained. cIMT was defined by the mean of the maximum of several common carotid artery measurements, measured at the far wall or the near wall. For most studies, this was an average of multiple measurements from both the left and right arteries. We also examined a carotid plaque phenotype, defined by atherosclerotic thickening of the carotid artery wall or the proxy measure of luminal stenosis greater than 25% (Supplementary Table 2).

Genotyping, imputation, and study-level quality control. Genotyping arrays and QC pre-imputation are shown in Supplementary Table 3. Each GWAS study conducted genome-wide imputation using a Phase 1 integrated (March 2012 release) reference panel from the 1000G Consortium using IMPUTE2[25] or MaCH/minimac[26], and used Human Reference Genome Build 37. Sample QC was performed with exclusions based on call rates, extreme heterozygosity, sex discordance, cryptic relatedness, and outlying ethnicity. SNP QC excluded variants based on call rates across samples and extreme deviation from Hardy–Weinberg equilibrium (Supplementary Table 3). Non-autosomal SNPs were excluded from imputation and association analysis.

Pre-meta-analysis GWAS study-level QC was performed using EasyQC software[27]. This QC excluded markers absent in the 1000G reference panel; non A/C/G/T/D/I markers; duplicate markers with low call rate; monomorphic SNPs and those with missing values in alleles, allele frequency, and beta estimates; SNPs with large effect estimates or standard error (SE) ≥10; and SNPs with allele frequency difference >0.3 compared to 1000G reference panel. There was a total of 9,574,088 SNPs for the cIMT meta-analysis and 8,578,107 SNPs for the carotid plaque meta-analysis.

Statistical analyses. Within each study, we used linear and logistic regression to model cIMT and carotid plaque, respectively, and an additive genetic model (SNP dosage) adjusted for age, sex, and up to 10 principal components. We combined summary estimates from each study and each trait using an inverse variance weighted meta-analysis. Additional filters were applied during meta-analyses including imputation quality (MACH $r^2 < 0.3$ and IMPUTE info <0.4), a minor allele frequency (MAF) <0.01, and SNPs that were not present in at least four studies. The genome-wide significance threshold was considered at $p < 5.0 \times 10^{-8}$.

To assess the evidence for independent associations at each locus attaining genome-wide significance, we performed conditional analysis in a 1-Mb genomic interval flanking the lead SNP using GCTA[28]. This approach uses summary meta-analysis statistics and a LD matrix from an ancestry-matched sample to perform approximate conditional SNP association analysis. The estimated LD matrix was based on 9713 unrelated individuals of European ancestry from the ARIC study, which was genotyped using an Affymetrix 6.0 array and imputed to the 1000G panel using IMPUTE2[25].

Gene expression analysis using GTEx. GTEx Analysis V6 (dbGaP Accession phs000424.v6.p1) eQTL results were downloaded from GTEx portal for 44 tissues, and then mapped to SNPs listed in Table 1. We used a false discovery rate (FDR) of ≤0.05.

Colocalization analyses using eQTLs. We integrated our GWAS results with *cis*-eQTL data using a Bayesian method (coloc)[10]. This method evaluates whether the GWAS and eQTL associations best fit a model in which the associations are due to a single shared variant (summarized by the posterior probability). We used gene expression datasets from multiple tissues from patients with CHD of the STARNET study, including blood, MAM, AOR, subcutaneous fat (SF), visceral fat (VAF), skeletal muscle (SKLM), and liver (LIV) obtained from 600 patients during open heart surgery[11]. Pairwise colocalization was tested between these expression disease tissue datasets and GWAS results from our cIMT/carotid plaque GWAS meta-analysis. We used GWAS and eQTL summary statistics of SNPs within a 200-kb window around each gene covered by the eQTL datasets. A posterior probability of colocalization ≥0.75 was considered a strong evidence for a causal gene. Next, we reported the gene(s) in the STARNET datasets that had the strongest evidence of sharing the same variant with cIMT or carotid plaque genome-wide. In an alternative analysis, we also tested loci with an SNP that reached a threshold of significant or suggestive genome-wide significance for cIMT or carotid plaque (reported in Table 1, Supplementary Table 5). For each region 200kb around the SNP with the lowest association p-value, we report the gene with the highest probability of being responsible for the GWAS signal (Supplementary Table 6).

Pairwise colocalization for these genes was also tested for publicly available GWAS for CHD case-controls (CARDIoGRAMPlusC4D) and stroke case-controls (MEGASTROKE consortium). The MEGASTROKE dataset uses genotypes imputed to the 1000G phase I haplotype panel. The European ancestry sample used to generate these results consisted of 40,585 stroke cases and 406,111 controls from 15 cohorts and two consortia: the METASTROKE and CHARGE consortia[29]. The phenotypes used in this analysis were any stroke ($n = 39,067$ cases, total $n = 442,142$), ischemic stroke (IS, $n = 32,686$ cases, total $n = 423,266$), and etiologic stroke subtypes:cardioembolic stroke (CE, $n = 6,820$ cases, total $n = 314,368$), large vessel disease ($n = 4,113$, total $n = 202,263$), and small vessel disease (SVD, $n = 4,975$, total $n = 242,250$). To explore multi-trait colocalizations, we used moloc[14] with prior probabilities of 10^{-4} for GWAS/GWAS/eQTL, 10^{-6} for GWAS+eQTL/GWAS or GWAS+GWAS/eQTL, and 10^{-7} for colocalization of all three association signals.

Functional annotation and epigenetic enrichment analyses. From the Epigenome Roadmap Project[30,31], we obtained regulatory information using broad classes of chromatin states ($n = 127$ tissues) capturing promoter-associated, transcription-associated, active intergenic, and large-scale repressed and repeat-associated states. From ENCODE[32], we obtained chromatin states, uniformly processed transcription factor (TF) Chip assays and DNaseI Hypersensitivity sites (DHS) for nine cells lines. From FANTOM5[33], we used information from expression of enhancers in each tissue ($n = 112$), and enhancers that are positively differentially expressed against any other tissue ($n = 110$).

We used fGWAS[13] to identify genomic annotations that are enriched within the cIMT results and to select the variants with support for a functional role based on the most informative annotations. We only considered cIMT for these analyses because of the small number of identified loci for carotid plaque. We first estimated the enrichment parameters for each annotation individually and identified the set of annotations with significant marginal associations. We then applied 10-fold cross-validation likelihood and forward selection to identify the set of annotations that significantly improve the model fit, and reverse selection of each annotation included in the model, as suggested in the fGWAS workflow. We reported the model with the highest cross-validation likelihood and SNPs that have regional posterior probability of association (PPA) >0.9 and directly overlap the genomic annotations considered.

Overall genetic correlation analysis. Genetic correlation between cIMT/carotid plaque, CHD, and stroke traits were calculated using LD score regression approach LD-score, which uses GWAS summary statistics and is not affected by sample overlap. This method relies on the fact that the χ^2 association statistic for a given SNP includes the effects of all SNPs that are in LD with it and it calculates genetic correlation by partitioning the SNP heritabilities[15]. Genetic correlations between stroke traits (IS, CE, large vessel disease, and SVD) and cIMT and carotid plaque were calculated using software available at http://github.com/bulik/ldsc with GWAS summary statistics for our cIMT/carotid plaque GWAS, CARDIO-GRAMPlusC4D data, and stroke GWAS. We used the LD-scores[15], which are based on the 1000 Genomes European population and estimated within 1-cM windows. Based on ten tests performed (two subclinical traits and five outcomes), we set the significance threshold to $p = 0.005$.

PATHWAY ANALYSES. Methods for GO Slim: The Ensembl identifiers of all protein-coding genes identified as in LD with the 12 variants for cIMT and 15 variants for carotid plaque (including variants from main and suggestive signals, Table 1 and Supplementary Table 5), and five genes for which there is strong evidence of colocalization (Table 3), were mapped to UniProt accession numbers, using the UniProt ID mapping service (http://www.uniprot.org/uploadlists/). A GO Slim analysis was performed on this list using QuickGO (www.ebi.ac.uk/QuickGO) and the Generic GO Slim. The GO terms used in the final slim analysis were further refined by adding/removing GO terms to provide more detailed information about the processes covered.

Methods for GO term enrichment analysis: The VLAD gene list analysis and visualization tool (http://proto.informatics.jax.org/prototypes/vlad/) was used to perform a GO term enrichment analysis on the same UniProt accessions as listed for the GO Slim. The background annotation set was obtained from the goa_human.gaf file (dated 21 November 2017, downloaded from ftp://ftp.ebi.ac.uk/pub/databases/GO/goa/HUMAN/) and the ontology data was obtained from the go-basic.obo file provided in the VLAD tool (analysis run 28 November 2017).

The LD block around top SNPs associated with cIMT and carotid plaque was constructed using LD information from the 1000 Genomes panel, as previously outlined in Finan et al.[16]. Briefly, the boundaries of the LD region were defined as the positions of the variants furthest upstream and downstream of a GWAS SNP with an r^2 value of ≥0.5 and within a 1-Mbp flank on either side of the GWAS variant. Associated variants that were not present in the 1000 Genomes panel that were not in LD with any other variants were given a nominal flank of 2.5 kbp on either side of the association. Gene annotations using Ensembl version 79 were then overlapped to the LD region.

Druggable genes. We examined the druggability status for the nearest coding genes identified in our GWAS analysis on cIMT and carotid plaque, including significant (novel and replicated) and suggestive ones, as well as genes identified through colocalization analysis. The druggable gene set was calculated using the previously described criteria: novel targets of first-in-class drugs licensed since 2005; the targets of drugs currently in late phase clinical development; pre-clinical phase small molecules with protein binding measurements reported in the ChEMBL database; and genes encoding secreted or plasma membrane proteins that are targets of monoclonal antibodies and other bio-therapeutics[16]. We defined three tiers of druggable gene sets based on their drug development. In Tier 1, 1427 genes were targets of approved small molecules and biotherapeutic drugs and clinical-phase drug candidates. Tier 2 comprised 682 genes encoding targets with known bioactive drug-like small molecule binding partners and those with significant sequence similarity to approved drug targets. Tier 3 contained 2370 genes encoding secreted or extracellular proteins, proteins with more distant similarity to approved drug targets, and druggable genes not included in Tier 1 or 2 such as GPCRs, nuclear hormone receptors, ion channels, kinases, and phosphodiesterases.

URLs. For GTEx, see http://gtexportal.org/. For Coloc, see https://cran.r-project.org/web/packages/coloc/coloc.pdf. For, Moloc, see https://github.com/clagiamba/

moloc/blob/master/man/moloc-package.Rd. For CARDIoGRAMPlusC4D, see www.cardiogramplusc4d.org/. For LD scores, www.broadinstitute.org/~bulik/eur_ldscores/. For UniProt ID, www.uniprot.org/uploadlists/. For QuickGO, www.ebi.ac.uk/QuickGO. For VLAD tool, see http://proto.informatics.jax.org/prototypes/vlad/.

Data availability

All relevant summary statistics data that support the findings of this study have been deposit in the database of Genotypes and Phenotypes (dbGaP) under the CHARGE acquisition number (https://www.ncbi.nlm.nih.gov/projects/gap/cgi-bin/study.cgi?study_id=phs000930.v6.p1; accession phs000930.v6.p1). GWAS data for most US studies are already available in dbGAP.

Received: 19 December 2017 Accepted: 24 September 2018
Published online: 03 December 2018

References

1. Mozaffarian, D. et al. Heart Disease and Stroke Statistics-2016 Update: a report from the American Heart Association. *Circulation* **133**, e38–e360 (2016).
2. Frieden, T. R. & Berwick, D. M. The Million Hearts initiative--preventing heart attacks and strokes. *N. Engl. J. Med.* **365**, e27 (2011).
3. O'Donnell, C. J. & Nabel, E. G. Genomics of cardiovascular disease. *N. Engl. J. Med.* **365**, 2098–2109 (2011).
4. Bis, J. C. et al. Meta-analysis of genome-wide association studies from the CHARGE consortium identifies common variants associated with carotid intima media thickness and plaque. *Nat. Genet.* **43**, 940–947 (2011).
5. Natarajan, P. et al. Multiethnic Exome-Wide Association Study of Subclinical Atherosclerosis. *Circ. Cardiovasc. Genet.* **9**, 511–520 (2016).
6. Pott, J. et al. Genome-wide meta-analysis identifies novel loci of plaque burden in carotid artery. *Atherosclerosis* **259**, 32–40 (2017).
7. Gertow, K. et al. Identification of the BCAR1-CFDP1-TMEM170A locus as a determinant of carotid intima-media thickness and coronary artery disease risk. *Circ. Cardiovasc. Genet.* **5**, 656–665 (2012).
8. 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).
9. Shah, S. et al. Causal relevance of blood lipid fractions in the development of carotid atherosclerosis: Mendelian randomization analysis. *Circ. Cardiovasc. Genet.* **6**, 63–72 (2013).
10. Giambartolomei, C. et al. Bayesian test for colocalisation between pairs of genetic association studies using summary statistics. *PLoS Genet.* **10**, e1004383 (2014).
11. Franzen, O. et al. Cardiometabolic risk loci share downstream cis- and trans-gene regulation across tissues and diseases. *Science* **353**, 827–830 (2016).
12. Pfeiffer, R. M., Mitchell, H. G. & Pee, D. On combining data from genome-wide associtio studies to discover disease-associated SNPs. *Stat. Sci.* **24**, 547–560 (2009).
13. Pickrell, J. K. Joint analysis of functional genomic data and genome-wide association studies of 18 human traits. *Am. J. Hum. Genet.* **94**, 559–573 (2014).
14. Giambartolomei, C. et al. A Bayesian framework for multiple trait colocalization from summary association statistics. *Bioinformatics* **34**, 2538–2545 (2018).
15. Finucane, H. K. et al. Partitioning heritability by functional annotation using genome-wide association summary statistics. *Nat. Genet.* **47**, 1228–1235 (2015).
16. Finan, C. et al. The druggable genome and support for target identification and validation in drug development. *Sci. Transl. Med.* **9**, eaag1166 (2017).
17. Zeggini, E. et al. Meta-analysis of genome-wide association data and large-scale replication identifies additional susceptibility loci for type 2 diabetes. *Nat. Genet.* **40**, 638–645 (2008).
18. Polfus, L. M. et al. Genome-wide association study of gene by smoking interactions in coronary artery calcification. *PLoS ONE* **8**, e74642 (2013).
19. Remus, E. W. et al. The role of lysyl oxidase family members in the stabilization of abdominal aortic aneurysms. *Am. J. Physiol. Heart Circ. Physiol.* **303**, H1067–H1075 (2012).
20. Peyot, M. L. et al. Extracellular adenosine induces apoptosis of human arterial smooth muscle cells via A(2b)-purinoceptor. *Circ. Res.* **86**, 76–85 (2000).
21. Erdmann, J. et al. Genome-wide association study identifies a new locus for coronary artery disease on chromosome 10p11.23. *Eur. Heart J.* **32**, 158–168 (2011).
22. Akashi, M., Higashi, T., Masuda, S., Komori, T. & Furuse, M. A coronary artery disease-associated gene product, JCAD/KIAA1462, is a novel component of endothelial cell-cell junctions. *Biochem. Biophys. Res. Commun.* **413**, 224–229 (2011).
23. Hara, T. et al. Targeted disruption of JCAD (Junctional Protein Associated With Coronary Artery Disease)/KIAA1462, a coronary artery disease-associated gene product, inhibits angiogenic processes in vitro and in vivo. *Arterioscler. Thromb. Vasc. Biol.* **37**, 1667–1673 (2017).
24. Hormozdiari, F., Kostem, E., Kang, E. Y., Pasaniuc, B. & Eskin, E. Identifying causal variants at loci with multiple signals of association. *Genetics* **198**, 497–508 (2014).
25. Howie, B. N., Donnelly, P. & Marchini, J. A flexible and accurate genotype imputation method for the next generation of genome-wide association studies. *PLoS Genet.* **5**, e1000529 (2009).
26. Howie, B., Fuchsberger, C., Stephens, M., Marchini, J. & Abecasis, G. R. Fast and accurate genotype imputation in genome-wide association studies through pre-phasing. *Nat. Genet.* **44**, 955–959 (2012).
27. Winkler, T. W. et al. Quality control and conduct of genome-wide association meta-analyses. *Nat. Protoc.* **9**, 1192–1212 (2014).
28. Yang, J., Lee, S. H., Goddard, M. E. & Visscher, P. M. GCTA: a tool for genome-wide complex trait analysis. *Am. J. Hum. Genet.* **88**, 76–82 (2011).
29. Malik, R. et al. Multiancestry genome-wide association study of 520,000 subjects identifies 32 loci associated with stroke and stroke subtypes. *Nat. Genet.* **50**, 524–537 (2018).
30. Roadmap Epigenomics, Consortium et al. Integrative analysis of 111 reference human epigenomes. *Nature* **518**, 317–330 (2015).
31. Zhu, L. J. Integrative analysis of ChIP-chip and ChIP-seq dataset. *Methods Mol. Biol.* **1067**, 105–124 (2013).
32. ENCODE Project Consortium. An integrated encyclopedia of DNA elements in the human genome. *Nature* **489**, 57–74 (2012).
33. Andersson, R. et al. An atlas of active enhancers across human cell types and tissues. *Nature* **507**, 455–461 (2014).

Acknowledgements

The work was supported by the following grants: National Institute of Health grants: R21HL123677, R21-HL140385, DK104806-01A1, R01-MD012765-01A1 (NF), National Institutes of Health awards R01HG009120, R01HG006399, U01CA194393, T32NS048004 (CG), the American Heart Association Grant #17POST33350042 (PV), the British Heart Foundation (RG/13/5/30112) and the National Institute for Health Research University College London Hospitals Biomedical Research Centre (RCL and RPH), the British Heart Foundation FS/14/55/30806 (JCH), the German Federal Ministry of Education and Research (BMBF) in the context of the e:Med program (e:AtheroSysMed), the DFG as part of the CRC 1123 (B3), and the FP7/2007-2103 European Union project CVgenes@target (grant agreement number Health-F2-2013-601456). We thank Li-Ming Gan for assistance with the STARNET study and Jon White for assistance with UCLEB analyses. Additional acknowledgements are included in Supplementary Note 2.

Author contributions

N.F., C.G., J.C.B., M.K., C.D., M.S., S.K.M., U.S., W.P., A.B.Z., A.H., A.T., A.G.U., A.B.N., B.W.P., C.D.L., E.T., F.R., H.V., I.J.D., L.J.L., M.D., O.R., O.H.F., R.S., R.M., T.B.H., T.L., U.F., W.P., A.D., A.S., A.H., C.M.D., D.A.L., D.O.M.-K., D.W.B., H.S., J.F.W., J.G.W., J.I.R., J.M.W., M.L., M.K.E., S.E.H., U.V., V.G., A.D.H., J.P.C., and C.J.O. contributed to study concept and design. C.D., M.S., S.K.M., U.S., W.P., A.I., A.C., A.B.Z., A.T., A.G.U., A.W., A.J.S., A.B., A.R., B.H., B.W.P., C.F., D.B., D.H.O., D.T., D.K., E.T., E.M., E.G., E.B., E.E.S., E.I., F.R., F.B., G.H., H.C., H.V., H.S.M., I.J.D., J.W.J., J.G., J.P., J.T., J.E., K.D.T., L.K., L.L., L.J.L., L.H., M.D., M.S., M.K., M.K., M.A.N., O.M., O.R., P.H.W., P.G., P.A., R.R., R.B., R.H., R.S., R.M., R.W.M.,. S.G.W., S.M.L., S.T., S.K., S.R.H., S.R., T.B.H., T.L., T.G., T.S., U.F., V.P., W.R., W.P., X.Z., A.S., A.H., B.M.P., C.M.D., D.A.L., D.O.M.-K., D.W.B., H.S., J.F.W., J.G.W., J.I.R., J.C.H., J.M.W., J.D., J.H., M.K.E., S.E.H., U.V., V.G., A.D.H., J.L.M.B., J.P.C., and C.J.O. contributed to acquisition of genotyping or phenotypic data. N.F., C.G., C.F., J.C.B., R.P.H., R.C.L., S.M.T., T.W.W., M.G., M.K., C.D., A.V.S., E.H., E.M.L., I.M.N., L.L., M.S., M.S., N.P., O.F., P.K.J., R.N., R.E.M., S.-J.H., S.K.M., U.S., A.I., A.T., K.R., A.J.C., B.S., C.D.L., C.W., F.V., G.C., H.S., J.P., J.L., K.G., L.M.R., M.T., M.A.N., O.M., P.R., P.A., Q.W., R.J.S., S.H., S.S., S.M.L., T.S., X.Z., X.Z., X.G., Y.S., and L.-.P.L. contributed to statistical analysis and interpretation of the data. N.F., C.G., P.S.V., J.C.B., M.K., S.K.M., A.D.H., and J.P.C. contributed to drafting of the manuscript. All authors contributed to the critical revision of the manuscript.

Additional information

Supplementary Information accompanies this paper at https://doi.org/10.1038/s41467-018-07340-5.

Competing interests: C.F. received a fee for speaking at a course by Springer Healthcare/Malesci. E.I. is a scientific advisor for Precision Wellness, Cellink and Olink Proteomics for work unrelated to the present project. M.A.N.'s participation in this project was supported by a consulting contract between Data Tecnica International and the National Institute on Aging, NIH, Bethesda, MD, USA. M.A.N. also consults for Illumina Inc., the Michael J. Fox Foundation, and University of California Healthcare. B.M.P. serves on the DSMB of a clinical trial funded by Zoll LifeCor and on the Steering Committee of the Yale Open Data Access Project funded by Johnson & Johnson. D.A.L. has received support from Roche Diagnostics and Medtronic for biomarker research unrelated to this paper. J.P.C. has received funding from GSK regarding methodological work around electronic health records, and -omics for drug discovery. All remaining authors declare no competing interests.

Reprints and permission information is available online at http://npg.nature.com/reprintsandpermissions/

Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit http://creativecommons.org/licenses/by/4.0/.

© The Author(s) 2018

Nora Franceschini[1], Claudia Giambartolomei[2], Paul S. de Vries[3], Chris Finan[4], Joshua C. Bis[5], Rachael P. Huntley[4], Ruth C. Lovering[4], Salman M. Tajuddin[6], Thomas W. Winkler[7], Misa Graff[1], Maryam Kavousi[8], Caroline Dale[9], Albert V. Smith[10,11], Edith Hofer[12,13], Elisabeth M. van Leeuwen[8], Ilja M. Nolte[14], Lingyi Lu[15], Markus Scholz[16,17], Muralidharan Sargurupremraj[18], Niina Pitkänen[19], Oscar Franzén[20,21], Peter K. Joshi[22], Raymond Noordam[23], Riccardo E. Marioni[24,25], Shih-Jen Hwang[26,27], Solomon K. Musani[28], Ulf Schminke[29], Walter Palmas[30], Aaron Isaacs[8,31], Adolfo Correa[28], Alan B. Zonderman[6], Albert Hofman[8,32], Alexander Teumer[33,34], Amanda J. Cox[35,36], André G. Uitterlinden[8,37], Andrew Wong[38], Andries J. Smit[39], Anne B. Newman[40], Annie Britton[41], Arno Ruusalepp[21,42,43], Bengt Sennblad[44,45], Bo Hedblad[46], Bogdan Pasaniuc[2,47], Brenda W. Penninx[48], Carl D. Langefeld[15], Christina L. Wassel[49], Christophe Tzourio[18], Cristiano Fava[46,50], Damiano Baldassarre[51,52], Daniel H. O'Leary[53], Daniel Teupser[17,54], Diana Kuh[38], Elena Tremoli[52,55], Elmo Mannarino[56], Enzo Grossi[57], Eric Boerwinkle[3,58], Eric E. Schadt[20,21], Erik Ingelsson[59,60,61], Fabrizio Veglia[52], Fernando Rivadeneira[8,37], Frank Beutner[62], Ganesh Chauhan[18,63], Gerardo Heiss[1], Harold Snieder[14], Harry Campbell[22], Henry Völzke[33,34], Hugh S. Markus[64], Ian J. Deary[24,65], J. Wouter Jukema[66], Jacqueline de Graaf[67], Jacqueline Price[22], Janne Pott[16,17], Jemma C. Hopewell[68], Jingjing Liang[69], Joachim Thiery[17,70], Jorgen Engmann[4], Karl Gertow[44], Kenneth Rice[71], Kent D. Taylor[72], Klodian Dhana[73], Lambertus A.L.M. Kiemeney[74], Lars Lind[75], Laura M. Raffield[76], Lenore J. Launer[6], Lesca M. Holdt[17,54], Marcus Dörr[34,77], Martin Dichgans[78,79], Matthew Traylor[64], Matthias Sitzer[80], Meena Kumari[41,81], Mika Kivimaki[41], Mike A. Nalls[82,83], Olle Melander[46], Olli Raitakari[19,84], Oscar H. Franco[8,85], Oscar L. Rueda-Ochoa[8,86], Panos Roussos[20,87,88], Peter H. Whincup[89], Philippe Amouyel[90,91,92], Philippe Giral[93], Pramod Anugu[28], Quenna Wong[94], Rainer Malik[78], Rainer Rauramaa[95,96], Ralph Burkhardt[17,97,98], Rebecca Hardy[38], Reinhold Schmidt[12], Renée de Mutsert[99], Richard W. Morris[100], Rona J. Strawbridge[44,101], S. Goya Wannamethee[102], Sara Hägg[103], Sonia Shah[4], Stela McLachlan[22], Stella Trompet[23,66], Sudha Seshadri[104], Sudhir Kurl[105], Susan R. Heckbert[5,106], Susan Ring[107,108], Tamara B. Harris[6], Terho Lehtimäki[109,110], Tessel E. Galesloot[74], Tina Shah[4], Ulf de Faire[111,112], Vincent Plagnol[113], Wayne D. Rosamond[1], Wendy Post[114], Xiaofeng Zhu[69], Xiaoling Zhang[27,115], Xiuqing Guo[72,116], Yasaman Saba[117], MEGASTROKE Consortium, Abbas Dehghan[8,118], Adrie Seldenrijk[119], Alanna C. Morrison[3], Anders Hamsten[44], Bruce M. Psaty[106,120], Cornelia M. van Duijn[8,68], Deborah A. Lawlor[107,108], Dennis O. Mook-Kanamori[99,121], Donald W. Bowden[122], Helena Schmidt[117], James F. Wilson[22,123], James G. Wilson[124], Jerome I. Rotter[72,116], Joanna M. Wardlaw[24,125], John Deanfield[4], Julian Halcox[126], Leo-Pekka Lyytikäinen[109,110], Markus Loeffler[16,17], Michele K. Evans[6], Stéphanie Debette[18], Steve E. Humphries[127], Uwe Völker[34,128], Vilmundur Gudnason[10,11], Aroon D. Hingorani[4], Johan L.M. Björkegren[20,21,42,129], Juan P. Casas[9] & Christopher J. O'Donnell[130,131,132]

