SUPPLEMENTAL MATERIAL

Online Data Supplemental Section S1: Detailed descriptions of recruitment procedures for discovery cohorts

The discovery cohorts included the Framingham Heart Study (FHS, n=6,634), the Precocious Coronary Artery Disease Study (PROCARDIS) cases (n=1,922) and controls (n=1,294), TwinsUK (n=2,017), the MONICA/KORA study (n=1,565), HealthABC (n=1,645), the Marseille Thrombosis Association (MARTHA) study (n=851) and the Prevention of Renal and Vascular End Stage Disease (PREVEND) study (n=3,671). Among these cohorts, five were population-based (FHS, KORA, TwinsUK, HealthABC, and PREVEND), one a case-control study of CAD (PROCARDIS) and the last one, MARTHA, a collection of patients with venous thromboembolism. All participants were of self-described European origin.

Considering the pronounced diurnal variation in plasma PAI-1 levels, subjects with blood sampling before 7 AM or after 11 AM were excluded. Subjects with type-1 diabetes were also excluded. In addition, subjects with low genome-wide call rates (<93-99%, varying with cohort) and non-European ancestry, as identified by genome-wide principal components analysis or multi-dimensional scaling, were not included in the analyses, and phenotype distribution outliers were excluded in PROCARDIS based on an extreme deviation (>5 SDs) from the cohort mean.

Characteristics of study participants and details of genotype acquisition and quality control in individual discovery cohorts are provided in **Supplemental Tables S1 and S2**.

The **Framingham Heart Study (FHS):** FHS started in 1948 with 5,209 randomly ascertained participants from Framingham, Massachusetts, US, who had undergone biannual examinations to investigate cardiovascular disease and its risk factors. In 1971, the Offspring cohort (comprising 5,124 children of the original cohort and the children's spouses) and in 2002, the Third Generation (consisting of 4,095 children of the Offspring cohort) were recruited. FHS participants in this study are of European ancestry. The methods of recruitment and data collection for the Offspring and Third Generation cohorts have been described.[1]

The **Precocious Coronary Artery Disease Study (PROCARDIS) cases and controls cohorts**: The PROCARDIS [2] study consists of coronary artery disease (CAD) cases and controls from four European countries (UK, Italy, Sweden and Germany). CAD (defined as myocardial infarction, acute coronary syndrome, unstable or stable angina, or need for coronary artery bypass surgery or percutaneous coronary intervention) was diagnosed before 66 years of age and 80% of cases had a sibling fulfilling the same criteria for CAD. Subjects with self-reported non-European ancestry were excluded. Among the "genetically-enriched" CAD cases, 70% had suffered myocardial infarction (MI).

In the UK, patients were identified from hospital records used previously to recruit patients for large-scale trials of cholesterol-lowering therapy. Patients were identified in Italy through hospitals that had collaborated in the GISSI studies, in Sweden through

existing registries of cases that had contracted MI at a young age or through the central database of the Stockholm County Council, and in Germany through the PROCAM and related databases. Controls with nopersonal or sibling history of CAD before age 66 years were contemporaneously recruited using the same infrastructure. For each of the CAD cases, one control was recruited of the same sex,ethnicity and within 5 years of age, with no personal or sibling history of CAD before age of 66 years.

The **Prevention of Renal and Vascular End stage Disease (PREVEND) study**: The PREVEND study is an ongoing prospective study investigating the natural course of increased levels of urinary albumin excretion and its relation to renal and cardiovascular disease. Inhabitants 28 to 75 years of age (n=85,421) in the city of Groningen, The Netherlands, were asked to complete a short questionnaire, 47% of whom responded. Individuals with a urinary albumin concentration of at least 10 mg/L (n=7,768) were then selected along with a randomly selected control group with a urinary albumin concentration less than 10 mg/L (n=3,395). Details of the protocol have been described elsewhere (www.prevend.org). A total of 3,671 subjects from the PREVEND study had both PAI-1 levels and GWA data available for this study.

The **MONICA/KORA Augsburg Study (KORA):** The KORA study [3] is a population-based study covering the city of Augsburg (Germany) and two adjacent counties, conducted initially to estimate the prevalence and distribution of cardiovascular risk factors among individuals aged 25 to 74 years as part of the WHO MONICA study. All participants were residents of German nationality who were identified through the registration office and underwent standardized examinations including blood sampling for preparation

of plasma and DNA. For the KORA genome-wide association (GWA) study, 1,644 subjects 45 to 69 years of age were selected from the KORA F3 samples obtained in 2004-06. After excluding subjects with no PAI-1 measurements available, 1,565 subjects remained for study.

Twins UK: The Twins UK Registry comprises unselected, mostly female volunteersascertained from the general population through national media campaigns in theUK[4]. Participants were excluded if the sample call rate was less than 95%, autosomal heterozygosity was outside the expected range or genotype concordance was over 97% with another sample and if they were of non-European ancestry, either self-identified or identified by cluster analysis in STRUCTURE or multidimensional scaling by comparison to the three HapMap phase 2 reference populations (CEU, YRI, CHB+JPT; www.hapmap.org), or if they showed unexplained relatedness (estimated proportion of allele shared identical by descent >0.05) to >120 samples from other individuals. This resulted in GWA data being available for 5,295 twins from the TwinsUK cohort, of whom 2,017 with PAI-1 data were used for the present study.

The **Marseille Thrombosis Association (MARTHA) study:** The MARTHA project[5] is composed of two independent samples of venous thrombosis (VT) patients, designated MARTHA08 (N=1,006) and MARTHA10 (N=586). MARTHA subjects are unrelated Caucasians consecutively recruited at the Thrombophilia center of the Timone hospital (Marseille, France) between January 1994 and October 2005. All patients had a documented history of VT and were free of established genetic risk factors, including deficiency of antithrombin III, protein C or protein S, homozygosity for Factor V Leiden or Factor II 20210A, and lupus anticoagulant. The

thrombotic events were confirmed by venography, Doppler ultrasound, spiral computed tomographic scanning angiography, and/or ventilation/perfusion lung scan. All subjects were of European origin, with the majority being of French ancestry. The MARTHA08 cohort was used as a discovery cohort while the MARTHA10 cohort was included in the replication analysis (see below). Individuals with genotyping success rates less than 95% were excluded from the analyses as well as individuals demonstrating close relatedness as detected by pairwise clustering of identity by state distances (IBS) and multi-dimensional scaling(MDS) implemented in PLINK software (v.1.04). Non-European ancestry was also investigated using the EIGENSTRAT program, leading to the final selection of 972 patients left for analysis in MARTHA08 among whom 851 have PAI-1 measurements.

The **Health ABC study**: HEALTH ABC is a prospective cohort study investigating the associations between body composition, weight-related health conditions, and incident functional limitation in older adults. Health ABC enrolled well-functioning, community-dwelling black (n=1,281) and white (n=1,794) men and women aged 70-79 years between April 1997 and June 1998. Participants were recruited from a random sample of all white and black Medicare eligible residents in the Pittsburgh, PA, and Memphis, TN, metropolitan areas. Of 3,075 participants at baseline, 1645 Caucasians had both genotype and phenotype data available for analysis.

Individuals were excluded from the dataset for the reasons of sample failure, genotypic sex mismatch, and status of first-degree relative of an included individual based on genotype data.

Online Data Supplemental Section S2: Detailed descriptions of recruitment procedures for replication cohorts

Replication sample genotypes were obtained by genotyping in PREVEND or by using *in silico* data from the Cardiovascular Health Study (CHS), the Atherosclerosis Risk in Communities Study (ARIC), the CHS Longitudinal Investigation of Thromboembolism Etiology (LITE) Study, GeneSTAR, MARTHA, the Multi-Ethnic Study of Atherosclerosis (MESA), the Ludwigshafen Risk and Cardiovascular Health (LURIC) Study, and the Rotterdam Study. Characteristics of study participants and details of genotype acquisition and quality control in individual replication cohorts are provided in **Supplemental Tables S2** and **S4**.

The **PREVEND study**: An independent part of the PREVEND study (described in its entirety in **Online Data Supplemental Section S1**), which lacked GWA data (n=4,539), was used for the replication stage.

The **Cardiovascular Health Study (CHS)**: **CHS** is a population-based cohort study of risk factors for CHD and stroke in adults ≥65 years conducted across 4 field centers [6]. The original predominantly Caucasian cohort of 5,201 persons was recruited in 1989-1990 from random samples of the Medicare eligibility lists. Subsequently, an additional predominantly African-American cohort of 687 persons was enrolled for a total sample of 5,888. In 2007-2008, genotyping was performed at the General Clinical Research Center's Phenotyping/Genotyping Laboratory at Cedars-Sinai on 3980 CHS participants who were free of CVD at baseline and consented to genetic testing. A total of 1,908 persons were excluded from the GWA study sample due to the presence at study baseline of coronary

heart disease, congestive heart failure, peripheral vascular disease, valvular heart disease, stroke or transient ischemic attack or lack of available DNA. Because the other cohorts were predominantly white, the African-American participants were excluded from this analysis. Participants were also excluded if they had a call rate<=95%. Genotyping has been attempted to date in 3,397 white participants, and was successful in 3,291 persons; the subset of the latter having data for PAI-1 (n=368) constitutes the CHS sample for this study.

The **Atherosclerosis Risk in Communities (ARIC) study**: ARIC recruited 15,792 predominately African American and European American adults aged 45 to 64 years by probability sampling in 1987-1989 from Forsyth County, North Carolina; Jackson, Mississippi; suburbs of Minneapolis, Minnesota; and Washington County, Maryland.[7] The Jackson site recruited African Americans only; samples from the other three sites included African Americans, European Americans, as well as a small proportion of participants of other ethnicity.[7] In the present study we included only European American adults. The sample included in this study was a stratified random sample of the entire ARIC cohort that was selected to serve as a reference cohort sample in a case-cohort study of CHD in ARIC.[8] Participants were excluded if they had prevalent CHD (or unknown status) at baseline, history of stoke or transient ischemic attack at baseline, or missing sampling or event information. The selection of the reference cohort sample was stratified by age and sex and oversampled non-cases with thin average carotid IMT measurements at baseline (<30 percentile).[8]

CHS Longitudinal Investigation of Thromboembolism Etiology (LITE) Controls: The LITE study is a prospective study of venous thromboembolism (VTE) occurrence in two pooled, multicenter, longitudinal population-based cohort studies: the ARIC study and the CHS.[9] The LITE study employed a nested case–control design to investigate prospective associations between VTE incidence and blood parameters measured in stored specimens, including PAI-1.[9]Only controls in the CHS LITE nested case-control sample were analyzed here. They were selected at random from the CHS cohort at a ratio of 2 controls per VTE case, frequency matched to the cases by age (5-year groupings), gender, race (African American, White), follow-up time (cases' event date within 2 years of controls' assigned date).[9] Only European American adults were included in the present study.

