Genome-wide association study for circulating levels of plasminogen activator inhibitor-1 (PAI-1) provides novel insights into the regulation of PAI-1

Running title: HUANG et al., GENETIC LOCI ASSOCIATED WITH PAI-1

Jie Huang1; Maria Sabater-Lleal20; Folkert W. Asselbergs26,27,28; David Tregouet12; So-Youn Shin34; Jingzhong Ding6; Jens Baumert15; Tiphaine Oudot-Mellakh12; Lasse Folkersen63; Andrew D. Johnson1; Nicholas L. Smith39,40,41; Scott M. Williams30; Mohammad A. Ikram35,55; Marcus E. Kleber56,59; Diane M. Becker48; Vinh Truong65; Josyf C. Mychaleckyj50; Weihong Tang36; Qiong Yang1,2; Bengt Sennblad2; Jason H. Moore29; Frances M.K. Williams33; Abbas Dehghan53,55; Günther Silbernagel62; Elisabeth M.C. Schrijvers5; Shelly Smith7; Mahir Karakas14; Geoffrey H. Toftel2; Angela Silveira20; Gerjan J. Nalis31; Kurt Lohman8; Ming-Huei Chen2; Annette Peters15; Anuj Goel21,22; MSc; Jemma C. Hopewell68; John C. Chambers69,70; Danish Saleheen71,72; Per Lundmark67; Bruce M. Psaty42,43; Rona J. Strawbridge20; Bernhard O. Boehm60; Angela M. Carter35; Christa Meisinger15; John F. Peden21,22; Joshua C. Bis44; Barbara McKnight45; John Òhrvik20; Kent Taylor46; Maria Grazia Franzosi23; Udo Seedorf24; Rory Collins68; Anders Franco-Cereceda64; Ann-Christine Syväänen67; Alison H. Goodall56; Lisa R. Yaneck48; Mary Cushman47,52; Martina Müller-Nurasyid16,17,18; Aaron R. Folsom36; Saonli Basu37; Nena Mattjévic38; Wieh H. van Gilst32; Jaspal S. Kooper70; Albert Hofman53,55; John Danesh72; Robert Clarke25; James B. Meigs77,78; DIAGRAM Consortium; Sekar Kathiresan73,74,75; Muredach P. Reilly76; CARDIOGRAM Consortium; Norman Klopp19; Tamara B. Harris9; Bernhard R. Winkelmann41; Peter J. Grant35; Hans L. Hillege32; Hugh Watkins21,22; C4D Consortium; Timothy D. Spector33; Lewis C. Becker49; Russell P. Tracy10; Winfried März56,57,58; Andre G. Uitterlinden54,55; Per Eriksson63; Francois Cambien65; CARDIOGENICS Consortium; Pierre-Emmanuel Morange13; Wolfgang Koenig14; Nicole Soranzo33,34; Pin van der Harst32; Yongmei Liu11,51; Christopher J. O’Donnell1,3,4,#; Anders Hamsten20; #

Framingham Heart Study: 1National Heart, Lung, and Blood Institute's (NHBII’ s)
Framingham Heart Study, Framingham, Mass, USA; 2Department of Biostatistics, Boston
University, Boston, Mass, USA; 3Division of Intramural Research, NHLBI, Bethesda, USA;
4Cardiology Division, Department of Medicine, Massachusetts General Hospital, Harvard
Medical School, Boston, Mass, USA; 5Royal North Shore Hospital, University of Sydney,
Australia.

Health ABC: 6Sticht Center on Aging, Wake Forest University School of Medicine;
7Department of Epidemiology & Prevention, Public Health Sciences, Wake Forest University
School of Medicine; 8Department of Biostatistical Sciences, Public Health Sciences, Wake
Forest University School of Medicine; 9Laboratory of Epidemiology, Demography, and
Biometry, National Institute on Aging; 10Department of Pathology and Biochemistry, University
of Vermont College of Medicine, Burlington, VT, USA; 11Department of Epidemiology &
Prevention, Public Health Sciences, Wake Forest University School of Medicine, Winston-
Salem, NC, USA.

Marseille Thrombosis Association Study: 12INSERM UMR_S 937, ICAN Institute, Université
Pierre et Marie Curie (UPMC, Paris 06), Paris, France; 13INSERM UMR_S 626, Université de la Méditerranée, Marseille, France.

MONICA/KORA Study: 14Department of Internal Medicine II –Cardiology, University of Ulm
Medical Center, Ulm, Germany; 15Institute of Epidemiology II, Helmholtz Zentrum München,
German Research Center for Environmental Health, Neuherberg, Germany; 16 Institute of Genetic Epidemiology, Helmholtz Zentrum München, German Research Center for Environmental Health, Neuherberg, Germany; 17 Institute of Medical Informatics, Biometry and Epidemiology, Chair of Epidemiology and Chair of Genetic Epidemiology, Ludwig-Maximilians-Universität München, Munich, Germany; 18 Department of Medicine I, University Hospital Grosshadern, Ludwig-Maximilians-Universität München, Munich, Germany; 19 Division of Molecular Epidemiology, Helmholtz Zentrum München, German Research Center for Environmental Health, Neuherberg, Germany.

**Precocious Coronary Artery Disease Study:** 20 Atherosclerosis Research Unit, Department of Medicine, Karolinska Institutet, Karolinska University Hospital Solna, Stockholm, Sweden; 21 Department of Cardiovascular Medicine, The Wellcome Trust Centre for Human Genetics, University of Oxford, Oxford, United Kingdom; 22 Department of Cardiovascular Medicine, University of Oxford, John Radcliffe Hospital, Headington, Oxford, United Kingdom; 23 Department of Cardiovascular Research, Istituto Mario Negri, Milano, Italy; 24 Gesellschaft für Arterioskleroseforschung V., Leibniz-Institut für Arterioskleroseforschung an der Universität Münster (LIFA), Münster, Germany; 25 Clinical Trial Service Unit, University of Oxford, United Kingdom.

**Prevention of Renal and Vascular End Stage Disease Study:** 26 Department of Cardiology, Division Heart and Lungs, University Medical Center Utrecht, Utrecht, the Netherlands; 27 Julius Center for Health Sciences and Primary Care, University Medical Center Utrecht, the Netherlands; 28 Department of Medical Genetics, University Medical Center, Utrecht, The Netherlands; 29 Department of Genetics and Department of Community and Family Medicine, Dartmouth Medical School, Lebanon, NH, USA; 30 Department of Molecular Physiology and Biophysics, Vanderbilt University, Nashville, TN, USA; 31 Department of Internal Medicine, University Medical Center Groningen, University of Groningen, Groningen, the Netherlands; 32 Department of Cardiology, University Medical Center Groningen, University of Groningen, Groningen, The Netherlands.

**Twins UK:** 33 Department of Twin Research & Genetic Epidemiology, King’s College London, London, United Kingdom; 34 Wellcome Trust Sanger Institute, Hinxton, United Kingdom; 35 Division of Cardiovascular and Diabetes Research, Leeds University, Leeds, UK.