[1]Department of Epidemiology, University of North Carolina, Chapel Hill, NC 27516, USA. [2]Department of Pathology and Laboratory Medicine, University of California (UCLA), Los Angeles, Los Angeles, CA 90095, USA. [3]Human Genetics Center, Department of Epidemiology, Human Genetics, and Environmental Sciences, School of Public Health, The University of Texas Health Science Center at Houston, Houston, TX 77030, USA. [4]Institute of Cardiovascular Science, University College London, London WC1 6BT, UK. [5]Cardiovascular Health Research Unit, Department of Medicine, University of Washington, Seattle, WA 98101, USA. [6]Laboratory of Epidemiology and Population Sciences, National Institute on Aging, National Institutes of Health, Baltimore, MD 20892, USA. [7]Department of Genetic Epidemiology, University of Regensburg, Regensburg 93053, Germany. [8]Department of Epidemiology, Erasmus Medical Center, Rotterdam 3015, The Netherlands. [9]Institute of Health Informatics, University College London, London WC1E 6BT, UK. [10]Icelandic Heart Association, Kopavogur IS-201, Iceland. [11]University of Iceland, Reykjavik 101, Iceland. [12]Clinical Division of Neurogeriatrics, Department of Neurology, Medical University of Graz, Graz 8036, Austria. [13]Institute for Medical Informatics, Statistics and Documentation, Medical University of Graz, Graz 8036, Austria. [14]Department of Epidemiology, University of Groningen, University Medical Center Groningen, Groningen 3015, The Netherlands. [15]Department of Biostatistical Sciences, Wake Forest University School of Medicine, Winston-Salem, NC 27157, USA. [16]Institute for Medical Informatics, Statistics and Epidemiology, , University of Leipzig, Leipzig 04107, Germany. [17]LIFE Research Center for Civilization Diseases, University of Leipzig, Leipzig 04107, Germany. [18]Univ. Bordeaux, Inserm, Bordeaux Population Health Research Center, UMR 1219, CHU Bordeaux, F-33000 Bordeaux, France. [19]Research Centre of Applied and Preventive Cardiovascular Medicine, University of Turku, Turku 20520, Finland. [20]Department of Genetics and Genomic Sciences, The Icahn Institute for Genomics and Multiscale Biology Icahn School of Medicine at Mount Sinai, New York, NY 10029, USA. [21]Clinical Gene Networks AB, Stockholm 104 62, Sweden. [22]Usher Institute of Population Health Sciences and Informatics, University of Edinburgh, Edinburgh EH8 9AG, UK. [23]Department of Internal Medicine, Section of Gerontology and Geriatrics, Leiden University Medical Center, Leiden 2300 RC, The Netherlands. [24]Centre for Cognitive Ageing and Cognitive Epidemiology, University of Edinburgh, Edinburgh EH8 9JZ, UK. [25]Medical Genetics Section, Centre for Genomic and Experimental Medicine, Institute of Genetics and Molecular Medicine, University of Edinburgh, Edinburgh EH4 2XU, UK. [26]Population Sciences Branch, Division of Intramural Research, NHLBI, NIH, Framingham, MA 01702-5827, USA. [27]National Heart, Lung and Blood Institute's Intramural Research Program, Framingham Heart Study, Framingham, MA 01702-5827, USA. [28]Department of Medicine, University of Mississippi Medical Center, Jackson, MS 39216, USA. [29]Department of Neurology, University Medicine Greifswald, Greifswald 17475, Germany. [30]Department of Medicine, Columbia University, New York, NY 10032, USA. [31]Department of Biochemistry, Maastricht Centre for Systems Biology (MaCSBio), CARIM School for Cardiovascular Diseases, Maastricht University, Maastricht 6229, The Netherlands. [32]Department of Epidemiology, Harvard T.H. Chan School of Public Health, Boston, MA 02115, USA. [33]Institute for Community Medicine, University Medicine Greifswald, Greifswald 17475, Germany. [34]DZHK (German Center for Cardiovascular Research), partner site Greifswald, Greifswald 17475, Germany. [35]Center for Diabetes Research, Wake Forest School of Medicine, Winston-Salem, NC 25157, USA. [36]Menzies Health Institute Queensland, Griffith University, Southport, QLD 4222, Australia. [37]Department of Internal Medicine, Erasmus Medical Center, University Medical Center Rotterdam, Rotterdam 3015, The Netherlands. [38]MRC Unit for Lifelong Health and Ageing at UCL, London WC1E 6BT, UK. [39]Department of Medicine, University of Groningen, University Medical Center Groningen, Groningen 2300, The Netherlands. [40]Department of Epidemiology, and School of Medicine, Division of Geriatric Medicine, University of Pittsburgh, Pittsburgh, PA 15213, USA. [41]Department of Epidemiology and Public Health, University College London, London WC1E 6BT, UK. [42]Department of Pathophysiology, Institute of Biomedicine and Translation Medicine, University of Tartu, Biomeedikum, Tartu 51010, Estonia. [43]Department of Cardiac Surgery, Tartu University Hospital, Tartu 51010, Estonia. [44]Cardiovascular Medicine Unit, Department of Medicine Solna, Karolinska Institutet, Stockholm 17177, Sweden. [45]Department of Cell and Molecular Biology, National Bioinformatics Infrastructure Sweden, Science for Life Laboratory, Uppsala University, Uppsala 75108, Sweden. [46]Department of Clinical Sciences in Malmö, Lund University, Malmö SE-205 02, Sweden. [47]Department of Human Genetics, University of California (UCLA), Los Angeles, CA 90095, USA. [48]Department of Psychiatry, EMGO Institute for Health and Care Research and Neuroscience Campus Amsterdam, VU University Medical Center, Amsterdam 1081 HL, The Netherlands. [49]Applied Sciences, Premier, Inc., Charlotte, NC 28277, USA. [50]Department of Medicine, University of Verona, Verona 37134, Italy. [51]Department of Medical Biotechnology and Translational Medicine, Università di Milano, Milan 20133, Italy. [52]Centro Cardiologico Monzino, IRCCS, Milan 20138, Italy. [53]St. Elizabeth's Medical Center, Tufts University School of Medicine, Boston, MA 02135, USA. [54]Institute of Laboratory Medicine, University Hospital Munich, LMU Munich 80539, Germany. [55]Dipartimento di Scienze Farmacologiche e Biomolecolari, Università di Milano, Milan 20133, Italy. [56]Department of Clinical and Experimental Medicine, Internal Medicine, Angiology and Arteriosclerosis Diseases, University of Perugia, Perugia 06123, Italy. [57]Centro Diagnostico Italiano, Milan 20147, Italy. [58]Human Genome Sequencing Center, Baylor College of Medicine, Houston, TX 77030-3411, USA. [59]Department of Medicine, Division of Cardiovascular Medicine, Stanford University School of Medicine, Stanford, CA 94309, USA. [60]Department of Medical Sciences, Molecular Epidemiology, Uppsala University, Uppsala 75185, Sweden. [61]Stanford Cardiovascular Institute, Stanford University, Stanford, CA G1120, USA. [62]Heart Center Leipzig, Leipzig 04103, Germany. [63]Centre for Brain Research, Indian Institute of Science, Bangalore 560012, India. [64]Stroke Research Group, Department of Clinical Neurosciences, University of Cambridge, Cambridge CB2 0QQ, UK. [65]Department of Psychology, University of Edinburgh, Edinburgh EH8 9JZ, UK. [66]Department of Cardiology, Leiden University Medical Center, Leiden 2300 RC, The Netherlands. [67]Department of Internal Medicine, Radboud University Medical Center, Nijmegen 6525 GA, The Netherlands. [68]Clinical Trial Service Unit and Epidemiological Studies Unit, Nuffield Department of Population Health, University of Oxford, Oxford OX3 7LF, UK. [69]Department of Population and Quantitative Health Sciences, School of Medicine, Case Western Reserve University, Cleveland, OH 44106, USA. [70]Institute for Laboratory Medicine, University of Leipzig, Leipzig 04109, Germany. [71]Department of Biostatistics, University of Washington, Seattle, WA 98105, USA. [72]Institute for Translational Genomics and Population Sciences, Los Angeles Biomedical Research Institute at Harbor-UCLA Medical Center, Torrance, CA 90502, USA. [73]Department of Internal Medicine, Rush University Medical Center, Chicago, IL 60612, USA. [74]Radboud Institute for Health Sciences, Radboud University Medical Center, Nijmegen, GA 6525, The Netherlands. [75]Department of Medical Sciences, Cardiovascular Epidemiology, Uppsala University, Uppsala 751 05, Sweden. [76]Department of Genetics, University of North Carolina, Chapel Hill, NC 27516, USA. [77]Department of Internal Medicine B, University Medicine Greifswald, Greifswald 17475, Germany. [78]Institute for Stroke and Dementia Research (ISD), University Hospital, Ludwig-Maximilians-University (LMU), Munich 80539, Germany. [79]Munich Cluster for Systems Neurology (SyNergy), Munich 81377, Germany. [80]Department of Neurology, Center for Neurology and Neurosurgery, Johann Wolfgang Goethe-University, Frankfurt am Main 60323, Germany. [81]Institute for Social and Economic Research, Essex University, Colchester CO4 3SQ, UK. [82]Laboratory of Neurogenetics, National Institute on Aging, National Institutes of Health, Bethesda, MD 20892, USA. [83]Data Tecnica International, Glen Echo, MD 20812, USA. [84]Department of Clinical Physiology and Nuclear

Medicine, Turku University Hospital, Turku 20521, Finland. [85]Institute of Social and Preventive Medicine (ISPM), University of Bern, Bern 3012, Switzerland. [86]Electrocardiography Research Group, School of Medicine, Universidad Industrial de Santander, Bucaramanga, Santander 680003, Colombia. [87]Department of Psychiatry and Friedman Brain Institute, Icahn School of Medicine at Mount Sinai, New York, NY 10029, USA. [88]Mental Illness Research Education and Clinical Center (MIRECC), James J. Peters VA Medical Center, Bronx, New York, NY 10468, USA. [89]Population Health Research Institute, St George's, University of London, London SW17 0RE, UK. [90]Inserm U1167, F-59000 Lille, France. [91]Institut Pasteur de Lille, U1167, F-59000 Lille, France. [92]Université de Lille, U1167 - RID-AGE & Centre Hospitalier Universitaire de Lille, U1167, F-59000 Lille, France. [93]Sorbonne Université, Cardiovascular Prevention Unit, Pitié Salpétrière Hospital, Paris 75013, France. [94]Collaborative Health Studies Coordinating Center, Department of Biostatistics, University of Washington, Seattle, WA 98195, USA. [95]Foundation for Research in Health Exercise and Nutrition, Kuopio Research Institute of Exercise Medicine, Kuopio 70100, Finland. [96]Department of Clinical Physiology and Nuclear Medicine, Kuopio University Hospital, Kuopio 70210, Finland. [97]Institute of Laboratory Medicine, University of Leipzig, Leipzig 04109, Germany. [98]Institute of Clinical Chemistry and Laboratory Medicine, University Hospital Regensburg, Regensburg 93053, Germany. [99]Department of Clinical Epidemiology, Leiden University Medical Center, Leiden 2333, The Netherlands. [100]Department of Population Health Sciences, Bristol Medical School, University of Bristol, Bristol BS8 1QU, UK. [101]Mental Health and Wellbeing, Institute of Health and Wellbeing, University of Glasgow, Glasgow G12 0XH, UK. [102]Department of Primary Care & Population Health, University College London, London WC1E 6BT, UK. [103]Department of Medical Epidemiology and Biostatistics, Karolinska Institutet, Stockholm SE-171 77, Sweden. [104]Department of Neurology, Boston University School of Medicine, Boston, MA 02118, USA. [105]Institute of Public Health and Clinical Nutrition, University of Eastern Finland, Kuopio Campus, Kuopio FI-70210, Finland. [106]Kaiser Permanente Washington Health Research Institute, Seattle, WA 98101, USA. [107]Population Health Science, Bristol Medical School, University of Bristol, Bristol BS8 1QU, UK. [108]MRC Integrative Epidemiology Unit at the University of Bristol, Bristol BS8 1TH, UK. [109]Department of Clinical Chemistry, Fimlab Laboratories, Tampere 33014, Finland. [110]Department of Clinical Chemistry, University of Tampere School of Medicine, Tampere 33014, Finland. [111]Division of Cardiovascular Epidemiology, Institute of Environmental Medicine, Karolinska Institutet, Stockholm S-171 77, Sweden. [112]Department of Cardiology, Karolinska University Hospital, Stockholm S-171 77, Sweden. [113]Genetics Institute, University College London, London WC1E 6BT, UK. [114]Departments of Medicine and Epidemiology, Johns Hopkins University, Baltimore, MD 21205, USA. [115]Section of Biomedical Genetics, School of Medicine, Boston University, Boston, MA 02215, USA. [116]Department of Pediatrics, Los Angeles Biomedical Research Institute at Harbor-UCLA Medical Center, Torrance, CA 90502, USA. [117]Institute of Molecular Biology and Biochemistry, Centre for Molecular Medicine, Medical University of Graz, Graz 8010, Austria. [118]Department of Epidemiology & Biostatistics, Imperial College London, London SW7 2AZ, UK. [119]GGZ inGeest and Amsterdam Public Health Research Institute, Department of Psychiatry, Amsterdam University Medical Center, Amsterdam 1081 HV, The Netherlands. [120]Cardiovascular Health Research Unit and Departments of Medicine, Epidemiology, and Health Services, University of Washington, Seattle, WA 98195, USA. [121]Department of Public Health and Primary Care, Leiden University Medical Center, Leiden 2333 ZA, The Netherlands. [122]Center for Human Genomics, Wake Forest University School of Medicine, Winston-Salem, NC 27157, USA. [123]MRC Human Genetics Unit, Institute of Genetics and Molecular Medicine, University of Edinburgh, Western General Hospital, Edinburgh EH4 2XU, UK. [124]Department of Physiology and Biophysics, University of Mississippi Medical Center, Jackson, MS 39216, USA. [125]Centre for Clinical Brain Sciences, and UK Dementia Research Institute at the University of Edinburgh, Edinburgh EH16 4SB, UK. [126]Swansea University Medical School, Swansea SA2 8PP, UK. [127]Centre for Cardiovascular Genetics, Institute Cardiovascular Science, University College London, London WC1E 6BT, UK. [128]Interfaculty Institute for Genetics and Functional Genomics, University Medicine Greifswald, Greifswald 17475, Germany. [129]Integrated Cardio Metabolic Centre, Department of Medicine, Karolinska Institutet, Karolinska Universitetssjukhuset, Huddinge SE-141 57, Sweden. [130]Intramural Administration Management Branch, National Heart, Lung, and Blood Institute, NIH, Bethesda, MD 20892, USA. [131]Cardiology Section, Boston Veteran's Administration Healthcare, Boston, MA 02130, USA. [132]Harvard Medical School, Boston, MA 02115, USA. These authors contributed equally: Nora Franceschini, Claudia Giambartolomei. A full list of consortium members can be found at the end of the article.

MEGASTROKE Consortium

Yukinori Okada[133,134,135], Aniket Mishra[136,137], Loes Rutten-Jacobs[138], Anne-Katrin Giese[139,140], Sander W. van der Laan[141], Solveig Gretarsdottir[142], Christopher D. Anderson[140,143,144], Michael Chong[145], Hieab H.H. Adams[8], Tetsuro Ago[146], Peter Almgren[147], Philippe Amouyel[148,149], Hakan Ay[140,150], Traci M. Bartz[5,71], Oscar R. Benavente[151], Steve Bevan[152], Giorgio B. Boncoraglio[153], Robert D. Brown Jr.[154], Adam S. Butterworth[155,156], Caty Carrera[157,158], Cara L. Carty[159,160], Daniel I. Chasman[132,161], Wei-Min Chen[162], John W. Cole[163], Ioana Cotlarciuc[164], Carlos Cruchaga[165,166], John Danesh[155,156,167], Paul I.W. de Bakker[168,169], Anita L. DeStefano[27,170], Marcel den Hoed[171], Qing Duan[172], Stefan T. Engelter[173,174], Guido J. Falcone[144,175], Rebecca F. Gottesman[176], Raji P. Grewal[177], Stefan Gustafsson[60], Jeffrey Haessler[178], Tamara B. Harris[6], Ahamad Hassan[179], Aki S. Havulinna[180,181], Elizabeth G. Holliday[182,183], George Howard[184], Fang-Chi Hsu[15], Hyacinth I. Hyacinth[185], M. Arfan Ikram[8], Marguerite R. Irvin[186], Xueqiu Jian[187], Jordi Jiménez-Conde[188], Julie A. Johnson[189,190], J. Wouter Jukema[66], Masahiro Kanai[2], Keith L. Keene[191,192], Brett M. Kissela[193], Dawn O. Kleindorfer[193], Charles Kooperberg[178], Michiaki Kubo[194], Leslie Lange[195], Carl D. Langefeld[196], Claudia Langenberg[172], Jin-Moo Lee[197], Robin Lemmens[198,199], Didier Leys[200], Cathryn M. Lewis[201,202], Wei-Yu Lin[203,204], Arne G. Lindgren[205,206], Erik Lorentzen[207], Patrik K. Magnusson[103], Jane Maguire[208], Ani Manichaikul[162], Patrick F. McArdle[209], James F. Meschia[210], Thomas H. Mosley[211,212], Toshiharu Ninomiya[213], Martin J. O'Donnell[145,214], Sara L. Pulit[215], Kristiina Rannikmäe[216], Alexander P. Reiner[178,217], Kathryn M. Rexrode[218], Stephen S. Rich[162], Paul M. Ridker[132,161], Natalia S. Rost[139,140], Peter M. Rothwell[219], Tatjana Rundek[220], Ralph L. Sacco[220], Saori Sakaue[3,221], Michele M. Sale[162], Veikko Salomaa[180], Bishwa R. Sapkota[222], Reinhold Schmidt[223], Carsten O. Schmidt[224], Ulf Schminke[217], Pankaj Sharma[164], Agnieszka Slowik[225], Cathie L.M. Sudlow[226], Christian Tanislav[227], Turgut Tatlisumak[228,229], Vincent N.S. Thijs[230,231], Gudmar Thorleifsson[142], Unnur Thorsteinsdottir[142], Steffen Tiedt[78], Stella Trompet[23], Matthew Walters[232], Nicholas J. Wareham[172], Sylvia Wassertheil-Smoller[233], Kerri L. Wiggins[5], Qiong Yang[170], Salim Yusuf[145], Tomi Pastinen[234], Arno Ruusalepp[21,42,43], Eric E. Schadt[20], Simon Koplev[20], Veronica Codoni[235,236], Mete Civelek[162,237], Nick Smith[5,106,238], David A. Trégouët[235,236], Ingrid E. Christophersen[144,239,240], Carolina Roselli[144], Steven A. Lubitz[144,239], Patrick T. Ellinor[144,239], E. Shyong Tai[241], Jaspal S. Kooner[242], Norihiro Kato[243], Jiang He[244], Pim van der Harst[245], Paul Elliott[246], John C. Chambers[118,247], Fumihiko Takeuchi[243], Andrew D. Johnson[27], Dharambir K. Sanghera[222,248,249], Olle Melander[46], Christina Jern[250], Daniel Strbian[251,252], Israel Fernandez-Cadenas[157,158], W.T. Longstreth Jr[5,253], Arndt Rolfs[147], Jun Hata[213], Daniel Woo[193], Jonathan Rosand[140,143,144], Guillaume Pare[145], Danish Saleheen[254], Kari Stefansson[142,255], Bradford B. Worrall[256], Steven J. Kittner[163], Joanna M.M. Howson[155] & Yoichiro Kamatani[133,257]

[133]Laboratory for Statistical Analysis, RIKEN Center for Integrative Medical Sciences, Yokohama 230-0045, Japan. [134]Department of Statistical Genetics, Osaka University Graduate School of Medicine, Osaka 565-0871, Japan. [135]Laboratory of Statistical Immunology, Immunology Frontier Research Center (WPI-IFReC), Osaka University, Suita 565-0871, Japan. [136]INSERM U1219 Bordeaux Population Health Research Center, Bordeaux F-33000, France. [137]University of Bordeaux, Bordeaux F-33000, France. [138]Stroke Research Group, Division of Clinical Neurosciences, University of Cambridge, Cambridge CB2 1TN, UK. [139]Department of Neurology, Massachusetts General Hospital, Harvard Medical School, Boston, MA 02114, USA. [140]J. Philip Kistler Stroke Research Center, Department of Neurology, MGH, Boston, MA 02215, USA. [141]Laboratory of Experimental Cardiology, Division of Heart and Lungs, University Medical Center Utrecht, Utrecht 3584 CX, Netherlands. [142]deCODE genetics/AMGEN inc, Reykjavik 101, Iceland. [143]Center for Genomic Medicine, Massachusetts General Hospital (MGH), Boston, MA 02114, USA. [144]Program in Medical and Population Genetics, Broad Institute, Cambridge, MA 02142, USA. [145]Population Health Research Institute, McMaster University, Hamilton L8L 2X2, Canada. [146]Department of Medicine and Clinical Science, Graduate School of Medical Sciences, Kyushu University, Fukuoka 819-0935, Japan. [147]Albrecht Kossel Institute, University Clinic of Rostock, Rostock 18147, Germany. [148]INSERM U1167, Institut Pasteur de Lille, Lille F-59000, France. [149]Department of Public Health, Lille University Hospital, Lille F-59000, France. [150]Department of Radiology, Massachusetts General Hospital, Harvard Medical School, AA Martinos Center for Biomedical Imaging, Boston, MA 02129, USA. [151]Division of Neurology, Faculty of Medicine, Brain Research Center, University of British Columbia, Vancouver 170-637, Canada. [152]School of Life Science, University of Lincoln, Lincoln LN6 7TS, UK. [153]Department of Cerebrovascular Diseases, Fondazione IRCCS Istituto Neurologico Carlo Besta, Milano 20133, Italy. [154]Department of Neurology, Mayo Clinic Rochester, Rochester, MN 55905, USA. [155]MRC/BHF Cardiovascular Epidemiology Unit, Department of Public Health and Primary Care, University of Cambridge, Cambridge CB2 1TN, UK. [156]The National Institute for Health Research Blood and Transplant Research Unit in Donor Health and Genomics, University of Cambridge, Cambridge CB2 1TN, UK. [157]Neurovascular Research Laboratory, Vall d'Hebron Institut of Research, Neurology and Medicine Departments-Universitat Autònoma de Barcelona, Vall d'Hebrón Hospital, Barcelona 08193, Spain. [158]Stroke Pharmacogenomics and Genetics, Fundacio Docència i Recerca MutuaTerrassa, Terrassa 08222, Spain. [159]Children's Research Institute, Children's National Medical Center, Washington, DC 20052, USA. [160]Center for Translational Science, George Washington University, Washington, DC 20052, USA. [161]Division of Preventive Medicine, Brigham and Women's Hospital, Boston, MA 02115, USA. [162]Department of Public Health Sciences, Center for Public Health Genomics, University of Virginia School of Medicine, Charlottesville, VA 22904-4259, USA. [163]Department of Neurology, University of Maryland School of Medicine and Baltimore VAMC, Baltimore, MD 21201, USA. [164]Institute of Cardiovascular Research, Royal Holloway University of London, Egham TW20 OEX, UK. [165]Department of Psychiatry,The Hope Center Program on Protein Aggregation and Neurodegeneration (HPAN), Washington University, School of Medicine, St. Louis, MO 98195, USA. [166]Department of Developmental Biology, Washington University School of Medicine, St. Louis, MO 98195, USA. [167]Wellcome Trust Sanger Institute, Hinxton CB10 1SA, UK. [168]Department of Medical Genetics, University Medical Center Utrecht, Utrecht 3584 CX, The Netherlands. [169]Department of Epidemiology, Julius Center for Health Sciences and Primary Care, University Medical Center Utrecht, Utrecht 3584 CX, The Netherlands. [170]Boston University School of Public Health, Boston, MA 02118, USA. [171]Department of Immunology, Genetics and Pathology and Science for Life Laboratory, Uppsala University, Uppsala 751 05, Sweden. [172]MRC Epidemiology Unit, University of Cambridge School of Clinical Medicine, Institute of Metabolic Science, Cambridge Biomedical Campus, Cambridge CB2 0SL, UK. [173]Department of Neurology and Stroke Center, Basel University Hospital, Basel 4031, Switzerland. [174]Neurorehabilitation Unit, University and University Center for Medicine of Aging and Rehabilitation Basel, Felix Platter Hospital, Basel 4055, Switzerland. [175]Department of Neurology, Yale University School of Medicine, New Haven, CT 06510, USA. [176]Department of Neurology, Johns Hopkins University School of Medicine, Baltimore, MD 21205, USA. [177]Neuroscience Institute, SF Medical Center, Trenton, NJ 08629, USA. [178]Division of Public Health Sciences, Fred Hutchinson Cancer Research Center, Seattle, WA 98109-1024, USA. [179]Department of Neurology, Leeds General Infirmary, Leeds Teaching Hospitals NHS Trust, Leeds LS1 3EX, UK. [180]National Institute for Health and Welfare, Helsinki FI-00271, Finland. [181]FIMM - Institute for Molecular Medicine Finland, Helsinki FI-00271, Finland. [182]Public Health Stream, Hunter Medical Research Institute, New Lambton NSW 2305, Australia. [183]Faculty of Health and Medicine, University of Newcastle, Newcastle 2308, Australia. [184]School of Public Health, University of Alabama at Birmingham, Birmingham, AL 35487, USA. [185]Aflac Cancer and Blood Disorder Center, Department of Pediatrics, Emory University School of Medicine, Atlanta, GA 30322, USA. [186]Epidemiology, School of Public Health, University of Alabama at Birmingham, Birmingham 35487, USA. [187]Brown Foundation Institute of Molecular Medicine, University of Texas Health Science Center at Houston, Houston, TX 77030, USA. [188]Neurovascular Research Group (NEUVAS), Neurology Department, Institut Hospital del Mar d'Investigació Mèdica, Universitat Autònoma de Barcelona, Barcelona 08193, Spain. [189]Department of Pharmacotherapy and Translational Research and Center for Pharmacogenomics, University of Florida, College of Pharmacy, Gainesville, FL 32611, USA. [190]Division of Cardiovascular Medicine, College of

Medicine, University of Florida, Gainesville, FL 32611, USA. [191]Department of Biology, East Carolina University, Greenville, NC 27858, USA. [192]Center for Health Disparities, East Carolina University, Greenville, NC 27858, USA. [193]University of Cincinnati College of Medicine, Cincinnati, OH 45220, USA. [194]RIKEN Center for Integrative Medical Sciences, Yokohama 230-0045, Japan. [195]University of Colorado, Denver, CO 80203, USA. [196]Center for Public Health Genomics and Department of Biostatistical Sciences, Wake Forest School of Medicine, Winston-Salem, NC 27157, USA. [197]Department of Neurology, Radiology, and Biomedical Engineering, Washington University School of Medicine, St. Louis, MO 98195, USA. [198]Department of Neurosciences, Experimental Neurology, KU Leuven – University of Leuven, Leuven 3000, Belgium. [199]VIB Center for Brain & Disease Research, University Hospitals Leuven, Department of Neurology, Leuven 3000, Belgium. [200]University of Lille, INSERM U1171, CHU Lille, Lille F-59000, France. [201]Department of Medical and Molecular Genetics, King's College London, London WC2R 2LS, UK. [202]SGDP Centre, Institute of Psychiatry, Psychology & Neuroscience, King's College London, London WC2R 2LS, UK. [203]Cardiovascular Epidemiology Unit, Department Public Health & Primary Care, University of Cambridge, Cambridge CB1 8RN, UK. [204]Northern Institute for Cancer Research, Paul O'Gorman Building, Newcastle University, Newcastle NE2 4AD, UK. [205]Department of Clinical Sciences Lund, Neurology, Lund University, Lund 221 00, Sweden. [206]Department of Neurology and Rehabilitation Medicine, Skåne University Hospital, Lund 222 29, Sweden. [207]Bioinformatics Core Facility, University of Gothenburg, Gothenburg 405 30, Sweden. [208]University of Technology Sydney, Faculty of Health, Ultimo NSW 2007, Australia. [209]Department of Medicine, University of Maryland School of Medicine, Baltimore, MD 21201, USA. [210]Department of Neurology, Mayo Clinic, Jacksonville, FL 32224, USA. [211]Division of Geriatrics, School of Medicine, University of Mississippi Medical Center, Jackson, MS 39216, USA. [212]Memory Impairment and Neurodegenerative Dementia Center, University of Mississippi Medical Center, Jackson, FL 39216, USA. [213]Department of Epidemiology and Public Health, Graduate School of Medical Sciences, Kyushu University, Fukuoka 819-0395, Japan. [214]Clinical Research Facility, Department of Medicine, NUI Galway, Galway H91 TK33, Ireland. [215]Department of Neurology, Brain Center Rudolf Magnus, University Medical Center Utrecht, Utrecht 3584, The Netherlands. [216]Centre for Clinical Brain Sciences, University of Edinburgh, Edinburgh EH4 2XU, UK. [217]Department of Neurology, University Medicine Greifswald, Greifswald 17489, Germany. [218]Department of Medicine, Brigham and Women's Hospital, Boston, MA 02115, USA. [219]Nuffield Department of Clinical Neurosciences, University of Oxford, Oxford OX3 9DU, UK. [220]Department of Neurology, Miller School of Medicine, University of Miami, Miami, FL 33136, USA. [221]Department of Allergy and Rheumatology, Graduate School of Medicine, the University of Tokyo, Tokyo 13-8654, Japan. [222]Department of Pediatrics, College of Medicine, University of Oklahoma Health Sciences Center, Oklahoma City, OK 73104, USA. [223]Department of Neurology, Medical University of Graz, Graz 8036, Austria. [224]University Medicine Greifswald, Institute for Community Medicine, SHIP-KEF, Greifswald 17489, Germany. [225]Department of Neurology, Jagiellonian University, Krakow 31-007, Poland. [226]University of Edinburgh, Edinburgh EH8 9JZ, UK. [227]Department of Neurology, Justus Liebig University, Giessen 35390, Germany. [228]Department of Clinical Neurosciences/Neurology, Institute of Neuroscience and Physiology, Sahlgrenska Academy at University of Gothenburg, Gothenburg SE-405, Sweden. [229]Sahlgrenska University Hospital, Gothenburg SE-405, Sweden. [230]Stroke Division, Florey Institute of Neuroscience and Mental Health, Heidelberg VIC 3084, Australia. [231]Austin Health, Department of Neurology, Heidelberg, Victoria 3084, Australia. [232]School of Medicine, Dentistry and Nursing at the University of Glasgow, Glasgow G12 8QQ, UK. [233]Department of Epidemiology and Population Health, Albert Einstein College of Medicine, Bronx, NY 10461, USA. [234]Department of Human Genetics, McGill University, Montreal H3A 0G4, Canada. [235]Sorbonne Universités, UPMC Univ. Paris 06, INSERM, UMR_S 1166, Team Genomics & Pathophysiology of Cardiovascular Diseases, Paris 75006, France. [236]ICAN, Institute for Cardiometabolism and Nutrition, Paris 75013, France. [237]Department of Biomedical Engineering, University of Virginia, Charlottesville, VA 22904-4259, USA. [238]Seattle Epidemiologic Research and Information Center, VA Office of Research and Development, Seattle, WA 98108, USA. [239]Cardiovascular Research Center, Massachusetts General Hospital, Boston, MA 02114, USA. [240]Department of Medical Research, Bærum Hospital, Vestre Viken Hospital Trust, Rud 3004, Norway. [241]Saw Swee Hock School of Public Health, National University of Singapore and National University Health System, Singapore 119077, Singapore. [242]National Heart and Lung Institute, Imperial College London, London SW7 2AZ, UK. [243]Department of Gene Diagnostics and Therapeutics, Research Institute, National Center for Global Health and Medicine, Tokyo 162-8655, Japan. [244]Department of Epidemiology, Tulane University School of Public Health and Tropical Medicine, New Orleans, LA 70112, USA. [245]Department of Cardiology, University Medical Center Groningen, University of Groningen, Groningen 9700 RB, Netherlands. [246]Department of Epidemiology and Biostatistics, Imperial College London, MRC-PHE Centre for Environment and Health, School of Public Health, London W2 1PG, UK. [247]Department of Cardiology, Ealing Hospital NHS Trust, Southall HA1 3UJ, UK. [248]Department of Pharmaceutical Sciences, College of Pharmacy, University of Oklahoma Health Sciences Center, Oklahoma City, OK 73104, USA. [249]Oklahoma Center for Neuroscience, Oklahoma City, OK 73104, USA. [250]Department of Pathology and Genetics, Institute of Biomedicine, The Sahlgrenska Academy at University of Gothenburg, Gothenburg SE-405, Sweden. [251]Department of Neurology, Helsinki University Hospital, Helsinki FI-00029, Finland. [252]Clinical Neurosciences, Neurology, University of Helsinki, Helsinki FI-00029, Finland. [253]Department of Neurology, University of Washington, Seattle, WA 98195, USA. [254]Department of Genetics, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA 19104, USA. [255]Faculty of Medicine, University of Iceland, Reykjavik 201, Iceland. [256]Departments of Neurology and Public Health Sciences, University of Virginia School of Medicine, Charlottesville, VA 22908, USA. [257]Center for Genomic Medicine, Kyoto University Graduate School of Medicine, Kyoto 606-8501, Japan

SUPPLEMENTARY INFORMATION

Genome-wide association study and colocalization analyses implicate carotid intima-media thickness and carotid plaque loci in cardiovascular outcomes

Franceschini, Giambartolomei et al.