The GeneSTAR study: GeneSTAR is a longitudinal study of CAD, stroke, and other vascular diseases in European-American and African-American families. Probands were identified with documented coronary disease prior to age 60 in one of ten Baltimore area hospitals and screened from 1983-2006. Their initially healthy siblings who were younger than age 60 were recruited and completed a baseline screening between 1983 and 2006, and were followed for incident disease at regular intervals. In addition, the offspring of either the probands or healthy siblings as well as the co-parents of the offspring who were between the ages of 21 and 80 were recruited and screened from 2003-2006. Participants were excluded from the genetic studies due to sample failure, sex discrepancies, or excessive (>5%) Mendelian inconsistencies. PAI-1 measures were available in 551 participants, and this constitutes the GeneSTAR sample in this study.

The **MARTHA10 study**: MARTHA10, an independent cohort (N=586) in the MARTHA project (described in Online Data Supplemental Section S1), was used as a replication cohort. For a total of 498 patients both PAI-1 measurements and genotype data were available.

The **Multi-Ethnic Study of Atherosclerosis (MESA)**: MESA is a longitudinal study of subclinical cardiovascular disease and risk factors that predict progression to clinically overt cardiovascular disease or progression of the subclinical disease. The first clinic visits occurred in 2000 in 6,814 participants recruited from six field centers across the United States. Approximately 38% of the recruited participants are White, 28% African-American, 22% Hispanic, and 12% Asian, the latter predominantly of Chinese descent. After testing for cryptic relatedness using KING (Manichaikul et al 2010) and randomly dropping 11 individuals who were first degree sibship relatives of others with blood PAI-1 measure, 433 individuals of self-reported white race, not of Hispanic descent, had both GWA genotype data and PAI-1 and were available for replication of the major results in this project[10].

The **Ludwigshafen Risk and Cardiovascular Health (LURIC) study:** The LURIC study is a prospective study of more than 3,300 individuals of German ancestry in whom cardiovascular and metabolic phenotypes (CAD, MI, dyslipidaemia, hypertension, metabolic syndrome and diabetes mellitus) have been defined or ruled out using standardised methodologies in all study participants. A 10-year clinical follow-up for total and cause specific mortality has been completed.[11] From 1997 to 2002 about 3,800 patients were recruited

at the Heart Center of Ludwigshafen (Rhein). Inclusion criteria were: German ancestry, clinical stability (except for acute coronary syndromes) and existence of a coronary angiogram. Exclusion criteria were: any acute illness other than acute coronary syndromes, any chronic disease where non-cardiac disease predominated and a history of malignancy within the last five years. A total of 2,928 participants had both genotyping and PAI-1 measurements available and were included in the analysis.

The **Rotterdam study:** The Rotterdam study is a population-based cohort study among inhabitants aged 55 years and over of Ommoord, a district of Rotterdam, the Netherlands (N=7,983, virtually all Caucasian). Details of the study have been described elsewhere.[12]Baseline examinations were conducted in 1990-1993, with follow-up examinations in 1993-1994, 1997-1999 and 2002-2004. All persons attending the baseline examination in 1990-1993 consented to genotyping and had DNA extracted. Genotyping was attempted in persons with high-quality extracted DNA (n=6,449). Samples with low call rate (<97.5%, n=209), with excess autosomal heterozygosity (> 0.336, n=21), with sex-mismatch (n=36), or defined as outliers identified by the IBS clustering analysis (>3 standard deviations from population mean, n=102 or IBS probabilities > 97%, n=129) were excluded from the study population with some persons meeting more than one exclusion criterion.[13] In total, 5974 samples were available with good quality genotyping data. PAI-1 measurements were performed in a random subset (n=971) of participants who took part in the third examination (1997-1999). A total of 840 participants had both genotyping and PAI-1 measurements available and were included in the analysis.

Online Data Supplemental Section S3: Generation of plasma samples and assay details for PAI-1 determinations

FHS: Fasting blood samples were collected between 8 and 9 a.m. in tubes with 3.8% sodium citrate (9:1 vol/vol). An ELISA method was used to measure levels of PAI-1 antigen (TintElize PAI-1, Biopool AB). The intra-assay coefficient of variation (CV) was 9.6% for PAI-1 antigen.

PROCARDIS: Blood samples were drawn between 8.00 and 11.00 a.m.at the time of recruitment of cases and controls.PAI-1 antigen levels were measured in EDTA plasma in PROCARDIS controls and in citrate plasma in PROCARDIS cases using an ELISA assay from Medinor (TriniLize PAI-1 Antigen). A coagulation reference from Technoclone GmbH (Vienna, Austria) was used for quality control. The inter-assay CV was 15.6%.

PREVEND: Blood samples were obtained in the morning hours in all subjects to minimize variability due to circadian rhythms. Plasma levels of PAI-1 were measured using an ELISA from TechnocloneGmbh (Vienna, Austria). Each plate included calibrators (1.0-25 ng/ml for PAI-1) and two quality control samples. The lower limit of quantification was 1.5 ng/ml Assays not meeting quality control standards were not included in the analyses.

KORA: PAI-1 activity was measured in EDTA plasma using the Technozym PAI-1 Actibind ELISA assay (Technoclone, Vienna, Austria) with standardization according to the WHO International Standard for PAI-1. For quality control, we used two control samples with high and low PAI-1 levels, revealing a mean of 46.5 U/ml (reference 31.8-53.1 U/ml, SD 3.43, coefficient of variation (CV) 7.38%) for the high level sample and a mean of 7.79 U/ml (reference 6.6-11.5 U/ml, SD 0.39, CV 5.04%) for the low level sample.

MARTHA: Venous blood samples were collected by antecubitalvenipuncture into Vacutainer® tubes 0.105 M trisodium citrate (ratio 9:1, Becton Dickinson) in the morning after subjects had fasted for 12 hours. Platelet-poor plasma was obtained after double centrifugation of citrated blood (3000 *g* for 10 min at 25°C) and kept frozen at -80°C until analysis. The plasma was then centrifuged for 30 minutes at 2000g and 4°C, and the resulting plasma was collected and stored in aliquots at -70°C until assayed. PAI-1 activity levels were measured using the TriniLIZE PAI-1 activity kit from Trinity Biotech (Bray, Ireland). Intra- and inter-assay CVs were 3.78% and 8.04%, respectively, for normal controls.

TwinsUK: An ELISA was used for analysis of PAI-1 (Biopool, Umeå, Sweden). A venous blood sample was taken in the morning after an overnight fast into 0.13 trisodium citrate vacutainers (Becton Dickinson, Oxford, UK), centrifuged at 2560g for 20 minutes to obtain platelet-poor plasma, frozen in aliquots in liquid nitrogen, and stored at –40°C until analysis.[14]

Health ABC: Blood samples were collected via venipuncture after an overnight fasting. The specimens were aliquoted into cyrovials, frozen at -70°C, and shipped to the Health ABC Core Laboratory at the University of Vermont. The PAI-1 assay was originally developed by Dr. DésiréCollen and colleagues,[15] and adapted for population use at the Laboratory for Clinical Biochemistry Research at the University of Vermont.[16] This two-site ELISA is sensitive to free PAI-1 (both latent and active) but not PAI-1 in complex with tissue plasminogen activator. Reagents were generously provided by Dr. Collen (Leuven, Belguim). The analytical CV for this assay is < 10%. The healthy reference range is 5 -66 ng/mL.

CHS: PAI-1 antigen level was measured in plasma drawn at baseline and stored at -70°C.[17, 18] Blood was collected in the morning, with minimal stasis after an 8- to 12-hour fast, into tubes containing sodium citrate. PAI-1 antigen was measured as described under Health ABC.

CHS_LITE: After selection of incident cases and controls, in a nested case control design, stored samples of plasma were retrieved from -70ºC storage freezers. If baseline plasma samples were limited, previously thawed, or exhausted for a participant, a sample was retrieved from the plasma repository for the next visit (approximately 3 years after baseline); if neither sample was available it was considered missing.[9] PAI-1 antigen was measured as described under Health ABC.

ARIC: Following a standard protocol, fasting blood was drawn in 1987-1989 at baseline into vacuum tubes containing sodium citrate. The blood samples werecentrifuged at -4°C. Then the samples were frozen at -70°C until analyzed in 1997-1998 for PAI-1.[8] PAI-1 antigen was measured in citrate plasma by ELISA with use of an IMUBIND Plasma PAI-1 kit (American Diagnostica).[15] This assay detects active and inactive forms of PAI-1 and complexes of tPA/PAI-1 and urokinase plasminogen activator/PAI-1. Its sensitivity level is less than 1 ng/mL.Pearson correlation coefficients were 0.76 for PAI-1 in blinded duplicate specimens from different tubes of single blood drawn (n=54 to 71 pairs).[8] An earlier ARIC study involving fresh samples taken 3 times at 1- to 2-week intervals yielded reliability coefficients of 0.72 for PAI-1.[19]

GeneSTAR: PAI-1 was measured in plasma collected between 1998 and 2002 from a morning blood draw after an overnight fast. PAI-1 was measured in citrate plasma by a chromogenic substrate assay, which is a functional method that measures the activity of PAI-1 in plasma.

MESA: PAI-1 was measured on fasting citrate plasma as described under Health ABC.

LURIC: Following an overnight fast, venous blood was sampled in the morning before coronary angiography between 6.10 a.m. and 10.30 a.m., with the participants in supine position for 5 to 10 minutes before phlebotomy. PAI-1 activity and antigen were measured

with a chromogenic substrate assay and a sandwich enzyme immunoassay, respectively, using reagents from Biopool (Umea, Sweden, Chromolize™ PAI-1and TintElize™ PAI-1, respectively).

Rotterdam study: Fasting blood samples were obtained at the research center at the third survey of the Rotterdam study. Citrate plasma (5 mL) was collected and stored at −80°C. In July 2008, 200 µL of citrate plasma from each participant was sent to Rules-Based Medicine, Austin, Texas (http://www.rulesbasedmedicine.com), where PAI-1 levels were analyzed by multiplex immunoassay on a human multi-analyte profile. The lowestdetectable level was 0.90 ng/mL. The intra- and inter-assay CVs were less than 11%.

