A meta-analysis of 120,246 individuals identifies 18 new loci for fibrinogen concentration

Paul S. de Vries[1], Daniel I. Chasman[2,3], Maria Sabater-Lleal[4], Ming-Huei Chen[5,6], Jennifer E. Huffman[7,6], Maristella Steri[8], Weihong Tang[9], Alexander Teumer[10], Riccardo E. Marioni[11-13], Vera Grossmann[14], Jouke J. Hottenga[15], Stella Trompet[16,17], Martina Müller-Nurasyid[18-20], Jing Hua Zhao[21], Jennifer A. Brody[22], Marcus E. Kleber[23], Xiuqing Guo[24,25], Jie Jin Wang[26], Paul L. Auer[27], John R. Attia[28,29], Lisa R. Yanek[30], Tarunveer S. Ahluwalia[31-33], Jari Lahti[34,35], Cristina Venturini[36,37], Toshiko Tanaka[38], Lawrence F. Bielak[39], Peter K. Joshi[40], Ares Rocanin-Arjo[41-43], Ivana Kolcic[44], Pau Navarro[45], Lynda M. Rose[2], Christopher Oldmeadow[28,29], Helene Riess[46], Johanna Mazur[47], Saonli Basu[48], Anuj Goel[49], Qiong Yang[50,6], Mohsen Ghanbari[1,51], Gonneke Willemsen[15], Ann Rumley[52], Edoardo Fiorillo[8], Anton J. M. de Craen[17], Anne Grotevendt[53], Robert Scott[21], Kent D. Taylor[24], Graciela E. Delgado[23], Jie Yao[24], Annette Kifley[26], Charles Kooperberg[54], Rehan Qayyum[30], Lorna M. Lopez[11,55,56], Tina L. Berentzen[57], Katri Räikkönen[34], Massimo Mangino[37], Stefania Bandinelli[58], Patricia A. Peyser[39], Sarah Wild[40], David-Alexandre Trégouët[41-43], Alan F. Wright[7], Jonathan Marten[7], Tatijana Zemunik[59], Alanna C. Morrison[60], Bengt Sennblad[4,61], Geoffrey Tofler[62], Moniek P. M. de Maat[63], Eco J. C. de Geus[15,64], Gordon D. Lowe[65], Magdalena Zoledziewska[8], Naveed Sattar[66], Harald Binder[67], Uwe Völker[68], Melanie Waldenberger[46,69], Kay-Tee Khaw[70], Barbara McKnight[71], Jie Huang[72], Nancy S. Jenny[73], Elizabeth G. Holliday[74], Lihong Qi[75], Mark G. McEvoy[76], Diane M. Becker[30], John M. Starr[11,77], Antti-Pekka Sarin[78,79], Pirro G. Hysi[37], Dena G. Hernandez[80], Min A. Jhun[39], Harry Campbell[40], Anders Hamsten[4], Fernando Rivadeneira[1,81], Wendy L. McArdle[82], P. Eline Slagboom[83], Tanja Zeller[84,85], Wolfgang Koenig[86], Bruce M. Psaty[87,88], Talin Haritunians[89], Jingmin Liu[54], Aarno Palotie[78], André G. Uitterlinden[1,81], David J. Stott[90], Albert Hofman[1], Oscar H. Franco[1], Ozren Polasek[91,92], Igor Rudan[40], Pierre-Emmanuel Morange[93-95], James F. Wilson[40], Sharon L. R. Kardia[39], Luigi Ferrucci[38], Tim D. Spector[37], Johan G. Eriksson[35,96-98], Torben Hansen[31], Ian J. Deary[11,99], Lewis C. Becker[30], Rodney J. Scott[100,101], Paul Mitchell[26], Winfried März[23,102,103], Nick J. Wareham[21], Annette Peters[46,20], Andreas Greinacher[104], Philipp S. Wild[14,105, 106], J. Wouter Jukema[16,107,108], Dorret I. Boomsma[15], Caroline Hayward[7], Francesco Cucca[8], Russell Tracy[109],

Hugh Watkins[49], Alex P. Reiner[54,110], Aaron R. Folsom[9], Paul M. Ridker[2,3], Christopher J. O'Donnell[6,111,112], Nicholas L. Smith[88,113,114], David P. Strachan[†115], and Abbas Dehghan[*†1]