Supplementary Note 1

Study Descriptions

This study includes data from the CHARGE and UCLEB Consortia. For all studies, each participant provided written informed consent. The Institutional Review Board at the parent institution for each respective study approved the study protocols.

CHARGE Consortium

The Aging Gene-Environment Susceptibility-Reykjavik Study (AGES) cohort originally comprised a random sample of 30,795 men and women born in 1907–1935 and living in Reykjavik in 1967.[1] A total of 19,381 individuals 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.

The Atherosclerosis Risk in Communities Study (ARIC) is a multi-center prospective investigation of atherosclerotic disease in a predominantly bi-racial population. [2] 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).

The Austrian Stroke Prevention Study (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.[3,4] A total of 2007 participants were randomly selected from the official community register stratified by gender and 5 year age groups. Individuals were excluded from the study if they had a history of neuropsychiatric disease, including previous stroke, transient ischemic attacks, and dementia, or an abnormal neurologic examination determined on the basis of a structured clinical interview and a physical and 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 MRI and neuropsychological testing was done in 1076 individuals aged 45 to 85 years randomly selected from the entire cohort: 509 from the first period and 567 from the second. In 1992, blood was drawn from all study participants for DNA extraction. They were all European Caucasians.

The Austrian Stroke Prevention Family Study (ASPS-Fam) is a prospective single-center, community-based study on the cerebral effects of vascular risk factors in the normal elderly population of the city of Graz, Austria.[5,6] The ASPS-Fam represents an extension of ASPS, which was established in 1991.[3,4] Between 2006 and 2013, study participants of the ASPS and their first grade relatives were invited to enter ASPS-Fam. Inclusion criteria were no history of previous stroke or dementia and a normal neurologic examination. A total of 381 individuals from 169 families were included into the study. The number of members per family ranged from two to six. The entire cohort underwent an extended diagnostic work- up including clinical history, blood tests, cognitive testing, and a thorough vascular risk factor assessment. The study protocol was approved by the ethics committee of the Medical University of Graz, Austria, and written informed consent was obtained from all subjects.

Carotid Atherosclerosis Progression Study (CAPS) is a community-based study from Germany. Details of the study have been published before. [7] 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.

The Cardiovascular Health Study (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.[8] 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.

Diabetes Heart Study (DHS) is a family-based observational cohort study of cardiovascular disease from a single research center in the United States. [9] The original predominantly (85%) European-ancestry cohort of 1443 persons was recruited in 1997-2005 from families with at least two type 2 diabetes affected siblings and, if possible, a non-diabetic sibling. Blood samples were drawn from all participants at their baseline examination and DNA was subsequently extracted from available samples.

The Erasmus Rucphen Family Study (ERF) is comprised of a family-based cohort embedded in the Genetic Research in Isolated Populations (GRIP) program in the southwest of the Netherlands. [10] 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.

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),[11] the Framingham Heart Study Offspring cohort (5,124 children of the original cohort, and spouses of those children, beginning in 1972[12] and the Third Generation cohort (4,095 children of the Offspring cohort, beginning in 2002).[13] 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. 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.

The Three-City Study (3C) is a prospective population-based cohort study conducted in three French cities, Bordeaux, Dijon, Montpellier, comprising 9,294 participants in total. [14] To be eligible participants had to live in the city, be registered on the electoral rolls in 1999, 65 years or older, and not institutionalized. The study protocol was approved by the Ethical Committee of the University Hospital of Kremlin-Bicêtre and each participant signed an informed consent. In the 3C-Dijon study 4,931 participants were recruited between March 1999 and March 2001. A carotid ultrasound examination was proposed to participants under the age of 85 (n=4,580), and performed with a high resolution B-mode system (Ultramark 9 High Definition Imaging) and a 5- to 10-MHz sounding. Owing to financial and logistic reasons, ultrasound examinations were not performed during the last 6 months of the baseline phase. In total 3,323 participants with ultrasound measures are available in 3C-Dijon. Using a standardized protocol both the left and right common carotid arteries, bifurcations and the internal carotid arteries (first 2cm) were scanned. [15] DNA samples of 3C-Dijon participants were genotyped at the Centre National de Génotypage, Evry, France (www.cng.fr), using Illumina Human610 Quad BeadChip systems on 4077 individuals. [16] After exclusion of individuals > 80 years, with a history of surgical procedure on the carotid artery, and without genome-wide genotypes, 2,518 participants had measurements of common carotid artery intima-media thickness and 2,473 participants had measurement of carotid plaque.

The Lothian Birth Cohort 1936 (LBC1936) is a longitudinal study of ageing, derived from the Scottish Mental Survey of 1947, where nearly all 11 year-old children in Scotland were given a test of general cognitive ability [17-19]. Survivors living in the Lothian area of Scotland were recruited in late-life at mean age 70 (n=1,091). Follow-up has taken place at ages 70, 73, 76, and 79 years. Collected data include genetic information, longitudinal epigenetic information, longitudinal brain imaging, and numerous blood biomarkers, anthropomorphic and lifestyle measures. CCA intima–media thickness (IMT) was measured manually with calipers [20]. This measures minimum, maximum and mean IMT over a 1 cm long segment of the common carotid artery and carotid bulb using the average of three measurements. The means of the maximum values were used with right and left measurements combined. carotid flow velocities, maximum stenosis affecting the internal carotid artery/bulb/CCA Plaques were defined by carotid stenosis of 25% or greater. Full measurement details are presented in Wardlaw et al. 2014[21]. Ethics permission for the study was obtained from the Multi-Centre Research Ethics Committee for Scotland (MREC/01/0/56) and from Lothian Research Ethics Committee (LBC1936: LREC/2003/2/29). The research was carried out in compliance with the Helsinki Declaration. All subjects gave written, informed consent.

MESA (Multi-Ethnic Study of Atherosclerosis) is a study of the characteristics of subclinical cardiovascular disease and the risk factors that predict progression to clinically overt cardiovascular disease or progression of the subclinical disease. [22] MESA consisted of a diverse, population-based sample of an initial 6,814 asymptomatic men and women aged 45-84. 38 percent of the recruited participants were white, 28 percent African American, 22 percent Hispanic, and 12 percent Asian, predominantly of Chinese descent. Participants were recruited from six field centers across the United States: Wake Forest University, Columbia University, Johns Hopkins University, University of Minnesota, Northwestern University and University of California - Los Angeles. Participants are being followed for identification and characterization of cardiovascular disease events, including acute myocardial infarction and other forms of coronary heart disease (CHD), stroke, and congestive heart failure; for cardiovascular disease interventions; and for mortality. The first examination took place over two years, from July 2000 - July 2002. It was followed by four examination periods that were 17-20 months in length. Participants have been contacted every 9 to 12 months throughout the study to assess clinical morbidity and mortality.

The Netherlands Epidemiology of Obesity (NEO) study was designed for extensive phenotyping to investigate pathways that lead to obesity-related diseases. [23] The NEO study is a population-based, prospective cohort study that includes 6,671 individuals aged 45–65 years, with an oversampling of individuals with overweight or obesity. At baseline, information on demography, lifestyle, and medical history have been collected by questionnaires. In addition, samples of 24-h urine, fasting and postprandial blood plasma and serum, and DNA were collected. Genotyping was performed using the Illumina HumanCoreExome chip, which was subsequently imputed to the 1000 genome reference panel. Participants underwent an extensive physical examination, including anthropometry, electrocardiography, spirometry, and measurement of the carotid artery intima-media thickness by ultrasonography. In random subsamples of participants, magnetic resonance imaging of abdominal fat, pulse wave velocity of the aorta, heart, and brain, magnetic resonance spectroscopy of the liver, indirect calorimetry, dual energy X-ray absorptiometry, or accelerometry measurements were performed. The collection of data started in September 2008 and completed at the end of September 2012. Participants are currently being followed for the incidence of obesity-related diseases and mortality.

The Netherlands Study of Depression and Anxiety (NESDA) is a multi-centre study designed to examine the long-term course and consequences of depressive and anxiety disorders (http://www.nesda.nl). [24] NESDA included both individuals with depressive and/or anxiety disorders and controls without psychiatric conditions. Inclusion criteria were age 18-65 years and self-reported western European ancestry, exclusion criteria were not being fluent in Dutch and having a primary diagnosis of another psychiatric condition (psychotic disorder, obsessive compulsive disorder, bipolar disorder, or severe substance use disorder). For all participants DNA was isolated from the baseline blood sample. Through funding from the NIH GAIN program (www.fnih.gov/gain), whole genome scan analysis was conducted for 1859 NESDA (1702 depressed cases and 157 controls) participants. A hundred subjects were excluded because of various quality control issues. [25]

The Orkney Complex Disease Study (ORCADES) is an ongoing family-based, cross-sectional study in the isolated Scottish archipelago of Orkney. [26] Genetic diversity in this population is decreased compared to Mainland Scotland, consistent with the high levels of endogamy historically. 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. Genotyping was performed with the Illumina HumanHap300 and Illumina Omni Express beadchips.

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.[27] 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.

The Study of Health in Pomerania (SHIP) and SHIP-TREND. The Study of Health in Pomerania (SHIP) is a population-based study in the North-East of Germany, which consists of two independent prospectively collected cohorts (SHIP and SHIP-TREND)[28]. Their aim is assessing the prevalence and incidence of common population-based diseases and their risk factors. The detailed study design has been published previously. In brief, 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. For SHIP, baseline examinations were performed between 1997 and 2001. The sample finally comprised 4,308 participants. SHIP-TREND finally comprised 4420 participants. Baseline examinations were conducted between 2008 and 2012. Individuals of both cohorts were analyzed separately. 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.

The Cardiovascular Risk In Young Finns (YFS) study. YFS is a Finnish multi-centre study that was initiated in 1980. [29] A total of 3596 children and adolescents aged 3-18 years participated in the first cross-sectional study. Study variables since childhood include serum lipids, blood pressure, obesity indices, insulin, glucose, life-style (diet, smoking, physical activity, alcohol), family risk and socioeconomic status. In addition, national register data on all hospitalizations with specific diagnoses is available from 1969 onwards. In adulthood, follow-up visits have been performed in 2001, 2007, and 2011, with a total of 2,800 individuals from childhood having at least one follow-up in adulthood. The follow-up studies in 2001 and 2007 have included non-invasive ultrasound measurements of arterial function and structure, which are indicative of subclinical atherosclerosis. [29] 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.

UCLEB (UNIVERSITY COLLEGE-LONDON SCHOOL-EDINBURGH-BRISTOL) CONSORTIUM

BRHS. From 1978 to 1980, 7735 men aged 40-59 were recruited from general practices across the UK.[30] A wide range of phenotypic measures is available for established risk markers such as lipids, blood pressure and inflammatory and hemostatic markers. Most of these measures were taken both at recruitment and re-examination, which occurred in 1998-2000 when men were aged 60-79. At this re-examination 4,252 participants attended and DNA was extracted for 3945. Data on important behavioral variables such as cigarette and alcohol consumption, as well as physical activity, have been regularly collected through follow up. Well validated outcome variables including major coronary heart disease and stroke, as well as cause-specific mortality, continue to be collected from medical records 30 years after recruitment.

The Edinburgh Artery Study (EAS) is an age-stratified random sample of men and women, aged 55-74 years, which was selected between August 1987 and September 1988 from the age-sex registers of ten general practices with a geographical and socio-economical catchment population spread throughout the city of Edinburgh, UK.[30] Subjects were excluded if they were unfit to participate (e.g. due to severe mental illness or terminal disease); excluded individuals were replaced by other randomly sampled subjects.

The Edinburgh Type 2 Diabetes Study (ET2DS) is based on an age-stratified random sample of men and women with type 2 diabetes, aged 60–74 years, which was selected between August 2006 and August 2007 from the Lothian Diabetes Register (LDR), a comprehensive database of subjects with known type 2 diabetes living in Lothian.[31] Subjects were excluded if they did not meet WHO criteria for type 2 diabetes, or if they were physically unable to complete the clinical and cognitive examination. The study population is almost exclusively European. DNA was extracted at baseline. Physical examinations were performed by specially trained research nurses using standardised operating procedures. The quality of measurements was checked using observation of research staff by study investigators and inter-observer variability assessments were made for key variables. Blood assays were performed in accredited laboratories using international standards. Retrospective data on cardiovascular disease and selected physical and biochemical variables were retrieved using record linkage for hospitalisations and deaths since 1985 and using data from the LDR. Subjects returned for further clinical examination after one year and were examined again after they had participated for 4 years.

MRC1946. The Medical Research Council (MRC) National Survey of Health and Development (NSHD; also known as MRC 1946 birth cohort) is an on-going prospective birth cohort study consisting of a sample of all singleton births, born to married mothers, in England, Scotland and Wales in one week in March 1946. [32] The sample includes all births whose fathers were in non-manual or agricultural occupations and a randomly selected one in four of all others, whose fathers were in manual occupations. The original cohort comprised 2,547 women and 2,815 men who have been followed up over 20 times since their birth. The data collected to date include cognitive function, physical, lifestyle and anthropomorphic measures as well as blood analytes and other measures. Through MRC Unit funding, a particularly intensive clinical assessment, with biological sampling, blood and urine sampling and analysis, and cardiac and vascular imaging has recently been completed when the cohort were aged 60-64 years.

The Whitehall II (WHII) Study recruited 10,308 participants (70% men) between 1985 and 1989 from 20 London based civil service departments.[33] In this longitudinal study blood pressure was recorded at phase 1 (1985-1988), phase 3 (1991-1993), phase 5 (1997-1999) and phase 7 (2003-2004). DNA was stored from phase 7 from over 6,000 participants. The study participants are all highly phenotyped for cardiovascular and other ageing related health outcomes.

IMPROVE is a multicentre, longitudinal, observational study, which involves seven recruiting centres in five European countries: Finland, France, Italy, the Netherlands, and Sweden.[34] Each recruiting centre was incorporated separately into the analysis. Recruitment of a total of 3598 patients (514 per centre) was targeted. Men and women, aged from 55 to 79 years, with at least three vascular risk factors, asymptomatic for cardiovascular diseases and free of any conditions that might limit longevity or IMT visualization were considered as eligible for the study. The primary objective of the IMPROVE study was to evaluate the association between C-IMT progression at 15 months and future vascular events (myocardial infarction, cardiovascular death, stroke, or any intervention in the carotid, coronary, or peripheral arterial districts occurring from the 15th to the 36th month of follow-up).

LIFE-Adult is a population-based cohort of 10,000 adult inhabitants of the city of Leipzig, Germany. [35] Participants were characterized regarding life-style and environmental risk factors and clinical and subclinical signs of diseases such as cardiovascular diseases, type 2 diabetes or cognition. Detailed description of the cohort can be found elsewhere. [35]LIFE-Adult meets the ethical standards of the Declaration of Helsinki. The study is approved by the Ethics Committee of the Medical Faculty of the University Leipzig, Germany (Reg. No 263-2009-14122009). Written informed consent including agreement with genetic analyses was obtained from all participants. High-resolution B-mode ultrasound images of carotid vessels were acquired using the GE Vivid ultrasound platform with a 12.0-MHz linear-array transducer (GE-Healthcare). For the assessments, subjects were in supine position. Genotyping was performed using the Affymetrix Axiom CEU1 SNP-array technology.

LIFE-Heart is a cohort of patients with suspected or confirmed stable coronary artery disease or myocardial infarction collected at the Heart Center of the University of Leipzig, Germany. Study details can be found elsewhere. [36] A total of about 7,000 patients were recruited. LIFE-Heart meets the ethical standards of the Declaration of Helsinki. The study is approved by the Ethics Committee of the Medical Faculty of the University Leipzig, Germany (Reg. No 276-2005) and is registered at ClinicalTrials.gov (NCT00497887). Written informed consent including agreement with genetic analyses was obtained from all participants. Patients with myocardial infarction were excluded from the present analysis. High-resolution B-mode ultrasound images of carotid vessels were acquired using the GE Vivid ultrasound platform with a 12.0-MHz linear-array transducer (GE-Healthcare). For the assessments, subjects were in supine position. Genotyping was performed with either Affymetrix Axiom CEU1 or Affymetrix Axiom CADLIFE. The latter is an array containing Axiom CEU as genome-wide backbone and an additional custom content of about 62,500 SNPs from CAD loci.

The Prospective Investigation of the Uppsala Seniors (PIVUS) cohort was randomly sampled from all men and women at age 70 living in Uppsala County in 2001 (n=1016; www.medsci.uu.se/PIVUS). Follow-ups were made at years 75 (n=827) and 80 (n=606). The participants underwent a medical examination with cognitive testing (MMSE), vascular status assessments (endothelium-dependent vasodilation, flow-mediated dilation and pulse-wave velocity), subclinical atherosclerosis measurements (intima-media thickness, grey-scale median and plaque occurrence) and blood sampling (low-density lipoprotein, high-density lipoprotein, triglycerides and total cholesterol) including a detailed questionnaire (medical history, exercise, smoking, alcohol, dietary habits and educational level). All participants were genotyped using the Ilumina MetaboChip genotyping array.

ALSPAC The Avon Longitudinal Study of Parents and Children (ALSPAC) (http://www.alspac.bristol.ac.uk/) recruited 14,541 pregnant women resident in Avon, UK with expected dates of delivery 1st April 1991 to 31st December 1992. Further details of the cohort and data collection are available in previous publications. [37,38] The study website (http://www.bristol.ac.uk/alspac/researchers/) contains details of all the data that is available through a fully searchable data dictionary. For this study data from a sub-sample of the women who were originally recruited when pregnant and who attended a follow-up clinic approximately 18-years after the birth of the study index child were included. All data collection and its use for research has been approved by the ALSPAC Ethics and Law Committee and/or UK National Health Service Research Ethics Committees. Participants provided informed written consent. cIMT measurements on these women were collected from both the left and right common carotid artery arteries, using high-resolution B ultrasound and scanning longitudinally 1 cm proximal to the carotid bifurcation following a standardized protocol. A ZONARE z.one Ultra convertible ultrasound system with L10-5 linear transducer was used. Images were focused on the posterior (far) wall of the artery and the zoom function was used to magnify the area. Ten-second cine loops were recorded in DICOM format and analyzed offline using Carotid Analyzer for Research (Vascular Research Tools 5, Medical Imaging Applications, LLC 2008). Three consecutive cardiac cycles were identified and three measures of cIMT were taken from end-diastolic frames and averaged. This was done for both right and left carotid arteries. Arterial distensibility was calculated as the difference between systolic and diastolic arterial diameter. The mean of the left- and right-sided readings was used in all analyses. The images were analyzed by a single trained reader.

The Nijmegen Biomedical Study (NBS) (http://www.nijmegenbiomedischestudie.nl) is a population-based survey conducted by the Department for Health Evidence and the Department of Laboratory Medicine of the Radboud University Medical Centre, Nijmegen, The Netherlands. A cohort profile description of the NBS is available. [39] Briefly, in 2002, 22,451 age and sex-stratified randomly selected adult inhabitants of Nijmegen, a city located in the eastern part of the Netherlands, received an invitation to fill out a postal questionnaire (QN) including questions about lifestyle, health status, and medical history, and to donate a blood sample for DNA isolation and biochemical studies. A total of 9350 (43%) persons filled out the QN, of which 6468 (69%) donated blood samples. A second, third and fourth questionnaire were sent out in 2005, 2008 and 2012, respectively. Approval to conduct the NBS was obtained from the Radboud university medical center Institutional Review Board. All participants gave written informed consent for participation in the NBS.

The Malmo Diet and Cancer (MDC) study is set in Malmö, Sweden's third largest city.[40] The background population consisted of all men born between 1923 and 1945 and all women born between 1923 and 1950 who were living in Malmö during the screening period 1991 to 1996 (n = 74,138). This population was identified through the Swedish national population registries. The final cohort consisted of 28,098 individuals (participation rate 40.8%). The subjects were recruited through advertisements in local media and through invitation by mail. The only exclusion criteria were inadequate Swedish language skills and mental incapacity. The Ethics Committee at Lund University approved the design of the MDC study (LU 51–90). Written informed consent was obtained from the participants.

Supplementary Note 2

Acknowledgements

AGES. This study has been funded by NIH contract N01-AG012100, the NIA Intramural Research Program, an Intramural Research Program Award (ZIAEY000401) from the National Eye Institute, an award from the National Institute on Deafness and Other Communication Disorders (NIDCD) Division of Scientific Programs (IAA Y2-DC_1004-02), Hjartavernd (the Icelandic Heart Association), and the Althingi (the Icelandic Parliament). The study is approved by the Icelandic National Bioethics Committee, VSN: 00-063. The researchers are indebted to the participants for their willingness to participate in the study.

ARIC. The ARIC study is carried out as a collaborative study supported by National Heart, Lung, and Blood Institute contracts (HHSN268201100005C, HHSN268201100006C, HHSN268201100007C, HHSN268201100008C, HHSN268201100009C, HHSN268201100010C, HHSN268201100011C, and HHSN268201100012C), R01HL087641, R01HL59367 and R01HL086694; National Human Genome Research Institute contract U01HG004402; and National Institutes of Health contract HHSN268200625226C. The authors thank the staff and participants of the ARIC study for their important contributions. Infrastructure was partly supported by Grant Number UL1RR025005, a component of the National Institutes of Health and NIH Roadmap for Medical Research.

ASPS/ASPS-Fam. The authors thank the staff and the participants for their valuable contributions. We thank Birgit Reinhart for her long-term administrative commitment, Elfi Hofer for the technical assistance at creating the DNA bank, Ing. Johann Semmler and Anita Harb for DNA sequencing and DNA analyses by TaqMan assays and Irmgard Poelzl for supervising the quality management processes after ISO9001 at the biobankingand DNA analyses. The research reported in this article was funded by the Austrian Science Fund (FWF) grant number P20545-P05, P13180, PI904 the Austrian National Bank Anniversary Fund, P15435, the Austrain Federal Ministry of Science, Research and Economy under the aegis of the EU Joint Programme-Neurodegenerative Disease Research (JPND)-www.jpnd.eu and by the Austrian Science Fund P20545-B05. The Medical University of Graz supports the databank of the ASPS.

Cardiovascular Health Study (CHS). Cardiovascular Health Study: This CHS research was supported by NHLBI contracts HHSN268201200036C, HHSN268200800007C, N01HC55222, N01HC85079, N01HC85080, N01HC85081, N01HC85082, N01HC85083, N01HC85086, N01HC85085, N01HC45133, HHSN268200960009C, N01HC85085, and N01HC45133; and NHLBI grants U01HL080295, R01HL087652, R01HL105756, R01HL103612, R01HL120393, and R01HL130114 with additional contribution from the National Institute of Neurological Disorders and Stroke (NINDS). Additional support was provided through R01AG023629 from the National Institute on Aging (NIA). A full list of principal CHS investigators and institutions can be found at CHS-NHLBI.org. The provision of genotyping data was supported in part by the National Center for Advancing Translational Sciences, CTSI grant UL1TR000124, and the National Institute of Diabetes and Digestive and Kidney Disease Diabetes Research Center (DRC) grant DK063491 to the Southern California Diabetes Endocrinology Research Center. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.Diabetes Heart Study (DHS). This study was supported by the National Institutes of Health through R01 HL6734, R01 HL092301, R01 AR48797, and F31 AG044879. DHS. The authors thank the Wake Forest School of Medicine investigators and staff and the participants of the Diabetes Heart Study for their valuable contributions.

Three City Study (3C) Dijon. The 3-City Study is conducted under a partnership agreement among the Institut National de la Santé et de la Recherche Médicale (INSERM), the University of Bordeaux, and Sanofi-Aventis. The Fondation pour la Recherche Médicale funded the preparation and initiation of the study. The 3C Study is also supported by the Caisse Nationale Maladie des Travailleurs Salariés, Direction Générale de la Santé, Mutuelle Générale de l'Education Nationale (MGEN), Institut de la Longévité, Conseils Régionaux of Aquitaine and Bourgogne, Fondation de France, and Ministry of Research–INSERM Programme "Cohortes et collections de données biologiques." This work was supported by the National Foundation for Alzheimer's Disease and Related Disorders, the Institut Pasteur de Lille, the Centre National de Génotypage and the LABEX (Laboratory of Excellence program investment for the future) DISTALZ - Development of Innovative Strategies for a Transdisciplinary approach to ALZheimer's disease. Ganesh Chauhan, Christophe Tzourio and

Stéphanie Debette are supported by a grant from the Fondation Leducq and grants from the Agence Nationale de la Recherche (ANR).

ERF (Erasmus Rucphen Family study). The ERF study was supported by CardioVasculair Onderzoek Nederland (CVON2012-03) of the Netherlands Heart Foundation and the Rotterdam Study by the European Union's Horizon 2020 research and innovation programme as part of the Common mechanisms and pathways in Stroke and Alzheimer's disease (CoSTREAM) project (www.costream.eu, grant agreement No 667375); European Union's Horizon 2020 research.