Supplemental Section S4: Haplotype study in PROCARDIS and analysis tools and software packages used in individual cohorts

Besides selecting the SNP with the lowest P-value in each region for replication, extra SNPs were selected to tag the main haplotypes in the candidate regions, using the haplotype structure derived from PROCARDIS. Haplotypes were estimated for 1290 individuals using Haploview v.4.2. Only haplotypes with frequencies higher that 2% were considered. Selection of haplotype tagSNPs was performed by using the corresponding application in Haploview. TagSNPs were substituted by proxies at an r-square >0.98 if the latter showed lower *P*-values in the PAI-1 discovery meta-analysis. Among the tagSNPs, those showing the lowest *P* values were prioritized.

To evaluate population substratification in the **FHS**, principal component analyses was conducted using EIGENSTRAT[20] on the genotypes from 882 unrelated participants. The first 10 principal components were estimated and the loadings of these components were applied to all genotyped participants. When one of the ten principal components was associated with a trait, that principal component was adjusted for in linear regression models.In **PROCARDIS,** familial clustering was taken into account using a robust (sandwich) estimator of the variance. GWA analyses in **KORA** using linear regression with adjustment for age were performed by ProbABEL. Non-European ancestry in **TwinsUK** was assessed by cluster analysis in STRUCTURE or multidimensional scaling by comparison to the three HapMap phase 2 reference populations (CEU, YRI, CHB+JPT; www.hapmap.org). To conduct association analyses for natural log transformed PAI-1, the score test implemented in the software Merlin was used, adjusting for age, sex, and

family structure. In **MARTHA**, relatedness was investigated by pairwise clustering of identity by state distances (IBS) and multi-dimensional scaling implemented in PLINK software (v.1.04). Non-European ancestry was also investigated using the EIGENSTRAT program. Statistical analyses in **Health ABC** and **GeneSTAR** used R version 2.10, taking familial correlations into account in GeneSTAR. SNPTEST was used in **MESA** for the analysis of both genotyped and imputed SNPs to calculate beta coefficients and standard errors, and for inference of statistical significance using the score test. In **LURIC**, principal component analyses were done with GoldenHelix. Relatedness was investigated by pairwise clustering of identity by state distances (IBS) in PLINK (v.1.07) and GWA analyses were also done using PLINK.

Online Data Supplemental Section S5: Detailed descriptions of the *ex vivo* studies of global gene expression in tissues and cells

Advanced Study of Aortic Pathology (ASAP): Tissue samples were taken from patients undergoing aortic valve surgery at the Karolinska University Hospital, Solna, as described.[21]Biopsies were obtained at open-heart surgery from liver, mammary artery and dilated and non-dilated ascending aorta. Biobank materials were generated after informed consent from all participants had been obtained according to the declaration of Helsinki and with approval by the ethics committee of the Karolinska Institute. The medial and adventitial layers of the vascular specimens were isolated by adventicectomy, incubated with RNAlater (Ambion) and homogenized with a FastPrep (Qbiogene, Irvine, CA) using Lysing Matrix D tubes (Invitro cat.no. 6913-100). Total RNA was prepared using Trizol (BRL-Life Technologies) and RNeasy Mini kit (Qiagen), including treatment with RNase-free DNase set (Qiagen) according the manufacturer's instructions. The quality of RNA was determined on an Agilent 2100 bioanalyzer (Agilent Technologies Inc., Paolo Alto, CA, USA) and quantity was measured by a NanoDrop (Thermo Scientific). Extracted RNA was hybridized to Affymetrix ST 1.0 Exon arrays at the Karolinska Institute microarray core facility. For probe set and meta probe set level investigations, i.e. the analysis of whole-gene variation with genotype, cel files were pre-processed using Robust Multichip Average (RMA) normalization[22], as implemented in the Affymetrix Power Tools 1.10.2 package apt-probeset-summarize. All expression measurements were log2 transformed as part of the RMA normalization. Investigations were performed in the *extended* subset of the Affymetrix meta probe sets.

DNA was extracted from circulating blood cells, and DNA samples were genotyped using Illumina Human 610W-Quad Beadarrays at the SNP technology platform at Uppsala University. The GenomeStudio™ software from Illumina was used for genotype calling and quality control. The average call rate per SNP was 99.84%. Replicate genotyping of twelve samples showed an overall concordance of 99.99 %.

Cardiogenics Study: The present study included monocyte and macrophage gene expression data from a total of 917 patients and healthy individuals of European descent recruited in five centers within the Cardiogenics consortium (http://www.cardiogenics.eu). Healthy individuals (n=458) were recruited in Cambridge (UK). CAD/MI patients (n=459) were recruited in Leicester (n=161), Lübeck (n=102), Regensburg (n=122) and Paris (n=74). The study was approved by the Institutional Ethical Committee of each participating center.

RNA preparation from monocytes and macrophages: Monocyte isolation, macrophage culture and RNA extraction were performed separately in each center according to standardized procedures. All RNA samples were subsequently sent to one center (Paris) for amplification, whole-genome microarray gene expression profiling and bioinformatics analysis. In brief, blood samples (30ml) from fasting subjects were collected into EDTA and monocytes were isolated with CD14 micro beads (Miltenyi) according to the manufacturer's instructions Monocyte purity was measured as the percentage of CD14+ve cells analyzed by flow cytometry. Half of the isolated cell preparation was immediately used for RNA extraction and the remaining cells were cultured for 7 days at 37°C in

Macrophage-SFM medium (Invitrogen, Cat # 12065-074) supplemented with glutamine (1mM; Invitrogen) and recombinant human M-CSF (R&D Systems), the latter taken from the same batch. At day 4 the medium was replaced and at day 7 the cells were harvested. Isolated monocytes and 7 day macrophages were lysed in Trizol and RNA was extracted in chloroform and ethanol, washed in RNeasy columns and incubated with DNase 1 before extracting in RNase-free water. RNA was quantified by the Nanodrop® method before transfer to Paris on dry ice.

Whole-genome transcriptional profiling**:** Gene expression profiling was performed using the *Illumina* Human Ref-8 Sentrix Bead Chip arrays (*Illumina* Inc., San Diego, CA) containing 24,516 probes corresponding to 18,311 distinct genes and 21,793 Ref Seq annotated transcripts. mRNA was amplified and labelled using the *Illumina* Total Prep RNA Amplification Kit (Ambion, Inc., Austin, TX). After hybridization, array images were scanned using the *Illumina*BeadArray Reader and probe intensities were extracted using the Gene expression module (version 3.3.8) of the *IlluminaBeadStudio* software (version 3.1.30). Raw intensities were processed in R statistical environment using the Lumi and beadarray packages. All array outliers were excluded and only arrays with high concordance in terms of gene expression measures (pairwise Spearman correlation coefficients within each cell type >0.85) were included in the analyses. After data quality control, 849 monocyte (414 from cases and 435 from controls) and 684 macrophage (348 from cases and 336 from controls) RNA samples were available for statistical analyses.

Genome-wide genotyping, samples and SNPs filtering: EDTA anticoagulated venous blood samples were collected from all participants. Genomic DNA was extracted from peripheral blood monocytes by standard procedures (Qiagen) and sent to the Sanger Institute for genotyping. Genome-wide genotyping was carried out using two *Illumina* arrays, the Sentrix Human Custom 1.2M array

(1,115,839 SNPs and 80,128 CNVs) and the Human 610 Quad Custom array (594,398 SNP s and 66,049 CNVs). After sample filtering, which was based on sample call rate, ancestry, duplications and genetic relatedness, 802 samples were kept for expression quantitative trait locus (eQTL) analyses.

Statistical analysis: Associations between SNP genotype and gene expression level were examined using additive linear models, with each genotype being coded as 0, 1 or 2. Disease status (in Cardiogenics only), age and gender were adjusted for in the linear regression models. *P* values for all genotype-gene expression combinations in all tissue types were included in a false discovery rate (FDR) calculation, which was conducted by using the Benjamini-Hochberg method, as implemented in the *multtestR-package* (http://www.bioconductor.org/packages/release/bioc/html/multtest.html). In instances with several probes per gene, only the most significant association was reported in table 4, although the FDR was calculated using all association tests.

Online Data Supplemental Section S6: Functional Studies

Studies in HuH7 Cells

Cell culture: HuH7 cells were cultured in low-glucose Dulbecco's modified essential medium (DMEM) from Gibco supplemented with 10% fetal bovine serum (FBS) and 1% penicillin/streptomycin (P/S) in a 5% CO2 atmosphere at 37 ◦C. The cells were plated on six-well plates at an approximate density of 10^5 cells per well, and transfected with 50 pg of siRNA in 5ul Lipofectamine2000 reagent (Invitrogen) in a total volume of 2ml of Opti-MEM transfection medium per well. After 15 hours from transfection the medium was changed to 1% FBS supplemented medium with 1% P/S.

Normalization of gene and protein expression measurements between individual experiments: Total cellular RNA was isolated with a commercial kit (Omega Bio-Tek) and specific expression of *MUC3* and *SERPINE1* was quantified by real-time quantitative PCR using expression probes obtained from Applied Biosystems. *MUC3* and *SERPINE1* expression were normalized in each experiment using *RPLPO* expression as a reference. Variation between experiments was accounted for by setting the mean expression of the negative control in each experiment to 100%. For each experiment, the mean normalized expression of the *MUC3*-silenced cells was shown as percentage relative to the mean normalized expression of the cells transfected with negative control of that experiment. Experiments that did not reach at least an average of 50% silencing of *MUC3* expression in the 3 replicas were excluded from the analysis. PAI-1 protein levels for the *MUC3*-silenced cell media were expressed as percentage relative to the protein levels detected in the media of the cells transfected with negative control of that experiment.

A pre-designed Silencer Select siRNA (siRNA ID#s226934) targeting *MUC3* (Ambion) was transiently transfected into the hepatic cell line HuH7 obtained from the Health Science Research Resources Bank (cell no. JCRB0403; Osaka, Japan). In addition, a negative control siRNA from Invitrogen was used in each experiment following the same transfection protocol to account for differences in expression caused by the transfection reagents. RNA extraction and expression quantification was performed in cells that were harvested after 48h from transfection. Each experiment (n=18) was performed in triplicates. Aliquots of the cell culture media were removed at 96 hours after transfection and PAI-1 antigen levels were quantified using a commercial ELISA Kit (T6003 TriniLize PAI Antigen Kit, Medinor). In 15 independent experiments (with 3 replicates in each), the mean silencing of *MUC3* was -61.0%, which was accompanied by a highly statistically significant increase in *SERPINE1* expression (+45.5%, $P<0.0001$) when RNA was extracted 48 hours after transfection.

When we quantified the amount of PAI-1 protein released into the media after *MUC3* silencing, the most pronounced differences between *MUC3*-silenced and control cells were observed at 96 hours after transfection with *MUC3*siRNA (data not shown).Subsequent experiments (n=10, with 3 replicates in each) demonstrated a highly significant increase in PAI-1 protein levels in the media collected from *MUC3*-silenced compared to control cells (increase of 88.8%, $P<0.0001$, **Figure 2C**).