**Atherosclerosis Risk in Communities Study:** 36 Division of Epidemiology and Community Health, University of Minnesota, Minneapolis, Minn, USA; 37 Division of Biostatistics, University of Minnesota, Minneapolis, MN, USA; 38 Hemostasis Laboratory, University of Texas Health Science Center at Houston, Houston, Texas, USA.

**Cardiovascular Health Study:** 39 Department of Epidemiology, University of Washington, Seattle, Washington, USA; 40 Group Health Research Institute, Group Health Cooperative, Seattle, WA, USA; 41 Seattle Epidemiologic Research and Information Center, VA Office of Research and Development, Seattle, WA, USA; 42 Departments of Medicine, Epidemiology, and Health Services, University of Washington, Seattle, Washington, USA; 43 Group Health Research Institute, Group Health Cooperative, Seattle, WA, USA; 44 Department of Medicine, University of Washington, Seattle, Washington, USA; 45 Department of Biostatistics, University of Washington, Seattle, Washington, USA; 46 Medical Genetics Institute, Cedars-Sinai Medical Center, Los Angeles, California, USA; 47 Departments of Medicine and Pathology, University of Vermont, Burlington, Vermont, USA.

**GeneSTAR:** 48 Division of Internal Medicine, Johns Hopkins School of Medicine, Baltimore, MD, USA; 49 Division of Cardiology, Johns Hopkins School of Medicine, Baltimore, MD, USA.
**MESA Study:** Center for Public Health Genomics, University of Virginia, Charlottesville, VA, USA; Department of Epidemiology & Prevention, Public Health Sciences, Wake Forest University School of Medicine, Winston-Salem, NC, USA; Departments of Medicine and Pathology, University of Vermont, Burlington, Vermont, USA.

**Rotterdam Study:** Department of Epidemiology, Erasmus MC Medical Center, Rotterdam, The Netherlands; Department of Internal Medicine, Erasmus MC Medical Center, Rotterdam, The Netherlands; Netherlands Consortium for Healthy Aging, The Netherlands.

The **Ludwigshafen Risk and Cardiovascular Health (LURIC) study:** Mannheim Institute of Public Health, Faculty of Medicine, University of Heidelberg, Mannheim, Germany; Synlab Academy, Mannheim, Germany; Clinical Institute of Medical and Chemical Laboratory Diagnostics, Medical University of Graz, Graz Austria; LURIC Study non profit LLC, Freiburg, Germany; Division of Endocrinology, Diabetes and Metabolism, Graduate School of Molecular Diabetology and Endocrinology, University of Ulm, Ulm, Germany; Cardiology Team Sachsenhausen, Frankfurt am Main, Germany; Division of Endocrinology, Diabetology, Nephrology, Vascular Disease, and Clinical Chemistry, Department of Internal Medicine, Eberhard-Karls-University Tübingen, Tübingen, Germany.

**Advanced Study of Aortic Pathology:** Atherosclerosis Research Unit, Department of Medicine, Karolinska Institutet, Karolinska University Hospital Solna, Stockholm, Sweden; Cardiothoracic Surgery Unit, Department of Molecular Medicine and Surgery, Karolinska Institutet, Stockholm, Sweden.

**Cardiogenics Consortium:** INSERM UMRS 937, ICAN Institute, Université Pierre et Marie Curie (UPMC, Paris 06), Paris, France; Department of Cardiovascular Sciences, University of Leicester, Leicester, United Kingdom; Molecular Medicine and Science for Life Laboratory, Department of Medical Sciences, Uppsala University, Sweden.

**The Coronary Artery Disease Genetics Consortium:** Department of Cardiovascular Medicine, The Wellcome Trust Centre for Human Genetics, University of Oxford, Oxford, United Kingdom; Clinical Trial Service Unit, University of Oxford, United Kingdom; Epidemiology and Biostatistics, Imperial College, London, London, United Kingdom; Cardiology, Ealing Hospital National Health Service (NHS) Trust, Middlesex, United Kingdom; Center for Non-Communicable Diseases Pakistan, Karachi, Pakistan; Department of Public Health and Primary Care, University of Cambridge, Strangeways Research Laboratory, Cambridge, United Kingdom.

**Coronary Artery Disease Genome Wide Replication and Meta-Analysis Consortium:** Cardiovascular Research Center and Cardiology Division, Massachusetts General Hospital, Boston, MA, USA; Center for Human Genetic Research, Massachusetts General Hospital, Boston, MA, USA; Program in Medical and Population Genetics, Broad Institute of Harvard and Massachusetts Institute of Technology (MIT), Cambridge, MA, USA; Cardiovascular Institute, the Perelman School of Medicine at the University of Pennsylvania, Philadelphia, PA, USA.

**Diabetes Genetics Replication And Meta-analysis Consortium:** General Medicine Division, Massachusetts General Hospital, Boston, Massachusetts, USA; Department of Medicine at Harvard Medical School, Boston, Massachusetts 02115, USA.

*These authors contributed equally as first authors.

# These authors contributed equally as last authors.
Corresponding authors:
1. Anders Hamsten MD FRCP, Atherosclerosis Research Unit, Center for Molecular Medicine, Building L8:03, Karolinska University Hospital Solna, S-171 76 Stockholm, Sweden; Fax: +46-8-311298; Phone: +46-8-51773222; E-mail: Anders.Hamsten@ki.se.
2. Christopher J. O'Donnell, MD MPH, National Heart, Lung and Blood Institute’s Framingham Heart Study, 73 Mt. Wayte Street, Suite #2, Framingham, MA 01702, USA; Fax: +1-508-626-1262; Phone: +1-508-935-3435; E-mail: odonnelle@nhlbi.nih.gov.
Abstract

We conducted a genome-wide association study to identify novel associations between genetic variants and circulating plasminogen activator inhibitor-1 (PAI-1) concentration, and examined functional implications of variants and genes that were discovered. A discovery meta-analysis was performed in 19,599 subjects, followed by replication analysis of genome-wide significant ($P<5\times10^{-8}$) single nucleotide polymorphisms (SNPs) in 10,796 independent samples. We further examined associations with type 2 diabetes (T2D) and coronary artery disease (CAD), assessed the functional significance of the SNPs for gene expression in human tissues, and conducted RNA silencing experiments for one novel association. We confirmed the association of the 4G/5G proxy SNP rs2227631 in the promoter region of SERPINE1 (7q22.1) and discovered genome-wide significant associations at 3 additional loci: chromosome 7q22.1 close to SERPINE1 (rs6976053, discovery $P=3.4\times10^{-10}$); chromosome 11p15.2 within ARNTL (rs6486122, discovery $P=3.0\times10^{-8}$); and chromosome 3p25.2 within PPARC (rs11128603, discovery $P=2.9\times10^{-8}$). Replication was achieved for the 7q22.1 and 11p15.2 loci. There was nominal association with T2D and CAD at ARNTL ($P<0.05$). Functional studies identified MUC3 as a candidate gene for the second association signal on 7q22.1. In summary, SNPs in SERPINE1, ARNTL, and a SNP associated with the expression of MUC3 were robustly associated with circulating levels of PAI-1.
**Introduction**