FHS (Framingham Heart Study). This research was conducted in part using data and resources from the Framingham Heart Study of the National Heart Lung and Blood Institute of the National Institutes of Health and Boston University School of Medicine. The analyses reflect intellectual input and resource development from the Framingham Heart Study investigators participating in the SNP Health Association Resource (SHARe) project. This work was partially supported by the National Heart, Lung and Blood Institute's Framingham Heart Study (Contract Nos. N01-HC-25195 and HHSN268201500001I) and its contract with Affymetrix, Inc for genotyping services (Contract No. N02-HL-6-4278). A portion of this research utilized the Linux Cluster for Genetic Analysis (LinGA-II) funded by the Robert Dawson Evans Endowment of the Department of Medicine at Boston University School of Medicine and Boston Medical Center.

GeneSTAR was supported by grants from the National Institutes of Health/National Heart, Lung, and Blood Institute (U01 HL72518, HL097698, HL49762, HL58625, HL071025, and HL092165), by a grant from the National Institutes of Health/National Institute of Nursing Research (NR0224103), and by a grant from the National Center for Research Resources and the National Center for Advancing Translational Sciences, National Institutes of Health to the Johns Hopkins Institute for Clinical & Translational Research (UL1 RR 025005).

JHS (Jackson Heart Study). We thank the Jackson Heart Study participants and staff for their contributions to this work. The JHS is supported by contracts HHSN268201300046C, HHSN268201300047C, HHSN268201300048C, HHSN268201300049C, HHSN268201300050C from the National Heart, Lung, and Blood Institute and the National Institute on Minority Health and Health Disparities. Dr. Wilson is supported by U54GM115428 from the National Institute of General Medical Sciences.

LBC1936. REM, IJD, and JMW are members of The University of Edinburgh Centre for Cognitive Ageing and Cognitive Epidemiology, part of the cross-council Lifelong Health and Wellbeing Initiative (MR/K026992/1). Funding from the Biotechnology and Biological Sciences Research Council (BBSRC) and Medical Research Council (MRC) is gratefully acknowledged.

MESA (Multi-Ethnic Study of Atherosclerosis). MESA and the MESA SHARe project are conducted and supported by the National Heart, Lung, and Blood Institute (NHLBI) in collaboration with MESA investigators. Support for MESA is provided by contracts HHSN268201500003I, N01-HC-95159, N01-HC-95160, N01-HC-95161, N01-HC-95162, N01-HC-95163, N01-HC-95164, N01-HC-95165, N01-HC-95166, N01-HC-95167, N01-HC-95168, N01-HC-95169, UL1-TR-000040, UL1-TR-001079, and UL1-TR-001420. This publication was developed under a STAR research assistance agreement, No. RD831697 (MESA Air), awarded by the U.S Environmental protection Agency. It has not been formally reviewed by the EPA. The views expressed in this document are solely those of the authors and the EPA does not endorse any products or commercial services mentioned in this publication. The provision of genotyping data was supported in part by the National Center for Advancing Translational Sciences, CTSI grant UL1TR001881, and the National Institute of Diabetes and Digestive and Kidney Disease Diabetes Research Center (DRC) grant DK063491 to the Southern California Diabetes Endocrinology Research Center. The authors thank the participants of the MESA study, the Coordinating Center, MESA investigators, and study staff for their valuable contributions. A full list of participating MESA investigators and institutions can be found at http://www.mesa-nhlbi.org.

NEO. The authors of the NEO study thank all individuals who participated in the Netherlands Epidemiology in Obesity study, all participating general practitioners for inviting eligible participants and all research nurses for collection of the data. We thank the NEO study group, Pat van Beelen, Petra Noordijk and Ingeborg de Jonge for the coordination, lab and data management of the NEO study. The genotyping in the NEO study was supported by the Centre National de

Génotypage (Paris, France), headed by Jean-Francois Deleuze. The NEO study is supported by the participating Departments, the Division and the Board of Directors of the Leiden University Medical Center, and by the Leiden University, Research Profile Area Vascular and Regenerative Medicine. Dennis Mook-Kanamori is supported by Dutch Science Organization (ZonMW-VENI Grant 916.14.023).

ORCADES was supported by the Chief Scientist Office of the Scottish Government (CZB/4/276, CZB/4/710), the Royal Society, the MRC Human Genetics Unit, Arthritis Research UK and the European Union framework program 6 EUROSPAN project (contract no. LSHG-CT-2006-018947). DNA extractions were performed at the Wellcome Trust Clinical Research Facility in Edinburgh. We would like to acknowledge the invaluable contributions of the research nurses in Orkney, the administrative team in Edinburgh and the people of Orkney.

YFS. Young Finns Study was financially supported by the Academy of Finland (134309 (Eye), 126925, 121584, 124282, 129378 (Salve), 117787 (Gendi), and 41071 (Skidi)); the Social Insurance Institution of Finland, Kuopio, Tampere; Turku University Hospital Medical Funds (grant 9M048 and 9N035 to T.L.); Juho Vainio Foundation; Paavo Nurmi Foundation; Finnish Foundation of Cardiovascular Research and Finnish Cultural Foundation; Tampere Tuberculosis Foundation; and Emil Aaltonen Foundation (to T.L.).

NESDA. Funding was obtained from the Netherlands Organization for Scientific Research (Geestkracht program grant 10-000-1002); the Center for Medical Systems Biology (CSMB, NWO Genomics), Biobanking and Biomolecular Resources Research Infrastructure (BBMRI-NL), University's Institutes for Health and Care Research (EMGO+) and Neuroscience Campus Amsterdam, University Medical Center Groningen, Leiden University Medical Center, National Institutes of Health (NIH, R01D0042157-01A, MH081802, Grand Opportunity grants 1RC2 MH089951 and 1RC2 MH089995). Part of the genotyping and analyses were funded by the Genetic Association Information Network (GAIN) of the Foundation for the National Institutes of Health. Computing was supported by BiG Grid, the Dutch e-Science Grid, which is financially supported by NWO. Statistical analyses were carried out on the Genetic Cluster Computer (http://www.geneticcluster.org) hosted by SURFsara and financially supported by the Netherlands Scientific Organization (NWO 480-05-003 PI: Posthuma) along with a supplement from the Dutch Brain Foundation and the VU University Amsterdam.

EAS. The Edinburgh Artery Study is funded by the British Heart Foundation (Programme Grant RG/98002), with Metabochip genotyping funded by a project grant from the Chief Scientist Office of Scotland (Project Grant CZB/4/672). ET2DS: The Edinburgh Type 2 Diabetes Study is funded by the Medical Research Council (Project Grant G0500877); the Chief Scientist Office of Scotland (Programme Support Grant CZQ/1/38); Pfizer plc (Unrestricted Investigator Led Grant); and Diabetes UK (Clinical Research Fellowship 10/0003985). Research clinics were held at the Welcome Trust Clinical Research Facility and Princess Alexandra Eye Pavilion in Edinburgh

Pivus. Swedish Research Council (grant no. 2015-02907), Göran Gustafsson Foundation, Swedish Heart-Lung Foundation (grant no. 20140422), Knut och Alice Wallenberg Foundation (grant no. 2013.0126)

ALSPAC. The UK Medical Research Council and Wellcome Trust (102215/2/13/2) and the University of Bristol provide core support for ALSPAC. A Wellcome Trust (WT088806) grant provided funds for completion of genome wide on the ALSPAC mothers. Phenotypic (cIMT) data collection was funded by the British Heart Foundation (SP/07 1008/24066), UK Medical Research Council (G1001357) and WellcomeTrust (WT092830M). DAL and SR work in a Unit that receives funds from the UK Medical Research Council (MC_UU_12013/5) and DAL is a National Institute of Health Research Senior Investigator (NF-SI-0611-10196). We are extremely grateful to all of the families who took part in this study, the midwives for their help in recruiting them, and the whole ALSPAC team, which includes interviewers, computer and laboratory technicians, clerical workers, research scientists, volunteers, receptionists, managers and nurses.

NBS. The Nijmegen Biomedical Study is a population-based survey conducted at the Department for Health Evidence, and the Department of Laboratory Medicine of the Radboud university medical center. Principal investigators of the Nijmegen Biomedical Study are L.A.L.M. Kiemeney, A.L.M. Verbeek, D.W. Swinkels and B. Franke

LIFE-Adult is funded by the Leipzig Research Center for Civilization Diseases (LIFE). LIFE is an organizational unit affiliated to the Medical Faculty of the University of Leipzig. LIFE is funded by means of the European Union, by the

European Regional Development Fund (ERDF) and by funds of the Free State of Saxony within the framework of the excellence initiative (project numbers 713-241202, 14505/2470, 14575/2470).

MALMO. This study was supported by grants from the European Research Council (StG-282255) Swedish Medical Research Council, the Swedish Heart and Lung Foundation, the Medical Faculty of Lund University, Malmö. University Hospital, the Albert Påhlsson Research Foundation, the Crafoord Foundation, the Ernhold Lundström Research Foundation, the Region Skane, Hulda and Conrad Mossfelt Foundation, King Gustaf V and Queen Victoria Foundation and the Lennart Hansson Memorial Fund. The authors acknowledge the Knut and Alice Wallenberg Foundation for its economic support of the SWEGENE DNA extraction facility.

BRHS. The British Regional Heart Study has been supported by Programme Grants from the British Heart Foundation (RG/08/013/25942 and RG/13/16/30528).

IMROVE. The authors wish to thank all the members of the IMPROVE group for their time and extraordinary commitment. The IMPROVE study was supported by the European Commission [Contract number: QLG1- CT- 2002-00896 to E.T., D.B., A.H., S.E.H., R.R., U.dF., A.J.S., P.G., S.K., E.M.], Ministero della Salute Ricerca Corrente, Italy [to E.T., D.B.], the Swedish Heart-Lung Foundation, the Swedish Research Council [projects 8691 to A.H. and 0593 to U.dF.], the Foundation for Strategic Research, the Stockholm County Council [project 562183 to A.H.], the Foundation for Strategic Research, the Academy of Finland [Grant #110413 to S.K.] and the British Heart Foundation [RG2008/008 to S.E.H.]. None of the aforementioned funding organizations or sponsors has had a specific role in design or conduct of the study, collection, management, analysis, or interpretation of the data, or preparation, review, or approval of the manuscript.

ORCADES was supported by the Chief Scientist Office of the Scottish Government (CZB/4/276, CZB/4/710), the Royal Society, the MRC Human Genetics Unit, Arthritis Research UK and the European Union framework program 6 EUROSPAN project (contract no. LSHG-CT-2006-018947). DNA extractions were performed at the Wellcome Trust Clinical Research Facility in Edinburgh. We would like to acknowledge the invaluable contributions of the research nurses in Orkney, the administrative team in Edinburgh and the people of Orkney.

The Rotterdam study is supported by the Erasmus MC and Erasmus University Rotterdam; the Netherlands Organisation for Scientific Research; the Netherlands Organisation for Health Research and Development (ZonMw: Zorg onderzoek Nederland Medische Wetenschappen); the Research Institute for Diseases in the Elderly; the Netherlands Genomics Initiative; the Ministry of Education, Culture and Science; the Ministry of Health, Welfare and Sports; the European Commission (Directorate-General XII); and the Municipality of Rotterdam. Maryam Kavousi is supported by the ZonMw Veni grant (Veni, 91616079). O.H. Franco works in ErasmusAGE, a center for aging research across the life course funded by Nestlé Nutrition (Nestec Ltd.); Metagenics Inc.; and AXA. None of the funders had any role in design and conduct of the study; collection, management, analysis, and interpretation of the data; and preparation, review, or approval of this article. The generation and management of GWAS genotype data for the Rotterdam study was executed by the Human Genotyping Facility of the Genetic Laboratory of the Department of Internal Medicine, Erasmus MC, Rotterdam, the Netherlands. The GWAS datasets are supported by the Netherlands Organisation of Scientific Research Investments (number: 175.010.2005.011, 911-03-012), the Genetic Laboratory of the Department of Internal Medicine, Erasmus MC, the Research Institute for Diseases in the Elderly (014-93-015), the Netherlands Genomics Initiative/Netherlands Organisation for Scientific Research Netherlands Consortium for Healthy Aging, project number: 050-060-810. We thank the Genetic Laboratory of the Department of Internal Medicine of the Erasmus MC and specifically Pascal Arp, Mila Jhamai, Marijn Verkerk, and Carolina Medina-Gomez for their help in creating the GWAS database and the creation and analysis of imputed data. The dedication, commitment, and contribution of inhabitants, general practitioners, and pharmacists of the Ommoord district to the Rotterdam Study are gratefully acknowledged

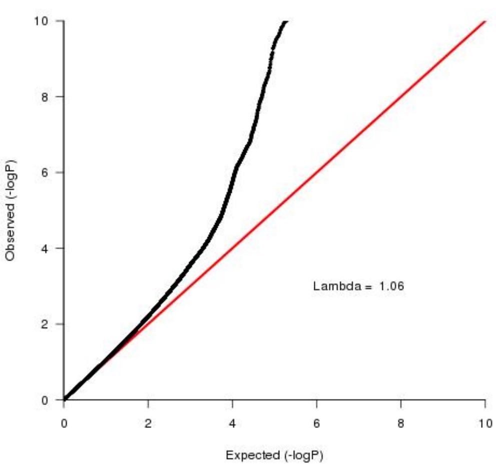
SHIP is part of the Community Medicine Research net of the University of Greifswald, Germany, which is funded by the Federal Ministry of Education and Research (grants no. 01ZZ9603, 01ZZ0103, and 01ZZ0403), the Ministry of Cultural Affairs as well as the Social Ministry of the Federal State of Mecklenburg-West Pomerania, and the network 'Greifswald Approach to Individualized Medicine (GANI_MED)' funded by the Federal Ministry of Education and Research (grant 03IS2061A). Genome-wide data have been supported by the Federal Ministry of Education and Research (grant no.

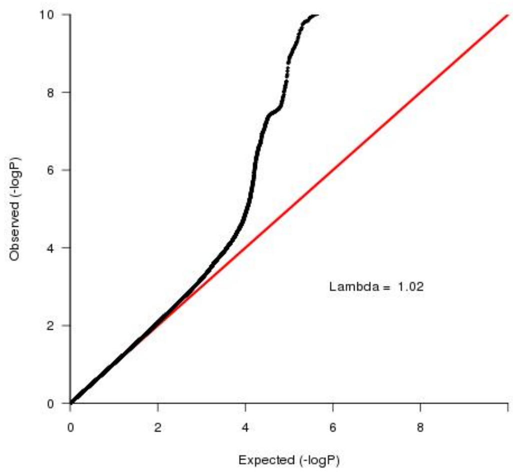
03ZIK012) and a joint grant from Siemens Healthcare, Erlangen, Germany and the Federal State of Mecklenburg- West Pomerania. The University of Greifswald is a member of the Caché Campus program of the InterSystems GmbH.

Supplementary Figures

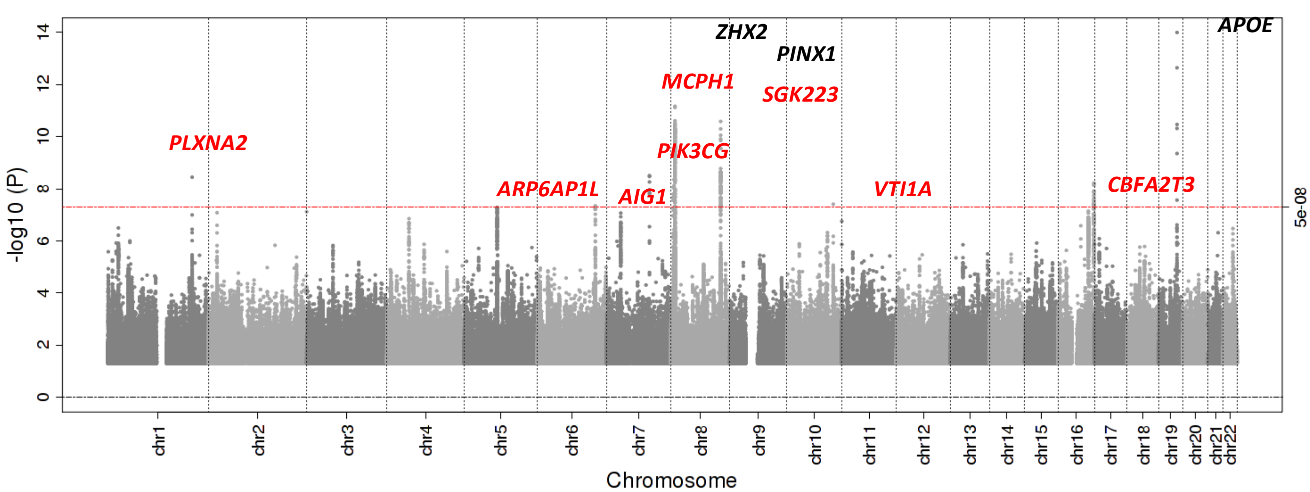
Supplementary Figure 1. QQ plots for meta-analyses of cIMT (A) and plaque (B) and Manhattan plots for cIMT (C) and plaque (D). Novel loci highlighted in red.

A **B**

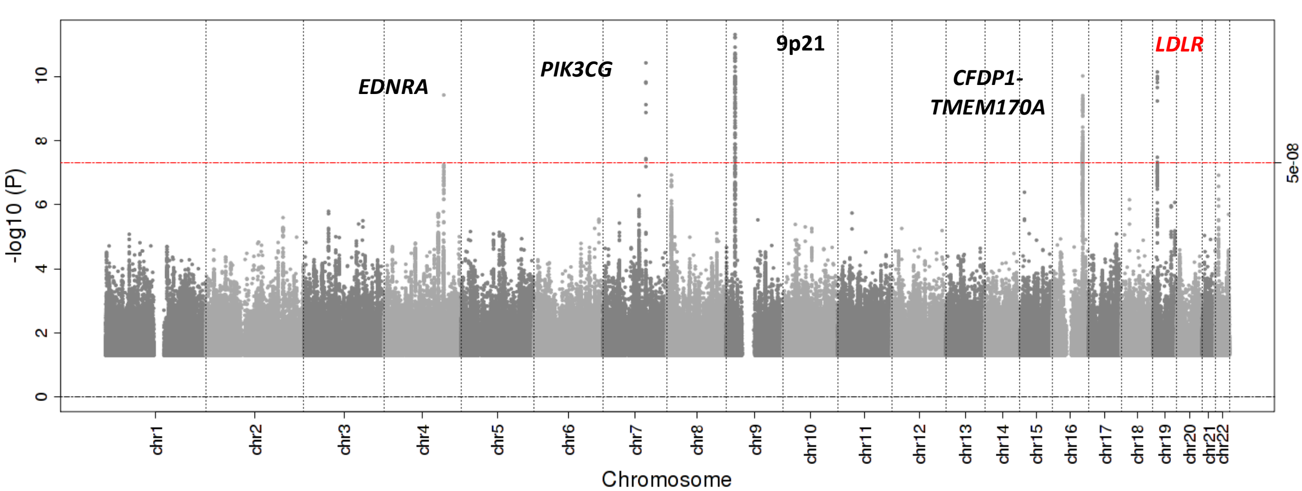




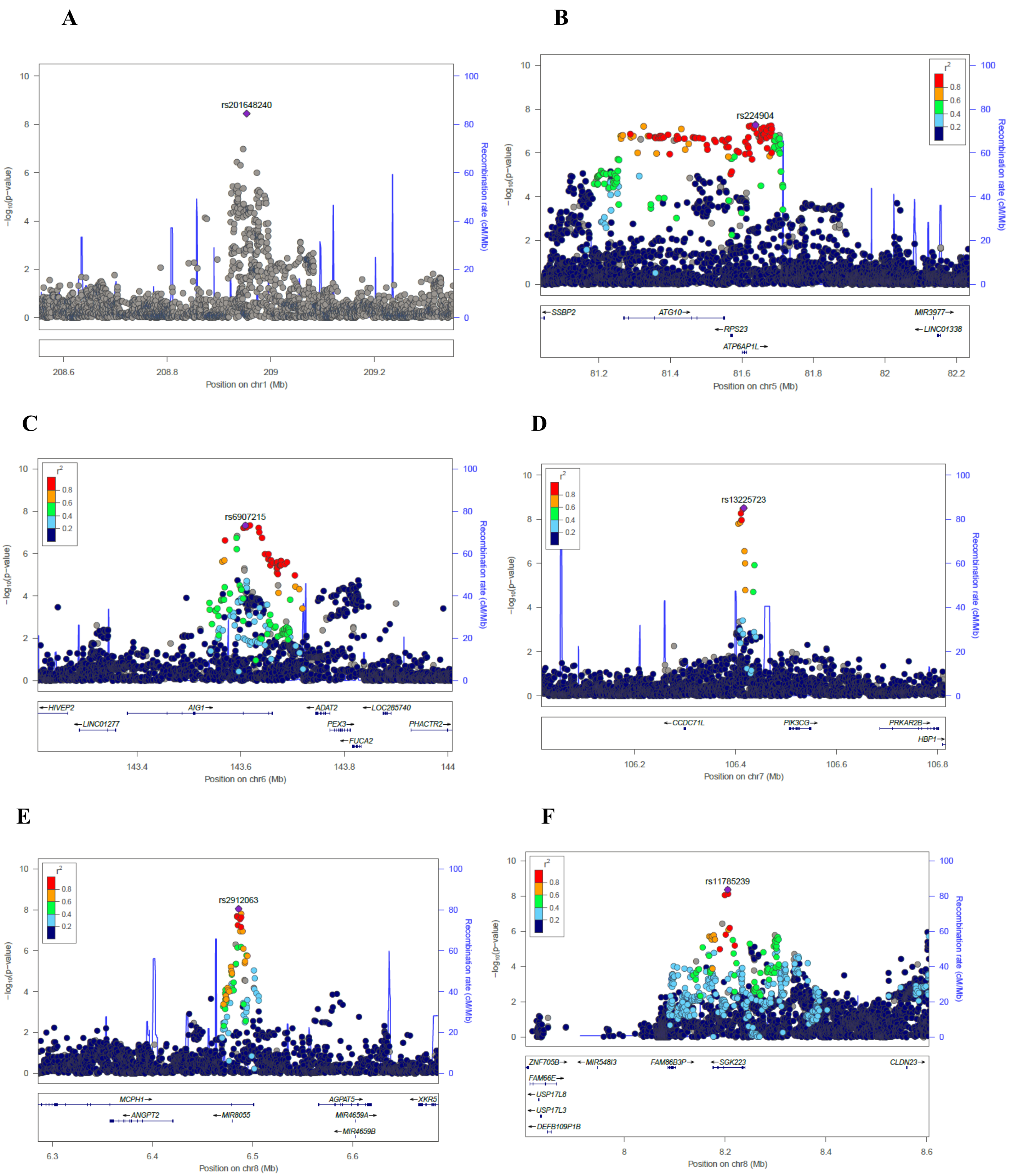
C

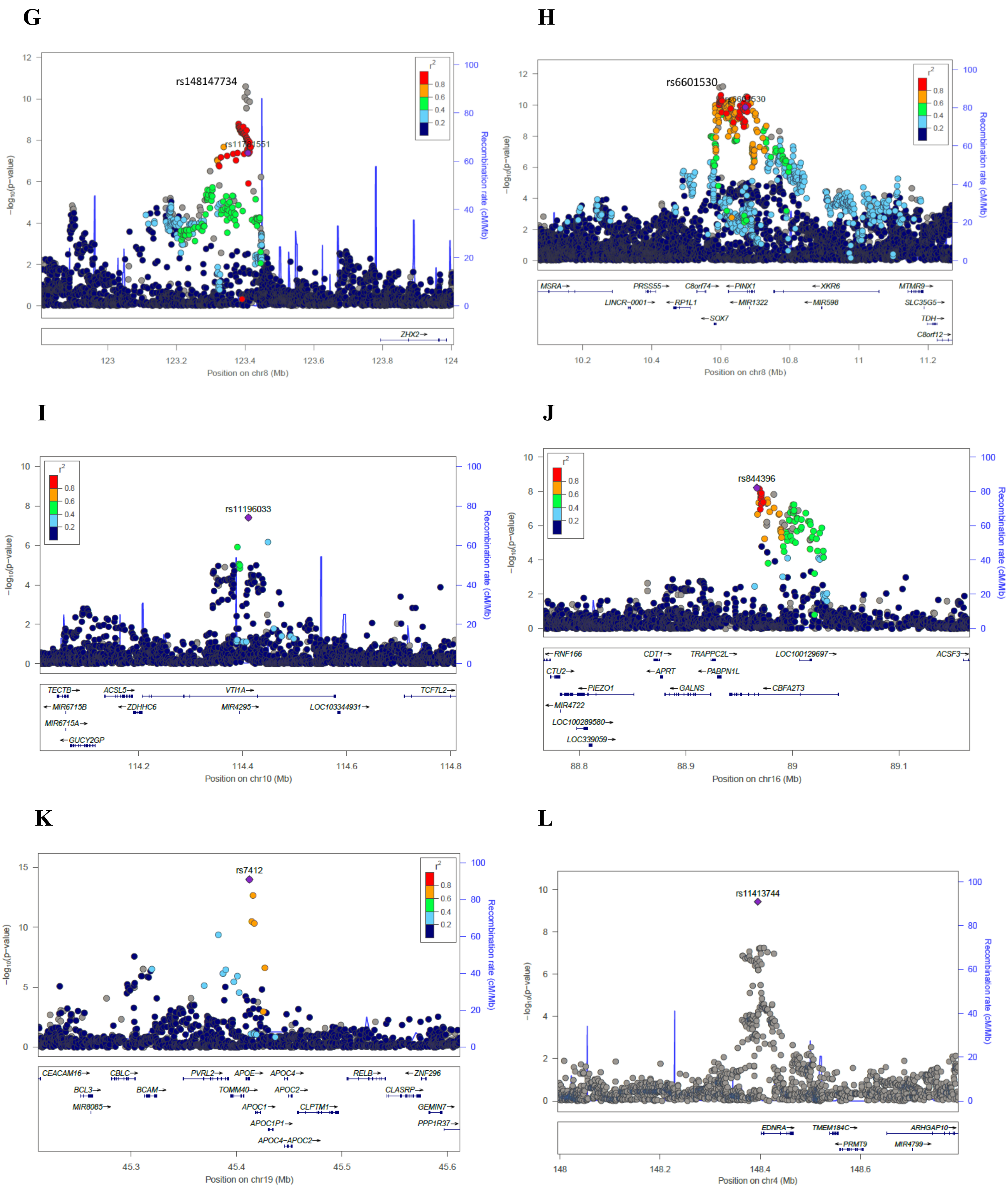


D



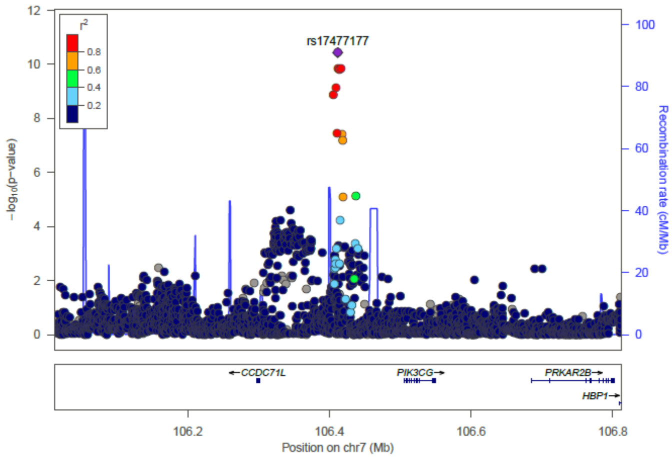
Supplementary Figure 2. Regional plots for significant loci for cIMT (A-K) and carotid plaque (L-P). Note the most significant SNP may not have LD with other SNPs.