For verification, experiments were repeated with two additional siRNAs(siRNA ID#s61561, siRNA ID#200302) targeting different parts of the *MUC3* transcript and 2 specific negative control siRNAs,using the same protocol as in the previous experiments. No filtering for silencing efficiency was performed in these additional experiments. In total, 6 extra experiments were performed for each

additional siRNA (3 using negative control #1(4390843) and 3 using negative control #3(AM4615)). RNA extraction and expression quantification was performed in cells that were harvested after 48h from transfection as described previously.

The mean reduction of *MUC3* expression in the experiments using siRNAID#s61561 was -41.3% (P=0.0001), and a concomitant statistically significant increase in *SERPINE1* expression (+32.3%, P<0.0005) was observed when normalized by the results obtained with negative control #1. A corresponding -41.1% reduction in *MUC3* expression (P=0.0001) and a statistically significant increase in *SERPINE1* expression (+13.9%, P<0.015) were found when normalized by results obtained with negative control #3.Consistent with these results, the mean silencing of *MUC3* expression in the experiments using siRNAID#s200302 was -24.6% (P=0.0003), and a concomitant statistically significant increase in *SERPINE1* expression (+22.1%, P<0.0001) was observed when normalized by the results obtained with negative control #1. A corresponding -34.3% silencing (P=0.0002) and a statistically significant increase in *SERPINE1* expression (+18.1%, P<0.0004) were found when normalized by results obtained with negative control #3.

Ten additional experiments were then performed for each siRNA (5 using negative control #1(4390843) and 3 using negative control #3(AM4615)) to quantify secretion of PAI-1 in the media. Consistent with previous results, cells silenced with siRNAID#s61561 showed a statistically significant increase in PAI-1 secretion in the media (+15.4%, P=0.004). No statistically significant differences could be found in media of cells silenced withsiRNA ID#200302.

Studies in THP-1 Cells

Cell culture: THP-1 cells were cultured in RPMI 1640 medium with L-glutamine and 25 mM HEPES (Invitrogen, Life Technologies) supplemented with 10% fetal bovine serum (FBS), 1% penicillin/streptomycin (P/S), and 1% sodium pyruvate in a 5% CO2 atmosphere at 37 ◦C. The cells were plated on six-well plates at an approximate density of 4x105 cells per well, and transfected with 50 nM of siRNA in 7.5ul LipofectamineRNAiMAX reagent (Invitrogen) in a total volume of 1ml of RPMI Medium with 1% sodium pyruvate (without FBS or P/S) per well. After 5 hours from transfection the medium was changed to 1% FBS supplemented medium with 1% P/S and 1% sodium pyruvate.

Normalization of gene and protein expression measurements between individual experiments: Total cellular RNA was isolated with a commercial kit (Omega Bio-Tek) and specific expressions of *TRIP6* and *SERPINE1* were quantified by real-time quantitative PCR using expression probes obtained from Applied Biosystems. *TRIP6* and *SERPINE1* expression were normalized in each experiment using *RPLPO* expression as a reference. Variation between experiments was accounted for by setting the mean expression of the negative control in each experiment to 100%. For each experiment, the mean normalized expression of the *TRIP6*-silenced cells was shown as percentage relative to the mean normalized expression of the cells transfected with negative control of that experiment. A two-sample t-test with unequal variances was performed to compare the means of the two groups (*TRIP6*-silenced versus negative control). A significance level of 0.05 was applied.

Results: A pre-designed Silencer Select siRNA (siRNA ID#s14403) targeting *TRIP6* (Ambion) was transiently transfected into the monocyte cell line THP-1, and a negative control siRNA from Invitrogen was used in each experiment, following the same transfection and PAI-1 mRNA and protein expression protocols as were used for *MUC3*. Each TRIP6 knock-down experiment (n=16)

was performed in triplicate. Ten experiments were then repeated using a different pre-designed Silencer Select siRNA (siRNA ID#s14404) targeting *TRIP6* (Ambion) but otherwise using the same protocol. The mean silencing of *TRIP6* expression in the 16 experiments with s14403 was -51.4% (P<0.0001) (**Figure A**). However, our data did not show any statistically significant downstream difference in *SERPINE1* expression (+3.3%, *P*=0.66) after *TRIP6* silencing (**Figure B**). No detectable amounts of PAI-1 protein were found in the medium.

Figure A

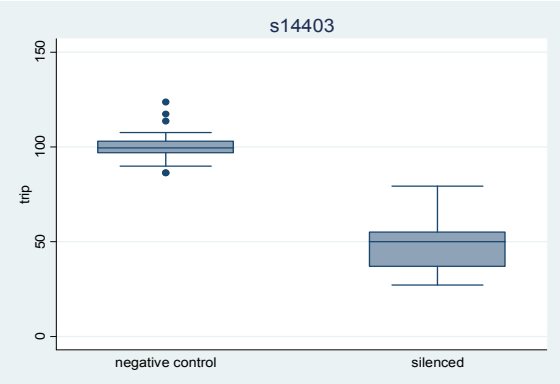
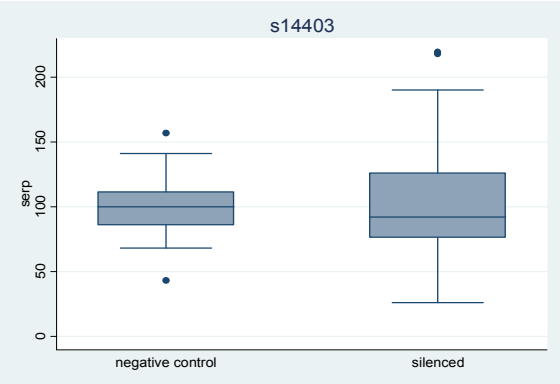


Figure B



Ten experiments were then repeated using siRNA s14404. Similarly, the mean silencing of *TRIP6* expression was -53.3% (P<0.0001) although no significant downstream differences in *SERPINE1* expression could be seen (-11.8%, *P*=0.12) and no detectable amounts of PAI-1 protein were found in the medium.

Online Data Supplemental Section S7: Acknowledgements

The authors acknowledge the essential role of the CHARGE (Cohorts for Heart and Aging Research in Genome Epidemiology) Consortium in providing a collaboration protocol and data sharing resource for this project. The authors fully acknowledge the thousands of study participants who volunteered their time to help advance science and the scores of research staff and scientists who have made this research possible. Twins UK would like to acknowledge the essential contribution of Peter Grant and Angela Carter from the Leeds Institute of Genetics, Health and Therapeutics, University of Leeds, UK for measurements of clotting factors phenotypes. The MONICA/KORA Augsburg studies thank GerlindeTrischler (University of Ulm) for excellent technical assistance. The following individuals contributed to the generation of the CARDIOGENICS monocyte and macrophage expression data set: Tony Attwood[1], Stephanie Belz[2], Peter Braund[3], Jessy Brocheton[4], François Cambien[4], Jason Cooper[5], Abi Crisp-Hihn[1], Patrick Diemert (formerly Linsel-Nitschke)[2], Panos Deloukas[6], Jeanette Eardman[2], Nicola Foad[1], Tiphaine Godefroy[4], Alison H Goodall[3,11], Jay Gracey[3], Emma Gray[6],, Rhian Gwilliams[6], Susanne Heimerl[7], Christian Hengstenberg[7], Jennifer Jolley[1], Unni Krishnan[3], Heather Lloyd-Jones[1], Ulrika Liljedahl[8], Ingrid Lugauer[7], Per Lundmark[8], Seraya Maouche[2,4], Jasbir S Moore[3] , Gilles Montalescot[4], David Muir[1], Elizabeth Murray[1], Chris P Nelson[3], Jessica Neudert[9], David Niblett[6], Karen O'Leary[1], Willem H Ouwehand[1,6], Helen Pollard[3], Carole Proust[4], Angela Rankin$_1$, Augusto Rendon[12], Catherine M Rice[6], Hendrik B Sager[2], Nilesh J Samani[3,11], Jennifer Sambrook[1], Gerd Schmitz[10], Michael Scholz[9], Laura Schroeder[2], Heribert Schunkert[2], Jonathan Stephens[1], Ann-Christine Syvannen[8], Stefanie Tennstedt (formerly Gulde)[2], Chris Wallace[5].

[1]Department of Haematology, University of Cambridge, Long Road, Cambridge, CB2 2PT, UK and National Health Service Blood and Transplant, Cambridge Centre, Long Road, Cambridge, CB2 2PT, UK; [2]Medizinische Klinik 2, Universität zu Lübeck, Lübeck Germany [3]Department of Cardiovascular Sciences, University of Leicester, Glenfield Hospital, Groby Road, Leicester, LE3 9QP, UK [4] INSERM UMRS 937, Pierre and Marie Curie University (UPMC, Paris 6) and Medical School, 91 Bd de l'Hôpital 75013, Paris, France [5]Juvenile Diabetes Research Foundation/Wellcome Trust Diabetes and Inflammation Laboratory, Department of Medical Genetics, Cambridge Institute for Medical Research, University of Cambridge, Wellcome Trust/MRC Building, Cambridge, CB2 0XY, UK [6]The Wellcome Trust Sanger Institute, Wellcome Trust Genome Campus, Hinxton, Cambridge CB10 1SA, UK; [7]Klinik und Poliklinik für Innere Medizin II, Universität Regensburg, Germany; [8]Molecular Medicine, Department of Medical Sciences, Uppsala University, Uppsala, Sweden; [9]Trium, Analysis Online GmbH, Hohenlindenerstr. 1, 81677, München, Germany; [10]Institut für Klinische Chemie und Laboratoriumsmedizin, Universität, Regensburg, D-93053 Regensburg, Germany; [11] Leicester NIHR Biomedical Research Unit in Cardiovascular Disease, Glenfield Hospital, Leicester, LE3 9QP, UK; [12] European Bioinformatics Institute, Wellcome Trust Genome Campus, Hinxton, Cambridge, CB10 1SD, UK.