Plasminogen activator inhibitor type 1 (PAI-1) is a serine protease inhibitor protein encoded by the *SERPINE1* gene. It is the principal inhibitor of tissue and urinary plasminogen activators, and therefore constitutes an important regulatory protein in fibrinolysis. PAI-1 is produced by vascular endothelium, liver, monocytes-macrophages, platelets and adipose tissue. High plasma levels of PAI-1 (antigen or activity) are associated with increased risk of atherothrombotic diseases, particularly coronary artery disease (CAD) and myocardial infarction (MI)\(^1\)-\(^4\), and PAI-1 accumulates in human atherosclerotic lesions\(^5\),\(^6\). PAI-1 dependent mechanisms are also implicated in the pathogenesis of obesity, insulin resistance, and type-2 diabetes (T2D)\(^7\)-\(^9\). The consistent association of PAI-1 with obesity and T2D contributes to a prevailing uncertainty regarding the role of PAI-1 as a causal factor in risk for cardiovascular disease.

Population based case-control studies, and family and twin studies have indicated that a major genetic component contributes to the variance of plasma PAI-1 concentration, with an estimated heritability of up to 0.83 in twin studies\(^10\),\(^11\). A number of common polymorphisms have been identified in *SERPINE1*, but only the 4G/5G insertion-deletion variant (rs1799889), located 675 base pairs (bp) from the transcriptional start site, has been consistently associated with the plasma level of PAI-1. The proportion of the variation explained by this *cis* variant is modest (generally around 1-3%), suggesting that yet unknown variation in *SERPINE1* or in genes in other pathways, or epistasis, may play a contributing role as well\(^12\)-\(^14\). Further, some but not all systematic overviews have found that the homozygous 4G/4G genotype is associated with a modest increase in MI risk\(^15\),\(^16\). Two high-coverage single nucleotide polymorphism (SNP) association studies on PAI-1, both with limited sample size, have been reported so far\(^13\),\(^17\), neither of which identified loci showing genome-wide significance.

We hypothesized that there are multiple *cis-* and *trans-*acting loci, in addition to the 4G allele variant, that contribute to the variation in circulating PAI-1 level and tested these hypotheses using data on 19,599 European-ancestry adults from 8 genome-wide association (GWA) studies, each with dense (~2.5 million) imputed SNP genotypes. We then attempted replication of genome-wide significant findings in independent samples from 9 cohorts with a total of 10,796 individuals of European ancestry. SNPs found to be robustly associated with plasma PAI-1 levels in the discovery meta-analysis were examined for association with clinically apparent CAD and T2D in very large clinical GWA study meta-analyses\(^18\),\(^19\). Finally, we performed expression analyses in human target tissues and silencing experiments in a relevant cell culture system to confirm the functional significance of identified candidate genes.

**Methods**

**Discovery and Replication Cohorts**

Separate GWA studies of plasma PAI-1 antigen concentration were conducted in 8 independent cohorts of European-ancestry individuals. These cohorts included the Framingham Heart Study (FHS, n=6,634), the Precocious Coronary Artery Disease Study (PROCARDIS, n=1,922 cases and 2,017 controls), Twins UK (n=1,294), MONICA/KORA (n=1,565), HealthABC (n=1,645), the Marseille Thrombosis Association study (MARTHA, n=851) and the Prevention of Renal and Vascular End Stage Disease study (PREVEND, n=3,671). Detailed descriptions of the
discovery cohorts and exclusion criteria adopted for the present study are given in **Online Data Supplemental Section S1**. Sample sizes, mean age and sex distribution of discovery study participants in each cohort at the time of the PAI-1 determination are summarized in **Supplemental Table S1**. Following meta-analysis on a total of 19,599 individuals, genome-wide significant loci ($P<5.0\times10^{-8}$) underwent replication in 10,796 independent samples from 9 cohorts of European ancestry (**Online Data Supplemental Section S2**), characteristics of which are provided in **Supplemental Table S2**. PAI-1 antigen levels (ng/mL) or PAI-1 activity (U/mL) were measured in plasma (EDTA or citrate) by ELISA and functional methods (**Supplemental Table S1** and **Online Data Supplemental Section S3**).

All participating cohorts were granted approval by the appropriate research ethics committees for the research, and all participants provided written informed consent for the use of their DNA.

**Genotyping and Imputation**
A description of the genotyping technologies used for the discovery cohorts is provided in **Supplemental Table S3** along with the quality control (QC) criteria for filtering and imputation methods. Briefly, approximately 2.5M autosomal SNPs were imputed for every cohort using the HapMap II Caucasian (CEU) sample from the Centre d’Etude du Polymorphisme Humain as reference panel$^{20, 21}$. Prior to imputation, every cohort applied SNP exclusions by call rate ($\leq 0.93$ to $0.99$ depending on the cohort), minor allele frequency (MAF) < 0.01 and deviation from Hardy-Weinberg equilibrium ($P<10^{-5}$ to $10^{-6}$). Details regarding genotyping, quality control and imputation characteristics of the replication cohorts are provided in **Supplemental Table S4**.

**Statistical Analysis of the Discovery Cohorts**
PAI-1 values were natural-logarithm transformed due to skewness of the distribution. Genotype-phenotype association analyses were performed independently in each cohort according to a pre-specified analysis plan. The association of measured and imputed SNPs with natural logarithm transformed PAI-1 values was assessed by a linear model assuming additive genetic effects. Standard linear regression analysis was used in non-familial studies. All analyses were adjusted for age and sex, and additionally on principal components or multi-dimensional scaling where appropriate. In PROCARDIS, analyses were further adjusted for antidiabetic medication in T2D. For family studies, FHS and PROCARDIS, a linear mixed-effects model was used to additionally account for relatedness$^{22}$. A genomic control coefficient was computed for each discovery cohort, estimating the GWA inflation coefficients, and applied subsequently to individual SNP association test statistics to correct for cryptic relatedness. Specialized analysis tools and software packages are listed in **Online Data Supplemental Section S4**.

Meta-analysis of all discovery GWA studies, using a total of 2,445,683 SNPs, was conducted using an inverse-variance weighted fixed-effects method as implemented in METAL (http://www.sph.umich.edu/csg/abecasis/Metal/index.html), and summary $P$ values and beta coefficients were calculated. The beta value represents the per-allele effect on the natural log-transformed PAI-1 level. To ensure accuracy and reliability of results, the meta-analysis was conducted independently by two investigators (JH, MSL) at two separate sites. The prespecified threshold of genome-wide significance was set at the conventional level of $P=5.0\times10^{-8}$.