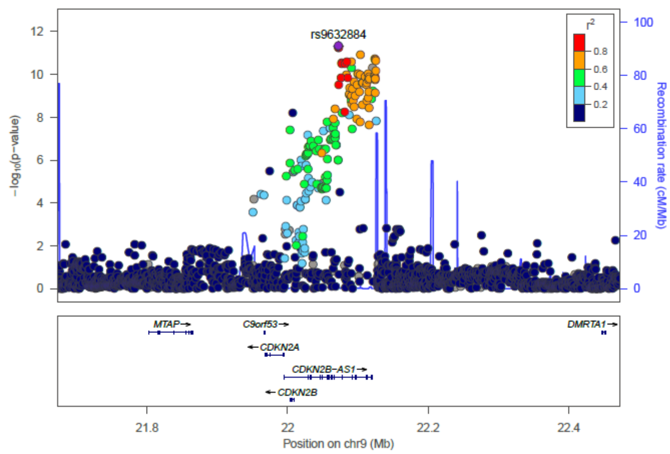


G
rs148147734
rs11781551
−log10(p−value)
Recombination rate (cM/Mb)
ZHX2
Position on chr8 (Mb)
H
rs6601530
MSRA
PRSS55
C8orf74
PINX1
XKR6
MTMR9
LINCR−0001
RP1L1
MIR1322
MIR598
SLC35G5
SOX7
TDH
C8orf12
Position on chr8 (Mb)
I
rs11196033
TECTB
ACSL5
VTI1A
TCF7L2
MIR6715B
ZDHHC6
MIR4295
LOC103344931
MIR6715A
GUCY2GP
Position on chr10 (Mb)
J
rs844396
RNF166
CDT1
TRAPPC2L
LOC100129697
ACSF3
CTU2
APRT
PABPN1L
PIEZO1
GALNS
CBFA2T3
MIR4722
LOC100289580
LOC339059
Position on chr16 (Mb)
K
rs7412
CEACAM16
CBLC
PVRL2
APOE
APOC4
RELB
ZNF296
BCL3
BCAM
TOMM40
APOC2
CLASRP
MIR8085
APOC1
CLPTM1
GEMIN7
APOC1P1
PPP1R37
APOC4−APOC2
Position on chr19 (Mb)
L
rs11413744
EDNRA
TMEM184C
ARHGAP10
PRMT9
MIR4799
Position on chr4 (Mb)

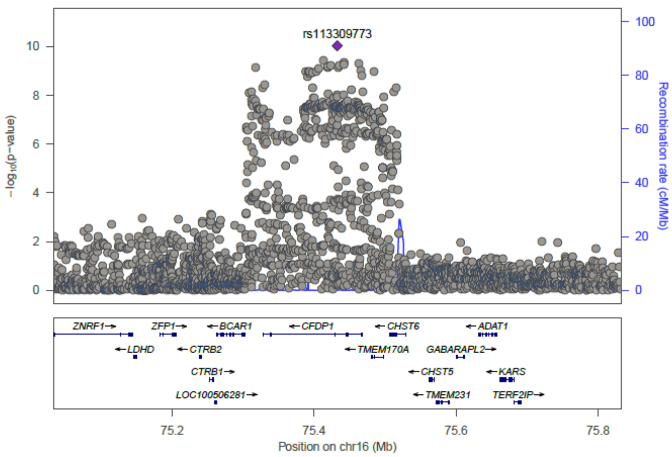
M


rs17477177
r²
0.8
0.6
0.4
0.2
−log10(p−value)
Recombination rate (cM/Mb)
CCDC71L
PIK3CG
PRKAR2B
HBP1
106.2
106.4
106.6
106.8
Position on chr7 (Mb)

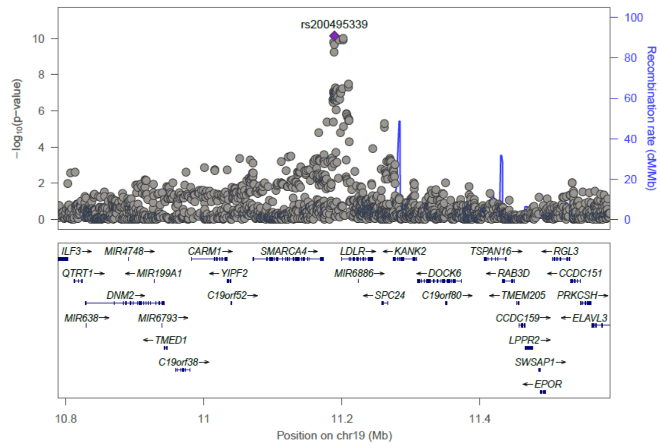
N


rs9632884
r²
0.8
0.6
0.4
0.2
−log10(p−value)
Recombination rate (cM/Mb)
MTAP
C9orf53
CDKN2A
CDKN2B−AS1
CDKN2B
DMRTA1
21.8
22
22.2
22.4
Position on chr9 (Mb)

O


rs113309773
−log10(p−value)
Recombination rate (cM/Mb)
ZNRF1
ZFP1
BCAR1
CFDP1
CHST6
ADAT1
LDHD
CTRB2
TMEM170A
GABARAPL2
CTRB1
CHST5
KARS
LOC100506281
TMEM231
TERF2IP
75.2
75.4
75.6
75.8
Position on chr16 (Mb)

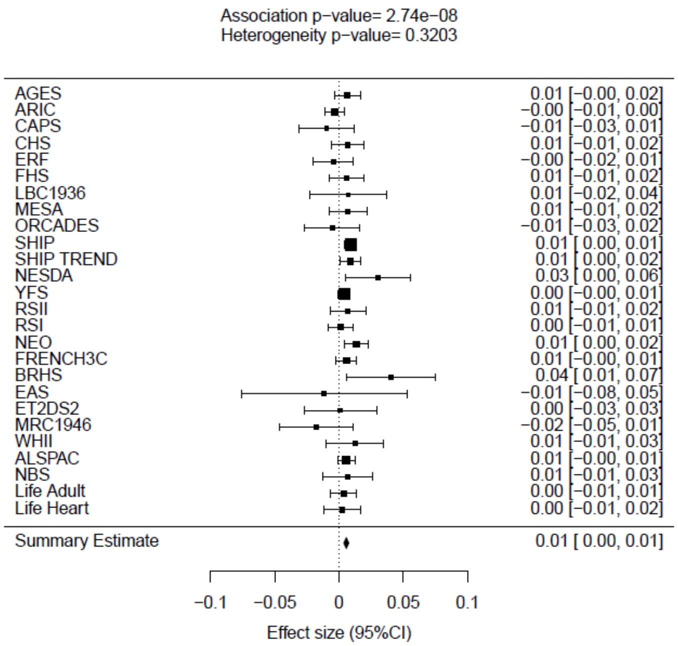
P


rs200495339
−log10(p−value)
Recombination rate (cM/Mb)
ILF3
MIR4748
CARM1
SMARCA4
LDLR
KANK2
TSPAN16
RGL3
QTRT1
MIR199A1
YIPF2
MIR6886
DOCK6
RAB3D
CCDC151
DNM2
C19orf52
SPC24
C19orf80
TMEM205
PRKCSH
MIR638
MIR6793
CCDC159
ELAVL3
TMED1
LPPR2
C19orf38
SWSAP1
EPOR
10.8
11
11.2
11.4
Position on chr19 (Mb)

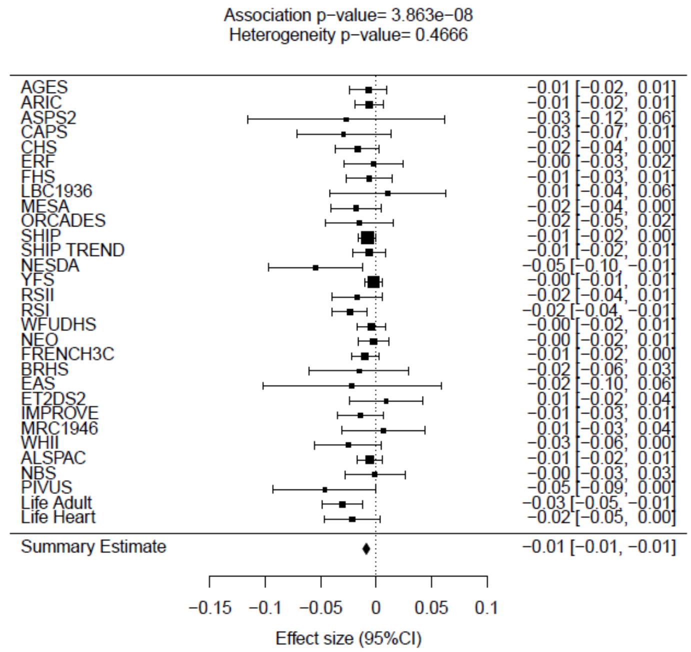
Supplementary Figure 3. Forest plots of SNPs significantly associated with cIMT or plaque

cIMT

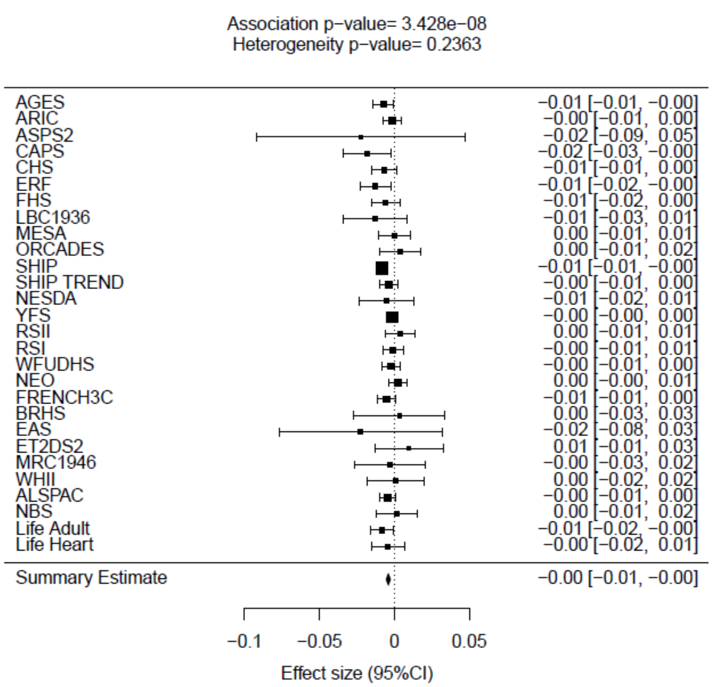
Chr1:208953176 (INDEL)



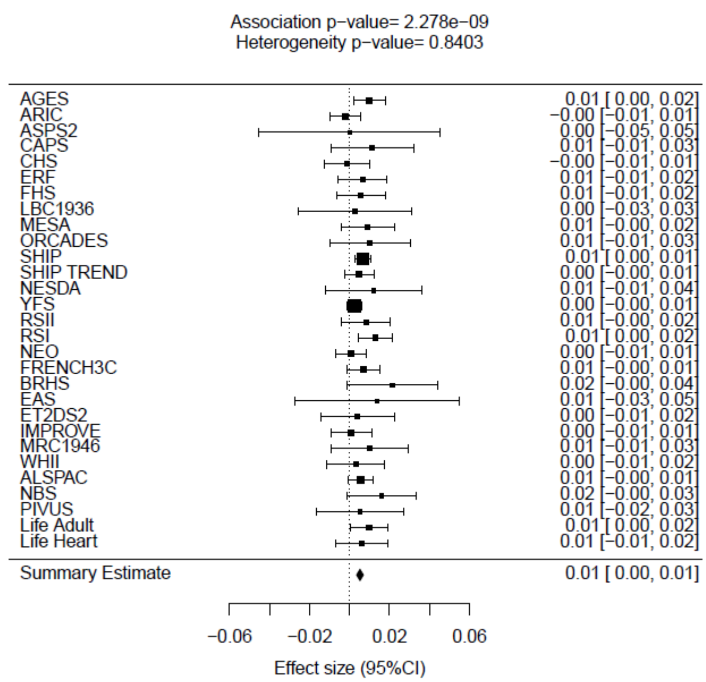
Chr5:81637916 (SNP)



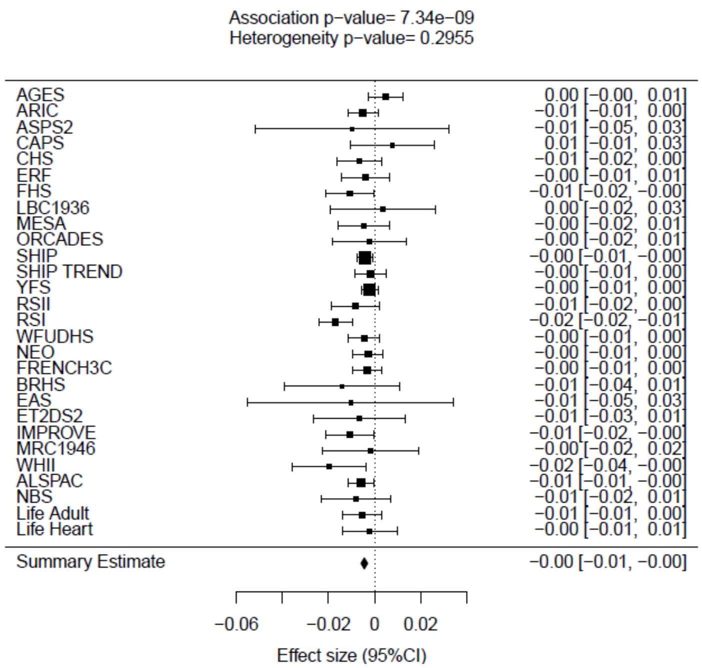
Chr6:143608968 (SNP)



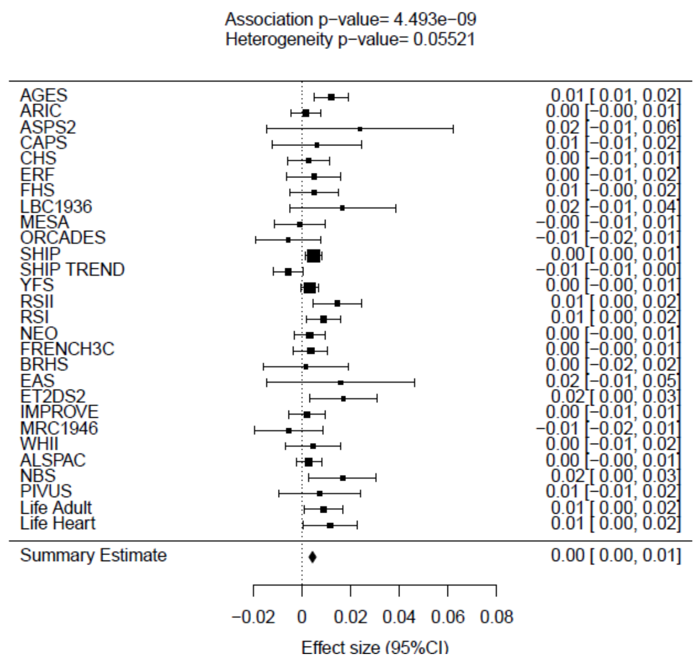
Chr7:106416467 (SNP)



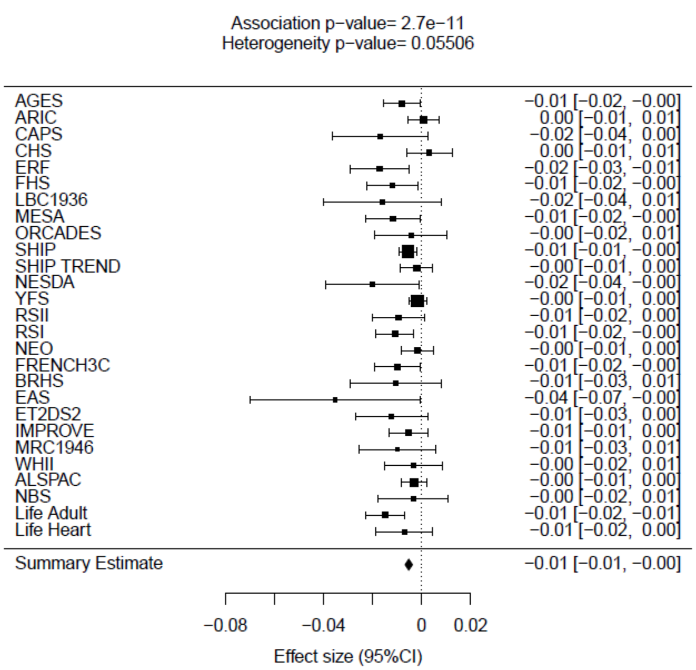
Chr8:6486033 (SNP)



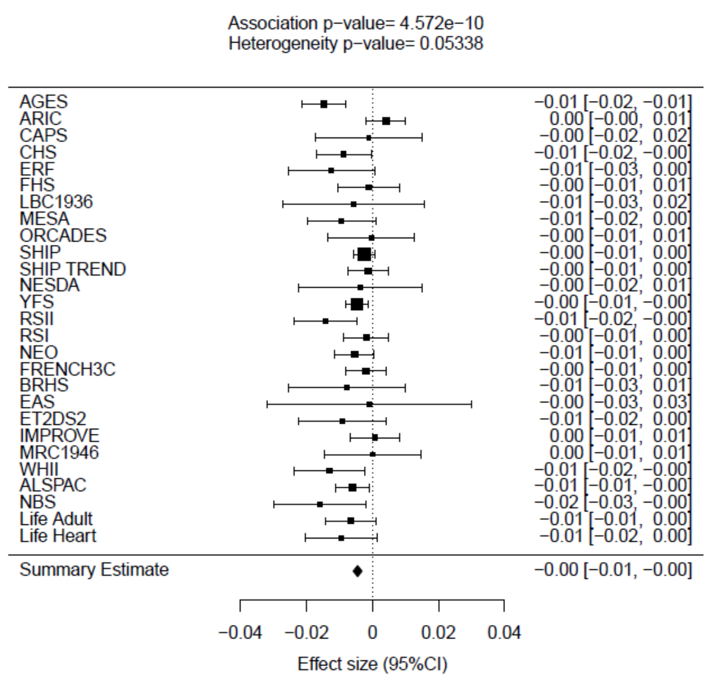
Chr8:8205010 (SNP)



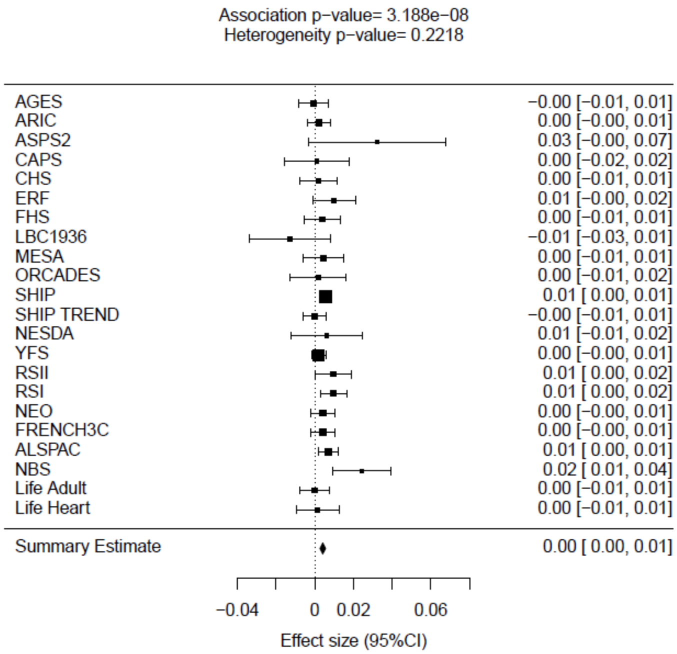
Chr8:10606223 (INDEL)



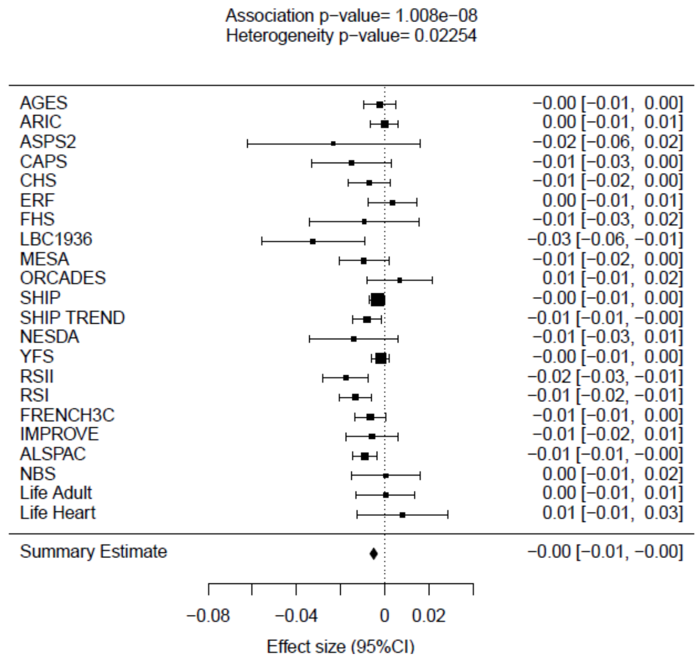
Chr8:123401537 (INDEL)



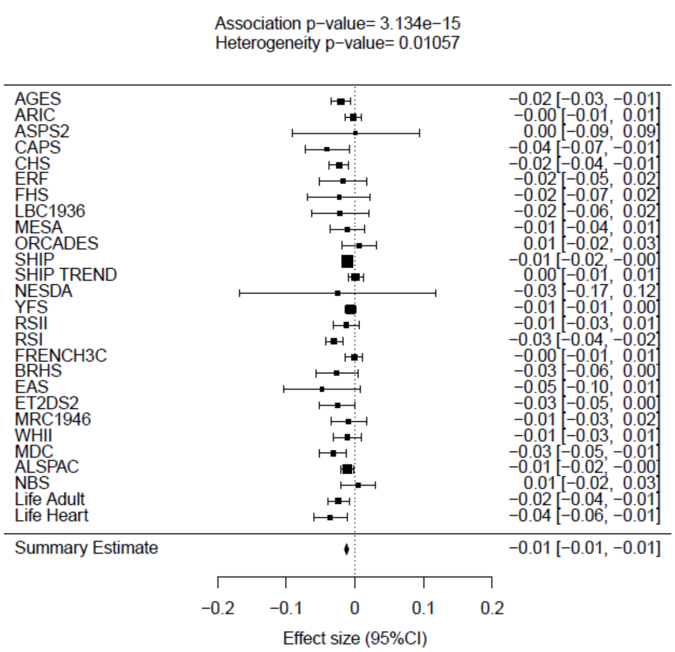
Chr10:114410998 (SNP)



Chr16:88966667 (SNP)

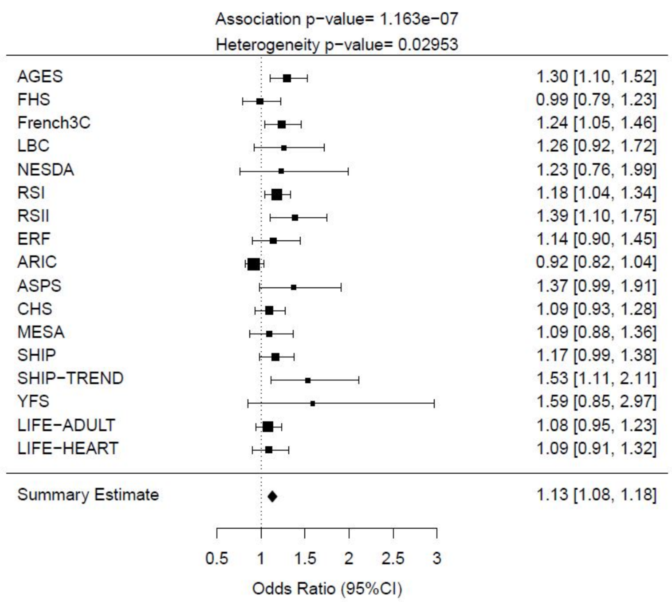


Chr19:45412079 (SNP)

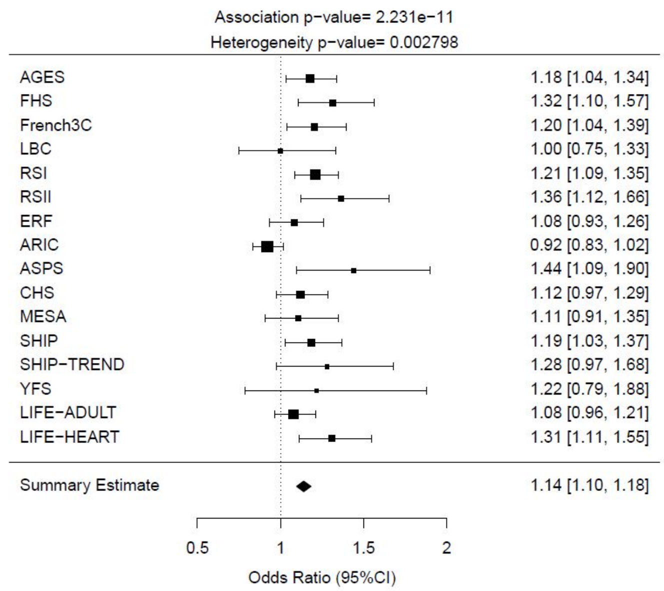


Plaque

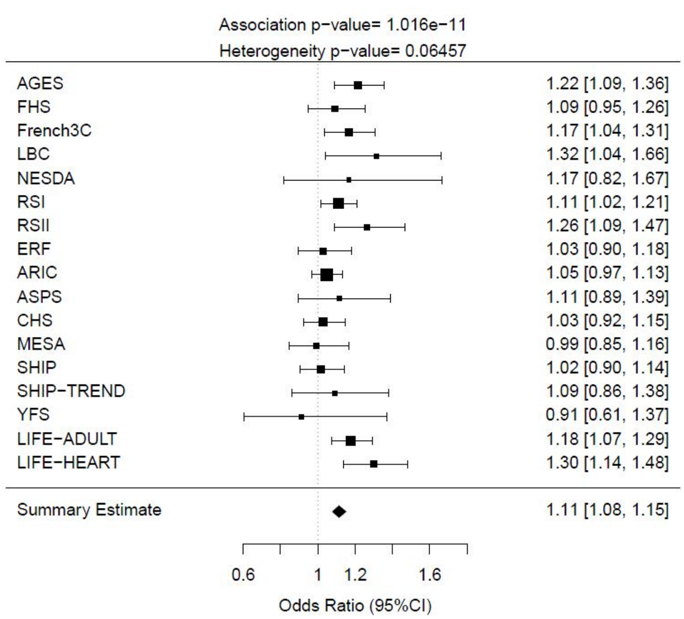
Chr4:148395284 (INDEL)



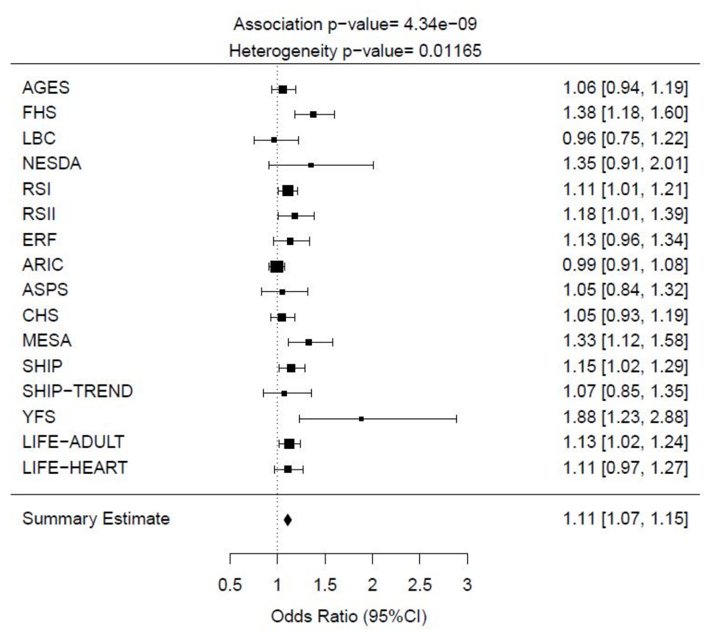
Chr7:106411858 (INDEL)



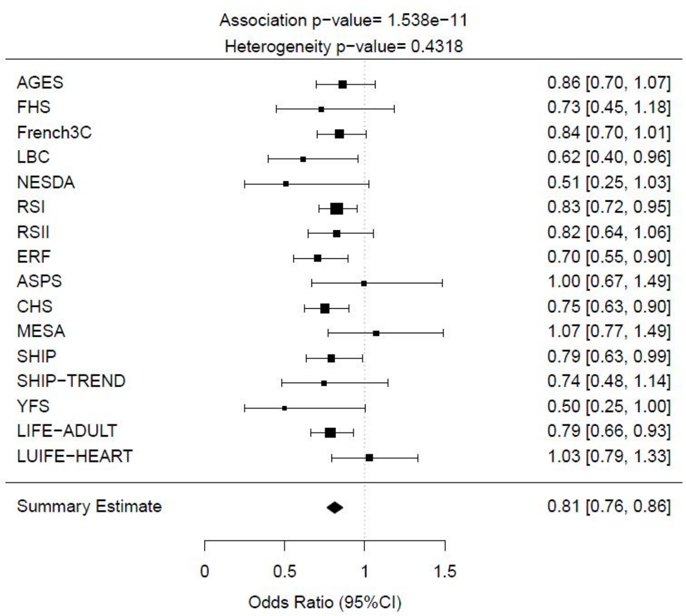
Chr9:22072301 (SNP)