Online Data Supplemental Section S8: Funding Sources

The **FHS** was partially supported by the National Heart, Lung, and Blood Institute's (NHLBI's) Framingham Heart Study (Contract No. N01-HC-25195) 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. The analyses reflect intellectual input and resource development from the Framingham Heart Study investigators participating in the SNP Health Association Resource (SHARe) project. Partial investigator support was provided by the National Institute of Diabetes and Digestive and Kidney DiseasesK24 DK080140 (JB Meigs). The **Health ABC** involvement wasfunded by the National Institutes of Aging (NIA contracts N01AG62101, N01AG62103, and N01AG62106). The genome-wide association study was funded by NIA grant 1R01AG032098-01A1 to Wake Forest University Health Sciences, and genotyping services were provided by the Center for Inherited Disease Research (CIDR). CIDR is fully funded through a federal contract from the National Institutes of Health to The Johns Hopkins University (Contract No. HHSN268200782096C). This research was also supported in part by the Intramural Research Program of the NIH, National Institute on Aging. The **MARTHA** study was supported by a grant from the Program Hospitalier de la Recherche Clinique. TiphaineOudot-Mellakh was supported by a grant from the Fondation pour la RechercheMédicale. The **MONICA/KORA** Augsburg studies were financed by the Helmholtz ZentrumMünchen, German Research Center for Environmental Health, Neuherberg, Germany and supported by grants from the German Federal Ministry of Education and Research (BMBF). Part of this work was financed by the German National Genome Research Network (NGFNPlus, project number 01GS0834) and through

additional funds from the University of Ulm. Furthermore, the research was supported within the Munich Center of Health Sciences (MC Health) as part of the Ludwig Maximilians University (LMU) initiative. **PROCARDIS** was supported by the European Community Sixth Framework Program (LSHM-CT- 2007-037273), AstraZeneca, the British Heart Foundation, the Wellcome Trust (Contract No. 075491/Z/04), the Swedish Research Council, the Knut and Alice Wallenberg Foundation, the Swedish Heart-Lung Foundation, the Torsten and Ragnar Söderberg Foundation, the Strategic Cardiovascular and Diabetes Programs of Karolinska Institutet and Stockholm County Council, the Foundation for Strategic Research and the Stockholm County Council. Maria Sabater-Lleal is a recipient of a Marie CurieIntra European Fellowship within the 7th Framework Programme of the European Union (PIEF-GA-2009-252361). **PREVEND** genetics was supported by the Dutch Kidney Foundation (Grant E033), The Netherlands Heart Foundation (Grant 2006B140, 2006T003), National Institutes of Health (grant LM010098, HL65234, HL67466, RR018787) and the EU project grant GENECURE (FP-6 LSHM CT 2006 037697). Pim van der Harst is supported by the NWO VENI grant 91676170 and the Dutch Inter University Cardiology Institute Netherlands (ICIN). Folkert W. Asselbergs is supported by a clinical fellowship from the Netherlands Organisation for Health Research and Development (ZonMw grant 90700342). The **Twins UK** study was funded by the Wellcome Trust; European Community's Sixth and Seventh Framework Programmes (FP-6/2005-2008) LIFE SCIENCES & HEALTH (Ref 005268 Genetic regulation of the end stage clotting process that leads to thrombotic stroke: The EuroClot Consortium and (FP7/2007-2013), ENGAGE project HEALTH-F4-2007-201413) and the FP-5 GenomEUtwin Project (QLG2-CT-2002-01254). The study also receives support from the Dept of Health via the National Institute for Health Research (NIHR) comprehensive Biomedical Research Centre award to Guy's & St Thomas' NHS Foundation Trust in partnership with King's

College London. Timothy D. Spector is an NIHR Senior Investigator. The project also received support from a Biotechnology and Biological Sciences Research Council (BBSRC) project grant. (G20234). The authors acknowledge the funding and support of the National Eye Institute via an NIH/CIDR genotyping project (PI: Terri Young). **ARIC** The Atherosclerosis Risk in Communities Study was supported by NHLBI contracts N01-HC-55015, N01-HC-55016, N01-HC-55018, N01-HC-55019, N01-HC-55020, N01-HC-55021, and N01-HC-55022 and grants R01-HL-087641, R01-HL-59367 and R01-HL-086694; National Human Genome Research Institute contract U01-HG-004402; and NIH contract HHSN268200625226C. The infrastructure was partly supported by grant number UL1-RR-025005, a component of the NIH and NIH Roadmap for Medical Research. Part of the work was supported by grant R01 HL095603. The authors thank the staff and participants of the ARIC study for their important contributions. Further, ARIC would like to acknowledge the University of Minnesota Supercomputing Institute for use of the blade supercomputers. **Cardiovascular Health Study**: This CHS research was supported by NHLBI contracts N01-HC-85239, N01-HC-85079 through N01-HC-85086; N01-HC-35129, N01 HC-15103, N01 HC-55222, N01-HC-75150, N01-HC-45133, HHSN268201200036C and NHLBI grants HL080295, HL087652, HL105756 with additional contribution from NINDS. Additional support was provided through AG-023629, AG-15928, AG-20098, and AG-027058 from the NIA. See also http://www.chs-nhlbi.org/pi.htm. DNA handling and genotyping was supported in part by National Center of Advancing Translational Technologies CTSI grant UL1TR000124 and National Institute of Diabetes and Digestive and Kidney Diseases grant DK063491 to the Southern California Diabetes Endocrinology Research Center. The **CHS LITE** was funded by grant R01-HL59367 from the National Heart, Lung, and Blood Institute. **GeneSTAR**was supported by grants 1U01 HL072518 and R01 HL59684 from the National Institutes of Health, National Heart, Lung, and Blood Institute, Bethesda, Maryland,

and the Johns Hopkins University School of Medicine General Clinical Research Center, National Institutes of Health Grant M01 RR00052. **MESA** and the MESA SHARe project are conducted and supported by contracts N01-HC-95159 through N01-HC-95169 and RR-024156 from the NHLBI. Funding for MESA SHARe genotyping was provided by NHLBI Contract N02-HL-6-4278. MESA Family is conducted and supported in collaboration with MESA investigators; support is provided by grants and contracts R01HL071051, R01HL071205, R01HL071250, R01HL071251, R01HL071252, R01HL071258, R01HL071259. A full list of participating MESA investigators and institutions can be found at http://www.mesa-nhlbi.org. The **Rotterdam Study** was supported by the Erasmus Medical Center and Erasmus University Rotterdam, the Netherlands Organisation of Scientific Research (NOW, NWO 918-46-615, 948-00-010 and 175.010.2005.011), the Netherlands Organisation for Health Research and Development (ZonMw), the Research Institute for Diseases in the Elderly, the Ministry of Education, Culture, and Science, the Ministry of Health, Welfare, and Sports, the European Commission (DG-XII), the Municipality of Rotterdam and the Netherlands Genomics Initiative (NWO 050-060-810). **The Coronary Artery Disease Genetics Consortium:** Jemma C Hopewell acknowledges support from the BHF Centre of Research Excellence, Oxford. Dr. John Danesh received funding from The Wellcome Trust, British Heart Foundation, UK Medical Research Council, and National Institute for Health Research. The **ASAP Study** was supported by the Swedish Research Council [12660], the Swedish Heart-Lung Foundation [20090541], the European Commission [FAD, Health F2 2008 200647] and a donation from Fredrik Lundberg. CARDIOGENICS was funded by the European Union FP6 program (LSHM-CT-2006-037593).

Supplemental Table S1.Characteristics of study participants in the discovery cohorts

Characteristic	FHS	PROCARDIS Cases	PROCARDIS Controls	KORA	Twins UK	MARTHA	Health ABC	PREVEND
Counts	6,634	1,922	1,294	1,565	2,017	851	1,645	3,671
Age, years	46.1 (11.6)	61.4 (7.0)	55.3 (7.2)	62.6 (10.1)	49.3 (12.4)	45.9 (14.8)	73.8 (2.8)	49.6 (12.5)
Male, %	46.4	78.5	74.3	49.3	4.8	29.5	53.0	49.5
European ancestry, %	100	100	100	100	100	100	100	100
BMI, kg/m^2	27.9 (5.1)	28.1 (4.7)	25.9 (3.6)	28.2 (4.5)	26.2 (4.9)	24.7 (4.4)	26.6 (4.1)	26.1 (4.3)
Arterial disease hx, %	10.8	100.0	0	11.8	4.9	1.0	26.0	6.8
Venous disease hx, %	NA	NA	NA	2.8	2.5	100	NA	NA
T2D, %	4.9	16.4	2.5	10.6	0.2	NA	20.0	4.1
Measurement	Citrated plasma	Citrated plasma	EDTA plasma	EDTA plasma	Citrated plasma	Citrated plasma	Citrated plasma	EDTA plasma
Assay Type	TintElize PAI-1 ELISA	TriniLIZE PAI-1 ELISA	TriniLIZE PAI-1 ELISA	Technozym PAI-1 actibind ELISA	TintElize PAI-1 ELISA	TriniLIZE PAI-1 activity	In-house ELISA	Technozym PAI-1 antigen ELISA
PAI-1, ng/ml or *U/ml	18.2 (14.9)	36.0 (19.7)	44.6 (18.0)	16.1 (21.9)*	11.7 (13.4)	10.4 (11.8)*	29.3 (24.2)	108.9 (119.3)
PAI-1 Quartiles	8.2; 14.2; 23.8	22.2; 32.0; 45.7	31.3; 43.0; 56.7	4.6; 8.7; 18.5	4.0; 7.2; 14.2	3.0; 6.0; 14.0	13.0; 22.0; 38.0	41.7; 73.3;126.6

Values shows as mean (SD) unless otherwise specified.
Arterial disease includes history of myocardial infarction (MI), angina, stroke, or transient ischemic attack except for KORA, where arterial disease is history of MI, angina or stroke, and for PREVEND, where arterial disease is history of MI and stroke as indicated on the questionnaire completed by the participant.
Venous disease includes history of deep vein thrombosis (DVT) or pulmonary embolism except for KORA where venous thrombosis is only DVT.
BMI, body mass index; T2D, type-2 diabetes; PAI-1, plasminogen activator inhibitor type-1; NA, not applicable.