**Replication**
Replication analysis was conducted for loci that contained at least one SNP passing the pre-specified genome-wide significance threshold in the discovery meta-analysis. A total of 10 SNPs
were selected for replication from the 4 loci (Table 1): the 4 lead SNPs (i.e. the SNP with the lowest $P$ value for each of the 4 genome-wide significant regions, also referred to as the index SNP), 2 non-synonymous SNPs within 2 of these regions showing suggestive evidence of association ($p < 5 \times 10^{-7}$) and 4 additional SNPs selected to tag the main haplotypes in 2 of the candidate regions. Procedures for selecting tag SNPs are described in Online Data Supplemental Section S4.

The replication sample size was estimated to have 80% power to detect an explained variance of 0.5% at a MAF of 20.0%. The same statistical methods were used as for discovery. The significance threshold was set to $P=0.005$ after Bonferroni correction for testing of 10 SNPs.

Conditional and Secondary Analysis

In order to confirm that 2 associated loci on the same chromosomal region (7q22.1) were independent, we conditioned on the sentinel SNP with the lowest $P$ value (rs2227631) and repeated the association analysis between genotype and PAI-1 levels using the same model as was used in the discovery analysis. Given the adipose production of PAI-1 and the widely reported association of body mass index (BMI) with plasma PAI-1 levels, the genotype-phenotype association analyses were further adjusted for BMI.

Association with Clinically Apparent Coronary Artery Disease and Type-2 Diabetes Mellitus

We explored associations of SNPs selected from the discovery meta-analysis with CAD and T2D to determine whether SNPs showing robust associations with PAI-1 levels may also be implicated in the pathogenesis of CAD or T2D, respectively. Genotype-phenotype association results for the selected SNPs were obtained from the Coronary Artery Disease Genome-wide Replication And Meta-analysis (CARDIoGRAM) and the Coronary Artery Disease (C4D) Genetics consortia. In all, CARDIoGRAM includes 22,233 cases of CAD and 64,762 controls, all of European ancestry, whereas C4D comprises 6,424 CAD cases and 7,268 controls of European ancestry. For T2D, genotype-phenotype association results for the selected SNPs were obtained from the GWAS meta-analysis of 8,130 individuals with T2D and 38,987 controls in the DIAbetes Genetics Replication And Meta-analysis (DIAGRAM) consortium. SNP associations with CAD and T2D were assessed by logistic regression analysis under an additive model, adjusting for age and sex. For the PROCARDIS component of C4D, a mixed-effects model was used to additionally account for relatedness. The significance threshold for disease association was set to $P<0.005$ after Bonferroni correction for testing of 10 SNPs.

Association with Gene Expression in Human Tissues

*Ex vivo* studies of global gene expression data from monocytes and macrophages were derived from the Cardiogenics study (http://www.cardiogenics.eu) and based on peripheral blood samples from both CAD patients and healthy controls. Global gene expression data from liver, mammary artery intima-media, aortic intima-media and aortic adventitia were obtained from the Advanced Study of Aortic Pathology (ASAP) study of patients undergoing aortic replacement and/or valve repair surgery. *In silico* lookup was also conducted on an eQTL transcriptome database of circulating monocytes (monocytes II) obtained from 1,490 unrelated individuals. Details of the biobanks and the methods for gene expression analysis, genotyping and statistical analysis are provided in Online Data Supplemental Section S5.
Functional Studies
Influence of MUC3 and TRIP6 on PAI-1 gene expression and protein secretion was studied in cultured HuH7 liver cells and the monocyte THP-1 cell line, respectively, using siRNA gene knock-down technology. Details on methods, cell culture and verification experiments are provided in Online Data Supplemental Section S6.

Results

Discovery Meta-Analysis
The results for the 2,445,683 meta-analyzed SNPs in the discovery cohorts are shown in Figure 1, organized by chromosome in a Manhattan plot. A quantile-quantile (QQ) plot is also shown in Supplemental Figure S1, with a lambda value of 1.012. A total of 24 SNPs passed the genome-wide significance threshold of \( P=5.0 \times 10^{-8} \). All genome-wide significant SNPs were located in three chromosomal regions 3p25.2, 7q22.1 (two loci), and 11p15.2. The \( P \) values for the lead (index) SNPs in these regions are listed in Table 1, and the regional plots are shown in Supplemental Figure S2. The MAFs of the top-SNPs ranged from 0.10 to 0.48.

The lead SNP in the region on chromosome 3 (rs11128603, discovery \( P=2.9 \times 10^{-8}, \beta =0.086 \)) is intronic to PPARG encoding peroxisome proliferator-activated receptor gamma (Supplemental Figure S2A).

The lead SNP on chromosome 7, rs2227631 (discovery \( P=7.8 \times 10^{-15}, \beta =0.076 \)), is a proxy for the previously and consistently reported 4G/5G polymorphism in SERPINE1, the structural gene for PAI-1. A total of 19 SNPs located in a wide region of 266.6 kilobases (kb) between position 100,289,668 bp and 100,556,258 bp were found to surpass the genome-wide significance threshold. Analysis of the LD structure in this region showed that there are two independent significant signals on 7q22.1, one exclusively represented by SNP rs2227631, located in the promoter region of SERPINE1 gene, and another comprising a large LD block spanning 60,387 bp located 206,203 bp upstream of SNP rs2227631 \( (r^2<0.10) \) (Supplemental Figures S2B and S2C, Supplemental Figure S3). The lead SNP in the second region on chromosome 7 is rs6976053 \( (P=3.4 \times 10^{-10}, \beta =0.054) \). The second region on chromosome 7 harbors several genes that contain SNPs with \( P \) values surpassing genome-wide significance: SLC12A9, TRIP6, SRRT, UFSP1 and ACHE (Supplemental Figure S2C). All significant SNPs in this region are in strong LD with each other, and are independent of rs2227631. This finding was confirmed by secondary analyses showing that the second region on chromosome 7 remained significant after conditional analysis on rs2227631, as shown in Supplemental Table S5. Amongst a total of 120 SNPs analysed in the 7q22.1 region, none had a significant interaction with rs2227631 (interaction \( P >0.05 \)); hence, we did not find support for a potential effect of rare variants that might drive the two observed association signals.

Finally, the strongest association on chromosome 11 was found for rs6486122, \( P=3.0 \times 10^{-8}, \beta =0.051 \). Four SNPs in this region showed genome-wide significant association with PAI-1 levels. These SNPs (rs6486122, rs10832027, rs2896635 and rs7947951) are in strong LD with each other and with the other SNPs exhibiting the lowest \( P \) values in this region (Supplemental Figure S2D), and they are all intronic variants located in the aryl hydrocarbon receptor nuclear translocator-like (ARNTL) gene.
Association results for these 10 SNPs in each individual GWA study are detailed in Supplemental Table S6. Secondary analyses with adjustment for BMI conducted in the discovery cohorts on the 10 SNPs selected for examination in the replication cohorts (Supplemental Table S5) resulted in attenuation of the statistical significance for most of the genotype-phenotype associations encountered in all four PAI-1 loci. It is notable that the association with the 4G/5G proxy SNP rs2227631 was considerably strengthened by adjustment for BMI (from $P=7.8 \times 10^{-15}$ to $P=3.0 \times 10^{-22}$). The total variation in PAI-1 levels in the discovery cohorts accounted for by the 10 SNPs, beyond that contributed by age and sex, ranged between 0.9% and 3.7%.