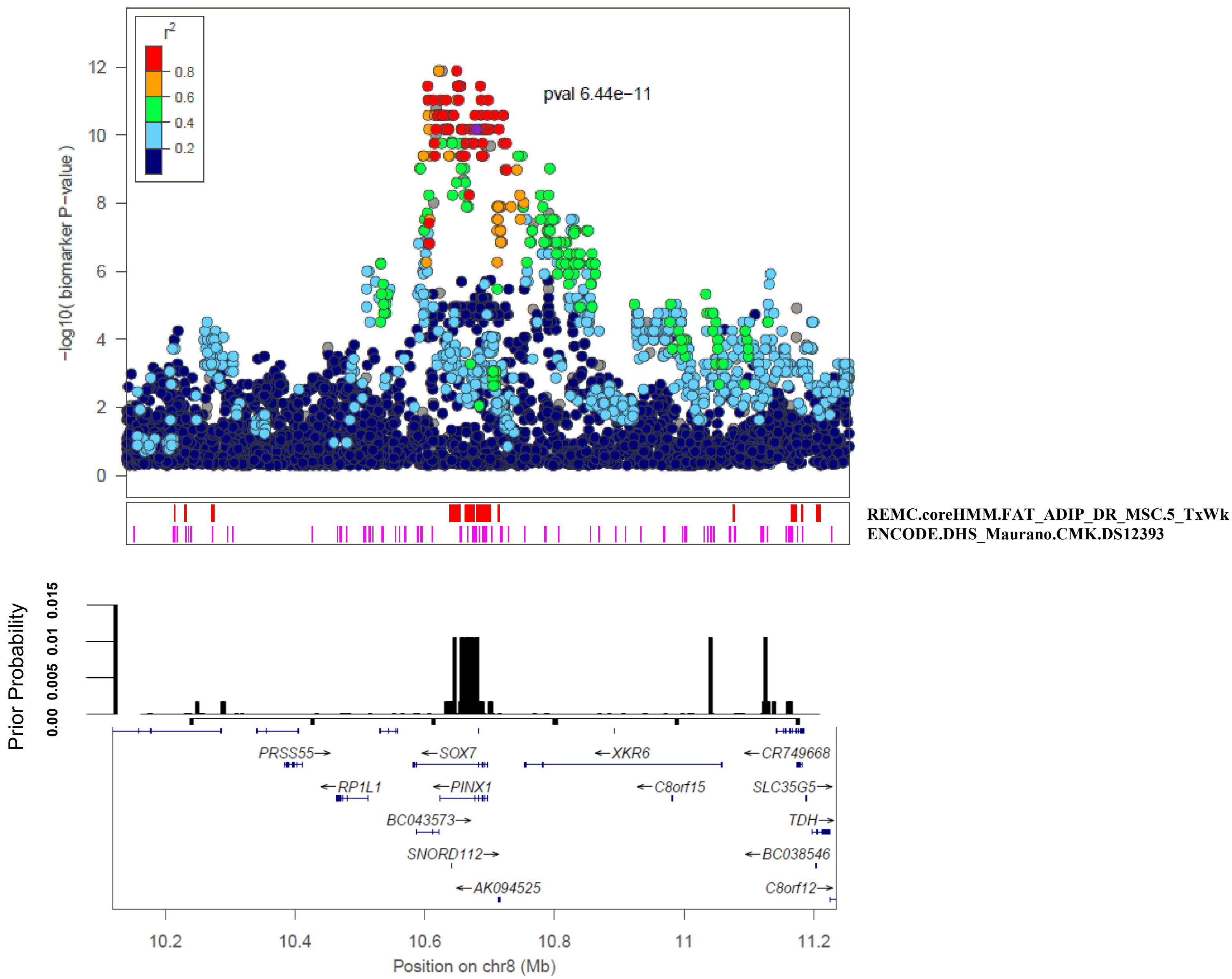
Chr16:75432688 (INDEL)



Chr19: 11189298 (INDEL)



Supplementary Figure 4. Regional plot surrounding the *PINX1* locus for cIMT. The top panel shows the P values for SNPs association with cIMT. The middle panel shows the overlaps of SNPs with annotations included in the combined model in fGWAS. The bottom panel shows the fitted empirical prior probability based on the fGWAS combined model. The SNP association shown in purple (chr8:10659406; P = 6.4 x 10^{-11}) falls within active transcription (REMC.coreHMM.FAT_ADIP_DR_MSC.5_TxWk) in Adipose Derived Mesenchymal Stem Cells and a DNaseI-hypersensitive site (ENCODE.DHS_Maurano.CMK.DS12393) leading the model to assign a higher probability compared to the index SNP (index SNP chr8:10606223:INDEL; P = 1.3 x 10^{-12}).

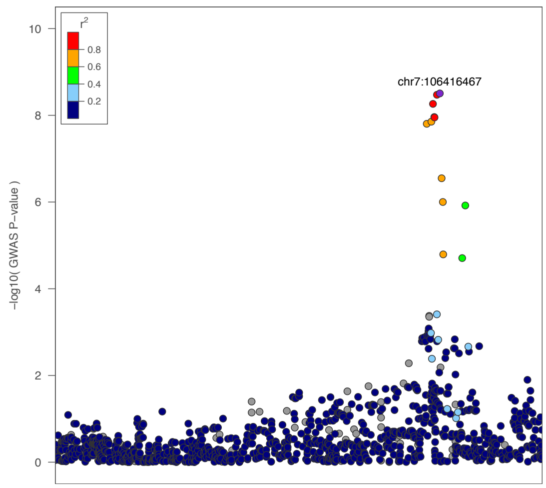


Supplementary Figure 5. Pairwise colocalization of GWAS SNPs and tissue eQTLs.

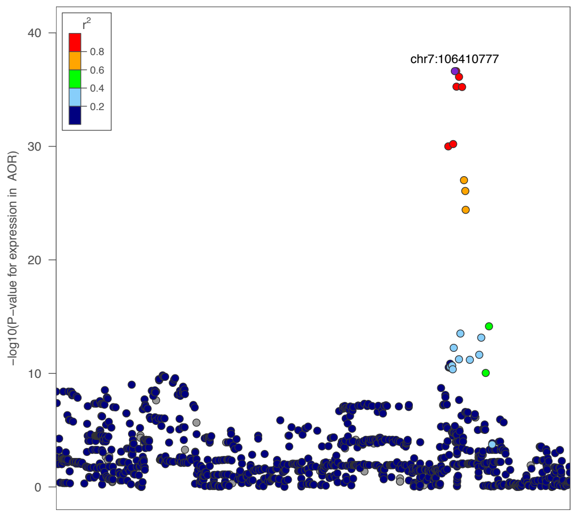
Colocalization results for cIMT (A) and AOR (B) and MAM (C) eQTLs

CCDC71L

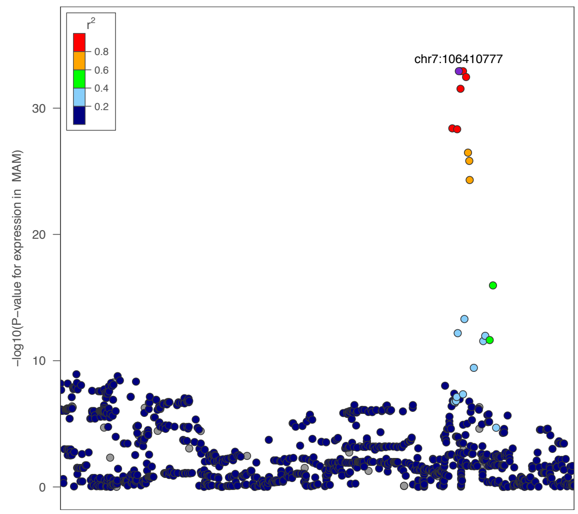
A



B

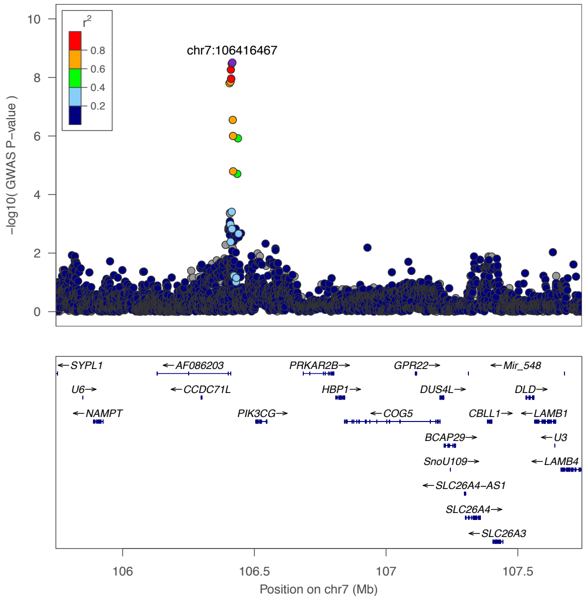


C

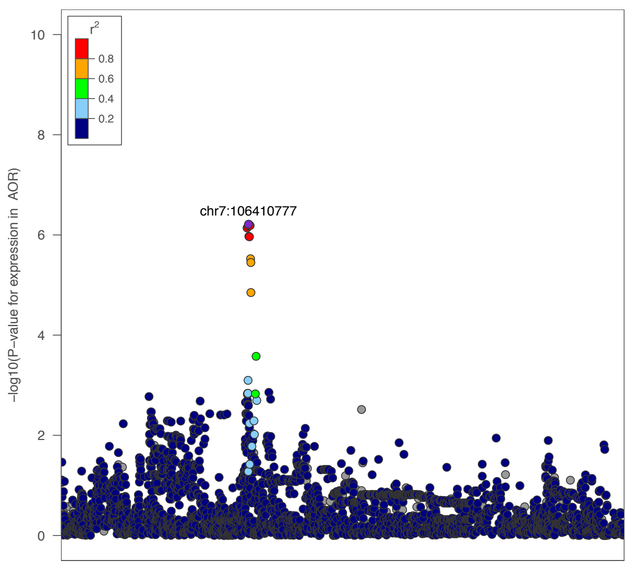


PRKAR2B

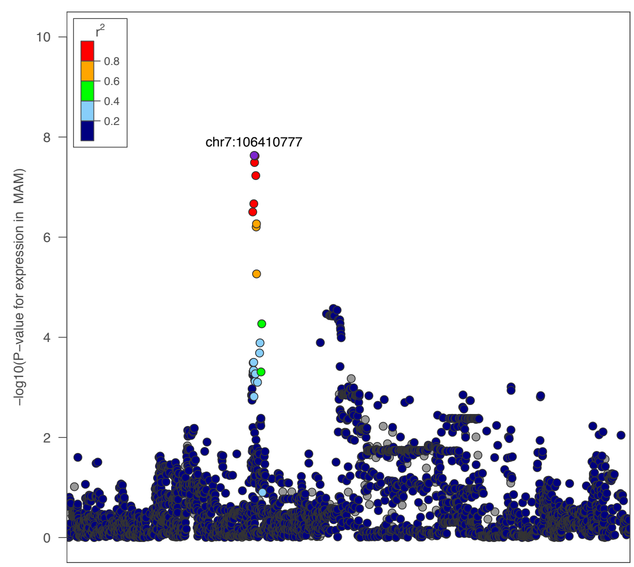
A


r2
0.8
0.6
0.4
0.2
chr7:106416467
−log10(GWAS P−value)
SYPL1
AF086203
PRKAR2B
GPR22
Mir_548
U6
CCDC71L
HBP1
DUS4L
DLD
NAMPT
PIK3CG
COG5
CBLL1
LAMB1
BCAP29
U3
SnoU109
LAMB4
SLC26A4−AS1
SLC26A4
SLC26A3
106
106.5
107
107.5
Position on chr7 (Mb)

B

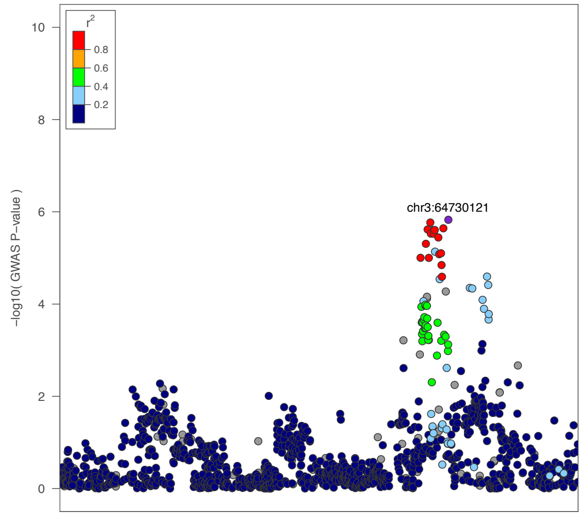

r2
0.8
0.6
0.4
0.2
chr7:106410777
−log10(P−value for expression in AOR)

C

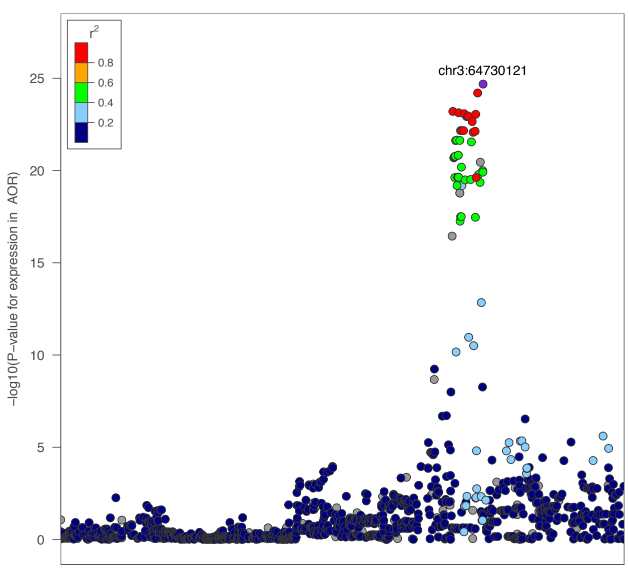

r2
0.8
0.6
0.4
0.2
chr7:106410777
−log10(P−value for expression in MAM)

ADAMTS9

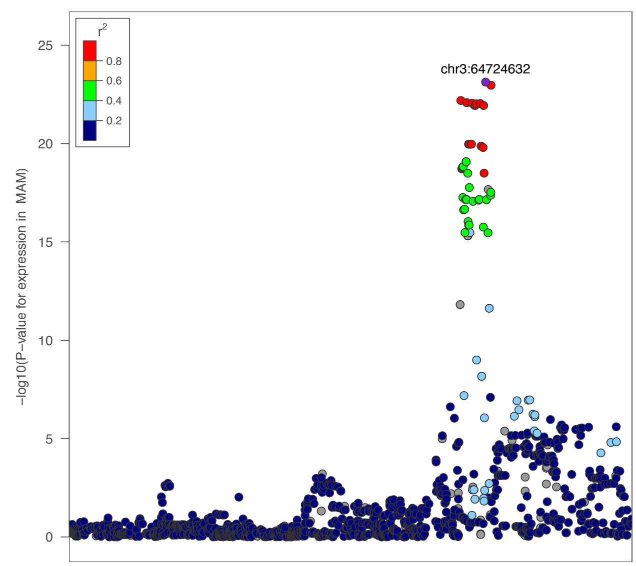
A


r2
0.8
0.6
0.4
0.2
chr3:64730121
−log10(GWAS P−value)
0
2
4
6
8
10

B

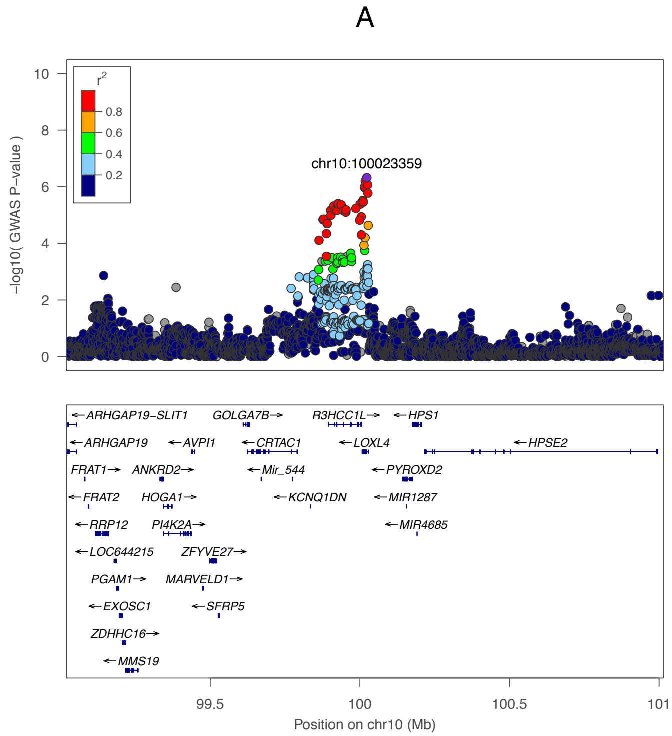

r2
0.8
0.6
0.4
0.2
chr3:64730121
−log10(P−value for expression in AOR)
0
5
10
15
20
25

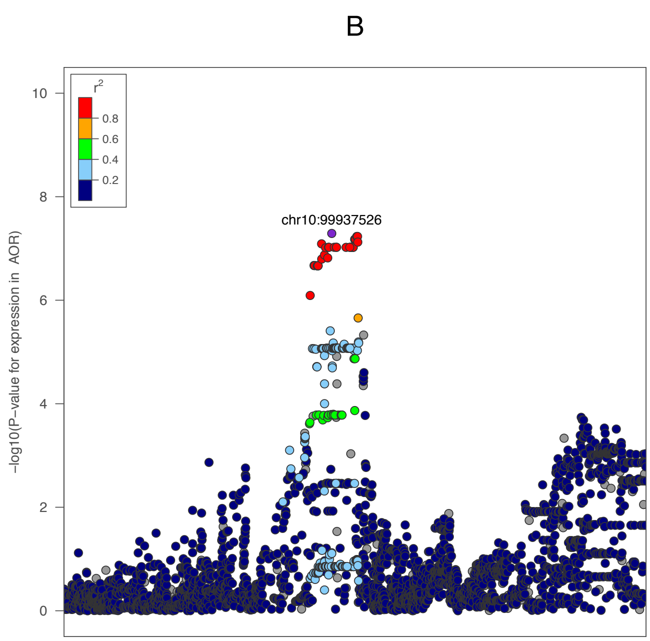
C


r2
0.8
0.6
0.4
0.2
chr3:64724632
−log10(P−value for expression in MAM)
0
5
10
15
20
25

Colocalization of cIMT (A), Aorta eQTL (B)

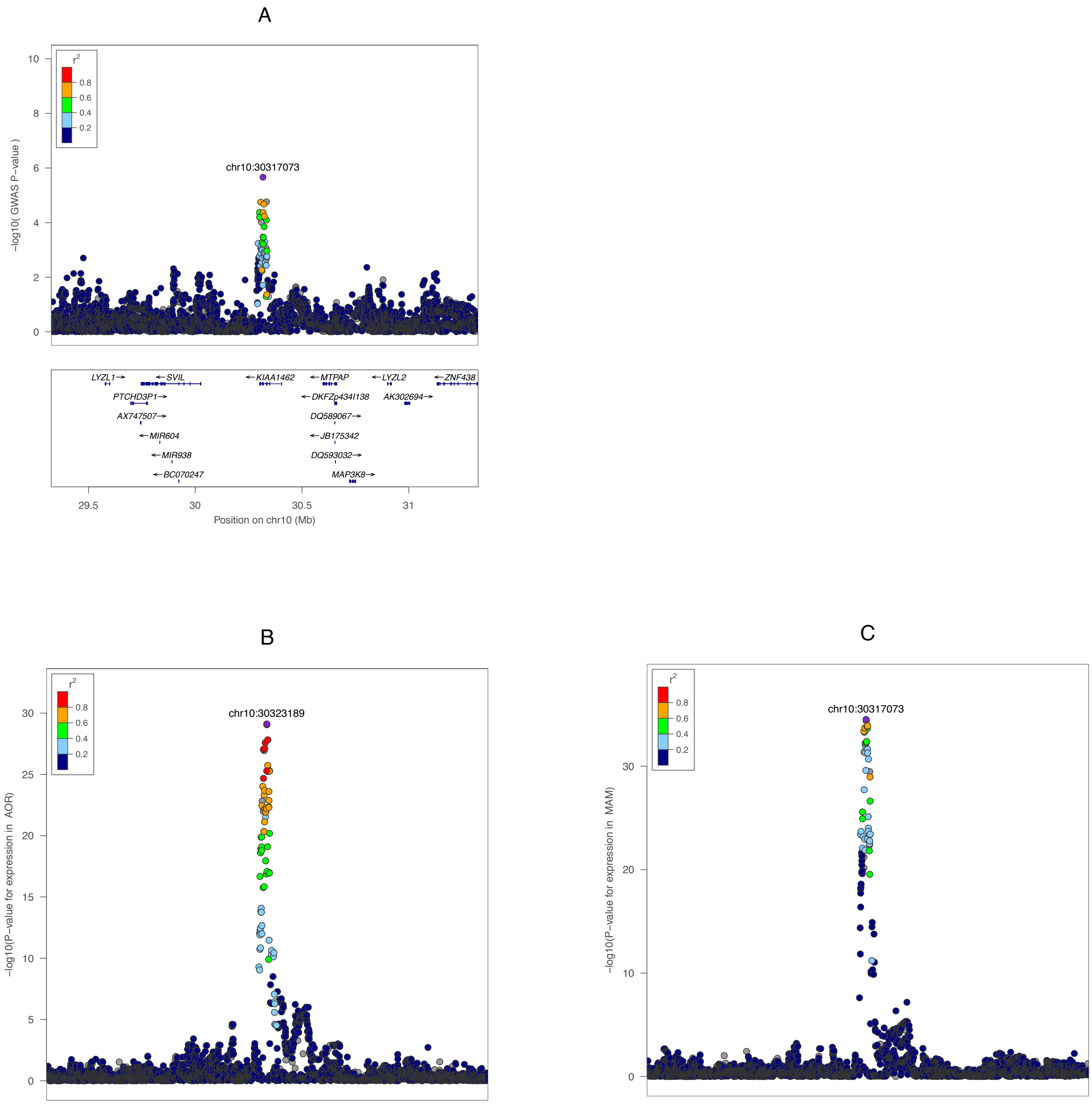
LOXL4



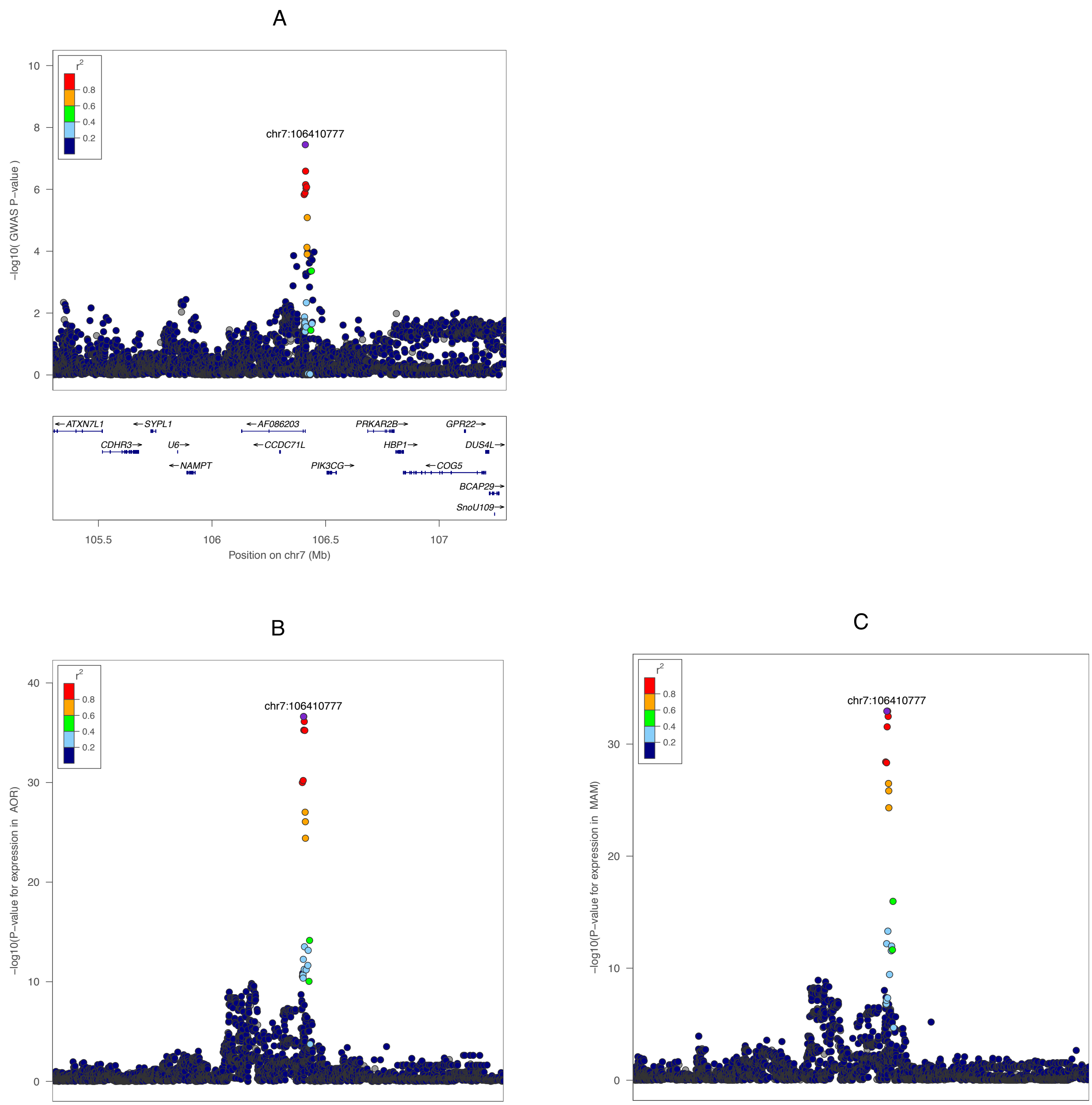


Colocalization of plaque (A), Aorta eQTL (B), and MAM eQTL (C)