Supplemental Table S2. Characteristics of study participants in the replication cohorts

Characteristic	PREVEND	CHS	CHS_LITE	ARIC	MARTHA10	MESA	Rotterdam	GeneSTAR	LURIC
Counts	4,539	368	210	410	498	433	840	551	2915
Age, years	49.1 (12.9)	70.0 (3.7)	73.4 (5.5)	56.0 (5.1)	49.4 (15.6)	60.9 (10.4)	73.1 (7.4)	55.5 (7.5)	62.6 (10.7)
Male, %	51.7	0.0	41.4	59.3	41.4	46.0	45.1	61.7	69.7
BMI, kg/m^2	26.1 (4.2)	25.8 (4.4)	26.4 (4.5)	25.9 (4.2)	25.6 (6.1)	28.0 (5.2)	26.7 (3.9)	29.1 (5.4)	27.5 (4.1)
Arterial disease hx, %	7.7	0.0	24.8	0.0	1.0	0.0	19.9	29.6	77.1
Venous disease hx, %	NA	3.8	3.5	2.8	100.0	4.4	4.5	--	6.0
T2D, %	3.9	1.1	13.8	4.9	--	4.2	13.2	14.0	39.5
Plasma	EDTA plasma	Citrate	Citrate	Citrate	Citrate	Citrate	Citrate	Citrate	Citrate
Measurement Assay Type	Technozym PAI-1 actibind ELISA	In-house ELISA	In-house ELISA	IMUBIND PAI-1 ELISA	TriniLIZE PAI-1 activity	In-house ELISA	RBM multiplexed immunoassay human multianalyte profile	Chromogenic substrate assay	Chromogenic substrate
PAI-1, ng/ml or *U/ml	103.4 (111.6)	25.5 (26.1)	45.3 (35.7)	20.8 (24.6)	13.7 (15.0)*	26.7 (30.4)	23.6 (19.2)	17.7 (12.9)*	30.5 (26.0)
Quartiles	41.3; 71.3; 122.8	9.0; 17.0; 34.0	20.1; 36.4; 59.4	7.0; 13.4; 26.1	4.0; 8.0; 19.0	9.0; 19.0; 34.0	10.6; 17.8; 29.4	8.6; 14.9; 24.5	15.9; 24.7; 37.0

Arterial diseases in ARIC and CHS LITE: self-reported prevalent coronary heart disease or stroke/transient ischemic attack at baseline; Venous disease in ARIC and CHS LITE: self-reported prevalent venous thromboembolic disease at baseline. T2D in CHS: known, treated diabetes.

BMI, body mass index; T2D, type-2 diabetes; PAI-1, plasminogen activator inhibitor type 1; NA, not applicable.

Supplemental Table S3. Methods, quality control criteria and results of genotyping and imputation in the discovery cohorts

	FHS	PROCARDIS	KORA	Twins UK	MARTHA08	Health ABC	PREVEND
Platform	Affymetrix	Illlumina	Affymetrix	Illumina	Illumina	Illlumina	Illumina
Array	500K (250K Nsp and 250K Sty) and 50K gene focused array	1M, 610 Quad	500K	300K, 300 Duo, 550K, and 610K Quad	610 Quad	1M Duo	HumanCyto SNP-12v1
SNPs investigated	490,700 (500K) 48,195 (50K)	573,015 (1M) 582,892 (610K)	490,700	582,892	567,589	1,151,215	281,748 Autosomal
SNP exclusion criteria							
Call rate	≤97%	<95%	<93%	≤97% (MAF > 0.05), ≤99% (MAF >0.01)	<99%	<97%	<98%
MAF	--	<0.01	--	<0.01	<0.01	<0.01	<0.01
HWE P-value	$<1.0x10^{-6}$	$<1.0x10^{-6}$	--	$<1.0x10^{-6}$	$<.0x10^{-5}$	$<1.0x10^{-6}$	$<.0x10^{-5}$
Variants included for imputation	343,361 (500K) 34,841 (50K)	498,717 (1M) 514,950 (610K)	490,032	874,733 (merged data set was imputed)	494,721	914,263	244,868
Percent of variants included	70% (500K) 72% (50K)	87% (1M) 88.3% (610K)	100%	--	87%	79.4%	87%
Imputation software	MACH 1.0.15	MACH 1.0.16	MACH 1.0.9	Impute v2	MACH 1.0.16	MACH 1.0.16	BEAGLE 3.2
Genome build	36.2	36	35.21	36	35	36	36r23a
Total # of SNPs	2,543,887	2,543,887	2,557,252	2,657,660	2,557,252	2,543,887	2,269,121

SNP, single nucleotide polymorphism; MAF, minor allele frequency; HWE, Hardy-Weinberg equilibrium.

Supplemental Table S4. Methods, quality control criteria and results of genotyping and imputation in the replication cohorts

	PREVEND	CHS	CHS_LITE	ARIC	GeneSTAR	MARTHA10	MESA	Rotterdam	LURIC
Platform	KASP*	Illumina	Illumina	Affymetrix	Illumina	Illumina	Affymetrix	IlluminaInfinium II	Affymetrix
Array	Wet-lab genotyping	370CNV BeadChip	370CNV BeadChip	Human SNP Array 6.0	1Mv1_C	660W	6.0	550K	500k; 6.0
SNPs investigated	10	346,831	346,831	841,820	1,004,036	557,124	909,622	547,458	500,568; 909,622
SNP exclusion criteria									
Call rate	<98%	< 97%	< 97%	< 95%	<90%	<99%	<95%	<98%	<95%
MAF	NA	NA	NA	<0.01	<0.01	<0.01	post- assoc testing	<1%	NA
HWE p-value	<0.05	$<10^{-5}$	$<10^{-5}$	$<10^{-5}$	$<10^{-6}$	$<10^{-5}$	post assoc testing	$<10^{-6}$	$<10^{-6}$
Variants included for imputation	NA	306,655	306,655	669,450	730,790	501,773	879,981	512,349	393,157; 893,909
Percent of variants included	100%	88.4%	88.4%	79.5%	73%	90%	96.7%	93.6%	78.5%; 98.3%
Imputation	NA	BimBam	BimBam	MACH	MACH	Mach	Impute	Mach	MaCH
Imputation software version	NA	v.0.99	v.0.99	v.1.0.16	v.1.0.16	v.1.0.16	v.2	v.1.00.15	v.1
Genome build	NA	Build 36	Build 36	Build 36	Build 36	Build 35	Build 36	Build 36	Build 36
Total # of SNPs	10	2,333,140	2,333,140	2,543,887	2,543,887	2,557,252	3,854,659	2,543,887	6,858,238

*KBioscience Competitive Allele-Specific PCR genotyping system (KASP, www.kbioscience.co.uk).
SNP, single nucleotide polymorphism; MAF, minor allele frequency; HWE, Hardy-Weinberg equilibrium; NA, not applicable.

Supplemental Table S5. Conditional and secondary analysis results for the 10 selected SNPs in the discovery phase.

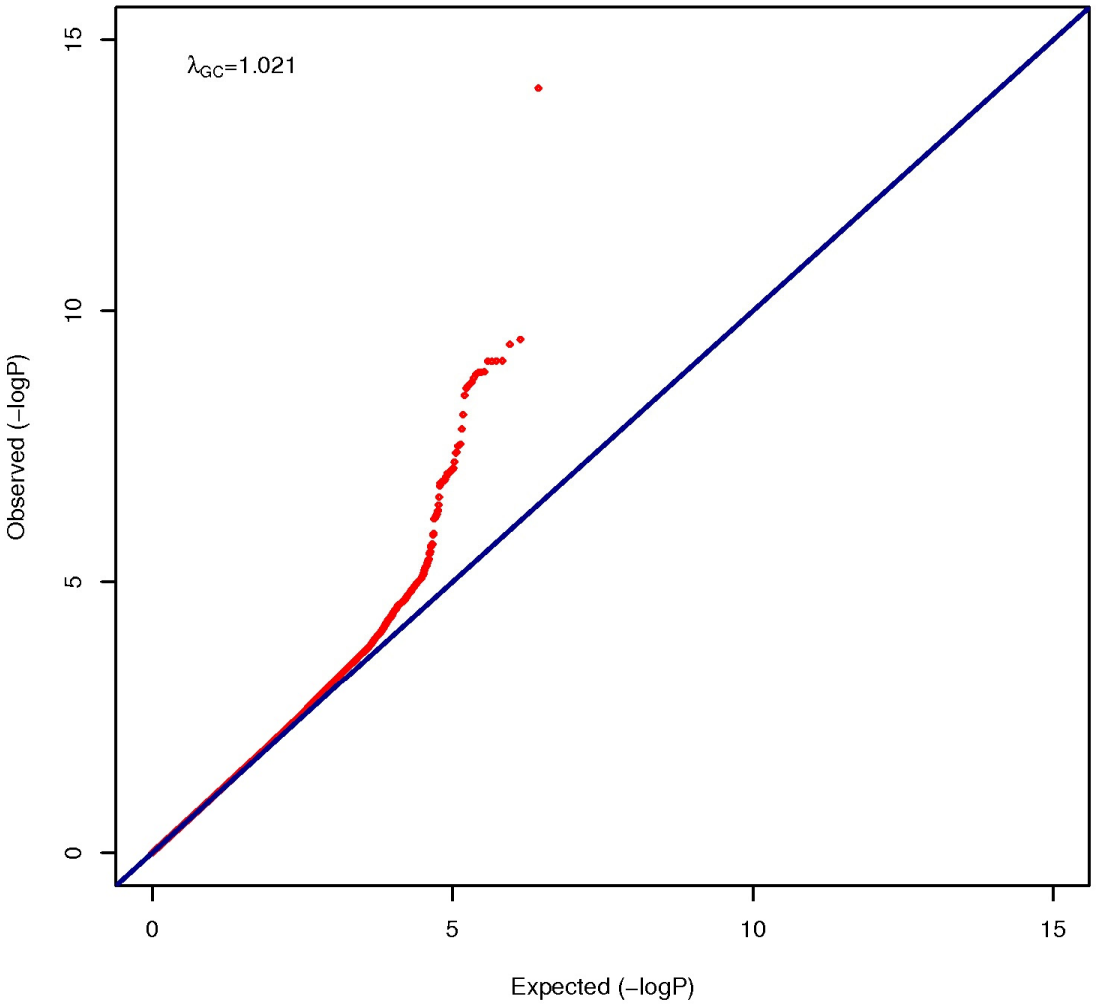
Region	SNP	Chr	Closest gene	Type	Alleles	Discovery		Adjusted for BMI		Conditional on 4G/5G proxy (rs2227631)	
						Beta	*P*-value	Beta	*P*-value	Beta	*P*-value
3p25.2	rs11128603	3	*PPARG*	Index SNP	A -> G	0.086	$2.9x10^{-8}$	0.058	$1.2x10^{-6}$	0.0585	$3.2x10^{-6}$
	rs1801282	3	*PPARG*	nsSNP	C -> G	0.069	$1.4x10^{-7}$	0.0534	$4.7x10^{-7}$	0.0538	$2.9x10^{-6}$
7q22.1	rs2227631	7	*SERPINE1*	Index SNP	A -> G	0.076	$7.8x10^{-15}$	0.0719	$3.0x10^{-22}$	--	--
7q22.1	rs6976053	7	*ACHE*	Index SNP	T -> C	0.054	$3.4x10^{-10}$	0.0379	$7.1x10^{-8}$	0.0473	$4.0x10^{-10}$
	rs2075756	7	*TRIP6*	nsSNP	A -> G	0.058	$2.7x10^{-9}$	0.0412	$1.4x10^{-7}$	0.0429	$4.2x10^{-7}$
	rs314376	7	*SLC12A9*	Tag SNP	G -> A	0.052	$2.4x10^{-9}$	0.0384	$4.0x10^{-8}$	0.0464	$1.2x10^{-9}$
	rs12672665	7	*SRRT*	Tag SNP	G -> A	0.047	$6.1x10^{-8}$	0.0272	$1.2x10^{-4}$	0.0442	$8.7x10^{-9}$
	rs3847067	7	*ACHE*	Tag SNP	A -> C	0.059	$4.1x10^{-10}$	0.042	$7.5x10^{-8}$	0.043	$3.9x10^{-7}$
11p15.2	rs6486122	11	*ARNTL*	index SNP	T -> C	0.051	$3.0x10^{-8}$	0.0441	$3.7x10^{-9}$	0.0435	$8.5x10^{-8}$
	rs3816360	11	*ARNTL*	Tag SNP	C -> T	0.048	$9.9x10^{-8}$	0.0447	$1.4x10^{-9}$	0.0424	$1.1x10^{-7}$