**Replication**

For all 10 SNPs followed up, the direction of the genetic effects was identical in the discovery and replication cohorts. The replication $P$ value was less than the prespecified threshold of independent replication, $P<0.005$, for seven SNPs: rs2227631 (in SERPINE1), all 5 SNPs contained in the second locus on chromosome 7 harboring SLC12A9, TRIP6, SRRT, UFSP1 and ACHE, and rs6486122 (in the ARNTL region). The combined discovery and replication $P$ value was less than $5.0 \times 10^{-8}$ for all SNPs except for rs11128603 and rs1801282 in the PPARG region (Table 1).

**Association with Coronary Artery Disease and Type-2 Diabetes**

For CAD association, we meta-analyzed the two European-ancestry cohorts from C4D together with CARDIoGRAM to calculate a combined $P$ value and effect size (Table 2). Two SNPs in ARNTL were nominally associated with CAD; the plasma PAI-1 raising alleles increased the risk of CAD (odds ratio (OR) and 95% confidence interval (95% CI)) per allele: 1.04 (1.01-1.07); 1.03 (1.01-1.06)). These two SNPs also showed significant association with T2D in an *in-silico* examination of the DIAGRAM results (rs6486122: 1.06 (1.01-1.10); rs3816360: 1.06 (1.02-1.11)). The two SNPs in PPARG were also significantly associated with T2D, as has been previously reported, but not with CAD. For both of the T2D associated loci, the alleles contributing to higher plasma PAI-1 levels increased the risk of T2D.

**Association with Expression in Target Tissues**

Global gene expression data from relevant tissues were used to explore the functional significance of candidate genes contained in the associated genomic regions. For each of the 10 identified PAI-1 associated SNPs, we investigated associations between genotype and expression levels of all genes located within ±200 kb of the SNP. A total of 87 genes were contained within these regions. Table 3 reports all significant associations at $P<0.00408$ (i.e. FDR = 0.05). In the candidate chromosome 3 region, the selected tag SNP rs1801282 showed a highly significant association with PPARG expression levels in macrophages ($P=1.7 \times 10^{-50}$). In addition, monocyte expression of 3 genes (TRIP6, UFSP and EPHB4) located in the chromosome 7 region were found to be significantly associated in 2 independent datasets with the SNPs selected from the meta-analysis of the GWA studies. Similar significant associations were found in macrophages for UFSP. These associations were however not observed in the other studied cell/tissue types. Conversely, in liver, SNPs located on chromosome 7 showed significant associations with expression levels of MUC3 ($P=1.3 \times 10^{-17}$) and acetylcholinesterase (ACHE, $P=0.00028$). In the chromosome 11 candidate region, the selected tag SNP rs10832027 showed a significant association with the ARNTL expression levels in monocytes ($P=4.1 \times 10^{-5}$).
Effects of Silencing MUC3 and TRIP6 on Gene Expression and PAI-1 Secretion

When prioritizing candidate genes for functional studies, we focused on genes with tissue expression levels in the top quartile of the overall gene expression distribution that showed strong allele-specific associations with the lead-SNP genotype (FDR<0.05, P<0.00408). Of the candidate genes contained in the chromosome 7 locus, only MUC3A expression in liver and TRIP6 expression in monocytes met these criteria. We therefore investigated MUC3 in liver cells and TRIP6 in monocytes using siRNA knock-down in vitro experiments. In 15 independent experiments (with 3 replicates in each), the mean silencing of MUC3 was -61.0% (Figure 2A), 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 (Figure 2B). Verification experiments using two additional MUC3 silencers showed a consistent significant increase in SERPINE1 RNA levels (Online Supplemental Section S6). 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). In contrast, despite effective silencing of expression of TRIP6 in monocytes, we observed no significant changes in SERPINE1 expression or PAI-1 production (details are provided in Online Supplemental Section S6).

Discussion

This is the first meta-analysis of GWA studies to examine genetic determinants of plasma PAI-1. Overall, in our discovery meta-analysis we identified four loci showing genome-wide significant associations with plasma PAI-1 antigen levels, located in three regions on chromosomes 3p25.2, 7q22.1 (two loci) and 11p15.2. Among the four loci, three contained obvious or plausible biological candidate genes for the regulation of PAI-1 (SERPINE1, PPARG and ARNTL), whereas one (harbouring SLC12A9, TRIP6, SRRT, UFSP1 and ACHE) represents a greater mechanistic challenge. Independent replication was obtained for the first chromosome 7 locus encompassing SERPINE1, the second locus on chromosome 7 containing SLC12A9, TRIP6, SRRT, UFSP1 and ACHE, and the third locus on chromosome 11 containing ARNTL. The 3p25.2 locus, however, was not replicated (P=0.1), which could be due to lack of power. Based on the observed MAF of 10% and effect size of 0.031 for the lead SNP at 3p25.2 in the replication samples, our replication sample size of 10,796 individuals only had 35.2% power to replicate the PPARG finding made in the discovery cohorts. Evidence of genotype associations with risk of CAD and T2D was demonstrated for ARNTL by in silico examinations of databases generated by the CARDIoGRAM, C4D and DIAGRAM consortia. Based on association studies between genotype and gene expression levels in relevant tissues and gene silencing in liver cells, we identified MUC3 as a strong candidate gene for the second, independent association signal on chromosome 7.

Genes at Chromosome 7q22.1

The strongest SNP association with plasma PAI-1 concentration was contributed by rs2227631, which is located in the structural gene for PAI-1 (SERPINE1). This SNP is in strong LD (D²=0.97, r²=0.78)12 with the widely reported 4G/5G polymorphism in the SERPINE1 promoter.
The strong LD between the 2 SNPs renders it exceedingly difficult to distinguish which of the 2 SNPs is functionally responsible for the association with plasma PAI-1. However, functional studies of the 4G/5G polymorphism suggest that the 5G allele contains a binding site for a repressor protein, which decreases PAI-1 transcription.