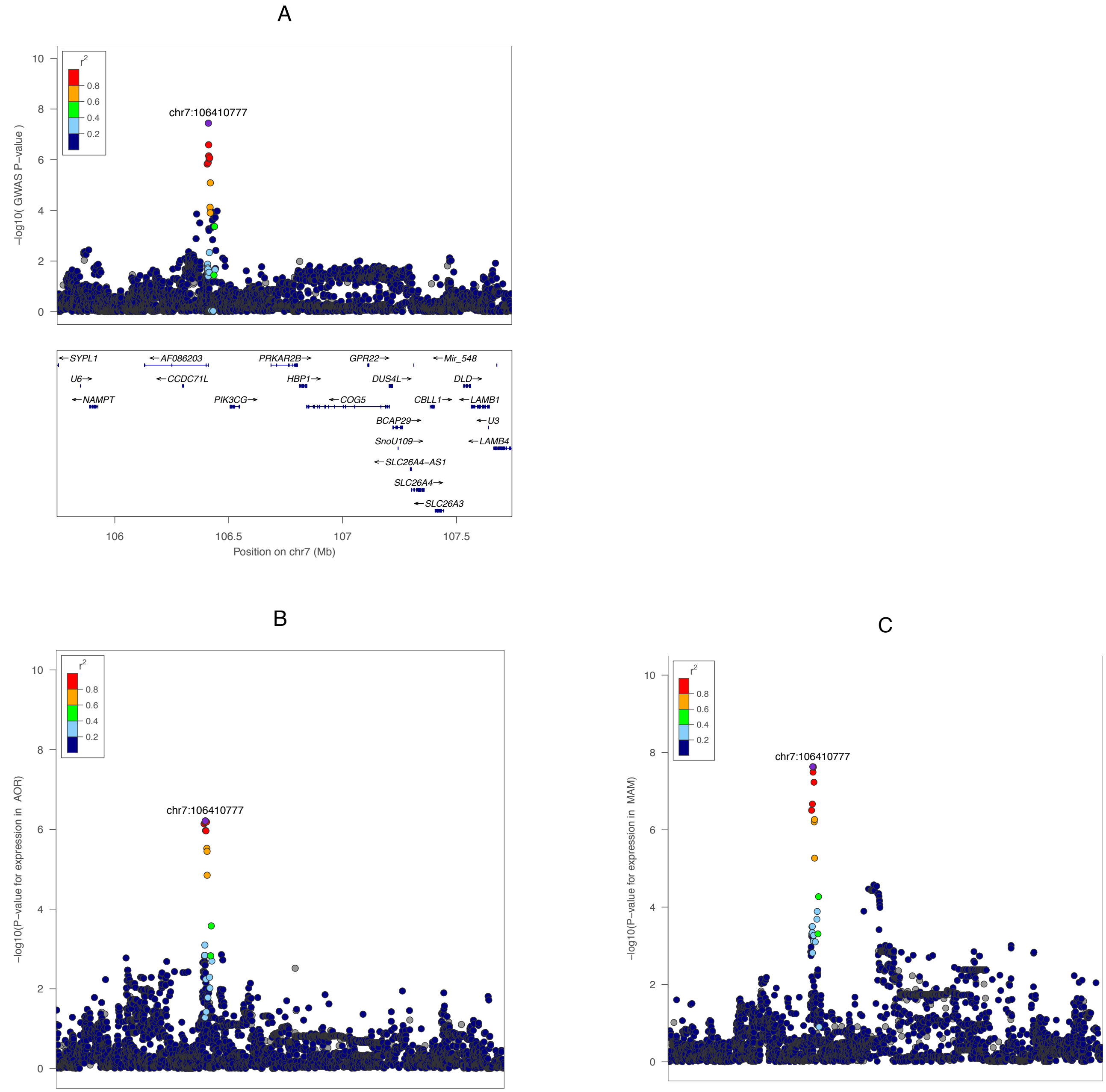
KIAA1462



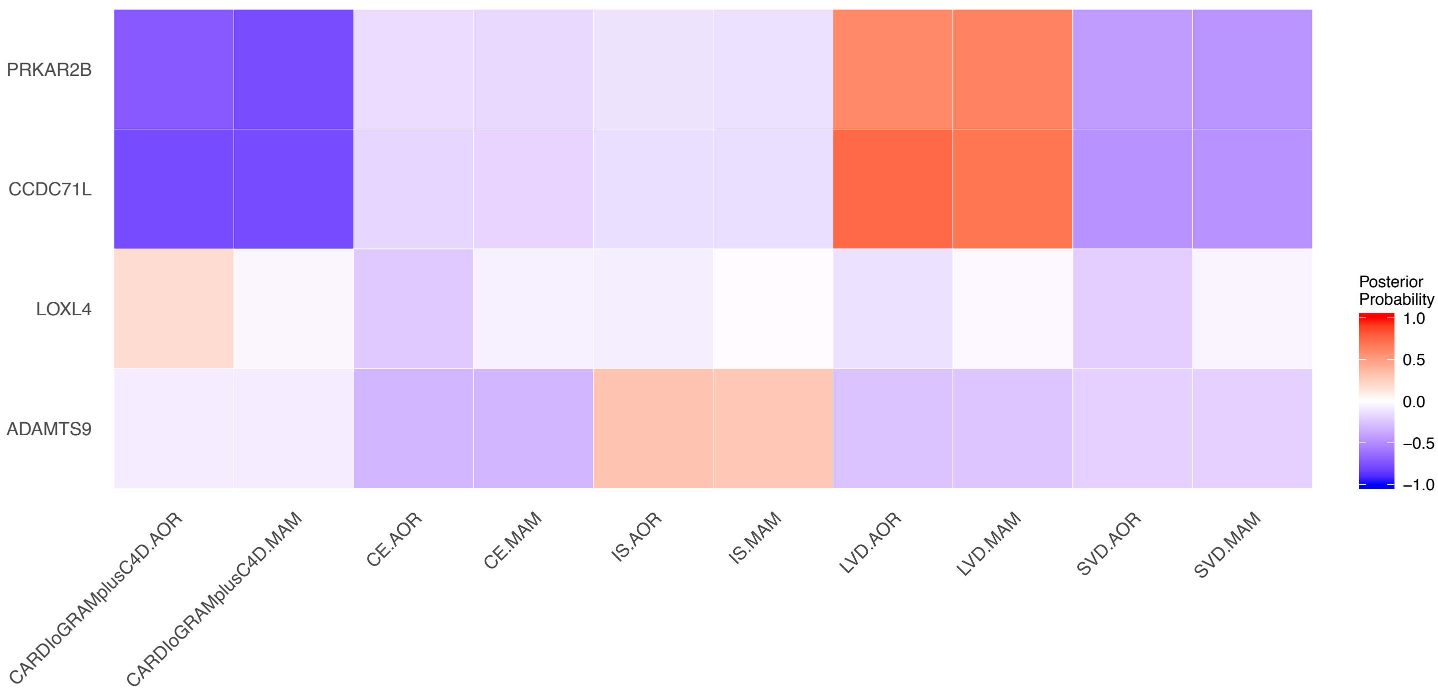
CCDC71L


A
r²
0.8
0.6
0.4
0.2
chr7:106410777
−log10(GWAS P−value)
ATXN7L1
SYPL1
AF086203
PRKAR2B
GPR22
CDHR3
U6
CCDC71L
HBP1
DUS4L
NAMPT
PIK3CG
COG5
BCAP29
SnoU109
105.5
106
106.5
107
Position on chr7 (Mb)
B
−log10(P−value for expression in AOR)
C
−log10(P−value for expression in MAM)

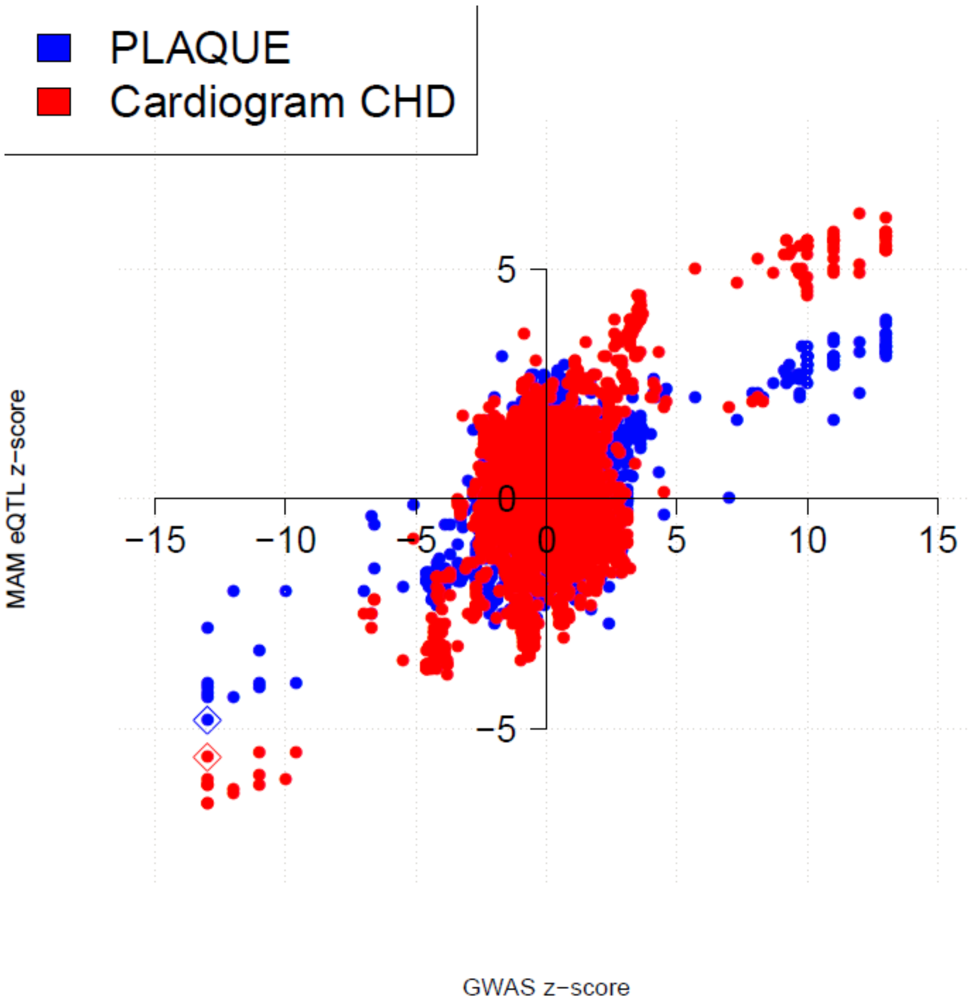
PRKAR2B


A
chr7:106410777
r2
0.8
0.6
0.4
0.2
−log10(GWAS P−value)
SYPL1
AF086203
PRKAR2B
GPR22
Mir_548
U6
CCDC71L
HBP1
DUS4L
DLD
NAMPT
PIK3CG
COG5
CBLL1
LAMB1
BCAP29
U3
SnoU109
LAMB4
SLC26A4−AS1
SLC26A4
SLC26A3
106
106.5
107
107.5
Position on chr7 (Mb)
B
chr7:106410777
−log10(P−value for expression in AOR)
C
chr7:106410777
−log10(P−value for expression in MAM)

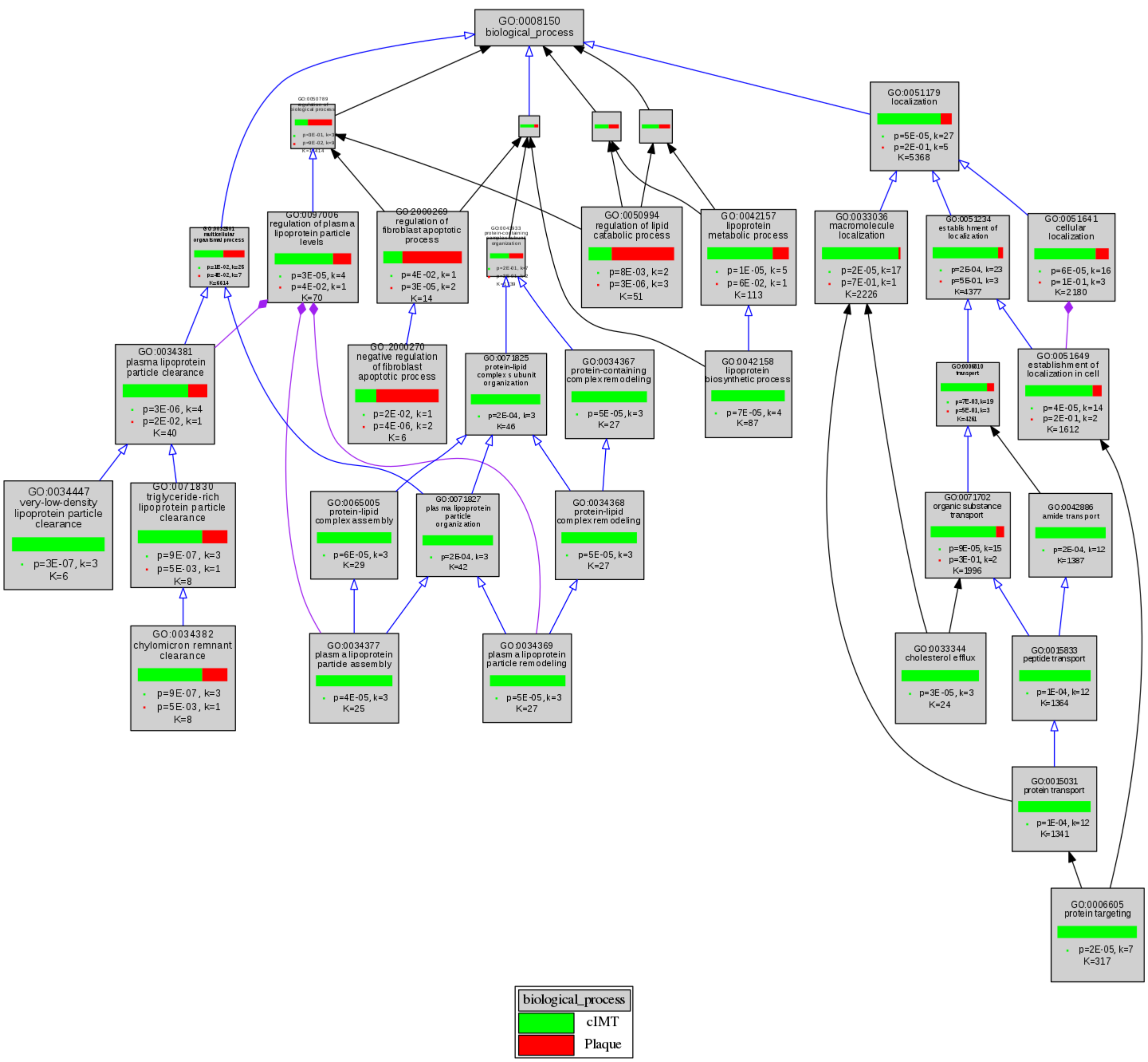
Supplementary Figure 6. Pairwise colocalization results for genes identified for cIMT and plaque GWAS meta-analysis with GWAS of coronary heart disease from CARDIOGRAMplusC4D and stroke subtypes from METASTROKE consortium. Posterior probability of colocalization is shown with red being a probability of colocalization of the same SNP and blue the high probability of no colocalization of the same SNP with clinical outcomes of coronary heart disease and stroke, or subtypes. Tissue expression in AOR (aortic root), MAM (mammary artery); stroke subtypes are IS (ischaemic stroke), CE (cardio-embolic stroke), LVD (large vessel disease), SVD (small vessel disease) as defined in Methods.



Supplementary Figure 7. Expression of *KIAA1462* gene in MAM, and plaque and CHD GWAS. Each dot represents effect size estimates from associations of gene expression of *KIAA1462* in MAM (x-axis) against associations of plaque (blue) or CHD (red) for SNPs at the *KIAA1462* locus. The diamond around the dots represent the SNP with the strongest association across all datasets (10:30317073, rs9337951).



Supplementary Figure 8. GO term enrichment analysis of the protein-coding genes identified as nearest to or in LD with the variants for cIMT and plaque (Table 1 and Supplementary Table 5), and genes identified in colocalization analyses (Table 2). GO term enrichment was performed using the VLAD tool to find GO terms that are significantly enriched in the list of genes identified compared to the human proteome. The larger the size of box, the more significantly enriched the term is; the significance is represented by the p-value after the term name. k = number of proteins in the input list that are annotated to the term, K = number of human proteins in total that are annotated to the term. The lines (edges) connecting the nodes in the graph represent the relationships between the terms. A purple line with a solid, diamond-shaped tip represents a “part-of” relationship between terms; a blue line with a hollow arrow tip represents an “is-a” relationship between the terms. A black line with a solid arrow tip indicates that several nodes in a multi-step path are not being displayed in order to simplify the graphic. See VLAD tool (https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4602057/).



Supplementary Tables

Supplementary Table 1. Characteristics of the study samples

Study	Sex (F/M)	Sample size cIMT	Age (years) mean (SD)	cIMT (mean, SD)	Sample size plaque	N plaques	Plaque frequency
AGES	1771/1297	3068	76.4 (5.4)	1.13 (0.16)	3053	2043	0.67
ARIC	4596/4067	8663	54.3 (5.7)	0.76 (0.18)	8857	1626	0.18
ASPS-FAM	176/127	303	65.5 (11.0)	0.84 (0.32)			
ASPS-FAM	439/334		65.9 (8.0)		773	490	0.63
CAPS	443/443	886	48.9 (13.3)	0.73 (0.19)			
CHS	1265/1975	3239	72.3 (5.4)	1.03 (0.20)	3125	2069	0.66
DHS	1 12/25	915	61.4 (9.5)	0.66 (0.12)			
ERF	1507/1214	2270	48.7 (14.4)	0.82 (0.20)	2443	1218	0.50
FHS	1601/1403	3004	58.5 (9.7)	1.02 (0.18)	3008	530	0.18
3C-Dijon	1581/937	2518	72.6 (4.0)	0.69 (0.11)	2473	1218	0.49
LBC1936	363/396	759	72.8 (0.8)	0.85 (0.19)	759	220	0.29
MESA	1309/1198	2500	62.6 (10.3)	0.87 (0.20	2492	393	0.16
NEO	2949/2726	5675	56.0 (5.9)	1.00 (0.16)			
NESDA	368/204	572	44.7 (12.2)	0.66 (0.16)	572	86	0.15
ORCADES	763/1128	1914	53.7 (14.9)	0.50 (0.10)			
RS I	2968/1978	4946	69.0 (8.8)	1.02 (0.21)	4910	2920	0.59
RS II	1079/901	1980	64.7 (7.9)	0.99 (0.17)	2016	1509	0.75
SHIP	1838/1781	3619	53.3 (13.7)	0.85 (0.20)	3666	1989	0.54
SHIP-TREND	551/432	983	50.1 (13.7)	0.73 (0.17)	985	338	0.34
ALSPAC	3200/0	3200	47.9 (4.5)	0.55 (0.11)			
YFS	1106 /909	2015	37.7 (5.0)	0.66 (0.10)	2013	48	0.02
BRHS	0/889	889	78.7 (4.8)	0.79 (0.18)			
EAS	378/353	731	69.8 (5.6)	0.75 (0.18)			
ET2DS	423/445	868	68.9 (4.2)	0.94 (0.11)			
IMPROVE	1753/1636	3389	64.5 (1.9)	0.85 (0.07)			
LIFE-Adult	1677/1531	3208	59.1 (11.9)	0.76 (0.15)	4534	2726	0.60
LIFE-Heart	684/1240	1924	62.5 (11.0)	0.78 (0.15)	2755	2117	0.77
MDC	1093/1050	2142	57.4 (6.0)	0.73 (0.144)			
MRC1946	655/603	1258	63.3 (1.1)	0.68 (0.18)			
NBS	281/268	549	57.8 (5.2)	0.86 (0.11)			
PIVUS	482/482	964	70.2 (0.2)	0.88 (0.16)			
WHII	508/1669	2177	60.8 (5.9)	0.77 (0.19)			

Supplementary Table 2. Study definition of carotid artery plaque		
Study	**Plaque definition**	**Reference (PMID)**
AGES	Of the left and right carotid bifurcation and internal carotid artery the presence of atherosclerotic lesions was be quantified during the ultrasound examination. The most severe lesion per segment was assessed in a semi-quantitative manner as none, minimal, moderate and severe lesion.	17351290
ARIC	Presence of a lesion defined by abnormal arterial wall thickness, shape, or texture. Acoustic shadowing defined as a reduction in amplitude of echoes caused by intervening structures with high attenuation.	9180252
ASPS	Plaque was graded according to the most severe visible changes in the CCA and ICA as 0, normal; 1,vessel wall thickening >1 mm; 2, minimal plaque (<2 mm); 3, moderate plaque (2 to 3 mm); 4, severe plaque (>3 mm), and 5, lumen completely obstructed	7800110;10408549
CHS	Largest focal lesion classified by surface characteristics, echogenicity, and texture. A discernible focal widening of the wall relative to adjacent segments with or without protrusion into the lumen was described according to the following criteria: surface—smooth, mildly irregular, markedly irregular, or ulcerated; morphology—homogeneous or heterogeneous; and densi- ty—hypodense, isodense, hyperdense, or calcified.	1669507
ERF	The cIMT 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.	15845033
FHS	Defined by carotid stenosis of 25% or greater.	
LBC1936	We measured carotid flow velocities, maximum stenosis affecting the internal carotid artery/bulb/CCA Plaques were defined by carotid stenosis of 25% or greater.	22253310
Life Adult & Life Heart	Carotid artery plaque was defined as echogenic thickening of intimal reflection that extends into the arterial lumen at least 0.5 mm or 50 % of the surrounding CCA-IMT value or an intimal + medial thickness of >1.5 mm. Plaque presence was documented as ‘present’ or ‘absent’ for the common part and bulb of the right and left carotid artery, respectively.	26362881
MESA	Defined by carotid stenosis of 25% or greater.	12397006
NESDA	Widening of the intimal and medial layers relative to adjacent segments, with the area of focal increased thickness ≥ 1.10 mm	18763692;19065144 21745125
RS-I	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.	19728115
RS-II	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.	19728115
SHIP/SHIP-TREND	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.	11565448; 20167617
3C-Dijon	The presence of plaques was defined as localized echo structures encroaching into the vessel lumen for which the distance between the media–adventitia interface and the internal side of the lesion was >1 mm on the common carotid arteries, the carotid bifurcations, and the internal carotid arteries.	14598854; 18063810

YFS	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.	18263651

Supplementary Table 3. Study-specific genotyping, quality control, imputation and analysis

Study	Genotyping array	Sample quality control		Imputation		Association analysis			
		Call rate	Other exclusions	software	Reference panel	software	covariates	λGC cIMT	λGC plaque
AGES	Illumina 370CNV BeadChip	< 0.97	HWE p-value < 10^-6	MaCH	1,000 Genomes Phase I integrated release March 2012 (v3)	ProbAbel	age,sex	0.984	1.173
ARIC	Affymetrix 6.0	<0.95	HWE p-value <10^-5, MAF<0.01	IMPUTE2	1,000 Genomes Phase I integrated release March 2012 (v3)	FAST	age, sex, region, 10 PCs	1.017	1.028
ASPS	Illumina Human610-Quad BeadChip	< 98%	HWE p-value 1<10^-6, MAF<0.01, sex mismatch, cryptic relatedness	IMPUTE2	1,000 Genomes Phase I integrated release March 2012 (v3)	plink	age, sex	1.000	1.013
ASPS-Fam	Affymetrix Genome-Wide Human SNP Array 6.0	< 98%	HWE p-value 5<10^-6, MAF<0.05, sex mismatch, cryptic relatedness	IMPUTE2	1,000 Genomes Phase I integrated release March 2012 (v3)	GWAF	age, sex	1.005	1.018
CAPS	Affymetrix 6.0	<0.90	HWE p-value 1<10^-6, MAF<0.01, sex mismatch, cryptic relatedness	SHAPEIT v2.778 (phasing) and IMPUTE2 2.3.0 (imputation)	1,000 Genomes Phase I integrated release March 2012 (v3)	plink2	age,sex,pc1,pc2,pc3,pc4	1.008	1.019
CHS	Illumina 370CNV BeadChip + Illumina IBC iSELECT	<0.95	HWE p-value <10^-5	MACH/miniMACH (whites) & IMPUTE v 2.2.2 (African Americans)	1,000 Genomes Phase I integrated release March 2012 (v3)	R	age, sex, clinic	1.026	NA
ERF	Illumina 318/370 K, Affymetrix 250 K, and Illumina 6 K	95%	HWE < 10^-6, MAF < 0.01, snp call rate < 98%, Mendelian errors	miniMACH	1,000 Genomes Phase I integrated release March 2012 (v3)	R, GenABEL, ProbABEL	age, sex (family structure)	0.997	1.057
FHS	Affymetrix 500K	<0.95	HWE p-value <10^-6	MaCH/mimimac	1,000 Genomes Phase I integrated release March 2012 (v3)	R, GEE for dichotomous, LME for continuous trait	Age at the examination cycle 6, sex, and 10 PCs	1.007	NA
3C-Dijon	Illumina Human610 Quad	≤0.95		minimac	1,000 Genomes Phase I integrated release March 2012 (v3)	ProbAbel/R	Age_baseline, Sex, PC1, PC2, PC3, PC4	1.018	1.006

LBC	Illumina 610 Quad V1	<95%	HWE P <10-3, relatedness, MAF<1%, gender mismatch, SNP call rate <98%	MiniMAC	1,000 Genomes Phase I integrated release March 2012 (v3)	Mach2qtl	age, sex, 4 PCs	1.025	1.001
MESA	Affymetrix 6.0	< 0.95	HWE p-value < 10^-8; MAF < 0.005	IMPUTE2	1,000 Genomes Phase I integrated release March 2012 (v3)	SNPTEST V2.4	age, gender, site, and 4 PCs	1.017	1.014
NEO	Illumina HumanCoreExome-24v1_A Beadchip	<0.98	HWE P < 1e-5	IMPUTE2	1,000 Genomes Phase I integrated release March 2012 (v3)	ProBABEL	age, sex, 4 PCs	1.009	1.03
NESDA	Affymetrix 6.0 907K	0.9	heterozygosity abs(PLINK F)>0.1; sex mismatch; unexpected relatedness	minimac	1,000 Genomes Phase I integrated release March 2012 (v3)	SNPTEST	age, sex	1.023	NA
ORCADES	Illumina HumanHap300	0.98	HWE p-value <10^-6	MaCH	1,000 Genomes Phase I integrated release March 2012 (v3)	ProbAbel	age, sex, 3 PCs	1.001	0.992
RS I	Illumina 550K	0.975	HWE p-value <10^-6, MAF<0.001	MaCH/minimac	1,000 Genomes Phase I integrated release March 2012 (v3)	ProbAbel		1.013	NA
RS II	Illumina 550K	0.975	HWE p-value <10^-6, MAF<0.001	MaCH/minimac	1,000 Genomes Phase I integrated release March 2012 (v3)	ProbAbel		0.996	NA
SHIP	Affymetrix 6.0	0.92	duplicates, reported/genotyped gender mismatch	IMPUTE2	1,000 Genomes Phase I integrated release March 2012 (v3)	Quicktest	age, sex	1.012	NA
SHIP-TREND	Illumina Human Omni 2.5	0.94	duplicates, reported/genotyped gender mismatch	IMPUTE2	1,000 Genomes Phase I integrated release March 2012 (v3)	Quicktest	age, sex	0.99	NA
DHS	Affymetrix Genome-Wide Human SNP Array 5.0	0.95	Excluded sample failures, sex discordance, unclear/unexpected sibling relationships (based on IBD)	IMPUTE2	Phase I 1000G Integrated Variant Set version 2, cosmopolitan (integrated) reference panel	SOLAR 6.3.6	age, sex, first two admixture PCs	1.054	NA
YFS	Illumina Human670-QuadCustom	0.95	HWE p-value <10^-6, MAF<0.01	SHAPEIT v1 and IMPUTE2	1,000 Genomes Phase 1 CEU haplotype set	SNPTEST		1.016	1.007
ALSPAC	Illumina human660W-quad array	0.95	HWE p-value <10^-6, MAF<0.01, gender mismatch (X	IMPUTE2	1,000 Genomes Phase I integrated release March 2012 (v3)	SNPTEST V2.5	age, 10 PCs	1.013	NA

			chromosome heterozygosity or extreme autosomal heterozygosity); unexpected relatedness (Pi hat of more than 0.125 which is expected to correspond to roughly 12.5% alleles shared IBD); population stratification (determined by IBS).						
BRHS	MetaboChip	0.95	HWE p-value <10^-6, MAF<0.01, gender mismatch and relatedness	MaCH/minimac	1,000 Genomes Phase I integrated release March 2012 (v3)	snpStats	age, sex	0.996	NA
EAS	MetaboChip	0.95	HWE p-value <10^-6, MAF<0.01, gender mismatch and relatedness	MaCH/minimac	1,000 Genomes Phase I integrated release March 2012 (v3)	snpStats	age, sex	1.012	NA
ET2DS	MetaboChip	0.95	HWE p-value <10^-6, MAF<0.01, gender mismatch and relatedness	MaCH/minimac	1,000 Genomes Phase I integrated release March 2012 (v3)	snpStats	age, sex	0.99	NA
IMPROVE	Combined Metabochip and Immunochip	0.95	HWE p-value <5^-7, MAF<0.01, gender mismatch and relatedness	MACH1/Minimac	1,000 Genomes Phase I integrated release March 2012 (v3)	Plink	age, sex, 3 PC	1.054	NA
LIFE-Adult	Affymetrix Axiom	< 0.97	HWE p-value <10^-6, MAF=0, gender mismatch and relatedness	IMPUTE2	1,000 Genomes Phase I integrated release March 2012 (v3)	SNPTEST v2.5	age, sex	1.016	1.007
LIFE-Heart	Affymetrix Axiom	< 0.97	HWE p-value <10^-6, MAF=0, gender mismatch and relatedness	IMPUTE2	1,000 Genomes Phase I integrated release March 2012 (v3)	SNPTEST v2.5	age, sex	1.010	1.004
MRC1946	MetaboChip	0.95	HWE p-value <10^-6, MAF<0.01, gender mismatch and relatedness	MaCH/minimac	1,000 Genomes Phase 1 CEU haplotype set	snpStats	age, sex	0.995	NA

MDC	MetaboChip	0.95	sex-mismatches, relatedness; SNP QC: callrate<95%; HWE p-value <10^-6	N/A	NA	Plink	age, sex	1.017	NA
NBS	Illumina HumanHapCNV 370-Duo BeadChip	0.95	HWE p-value <10^-4, MAF<0.01, gender mismatch and relatedness	IMPUTE2	1000 Genomes phase1 v3 together with Genome of The Netherlands (GoNL) release 5	snpStats	age, sex	0.999	NA
PIVUS	MetaboChip	0.95	HWE p-value <10^-6, MAF<0.01, gender mismatch and relatedness	IMPUTE2	HapMap2	Plink	age, sex	1.013	NA
WHII	MetaboChip	0.95	HWE p-value <10^-6, MAF<0.01, gender mismatch and relatedness	MaCH/minimac	1,000 Genomes Phase I integrated release March 2012 (v3)	snpStats	age, sex	0.995	NA

NA, not available

Supplementary Table 4. Conditional analysis using GCTA for cIMT and plaque

Trait	SNP	chr:position	RefAllele	Freq	Beta	SE	*p*	n	freq_geno	bJ	bJ_se	pJ	LD_r
cIMT	rs2912064	8:6488710	T	0.36	-0.0041	0.0007	1.58E-08	76752	0.64775	-0.0041	0.0007	4.30E-09	0.00333
cIMT	rs11785239	8:8205010	T	0.65	0.0044	0.0008	4.48E-09	59522.7	0.35388	0.00441	0.0008	3.47E-08	0

Columns are: freq_geno: frequency of the effect allele in the reference sample;
bJ, bJ_se, pJ: effect size, standard error and p-value from a joint analysis of all the selected SNPs;
LD_r: LD correlation between the SNP i and SNP i + 1 for the SNPs on the list.