Supplemental Table S6. Association statistics for each individual cohort for the 10 selected SNPs in the discovery phase

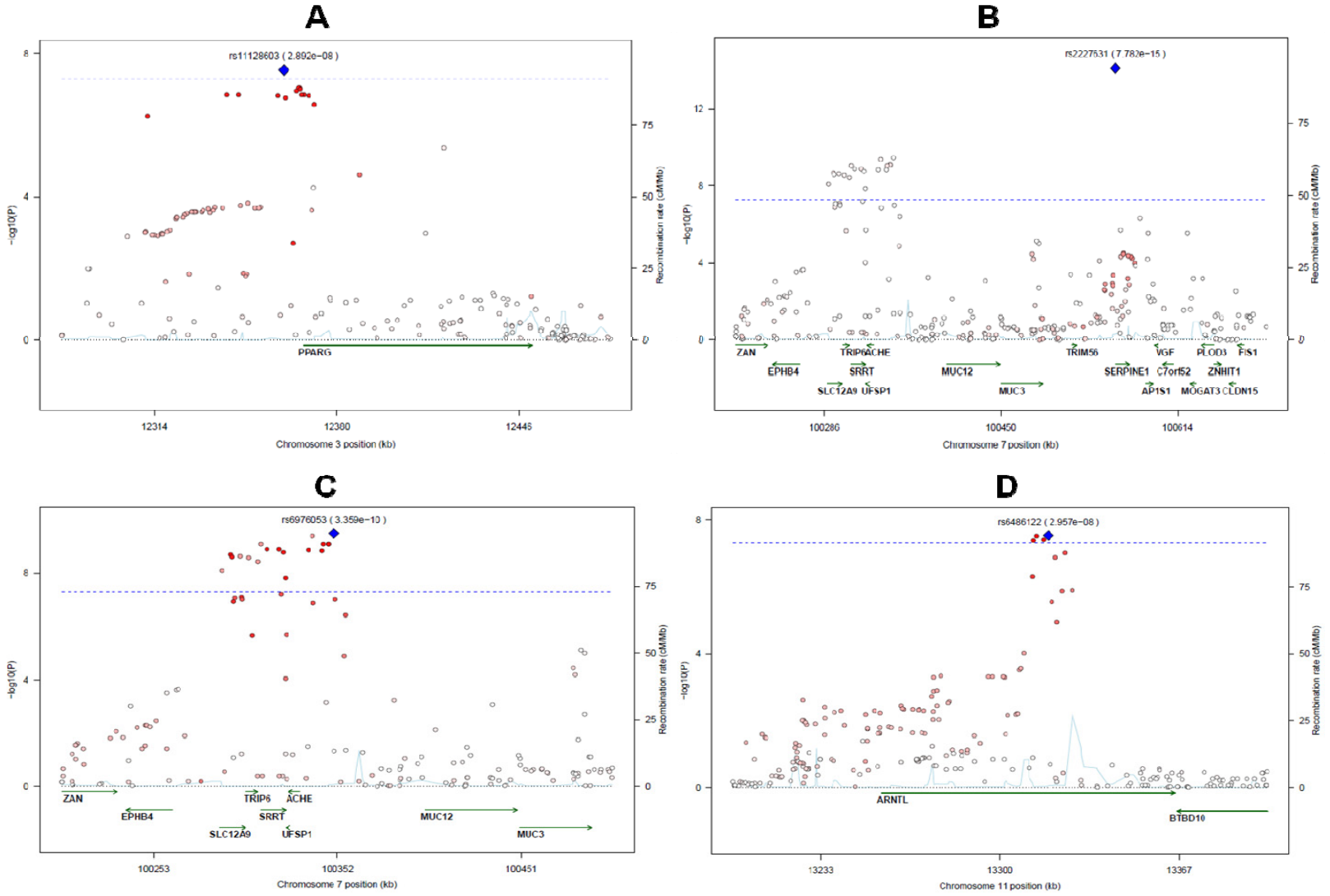
		FHS	PROCARDIS Cases	PROCARDIS Controls	KORA	Twins UK	MARTHA	Health ABC	PREVEND
GWAS lambda		1.029	1.023	1.013	1.000	1.032	1.000	1.016	1.058
3p25.2	rs11128603	0.09, 0.070, 0.004, 0.84	0.11, 0.070, 0.2, 0.87	0.12, 0.158, 0.009, 0.87	0.15, 0.163, 0.006, 0.87	0.06, 0.041, 0.5, 0.85	0.08, 0.116, 0.2, 0.85	0.10, 0.052, 0.3, 0.87	0.11, 0.102, 0.002, 0.78
	rs1801282	0.11, 0.059, 0.004, 0.99	0.12, 0.019, 0.7, 0.98	0.14, 0.154, 0.007, 0.98	0.19, 0.127, 0.01, 1.00	0.12, 0.061, 0.2, 0.98	0.09, 0.098, 0.2, 1.00	0.13, 0.030, 0.4, 1.00	0.13, 0.083, 0.003, 0.96
7q22.1	rs2227631	0.40, 0.082, 4.6E-8, 0.75	0.42, 0.071, 0.04, 1.00	0.40, 0.028, 0.5,0.9986	0.42, 0.142, 4E-3, 0.76	0.39, 0.103, 0.003, 0.96	0.41, 0.057, 0.3, 1	0.41, 0.115, 1.6E-05, 1	0.45, 0.003, 0.9, 0.50
7q22.1	rs6976053	0.47, 0.042, 0.001, 1.00	0.49, 0.069, 0.04, 1.00	0.50, 0.012, 0.8, 1	0.49, 0.054, 0.1, 0.93	0.47, 0.053, 0.09, 0.99	0.46, 0.157, 0.001, 1.00	0.47, 0.089, 8.4E-4, 1.00	0.49, 0.050, 0.008, 1.00
	rs2075756	0.28, 0.050, 6.6E-4, 0.94	0.29, 0.027, 0.47, 1.00	0.29, 0.012, 0.8, 1	0.26, 0.099, 0.03, 0.65	0.28, 0.079, 0.02, 0.99	0.30, 0.031, 0.5, 1.00	0.29, 0.142, 1.0E-6, 1.00	0.28, 0.037, 0.08, 0.99
	rs314376	0.47, 0.041, 0.002,0.97	0.49, 0.064, 0.06, 1.00	0.50, 0.008, 0.8, 1	0.50, 0.062, 0.1, 0.77	0.47, 0.05, 0.1, 0.99	0.47, 0.143, 0.004, 1.00	0.48, 0.096, 3.4E-4, 1.00	0.50, 0.042, 0.02, 0.97
	rs12672665	0.48, 0.034, 0.009, 0.98	0.47, 0.054, 0.1, 1.00	0.47, 0.008, 0.8, 0.99	0.45, 0.063, 0.09, 0.82	0.50, 0.052, 0.1, 0.96	0.48, 0.124, 0.01, 1.00	0.48, 0.088, 9.8E-4, 1.00	0.48, 0.042, 0.02, 0.98
	rs3847067	0.29, 0.053, 2.3E-4, 0.98	0.30, 0.030, 0.4, 0.99	0.30, 0.006, 0.9, 0.99	0.29, 0.088, 0.04, 0.70	0.29, 0.079, 0.02, 0.98	0.30, 0.026, 0.6, 1.00	0.30, 0.137, 2.4E-6, 1.00	0.29, 0.047, 0.02, 0.98
11p15.2	rs6486122	0.30, 0.048, 5.6E-4, 0.99	0.33, 0.041, 0.3, 0.99	0.35, 0.109, 0.009, 0.99	0.30, 0.035, 0.4, 0.98	0.32, 0.018, 0.6, 0.99	0.29, 0.055, 0.3, 0.99	0.31, 0.043, 0.1, 0.98	0.32, 0.067, 0.001, 1.00
	rs3816360	0.32, 0.043, 0.002, 1.00	0.34, 0.056, 0.1, 1.00	0.36, 0.120, 0.004, 0.99	0.33, 0.022, 0.6, 1.00	0.34, 0.014, 0.7, 0.99	0.30, 0.041, 0.4484, 1.00	0.33, 0.055, 0.05, 1.00	0.35, 0.059, 0.002, 1.02

The association statistics includes MAF, beta of effect allele, *P*-value, and imputation R2, separated by comma.

Supplemental Figure S1. QQ plot showing the results for the 2,445,683 meta-analyzed SNPs in the discovery cohorts (a total of 19,683 individuals).