Interestingly, we found another association signal on chromosome 7, located 206.2 kb upstream of SNP rs2227631, which had not been previously reported to associate with PAI-1. In spite of the proximity of this locus to SERPINE1, analysis of the LD structure in this region and secondary analyses conditional on rs2227631 confirmed that this locus is independent of known variation in SERPINE1. Among the genes contained in the associated region are SLC12A9, TRIP6, SRRT, UFSP1 and ACHE, none of which have been linked to PAI-1 regulation. To refine the association signal, we examined the relationship between PAI-1 associated SNPs and expression levels of genes in this region and nearby, using human gene expression data generated in relevant cells and tissues. Significant associations were found for expression levels of TRIP6, EPHB4 and UFSP1 in monocytes, for UFSP1 in macrophages, and for ACHE and MUC3 in liver. According to our ex vivo target tissue gene expression analyses, MUC3 expression in the liver is strongly associated with the top PAI-1 associated SNPs in the second chromosome 7 locus (P=1.3x10^{-17}), constituting the strongest reported association in the liver. This finding suggests that MUC3 could have an important role in the regulation of PAI-1 synthesis in the liver. Moreover, our in vitro analyses in HuH7 liver cells, showing that silencing of MUC3 increases SERPINE1 expression both at the RNA and protein levels, are consistent with this hypothesis.

MUC3 belongs to a gene family encoding mucins. Mucins are large, heavily glycosylated glycoproteins that exert important functions in the protection of the epithelium. They are also thought to be implicated in signal transduction, regulation of cell growth, cell-cell adhesion, and modulation of the immune system, and MUC3 has been implicated in common complex diseases such as rheumatoid arthritis and inflammatory bowel diseases. The mechanisms by which MUC3 is implicated in PAI-1 regulation remain to be elucidated. It is notable in this context that there are two transcripts from MUC3. Due to the high level of homology between MUC3A and MUC3B (the unique exonic sequences range from 94 to 100% identity at the nucleotide level and the introns show around 95% identity), the probes used to detect and silence MUC3 expression in this study would recognize either MUC3A or MUC3B. For this reason, we chose to use the generic name for the gene (MUC3). However, it is worth mentioning that the tissue expression patterns of MUC3A and MUC3B are different. While expression of MUC3B has detected exclusively in the small and large intestines, that of MUC3A has been observed in the intestines, in heart, liver, thymus, prostate and pancreas. Thus, it is plausible to conclude that only MUC3A was detected and silenced in liver. It should also be emphasized that we cannot exclude that other genes contained in the second region on chromosome 7 are involved in the regulation of PAI-1, e.g. the ones showing allele-specific expression in monocytes/macrophages. However, we observed no significant changes in SERPINE1 expression when silencing TRIP6, the most likely candidate gene expressed in monocytes/macrophages.

**Chromosome 11p15.2 Locus Containing ARNTL**

The locus on chromosome 11 harbors the ARNTL gene, which encodes an important activator protein for circadian rhythm-associated genes. This is of major interest since plasma PAI-1 levels fluctuate in a circadian manner resulting in reduced fibrinolytic activity during the early morning, with the peak in plasma PAI-1 concentration preceding an increased incidence of
Furthermore, circadian variation in SERPINE1 expression in vitro is regulated by heterodimers formed from the circadian clock proteins CLOCK:ARNTL1 and CLOCK:ARNTL2 that bind to the SERPINE1 promoter and upregulate its expression. We demonstrated highly significant associations of SNPs located in the ARNTL locus with both plasma PAI-1 concentrations and ARNTL mRNA levels in monocytes (P= 4.1x10^{-5}). These findings confirm previous results obtained in vitro and show for the first time that ARNTL is implicated in PAI-1 regulation in vivo in humans. Furthermore, ARNTL proved to influence CAD risk as well as proneness to T2D. The CAD finding is in agreement with the previous observation that circadian variation in plasma PAI-1 levels is linked to variation in the incidence of MI and indicates that PAI-1 and/or pathways involving ARNTL may be causally related to the risk of CAD in some individuals.

Chromosome 3p25.2 Locus Containing PPARG

Although not significant in the replication studies and thus not demonstrated as being robustly associated with plasma PAI-1 level, the locus on chromosome 3 identified at a genome-wide significant level at the discovery stage is of considerable interest. It contains several SNPs located in the PPARG gene encoding peroxisome proliferator-activated receptor gamma (PPARγ). PPARs are ligand-activated transcription factors that form functional heterodimers with retinoid X receptors and bind to specific peroxisome proliferator response elements in the promoter regions of their target genes. PPARγ is a regulator of adipocyte differentiation that has been implicated in the pathogenesis of obesity, T2D, atherosclerosis and cancer. It is also important for inflammatory responses including macrophage function and regulates PAI-1 expression in endothelial cells, adipose tissue and the liver. Interestingly, secondary analysis with adjustment for BMI revealed that variation in BMI had only minor effects on the genotype-phenotype associations at the PPARG locus. In addition, expression analyses showed that the SNPs in the PPARG locus most strongly associated with plasma PAI-1 are also distinctly associated with expression levels of the PPARG transcript in macrophages (P =1.70x10^{-50}). Overall, our results support that further studies be conducted of PPARG as a potentially important regulator of PAI-1 expression independent of the degree of adiposity.

Relevance for Cardiovascular Disease

Besides its important inhibitory role in fibrinolysis, PAI-1 is involved in the pathogenesis of obesity, insulin resistance and T2D, and appears to exert a wide range of pleiotropic functions in vascular remodelling, wound healing, fibrosis, tumour angiogenesis, bone remodelling, asthma, rheumatoid arthritis, glomerulonephritis and sepsis. This wide array of functions has contributed to the prevailing uncertainty as to whether PAI-1 is a causal factor in relation to risk for cardiovascular disease. It is notable in this context that SERPINE1, the gene encoding PAI-1, was not associated with risk of CAD despite showing the strongest SNP association with circulating PAI-1 levels, a finding speaking against a causal involvement of PAI-1 in CAD. Against this background, the CAD association demonstrated for ARNTL may be of major significance and warrants further investigation to elucidate whether pathways involving ARNTL, rather than the PAI-1 molecule itself, are causally related to the risk of CAD.

Strengths and Limitations

The strengths of this study include the large size (n=19,599) of the sample used for the discovery meta-analysis, access to gene expression data in target tissues from humans subjects and
inclusion of functional studies in cell culture systems as a complementary approach to gene identification. Conversely, it should be emphasized that the GWA study strategy based on the SNP arrays used here primarily identifies marginal to small effects of common genetic variation (MAF>0.10). Low-frequency and rare variants that may influence plasma PAI-1 levels are likely to have escaped detection in the present study and hence contribute to the missing heritability, and additional common variants with real effects on plasma PAI-1 levels but smaller effect sizes did not attain the conservative threshold for genome-wide significance. Furthermore, gene-gene and gene-environment interactions, which were not examined, may play an important role in the regulation of PAI-1. It should also be emphasized that the results were generated in European ancestry populations and that additional studies are needed in other ethno-geographic groups. Two of our discovery cohorts (KORA and MARTHA) measured PAI-1 activity while the other cohorts measured PAI-1 antigen. It is notable in this context that these 2 cohorts comprised only about 10% of the total discovery samples, and the effect sizes for all 10 SNPs were consistent and comparable across all 8 discovery cohorts (heterogeneity $P>0.05$). Finally, very few clinical or population-based studies with adequate numbers of prevalent or incident CVD/T2D events have included measurements of plasma PAI-1 levels. Accordingly, the power of a conventional Mendelian randomisation study of PAI-1 would unfortunately be low. The best substitute available is the look-up of PAI-1-associated SNPs in adequately powered genetic case control studies, such as C4D and Cardiogram, which was conducted in our study.