[*]Corresponding author

[†]DP Strachan and A Dehghan contributed equally to this study as senior authors

Affiliations

[1]Department of Epidemiology, Erasmus MC, Rotterdam, the Netherlands

[2]Division of Preventive Medicine, Brigham and Women's Hospital, Boston, MA, USA

[3]Harvard Medical School, Boston, MA, USA

[4] Cardiovascular Medicine Unit, Department of Medicine, Karolinska Institutet, Stockholm, Sweden

[5]Department of Neurology, Boston University School of Medicine, Boston, MA, USA

[6]The National Heart, Lung, Blood Institute’s Framingham Heart Study, Framingham, MA, USA

[7]MRC Human Genetics Unit, Institute of Genetics and Molecular Medicine, University of Edinburgh, Edinburgh , UK

[8]Istituto di Ricerca Genetica e Biomedica, Consiglio Nazionale delle Ricerche, Monserrato, Cagliari, Italy

[9]Division of Epidemiology and Community Health, University of Minnesota, Minneapolis, MN, USA

[10]Institute for Community Medicine, University Medicine Greifswald, Greifswald, Germany

[11]Centre for Cognitive Ageing and Cognitive Epidemiology, University of Edinburgh, Edinburgh, UK

[12]Centre for Genomic and Experimental Medicine, University of Edinburgh, Edinburgh, UK

[13]Queensland Brain Institute, University of Queensland, Brisbane, Australia

[14]Center for Thrombosis and Hemostasis (CTH), University Medical Center of the Johannes Gutenberg University Mainz, Germany

[15]Department of Biological Psychology, Netherlands Twin Register, VU University, Amsterdam, the Netherlands

[16]Department of Cardiology, Leiden University Medical Center, Leiden, the Netherlands

[17]Department of Gerontology and Geriatrics, Leiden University Medical Center, Leiden, the Netherlands
[18]Institute of Genetic Epidemiology, Helmholtz Zentrum München - German Research Center for Environmental Health, Neuherberg, Germany
[19]Department of Medicine I, Ludwig-Maximilians-University Munich, Munich, Germany
[20]DZHK (German Centre for Cardiovascular Research), partner site Munich Heart Alliance, Munich, Germany
[21]MRC Epidemiology Unit, School of Clinical Medicine, University of Cambridge, Cambridge, UK
[22]Department of Medicine, University of Washington, Seattle WA, USA
[23]Vth Department of Medicine, Medical Faculty Mannheim, Heidelberg University, Mannheim, Germany
[24]Institute for Translational Genomics and Population Sciences, Los Angeles Biomedical Research Institute at Harbor/UCLA Medical Center, Torrance, CA, USA
[25]Department of Pediatrics, Los Angeles Biomedical Research Institute at Harbor/UCLA Medical Center, Torrance, CA, USA
[26]Centre for Vision Research, Department of Ophthalmology, and Westmead Millennium Institute, University of Sydney, Sydney, Australia
[27]Zilber School of Public Health, University of Wisconsin-Milwaukee, Milwaukee, WI, USA
[28]Public Health Stream, Hunter Medical Research Institute, New Lambton Heights, Australia
[29]School of Medicine and Public Health, University of Newcastle, New Lambton Heights, Australia
[30]Department of Medicine, Johns Hopkins University School of Medicine, Baltimore, MD, USA
[31]Novo Nordisk Foundation Centre for Basic Metabolic Research, Section of Metabolic Genetics, Faculty of Health and Medical Sciences, University of Copenhagen, Copenhagen, Denmark
[32]Copenhagen Prospective Studies on Asthma in Childhood, Faculty of Health and Medical Sciences, University of Copenhagen, Copenhagen, Denmark
[33]The Danish Pediatric Asthma Center, Gentofte Hospital, The Capital Region, Copenhagen, Denmark
[34]Institute of Behavioural Sciences, University of Helsinki, Helsinki, Finland
[35]Folkhälsan Research Centre, Helsinki, Finland

[36]Institute of Opthalmology, UCL, London, UK

[37]Department of Twin Research and Genetic Epidemiology, Kings College London, London, UK

[38]Translational Gerontology Branch, National Institute on Aging, Baltimore, MD, USA

[39]Department of Epidemiology, University of Michigan, Ann Arbor, MI, USA

[40]Centre for Population Health Sciences, University of Edinburgh, Edinburgh, UK

[41]Institut National pour la Santé et la Recherche Médicale (INSERM), Unité Mixte de Recherche en Santé (UMR_S) 1166, F-75013, Paris, France