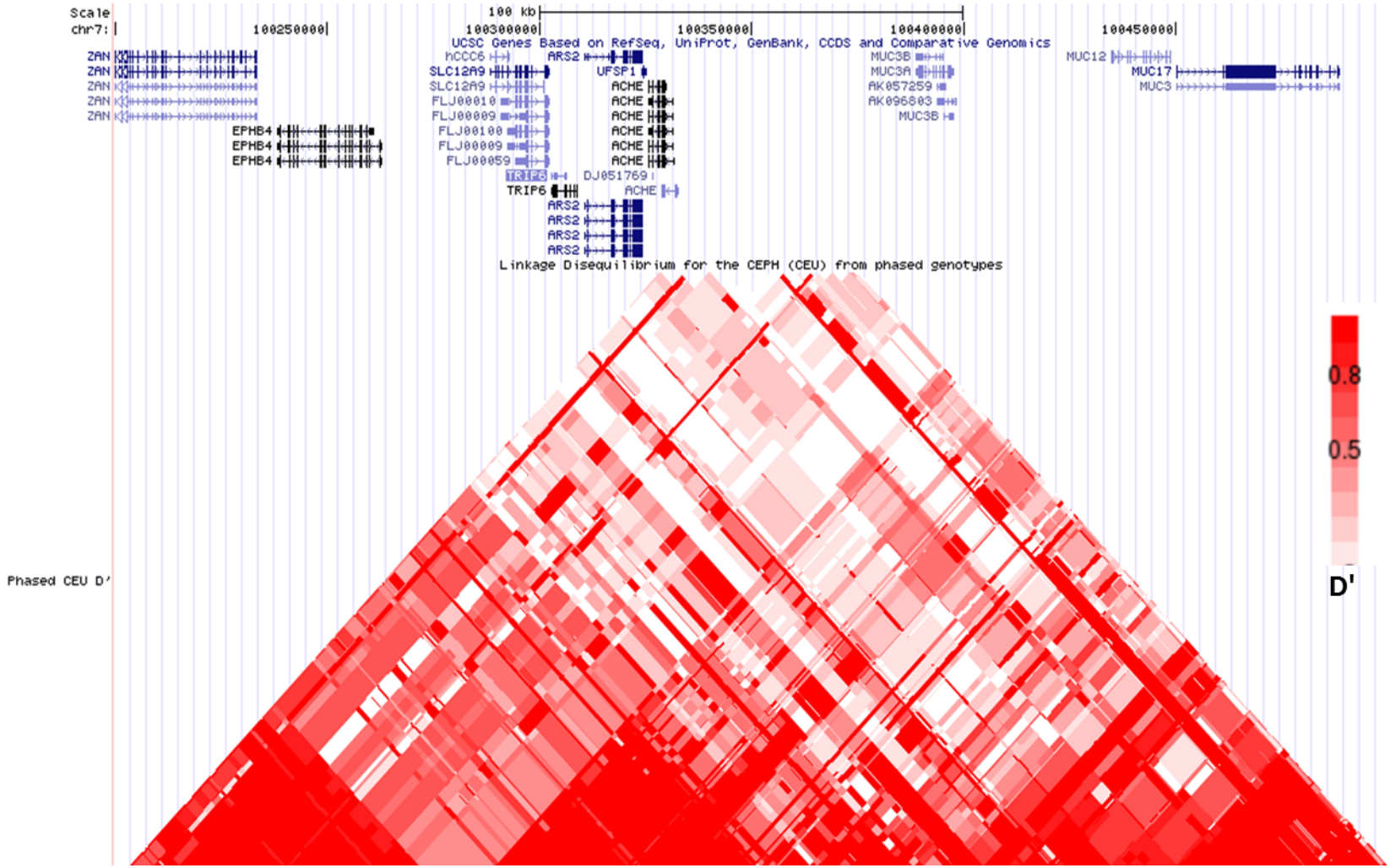


Supplemental Figure S2. Regional plots of 4 chromosomal loci (A-D) reaching genome-wide significance in discovery.



Each panel shows SNPs plotted by position (build 36) on chromosome against −log10(*P*),with estimated recombination rates from HapMap rel27 (CEU populations) plotted in light blue to reflect the local LD structure on a secondary Y axis. The most significant SNP is shown as a blue diamond. The surrounding SNPs are colour-coded to reflect their LD with the lead SNP, measured in R2. The redder the colour, the higher the LD. Green bars represent RefSeq genes in the region. Figure S1B shows that the two loci on chromosome 7 are about 200kb apart and independent. Figure S1C shows the association signals for the first locus on chromosome 7.

Supplemental Figure S2. LD plot of the novel chromosome 7 loci.



LD plot is generated from UCSC genome browser, based on HapMap CEU data. The color reflects the extent of LD level, the redder the color, the higher the LD (measured in D-prime).

Supplemental References:

1. Feinleib M, Kannel WB, Garrison RJ, McNamara PM, Castelli WP. The Framingham Offspring Study. Design and preliminary data. *Prev Med.* 1975;4(4):518-525.

2. Clarke R, Peden JF, Hopewell JC, Kyriakou T, Goel A, Heath SC, Parish S, Barlera S, Franzosi MG, Rust S, Bennett D, Silveira A, Malarstig A, Green FR, Lathrop M, Gigante B, Leander K, de Faire U, Seedorf U, Hamsten A, Collins R, Watkins H, Farrall M. Genetic variants associated with Lp(a) lipoprotein level and coronary disease. *N Engl J Med.* 2009;361(26):2518-2528.

3. Wichmann HE, Gieger C, Illig T. KORA-gen--resource for population genetics, controls and a broad spectrum of disease phenotypes. *Gesundheitswesen.* 2005;67 Suppl 1:S26-30.

4. Spector TD, Williams FM. The UK Adult Twin Registry (TwinsUK). *Twin Res Hum Genet.* 2006;9(6):899-906.

5. Tregouet DA, Heath S, Saut N, Biron-Andreani C, Schved JF, Pernod G, Galan P, Drouet L, Zelenika D, Juhan-Vague I, Alessi MC, Tiret L, Lathrop M, Emmerich J, Morange PE. Common susceptibility alleles are unlikely to contribute as strongly as the FV and ABO loci to VTE risk: results from a GWAS approach. *Blood.* 2009;113(21):5298-5303.

6. Fried LP, Borhani NO, Enright P, Furberg CD, Gardin JM, Kronmal RA, Kuller LH, Manolio TA, Mittelmark MB, Newman A, O'Leary DH, Psaty BM, Rautaharju P, Tracy RP, Weiler PG. The Cardiovascular Health Study: design and rationale. *Ann Epidemiol.* 1991;1(3):263-276.

7. The Atherosclerosis Risk in Communities (ARIC) Study: design and objectives. The ARIC investigators. *Am J Epidemiol.* 1989;129(4):687-702.

8. Folsom AR, Aleksic N, Park E, Salomaa V, Juneja H, Wu KK. Prospective study of fibrinolytic factors and incident coronary heart disease: the Atherosclerosis Risk in Communities (ARIC) Study. *Arterioscler Thromb Vasc Biol.* 2001;21(4):611-617.

9. Folsom AR, Cushman M, Heckbert SR, Rosamond WD, Aleksic N. Prospective study of fibrinolytic markers and venous thromboembolism. *J Clin Epidemiol.* 2003;56(6):598-603.

10. Bild DE, Bluemke DA, Burke GL, Detrano R, Diez Roux AV, Folsom AR, Greenland P, Jacob DR, Jr., Kronmal R, Liu K, Nelson JC, O'Leary D, Saad MF, Shea S, Szklo M, Tracy RP. Multi-ethnic study of atherosclerosis: objectives and design. *Am J Epidemiol.* 2002;156(9):871-881.

11. Winkelmann BR, Marz W, Boehm BO, Zotz R, Hager J, Hellstern P, Senges J. Rationale and design of the LURIC study--a resource for functional genomics, pharmacogenomics and long-term prognosis of cardiovascular disease. *Pharmacogenomics.* 2001;2(1 Suppl 1):S1-73.

12. Hofman A, Breteler MM, van Duijn CM, Janssen HL, Krestin GP, Kuipers EJ, Stricker BH, Tiemeier H, Uitterlinden AG, Vingerling JR, Witteman JC. The Rotterdam Study: 2010 objectives and design update. *Eur J Epidemiol.* 2009;24(9):553-572.

13. Seshadri S, Fitzpatrick AL, Ikram MA, DeStefano AL, Gudnason V, Boada M, Bis JC, Smith AV, Carassquillo MM, Lambert JC, Harold D, Schrijvers EM, Ramirez-Lorca R, Debette S, Longstreth WT, Jr., Janssens AC, Pankratz VS, Dartigues JF, Hollingworth P, Aspelund T, Hernandez I, Beiser A, Kuller LH, Koudstaal PJ, Dickson DW, Tzourio C, Abraham R, Antunez C, Du Y, Rotter JI, Aulchenko YS, Harris TB, Petersen RC, Berr C, Owen MJ, Lopez-Arrieta J, Varadarajan BN, Becker JT, Rivadeneira F, Nalls MA, Graff-Radford NR,

Campion D, Auerbach S, Rice K, Hofman A, Jonsson PV, Schmidt H, Lathrop M, Mosley TH, Au R, Psaty BM, Uitterlinden AG, Farrer LA, Lumley T, Ruiz A, Williams J, Amouyel P, Younkin SG, Wolf PA, Launer LJ, Lopez OL, van Duijn CM, Breteler MM. Genome-wide analysis of genetic loci associated with Alzheimer disease. *JAMA.* 2010;303(18):1832-1840.

14. de Lange M, Snieder H, Ariens RA, Andrew T, Grant PJ, Spector TD. The relation between insulin resistance and hemostasis: pleiotropic genes and common environment. *Twin Res.* 2003;6(2):152-161.

15. Declerck PJ, Alessi MC, Verstreken M, Kruithof EK, Juhan-Vague I, Collen D. Measurement of plasminogen activator inhibitor 1 in biologic fluids with a murine monoclonal antibody-based enzyme-linked immunosorbent assay. *Blood.* 1988;71(1):220-225.

16. Macy EM, Meilahn EN, Declerck PJ, Tracy RP. Sample preparation for plasma measurement of plasminogen activator inhibitor-1 antigen in large population studies. *Arch Pathol Lab Med.* 1993;117(1):67-70.

17. Cushman M, Cornell ES, Howard PR, Bovill EG, Tracy RP. Laboratory methods and quality assurance in the Cardiovascular Health Study. *Clin Chem.* 1995;41(2):264-270.

18. Cushman M, Lemaitre RN, Kuller LH, Psaty BM, Macy EM, Sharrett AR, Tracy RP. Fibrinolytic activation markers predict myocardial infarction in the elderly. The Cardiovascular Health Study. *Arterioscler Thromb Vasc Biol.* 1999;19(3):493-498.

19. Nguyen ND, Ghaddar H, Stinson V, Chambless LE, Wu KK. ARIC hemostasis study--IV. Intraindividual variability and reliability of hemostatic factors. The Atherosclerosis Risk in Communities (ARIC). *Thromb Haemost.* 1995;73(2):256-260.

20. Price AL, Patterson NJ, Plenge RM, Weinblatt ME, Shadick NA, Reich D. Principal components analysis corrects for stratification in genome-wide association studies. *Nat Genet.* Vol 38; 2006:904-909.

21. Folkersen L, van't Hooft F, Chernogubova E, Agardh HE, Hansson GK, Hedin U, Liska J, Syvanen AC, Paulssson-Berne G, Franco-Cereceda A, Hamsten A, Gabrielsen A, Eriksson P. Association of genetic risk variants with expression of proximal genes identifies novel susceptibility genes for cardiovascular disease. *Circ Cardiovasc Genet.* 2010;3(4):365-373.

22. Bolstad BM, Irizarry RA, Astrand M, Speed TP. A comparison of normalization methods for high density oligonucleotide array data based on variance and bias. *Bioinformatics.* 2003;19(2):185-193.