Conclusions
Using genome-wide, high-coverage genotype data from a total of 19,599 subjects of European ancestry, we identified 4 loci showing genome-wide significant associations with plasma PAI-1 antigen levels and successfully replicated findings for 3 of the loci. Among the 4, 3 loci contained obvious or plausible biological candidate genes (SERPINE1, PPARG and ARNTL), whereas one, for which MUC3 was identified as a strong candidate gene in functional studies, represents an entirely new mechanism. Although only a fairly small sample was available for formal replication of the discovery GWA meta-analysis findings, independent replication was obtained for SERPINE1, ARNTL and the locus on chromosome 7 for which MUC3 was identified as a candidate gene. Evidence of genotype associations with risk of CAD was demonstrated for ARNTL, but not for SERPINE1 and MUC3, indicating that pathways involving ARNTL, rather than the PAI-1 molecule itself, may be causally related to the risk of CAD.


Acknowledgments and Funding Sources
See Online Data Supplemental Sections S7 and S8.
Authorship

Contribution: A.H and C.O.J designed the study and organized the project. J.H. and M.S-L collected the data and ran the meta-analysis. M.S-L conducted the in vitro cell study. J.H., M.S-L., F.W.A., D.T., A.D.J., N.L.S., S.M.W., W.T., P.E.M., P.v.d.H., Y.L, C.J.O., A.H. wrote the manuscript; and all authors involved in individual GWAS analysis, reviewed and approved the final manuscript.

Conflict-of-interest disclosure: None of the authors for this manuscript has disclosed any conflict of interest related to the manuscript.

References


<table>
<thead>
<tr>
<th>Region</th>
<th>SNP</th>
<th>Chr</th>
<th>Closest gene</th>
<th>Type</th>
<th>Alleles</th>
<th>MAF</th>
<th>Discovery</th>
<th>Replication</th>
<th>Combined</th>
</tr>
</thead>
<tbody>
<tr>
<td>3p25.2</td>
<td>rs11128603</td>
<td>3</td>
<td>PPARG</td>
<td>Index SNP</td>
<td>A -&gt; G</td>
<td>0.102</td>
<td>0.086</td>
<td>2.9x10^{-8}</td>
<td>0.031</td>
</tr>
<tr>
<td></td>
<td>rs1801282</td>
<td>3</td>
<td>PPARG</td>
<td>nsSNP</td>
<td>C -&gt; G</td>
<td>0.127</td>
<td>0.069</td>
<td>1.4x10^{-7}</td>
<td>0.026</td>
</tr>
<tr>
<td>7q22.1</td>
<td>rs2227631</td>
<td>7</td>
<td>SERPINE1</td>
<td>Index SNP</td>
<td>A -&gt; G</td>
<td>0.409</td>
<td>0.076</td>
<td>7.8x10^{-15}</td>
<td>0.069</td>
</tr>
<tr>
<td>7q22.1</td>
<td>rs6976053</td>
<td>7</td>
<td>ACHE</td>
<td>Index SNP</td>
<td>T -&gt; C</td>
<td>0.478</td>
<td>0.054</td>
<td>3.4x10^{-10}</td>
<td>0.039</td>
</tr>
<tr>
<td></td>
<td>rs2075756</td>
<td>7</td>
<td>TRIP6</td>
<td>nsSNP</td>
<td>A -&gt; G</td>
<td>0.282</td>
<td>0.058</td>
<td>2.7x10^{-9}</td>
<td>0.043</td>
</tr>
<tr>
<td></td>
<td>rs314376</td>
<td>7</td>
<td>SLC12A9</td>
<td>Tag SNP</td>
<td>G -&gt; A</td>
<td>0.483</td>
<td>0.052</td>
<td>2.4x10^{-9}</td>
<td>0.041</td>
</tr>
<tr>
<td></td>
<td>rs12672665</td>
<td>7</td>
<td>SRRT</td>
<td>Tag SNP</td>
<td>G -&gt; A</td>
<td>0.479</td>
<td>0.047</td>
<td>6.1x10^{-8}</td>
<td>0.033</td>
</tr>
<tr>
<td></td>
<td>rs3847067</td>
<td>7</td>
<td>ACHE</td>
<td>Tag SNP</td>
<td>A -&gt; C</td>
<td>0.289</td>
<td>0.059</td>
<td>4.1x10^{-10}</td>
<td>0.038</td>
</tr>
<tr>
<td>11p15.2</td>
<td>rs6486122</td>
<td>11</td>
<td>ARNTL</td>
<td>index SNP</td>
<td>T -&gt; C</td>
<td>0.311</td>
<td>0.051</td>
<td>3.0x10^{-8}</td>
<td>0.038</td>
</tr>
<tr>
<td></td>
<td>rs3816360</td>
<td>11</td>
<td>ARNTL</td>
<td>Tag SNP</td>
<td>C -&gt; T</td>
<td>0.333</td>
<td>0.048</td>
<td>9.9x10^{-8}</td>
<td>0.031</td>
</tr>
<tr>
<td>SNP</td>
<td>Chr</td>
<td>Closest Gene</td>
<td>Type</td>
<td>Alleles</td>
<td>CAD OR (95% CI)</td>
<td>CAD P-value</td>
<td>CAD OR (95% CI)</td>
<td>P-value</td>
<td></td>
</tr>
<tr>
<td>-----------</td>
<td>-----</td>
<td>--------------</td>
<td>----------</td>
<td>---------</td>
<td>----------------</td>
<td>-------------</td>
<td>----------------</td>
<td>---------</td>
<td></td>
</tr>
<tr>
<td>rs11128603</td>
<td>3</td>
<td>PPARG</td>
<td>Index SNP</td>
<td>A -&gt; G</td>
<td>0.99 (0.95-1.03)</td>
<td>0.7</td>
<td>1.16 (1.09-1.24)</td>
<td>1.3x10^-5</td>
<td></td>
</tr>
<tr>
<td>rs1801282</td>
<td>3</td>
<td>PPARG</td>
<td>nsSNP</td>
<td>C -&gt; G</td>
<td>1.00 (0.96-1.04)</td>
<td>1.0</td>
<td>1.15 (1.08-1.22)</td>
<td>8.0x10^-6</td>
<td></td>
</tr>
<tr>
<td>rs2227631</td>
<td>7</td>
<td>SERPINE1</td>
<td>Index SNP</td>
<td>A -&gt; G</td>
<td>1.00 (0.97-1.03)</td>
<td>1.0</td>
<td>1.02 (0.98-1.06)</td>
<td>0.4</td>
<td></td>
</tr>
<tr>
<td>rs6976053</td>
<td>7</td>
<td>ACHE</td>
<td>Index SNP</td>
<td>T -&gt; C</td>
<td>1.00 (0.98-1.03)</td>
<td>0.7</td>
<td>0.99 (0.95-1.03)</td>
<td>0.6</td>
<td></td>
</tr>
<tr>
<td>rs2075756</td>
<td>7</td>
<td>TRIP6</td>
<td>nsSNP</td>
<td>A -&gt; G</td>
<td>1.00 (0.97-1.03)</td>
<td>0.8</td>
<td>0.98 (0.94-1.03)</td>
<td>0.5</td>
<td></td>
</tr>
<tr>
<td>rs314376</td>
<td>7</td>
<td>SLC12A9</td>
<td>Tag SNP</td>
<td>G -&gt; A</td>
<td>1.01 (0.98-1.03)</td>
<td>0.6</td>
<td>0.99 (0.95-1.03)</td>
<td>0.7</td>
<td></td>
</tr>
<tr>
<td>rs12672665</td>
<td>7</td>
<td>SRTT</td>
<td>Tag SNP</td>
<td>G -&gt; A</td>
<td>1.00 (0.97-1.02)</td>
<td>0.7</td>
<td>0.99 (0.95-1.03)</td>
<td>0.6</td>
<td></td>
</tr>
<tr>
<td>rs3847067</td>
<td>7</td>
<td>ACHE</td>
<td>Tag SNP</td>
<td>A -&gt; C</td>
<td>0.99 (0.96-1.02)</td>
<td>0.5</td>
<td>0.98 (0.94-1.03)</td>
<td>0.5</td>
<td></td>
</tr>
<tr>
<td>rs6486122</td>
<td>11</td>
<td>ARNTL</td>
<td>index SNP</td>
<td>T -&gt; C</td>
<td>1.04 (1.01-1.07)</td>
<td>5.2x10^-3</td>
<td>1.06 (1.01-1.10)</td>
<td>1.0x10^-2</td>
<td></td>
</tr>
<tr>
<td>rs3816360</td>
<td>11</td>
<td>ARNTL</td>
<td>Tag SNP</td>
<td>C -&gt; T</td>
<td>1.03 (1.01-1.06)</td>
<td>3.0x10^-2</td>
<td>1.06 (1.02-1.11)</td>
<td>3.9x10^-3</td>
<td></td>
</tr>
</tbody>
</table>