Supplementary Table 5.Loci associated with cIMT and plaque GWAS at p< 10^{-7} among individuals of European ancestry

SNP	chr:position	Nearest Coding Gene	Alleles Effect/ Other	Effect allele frequency	Beta (SE)	*p*	N
cIMT							
rs515135	chr2:21286057	*APOB*	T/C	0.18	-0.0487 (0.0009)	8.2 x 10-8	65,428
rs139302128	chr2:242594226	*ATG4B*	T/C	0.03	0.0487 (0.0091)	7.6 x 10-8	17,713
Plaque							
rs4779614	chr15:33540117	*TMCO5B*	T/C	0.35	0.0869 (0.0171)	4.0 x 10-7	48,434
rs259140	chr7:89624347	*STEAP2-AS1*	T/G	0.30	0.0889 (0.0177)	5.1 x 10-7	47,862

Supplementary Table 6. Nearest gene from top GWAS SNP of cIMT and plaque, and best colocalizing gene intersecting a region of 200kb from the listed SNP using STARNET tissue eQTL. Main association are SNPs with p-value < $5x10^{-8}$ (Table 1); Suggestive association are $p<10^{-7}$ (Table S5). Best colocalizing gene is the gene with the largest posterior probability of colocalization in the joint GWAS and eQTL analysis. ngenes is the max number of genes considered across the tissues in a region of +/-200kb around the GWAS SNP.

SNP	chr:position	Nearest Coding Genes	Max number of genes across tissues considerd in region (+/- 200Kb from GWAS SNP)	Max number of genes across tissues suggestive of association (PP3 + PP4 ^{3} 50%) with both GWAS and STARNET eQTLs in region (+/- 200Kb from GWAS SNP)	Best co-localizing gene (eQTL STARNET data, PP4)
cIMT					
Main					
rs13225723	chr7:106416467	*PIK3CG, CCDC71L, PRKAR2B*	20	6	*CCDC71L* (AOR, PP4=97.48)
rs148147734	chr8: 123401537	*ZHX2*	10	4	*HAS2* (AOR, PP4=53.58)
rs6907215	chr6:143608968	*AIG1*	18	7	*ENSG00000217648* (MAM, PP4=47.77)
rs7412	chr19: 45412079	*APOE*	61	28	*APOE* (SF, PP4=24.84)
rs2912063	chr8: 6486033	*MCPH1, ANGPT2*	5	4	*ENSG00000249898* (AOR, PP4=27.36)
rs844396	chr16:88966667	*CBFA2T3*	41	23	*RPL13* (AOR, PP4=10.5)
rs200482500	chr8: 10606223	*PINX1, SOX7*	13	10	*ENSG00000258724** (LIV, PP4=4.29)
rs11785239	chr8:8205010	*SGK223****	8	7	*PPP1R3B* (VAF, PP4=3.29)
rs201648240	chr1:208953176:INDEL		11	4	*TRAF3IP3* (MAM, PP4=3.74)
rs11196033	chr10: 114410998	*VTI1A*	10	4	*VTI1A* (LIV, PP4=2.44)
rs224904	chr5:81637916	*ATP6AP1L, ATG10*	15	10	*ENSG00000248870* (MAM, PP4=0.81)
Suggestive					
rs515135	chr2:21286057	*APOB*	18	7	*HS1BP3* (MAM, PP4=5.27)
rs139302128	chr2:242594226	*ATG4B*	29	0	*ENSG00000237940* (Blood, PP4=2.26)
Plaque					
Main					
rs113309773	16:75432686	*CFDP1- TMEM170A*	19	10	*BCAR1* (AOR, PP4=26)
rs11413744	4:148395284:INDEL	*EDNRA*	9	3	*EDNRA* (AOR, PP4=49)
rs17477177	7:106411858	*PIK3CG*	21	6	*CCDC71L* (AOR, PP4=97)
rs200495339	19:11189298:INDEL	*LDLR*	69	21	*ELOF1* (MAM, PP4=5.5)
rs9632884	9:22072301	*9p21*	6	6	*CDKN2B* (AOR, PP4=39)
Suggestive					
rs259140	7:89624347	*STEAP2-AS1*	7	6	*CLDN12* (LIV, PP4=13)

rs4779614	15:33540117	*TMCO5B*	15	6	*GREM1* (SKLM, PP4=21)

**ENSG00000258724* is a long transcript that has exons derived from both *PINX1* and *SOX7*, the encoded protein is 440 aa long, with approx. 310 aa derived from *SOX7* exons (*SOX7* is 388aa) and 130 aa derived from *PINX1* exons. UniProt has included ENSG00000258724 within the SOX7, describing it as an alternative spliced product Q9BT81-2. Although it has aa sequence from both genes.

** *LDLR* has an eQTL only in LIV, with p-value 1.73e-05. However, there is no evidence of colocalization with GWAS (PP4=0.5%)

****SGK223*, *SCEL* not covered in STARNET

Supplementary Table 7. Multiple trait colocalization of cIMT and plaque with AOR/MAM eQTLs (STARNET) and CHD (CARDIoGRAMPlusC4D), or stroke subtypes (MEGASTROKE) with probability of colocalization across three traits ≥ 75%.

Gene.name	Chr	Start-Stop	Data a	Data b	Data c	N snps	PPA abc	Best snp abc	Min p_value data a	Min p_value data b	Min p_value data c	Min p-value SNP data a	Min p-value SNP data b	Min p-value SNP data c
ADAMTS9	3	63588304-65587494	AOR	cIMT	Stroke AS	4415	0.80	rs17676309	2.06E-25	1.49E-06	1.11E-05	rs17676309	rs17676309	rs28546794
ADAMTS9	3	63588304-65587494	MAM	cIMT	Stroke AS	4415	0.77	rs17676309	7.48E-24	1.49E-06	1.11E-05	rs6775974	rs17676309	rs28546794
ADAMTS9-AS1	3	63561280-65561009	AOR	cIMT	Stroke AS	4411	0.80	rs17676309	4.03E-15	1.49E-06	1.11E-05	rs17676309	rs17676309	rs28546794
ADAMTS9-AS2	3	63841395-65833136	MAM	cIMT	Stroke AS	4533	0.76	rs17676309	6.41E-13	1.49E-06	1.11E-05	rs17676309	rs17676309	rs28546794
CCDC71L	7	105299372-107298840	AOR	Plaque	Stroke LAS	3929	0.81	rs17477177	2.37E-37	3.71E-11	0.000362	rs12705390	rs17477177	rs17398575
CCDC71L	7	105299372-107298840	AOR	cIMT	Stroke LAS	3910	0.81	rs12705390	2.37E-37	3.12E-09	0.000362	rs12705390	rs13225723	rs17398575
CCDC71L	7	105299372-107298840	MAM	cIMT	Stroke LAS	3910	0.80	rs12705390	1.17E-33	3.12E-09	0.000362	rs12705390	rs13225723	rs17398575
CCDC71L	7	105299372-107298840	MAM	Plaque	Stroke LAS	3929	0.79	rs17477177	1.17E-33	3.71E-11	0.000362	rs12705390	rs17477177	rs17398575
CDH13	16	82245780-84245226	AOR	cIMT	CHD	7999	0.94	16:83045790	6.81E-71	1.71E-05	2.11E-06	16:83045790	16:83045790	16:83045790
CDH13	16	82245780-84245226	MAM	cIMT	CHD	7999	0.94	16:83045790	2.47E-46	1.71E-05	2.11E-06	16:83045790	16:83045790	16:83045790
EDNRA	4	147434590-149433978	AOR	Plaque	Stroke LAS	2955	0.8	rs17612742	3.16E-05	5.68E-08	1.05E-06	rs6841581	rs10305839	rs17612742
ENSG00000232821.1	7	18153930-20152801	MAM	Plaque	CHD	4244	0.84	7:19049388	3.00E-14	0.00024	8.00E-11	7:19049388	7:18843808	7:19049388
ENSG00000232821.1	7	18153930-20152801	MAM	Plaque	Stroke AS	3874	0.84	rs2107595	2.97E-14	0.00024	3.59E-11	rs2107595	rs2520343	rs2107595
ENSG00000232821.1	7	18153930-20152801	MAM	Plaque	Stroke IS	3878	0.84	rs2107595	2.97E-14	0.00024	2.33E-11	rs2107595	rs2520343	rs2107595
ENSG00000232821.1	7	18153930-20152801	MAM	Plaque	Stroke LAS	3894	0.84	rs2107595	2.97E-14	0.00024	1.44E-13	rs2107595	rs2520343	rs2107595
ENSG00000260228.1	16	82832549-84832129	AOR	cIMT	CHD	8667	0.94	16:83045790	8.36E-18	1.71E-05	2.11E-06	16:83045790	16:83045790	16:83045790
ENSG00000260523.1	16	82832949-84832129	AOR	cIMT	CHD	8664	0.94	16:83045790	7.83E-30	1.71E-05	2.11E-06	16:83045790	16:83045790	16:83045790
ENSG00000260523.1	16	82832949-84832129	MAM	cIMT	CHD	8664	0.94	16:83045790	4.54E-17	1.71E-05	2.11E-06	16:83045790	16:83045790	16:83045790
ENSG00000260788.1	16	82780756-84780613	AOR	cIMT	CHD	8774	0.94	16:83045790	1.15E-33	1.71E-05	2.11E-06	16:83045790	16:83045790	16:83045790
ENSG00000260788.1	16	82780756-84780613	MAM	cIMT	CHD	8774	0.94	16:83045790	5.39E-20	1.71E-05	2.11E-06	16:83045790	16:83045790	16:83045790

ENSG00000260832.1	16	82006107-84004822	AOR	cIMT	CHD	7626	0.94	16:83045790	3.95E-36	1.71E-05	2.11E-06	16:83045790	16:83045790	16:83045790
ENSG00000260832.1	16	82006107-84004822	MAM	cIMT	CHD	7626	0.94	16:83045790	5.54E-17	1.71E-05	2.11E-06	16:83045790	16:83045790	16:83045790
ENSG00000261103.1	16	82748233-84747503	AOR	cIMT	CHD	8809	0.94	16:83045790	9.70E-19	1.71E-05	2.11E-06	16:83045790	16:83045790	16:83045790
ENSG00000261410.1	16	82425634-84423396	AOR	cIMT	CHD	8403	0.94	16:83045790	1.62E-18	1.71E-05	2.11E-06	16:83045790	16:83045790	16:83045790
ENSG00000261410.1	16	82425634-84423396	MAM	cIMT	CHD	8403	0.93	16:83045790	2.39E-12	1.71E-05	2.11E-06	16:83017777	16:83045790	16:83045790
KIAA1462	10	29325616-31325032	MAM	Plaque	CHD	6180	0.83	10:30321598	3.00E-35	4.10E-06	4.40E-11	10:30317073	10:30317073	10:30323892
KIAA1462	10	29325616-31325032	MAM	cIMT	CHD	6222	0.84	10:30323892	3.00E-35	1.29E-06	4.41E-11	10:30317073	10:30333622	10:30323892
PRKAR2B	7	105745093-107743409	MAM	cIMT	Stroke LAS	3679	0.77	rs12705390	2.33E-08	3.12E-09	0.000362	rs12705390	rs13225723	rs17398575
PRKAR2B	7	105745093-107743409	AOR	Plaque	Stroke LAS	3697	0.76	rs17477177	6.12E-07	3.71E-11	0.000362	rs12705390	rs17477177	rs17398575
PRKAR2B	7	105745093-107743409	MAM	Plaque	Stroke LAS	3697	0.76	rs17477177	2.33E-08	3.71E-11	0.000362	rs12705390	rs17477177	rs17398575
PRKAR2B	7	105745093-107743409	AOR	cIMT	Stroke LAS	3679	0.76	rs12705390	6.12E-07	3.12E-09	0.000362	rs12705390	rs13225723	rs17398575
TWIST1	7	18157302-20154816	AOR	Plaque	CHD	4240	0.84	7:19049388	1.50E-10	0.00024	8.00E-11	7:19049388	7:18843808	7:19049388
TWIST1	7	18157302-20154816	MAM	Plaque	CHD	4240	0.84	7:19049388	1.60E-37	0.00024	8.00E-11	7:19049388	7:18843808	7:19049388
TWIST1	7	18157302-20154816	AOR	Plaque	Stroke AS	3870	0.84	rs2107595	1.46E-10	0.00024	3.59E-11	rs2107595	rs2520343	rs2107595
TWIST1	7	18157302-20154816	MAM	Plaque	Stroke AS	3870	0.84	rs2107595	1.58E-37	0.00024	3.59E-11	rs2107595	rs2520343	rs2107595
TWIST1	7	18157302-20154816	AOR	Plaque	Stroke IS	3874	0.84	rs2107595	1.46E-10	0.00024	2.33E-11	rs2107595	rs2520343	rs2107595
TWIST1	7	18157302-20154816	MAM	Plaque	Stroke IS	3874	0.84	rs2107595	1.58E-37	0.00024	2.33E-11	rs2107595	rs2520343	rs2107595
TWIST1	7	18157302-20154816	AOR	Plaque	Stroke LAS	3890	0.84	rs2107595	1.46E-10	0.00024	1.44E-13	rs2107595	rs2520343	rs2107595
TWIST1	7	18157302-20154816	MAM	Plaque	Stroke LAS	3890	0.84	rs2107595	1.58E-37	0.00024	1.44E-13	rs2107595	rs2520343	rs2107595

Supplementary Table 8. Druggability of genes in loci genome-wide significantly associated with cIMT or plaque. Tier 1, 2 and 3 druggability are highlighed: Tier 1=approved drugs and drugs in clinical development; Tier 2= proteins closely related to drug targets or associated with drug-compounds; Tier 3: extracellular proteins and members of key drug-target families

Gene.name	Drug tier	drug_type	Distance to gene	variant	trait	strength	chr.position	Nearest coding gene	Max number of genes across tissues considered in region	Max number of genes across tissues suggestive assoc
ATG4B	2	SMALL_MOL	-1	rs139302128	cIMT	Suggestive	chr2:242594226	*ATG4B*	29	0
ALPL	2	SMALL_MOL,BIO_MOL	-1	rs147771110	Plaque	Suggestive	chr1:21868723	*ALPL*	23	0
LDLR	2	SMALL_MOL,BIO_MOL	10739	rs200495339	Plaque	Main	chr19:11189298	*LDLR*	70	22
APOB	1	SMALL_MOL,BIO_MOL	19112	rs515135	cIMT	Suggestive	chr2:21286057	*APOB*	18	7
EDNRA	1	SMALL_MOL,BIO_MOL	-1	rs6841473	Plaque	Suggestive	chr4:148407652	*EDNRA*	9	3
APOE	3	BIO_MOL	-1	rs7412	cIMT	Main	chr19: 45412079	*APOE*	61	28

Best coloc	variant	alleles	chr	consequence_types	Rsq EUR
ENSG00000237940 (Blood, PP4=2.26)	rs139302128	C/T	2	non_coding_transcript_variant,non_coding_transcript_exon_variant,NMD_transcript_variant,intron_variant,downstream_gene_variant	0.5
DDOST (SF, PP4=4)	rs147771110	-/C	1	non_coding_transcript_variant,intron_variant,regulatory_region_variant	0.5
ELOF1 (MAM, PP4=5.1)	rs200495339	G/-	19	intergenic_variant	0.5
HS1BP3 (MAM, PP4=5.27)	rs515135	T/C	2	intergenic_variant	0.5
EDNRA (AOR, PP4=37.62)	rs6841473	C/T	4	non_coding_transcript_variant,NMD_transcript_variant,intron_variant	0.5
APOE (SF, PP4=24.84)	rs7412	C/T	19	missense_variant,downstream_gene_variant	0.5

description	biotype	gene	Gene start pos	Gene end pos	Gene window overlap
autophagy related 4B, cysteine peptidase [Source:HGNC Symbol;Acc:20790]	protein_coding	KNOWN	242576628	242613272	36645
alkaline phosphatase, liver/bone/kidney [Source:HGNC Symbol;Acc:438]	protein_coding	KNOWN	21835858	21904905	61570
low density lipoprotein receptor [Source:HGNC Symbol;Acc:6547]	protein_coding	KNOWN	11200038	11244492	10875
apolipoprotein B [Source:HGNC Symbol;Acc:603]	protein_coding	KNOWN	21224301	21266945	3307
endothelin receptor type A [Source:HGNC Symbol;Acc:3179]	protein_coding	KNOWN	148402069	148466106	40938
apolipoprotein E [Source:HGNC Symbol;Acc:613]	protein_coding	KNOWN	45409011	45412650	572

Supplementary Table 9. Druggability of genes identified in colocalization analyses

Gene	Chr	Position	nsnps	data1	PPA.abc	data3	data2	Gene description
CDH13	16	82245780-84245226	7999	AOR	0.94	CARDIoGRAMplusC4D	cIMT	cadherin 13 [Source:HGNC Symbol;Acc:1753]
CDH13	16	82245780-84245226	7999	MAM	0.94	CARDIoGRAMplusC4D	cIMT	cadherin 13 [Source:HGNC Symbol;Acc:1753]
ADAMTS9	3	63588304-65587494	4415	AOR	0.80	AS	cIMT	ADAM metallopeptidase with thrombospondin type 1 motif, 9 [Source:HGNC Symbol;Acc:13202]
ADAMTS9	3	63588304-65587494	4415	MAM	0.77	AS	cIMT	ADAM metallopeptidase with thrombospondin type 1 motif, 9 [Source:HGNC Symbol;Acc:13202]
EDNRA	4	147434590-149433978	2955	AOR	0.80	LAS	PLAQUE	endothelin receptor type A [Source:HGNC Symbol;Acc:3179]

Gene	Drug tier	Drug type	Compound activities	Compound activities
CDH13	tier 3	BIO_MOL	0	0
CDH13	tier 3	BIO_MOL	0	0
ADAMTS9	tier 3	BIO_MOL	0	0
ADAMTS9	tier 3	BIO_MOL	0	0
EDNRA	tier 1	BIO_MOL | SMALL_MOL	46	1*

***Drugs and indications:** AMBRISENTAN (Andes disease,SCD,Asma,HT,HYPERTENSION PULM,Vasc,HPAH,UIP,PULMONARY HYPERTENSION, PRIMARY, DEXFENFLURAMINE-ASSOCIATED,Pulmonary Hypertension, Primary, Fenfluramine-Associated,PAH,Pph1 With Hht,Mountain Sickness,Altitude Hypoxia)|BOSENTAN (Asma,MOD,HYPERTENSION PULM,COLD,melanoma,PSS,Optic Nerve Ischemia,CDH,Morgagni hernia,Bochdalek hernia,HPAH,Anterior Ischemic Optic Neuropathy,Posterior Ischemic Optic Neuropathy,Chronic Airflow Obstruction,UIP,PULMONARY HYPERTENSION, PRIMARY, DEXFENFLURAMINE-ASSOCIATED,Pulmonary Hypertension, Primary, Fenfluramine-Associated,PAH,Pph1 With Hht)|CLAZOSENTAN (Subarachnoid bleeding,Perinatal Subarachnoid Hemorrhage,Spontaneous subarachnoid hemorrhage,Aneurysmal Subarachnoid Hemorrhage,INTRACRANIAL SUBARACHNOID HEMORRHAGE)|DARUSENTAN (HT)|MACITENTAN (HYPERTENSION PULM,PSS,HPAH,UIP,PULMONARY HYPERTENSION, PRIMARY, DEXFENFLURAMINE-ASSOCIATED,Pulmonary Hypertension, Primary, Fenfluramine-Associated,PAH,Pph1 With Hht)|SPARSENTAN (FSGS,HT,Segmental hyalinosis)|TEZOSENTAN (Weak heart,CHF,HYPERTENSION PULM,LVF,rvf,Myocardial Failure,Heart Decompensation)|ZIBOTENTAN (Ca breast,CA,Liver,Tumor,prostate tumor,Liver Dysfunction,BENIGN TUMOR,Ca prostate,Human Mammary Neoplasm,Breast tumor)

Supplementary References

1. Harris, T.B. *et al.* Age, Gene/Environment Susceptibility-Reykjavik Study: multidisciplinary applied phenomics. *Am J Epidemiol* **165**, 1076-87 (2007).
2. The Atherosclerosis Risk in Communities (ARIC) Study: design and objectives. The ARIC investigators. *Am J Epidemiol* **129**, 687-702 (1989).
3. Schmidt, R., Fazekas, F., Kapeller, P., Schmidt, H. & Hartung, H.P. MRI white matter hyperintensities: three-year follow-up of the Austrian Stroke Prevention Study. *Neurology* **53**, 132-9 (1999).
4. 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).
5. Ghadery, C. *et al.* R2* mapping for brain iron: associations with cognition in normal aging. *Neurobiol Aging* **36**, 925-32 (2015).
6. Seiler, S. *et al.* Magnetization transfer ratio relates to cognitive impairment in normal elderly. *Front Aging Neurosci* **6**, 263 (2014).
7. Sitzer, M. *et al.* C-reactive protein and carotid intimal medial thickness in a community population. *J Cardiovasc Risk* **9**, 97-103 (2002).
8. Fried, L.P. *et al.* The Cardiovascular Health Study: design and rationale. *Ann Epidemiol* **1**, 263-76 (1991).
9. Bowden, D.W. *et al.* Review of the Diabetes Heart Study (DHS) family of studies: a comprehensively examined sample for genetic and epidemiological studies of type 2 diabetes and its complications. *Rev Diabet Stud* **7**, 188-201 (2010).
10. Aulchenko, Y.S. *et al.* Linkage disequilibrium in young genetically isolated Dutch population. *Eur J Hum Genet* **12**, 527-34 (2004).
11. Dawber, T.R. & Kannel, W.B. The Framingham study. An epidemiological approach to coronary heart disease. *Circulation* **34**, 553-5 (1966).
12. 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).
13. 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).
14. Group, C.S. Vascular factors and risk of dementia: design of the Three-City Study and baseline characteristics of the study population. *Neuroepidemiology* **22**, 316-25 (2003).
15. Debette, S. *et al.* Tea consumption is inversely associated with carotid plaques in women. *Arterioscler Thromb Vasc Biol* **28**, 353-9 (2008).
16. Lambert, J.C. *et al.* Meta-analysis of 74,046 individuals identifies 11 new susceptibility loci for Alzheimer's disease. *Nat Genet* **45**, 1452-8 (2013).
17. Deary, I.J., Gow, A.J., Pattie, A. & Starr, J.M. Cohort profile: the Lothian Birth Cohorts of 1921 and 1936. *Int J Epidemiol* **41**, 1576-84 (2012).
18. Deary, I.J. *et al.* The Lothian Birth Cohort 1936: a study to examine influences on cognitive ageing from age 11 to age 70 and beyond. *BMC Geriatr* **7**, 28 (2007).
19. Deary, I.J., Whiteman, M.C., Starr, J.M., Whalley, L.J. & Fox, H.C. The impact of childhood intelligence on later life: following up the Scottish mental surveys of 1932 and 1947. *J Pers Soc Psychol* **86**, 130-47 (2004).
20. Wardlaw, J.M. *et al.* Brain aging, cognition in youth and old age and vascular disease in the Lothian Birth Cohort 1936: rationale, design and methodology of the imaging protocol. *Int J Stroke* **6**, 547-59 (2011).
21. Wardlaw, J.M. *et al.* Vascular risk factors, large-artery atheroma, and brain white matter hyperintensities. *Neurology* **82**, 1331-8 (2014).
22. Bild, D.E. *et al.* Multi-Ethnic Study of Atherosclerosis: objectives and design. *Am J Epidemiol* **156**, 871-81 (2002).
23. de Mutsert, R. *et al.* The Netherlands Epidemiology of Obesity (NEO) study: study design and data collection. *Eur J Epidemiol* **28**, 513-23 (2013).
24. Penninx, B.W. *et al.* The Netherlands Study of Depression and Anxiety (NESDA): rationale, objectives and methods. *Int J Methods Psychiatr Res* **17**, 121-40 (2008).
25. Sullivan, P.F. *et al.* Genome-wide association for major depressive disorder: a possible role for the presynaptic protein piccolo. *Mol Psychiatry* **14**, 359-75 (2009).
26. McQuillan, R. *et al.* Runs of homozygosity in European populations. *Am J Hum Genet* **83**, 359-72 (2008).

27. Hofman, A. *et al.* The Rotterdam Study: 2010 objectives and design update. *Eur J Epidemiol* **24**, 553-72 (2009).
28. Volzke, H. *et al.* Cohort profile: the study of health in Pomerania. *Int J Epidemiol* **40**, 294-307 (2011).
29. Raitakari, O.T. *et al.* Cohort profile: the cardiovascular risk in Young Finns Study. *Int J Epidemiol* **37**, 1220-6 (2008).
30. Lawlor, D.A., Bedford, C., Taylor, M. & Ebrahim, S. Geographical variation in cardiovascular disease, risk factors, and their control in older women: British Women's Heart and Health Study. *J Epidemiol Community Health* **57**, 134-40 (2003).
31. Price, J.F. *et al.* The Edinburgh Type 2 Diabetes Study: study protocol. *BMC Endocr Disord* **8**, 18 (2008).
32. Wadsworth, M., Kuh, D., Richards, M. & Hardy, R. Cohort Profile: The 1946 National Birth Cohort (MRC National Survey of Health and Development). *Int J Epidemiol* **35**, 49-54 (2006).
33. Marmot, M.G. *et al.* Health inequalities among British civil servants: the Whitehall II study. *Lancet* **337**, 1387-93 (1991).
34. Baldassarre, D. *et al.* Cross-sectional analysis of baseline data to identify the major determinants of carotid intima-media thickness in a European population: the IMPROVE study. *Eur Heart J* **31**, 614-22 (2010).
35. Loeffler, M. *et al.* The LIFE-Adult-Study: objectives and design of a population-based cohort study with 10,000 deeply phenotyped adults in Germany. *BMC Public Health* **15**, 691 (2015).
36. Beutner, F. *et al.* Rationale and design of the Leipzig (LIFE) Heart Study: phenotyping and cardiovascular characteristics of patients with coronary artery disease. *PLoS One* **6**, e29070 (2011).
37. Boyd, A. *et al.* Cohort Profile: the 'children of the 90s'--the index offspring of the Avon Longitudinal Study of Parents and Children. *Int J Epidemiol* **42**, 111-27 (2013).
38. Fraser, A. *et al.* Cohort Profile: the Avon Longitudinal Study of Parents and Children: ALSPAC mothers cohort. *Int J Epidemiol* **42**, 97-110 (2013).
39. Galesloot, T.E. *et al.* Cohort Profile: The Nijmegen Biomedical Study (NBS). *Int J Epidemiol* (2017).
40. Berglund, G., Elmstahl, S., Janzon, L. & Larsson, S.A. The Malmo Diet and Cancer Study. Design and feasibility. *J Intern Med* **233**, 45-51 (1993).