*Individual SNP results from the CARDIoGRAM and C4D consortia.
†Individual SNP results from the DIAGRAM Consortium.
CAD, coronary artery disease; T2D, type-2 diabetes.
Table 3. Association between PAI-1 associated SNPs and proximal gene expression in human cells and tissues.

<table>
<thead>
<tr>
<th>Chr</th>
<th>SNP (Alleles; β; P)</th>
<th>Gene Expression</th>
<th>Monocyte (n=849)</th>
<th>Monocyte II (n=1490)</th>
<th>Macrophage (n=614)</th>
<th>Mammary artery intima-media (n=88)</th>
<th>Liver (n=152)</th>
<th>Aortic intima-media (n=131)</th>
<th>Aortic adventitia (n=113)</th>
</tr>
</thead>
<tbody>
<tr>
<td>3</td>
<td>rs1801282 (C-&gt;G; 0.069; 1.4x10^{-5})</td>
<td>PPARG</td>
<td>0.061.7x10^{-2}</td>
<td>NS</td>
<td>0.38;1.7x10^{-50}</td>
<td>NS</td>
<td>1.11; 2.0x10^{-7}</td>
<td>NS</td>
<td>NS</td>
</tr>
<tr>
<td>7</td>
<td>rs314376 (G-&gt;A; 0.052; 2.4x10^{-9})</td>
<td>TRIP6</td>
<td>-0.08;4.7x10^{-15}</td>
<td>NA;8.3x10^{-15}</td>
<td>-0.04;6.2x10^{-4}</td>
<td>NS</td>
<td>NS</td>
<td>NS</td>
<td>NS</td>
</tr>
<tr>
<td></td>
<td>UFSP1</td>
<td></td>
<td>-0.04;3.5x10^{-9}</td>
<td>NA;1.8x10^{-12}</td>
<td>-0.03;6.7x10^{-7}</td>
<td>NS</td>
<td>1.04; 4.0x10^{-2}</td>
<td>NS</td>
<td>NS</td>
</tr>
<tr>
<td></td>
<td>EPHB4</td>
<td></td>
<td>-0.14;2.9x10^{-24}</td>
<td>NA;3.4x10^{-25}</td>
<td>NS</td>
<td>NS</td>
<td>NS</td>
<td>NS</td>
<td>1.08; 5.0x10^{-2}</td>
</tr>
<tr>
<td></td>
<td>ACHE</td>
<td></td>
<td>NS</td>
<td>NA</td>
<td>NS</td>
<td>NS</td>
<td>1.06; 2.8x10^{-4}</td>
<td>NS</td>
<td>NS</td>
</tr>
<tr>
<td></td>
<td>MUC3A</td>
<td></td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NS</td>
<td>1.16; 1.3x10^{-17}</td>
<td>NS</td>
<td>NS</td>
</tr>
<tr>
<td>11</td>
<td>rs10832027 (A-&gt;G; 0.051; 3.1x10^{-8})</td>
<td>ARNTL</td>
<td>-0.07;4.1x10^{-3}</td>
<td>NA</td>
<td>-0.04;9.2x10^{-3}</td>
<td>NS</td>
<td>NS</td>
<td>NS</td>
<td>NS</td>
</tr>
</tbody>
</table>

Associations between 10 SNPs (see table 2) and proximal genes (<200 kb) were tested. For each cell and tissue type, β and P-values for association are shown calculated according to a linear additive model. Only SNP-gene combinations with significant associations are reported (FDR<0.05, P<0.00408). NA indicates that the gene and/or SNP is not available in the data set; NS indicates that P> 0.05 for the association.
**Figure Legends**

**Figure 1:** Manhattan plot showing the results for the 2,445,683 meta-analyzed SNPs in the discovery cohorts (a total of 19,683 individuals). SNPs are represented on the X-axis organized by chromosome. On the Y-axis, statistical significance is expressed as $-\log_{10}$ of the $P$ values. The horizontal line marks the $P=5.0 \times 10^{-8}$ threshold of genome-wide significance.

**Figure 2:** Box plots (median, inter-quartile range and 95% confidence intervals) showing the differences in expression levels of *MUC3* (Figure 2a) and *PAI-1* (Figure 2b) when comparing *MUC3*-silenced cells with control cells. Figure 2c shows the differences in PAI-1 released into the media when comparing *MUC3*-silenced cells with control cells. Gene expression and protein levels are normalized to control (non-silenced) cells (100%).
Prepublished online September 18, 2012; doi:10.1182/blood-2012-06-436188

Genome-wide association study for circulating levels of plasminogen activator inhibitor-1 (PAI-1) provides novel insights into the regulation of PAI-1