[42]Sorbonne Universités, Université Pierre et Marie Curie (UPMC Univ Paris 06), UMR_S 1166, Team Genomics & Pathophysiology of Cardiovascular Diseases, F-75013, Paris, France

[43]Institute for Cardiometabolism and Nutrition (ICAN), F-75013, Paris, France

[44]Department of Public Health, Faculty of Medicine, University of Split, Split, Croatia

[45]MRC Human Genetics Unit, Institute of Genetics and Molecular Medicine, University of Edinburgh, Edinburgh, UK

[46]Institute of Epidemiology II, Helmholtz Zentrum München - German Research Center for Environmental Health, Neuherberg, Germany

[47]Institute of Medical Biostatistics, Epidemiology and Informatics, University Medical Center of the Johannes Gutenberg University Mainz, Germany

[48]Division of Biostatistics, University of Minnesota, Minneapolis, MN, USA

[49]Cardiovascular Medicine Dept/Radcliffe Dept of Medicine, Wellcome Trust Centre for Human Genetics, University of Oxford, Oxford, UK

[50]Department of Biostatistics, Boston University School of Public Health, Boston, MA, USA

[51]Department of Genetics, School of Medicine, Mashhad University of Medical Sciences, Mashhad, Iran

[52]Institute of Cardiovascular and Medical Sciences, University of Glasgow, UK

[53]Institute of Clinical Chemistry and Laboratory Medicine, University Medicine Greifswald, Greifswald, Germany

[54]Fred Hutchinson Cancer Research Center, Seattle, WA, USA

[55]Royal College of Surgeons in Ireland, Department of Psychiatry, Education and Research Centre, Beaumont Hospital, Dublin, Ireland

[56]University College Dublin, UCD Conway Institute, Centre for Proteome Research, UCD, Belfield, Dublin, Ireland

[57]Institute of Preventive Medicine, Bispebjerg and Frederiksberg Hospital, The Capital Region, Copenhagen, Denmark

[58]Geriatric Unit, Azienda Sanitaria Firenze (ASF), Florence, Italy

[59]Department of Biology, University of Split, School of Medicine, Split, Croatia

[60]Human Genetics Center, University of Texas Health Science Center at Houston, Houston, TX, USA

[61]Science for Life Laboratory, Karolinska Institutet, Stockholm, Sweden

[62]Royal North Shore Hospital, Sydney University, Sydney, Australia

[63]Department of Hematology, Erasmus MC, Rotterdam, the Netherlands

[64]EMGO+ institute, VU University & VU medical center, Amsterdam

[65]Institute of Cardiovascular and Medical Sciences, University of Glasgow, Glasgow, UK

[66]BHF Glasgow Cardiovascular Research Centre, Faculty of Medicine, Glasgow, UK

[67]Institute of Medical Biostatistics, Epidemiology and Informatics, University Medical Center of the Johannes Gutenberg University Mainz, Mainz, Germany

[68]Interfaculty Institute for Genetics and Functional Genomics, University Medicine Greifswald, Greifswald, Germany

[69]Research Unit of Molecular Epidemiology, Helmholtz Zentrum München - German Research Center for Environmental Health, Neuherberg, Germany

[70]Clinical Gerontology Unit, School of Clinical Medicine, University of Cambridge, Cambridge, UK

[71]Department of Biostatistics, University of Washington, Seattle, WA, USA

[72]Department of Human Genetics, Wellcome Trust Sanger Institute, Hinxton, Cambridge, UK

[73]Department of Pathology and Laboratory Medicine, University of Vermont College of Medicine, Colchester, VT, USA

[74]Public Health Stream, Hunter Medical Research Institute, and School of Medicine and Public Health, University of Newcastle, Newcastle, Australia

[75]UC Davis, Davis, CA, USA

[76]School of Medicine and Public Health, University of Newcastle, Australia

[77]Alzheimer Scotland Dementia Research Centre, University of Edinburgh, Edinburgh, UK

[78]Institute for Molecular Medicine Finland (FIMM), University of Helsinki, Helsinki, Finland

[79]Public Health Genomics Unit, National Institute for Health and Welfare, Finland

[80]Laboratory of Neurogenetics, National Institute on Aging, Bethesda, MD, USA

[81]Department of Internal Medicine, Erasmus MC, Rotterdam, the Netherlands

[82]School of Social and Community Medicine, University of Bristol, Bristol, UK

[83]Department of Molecular Epidemiology, Leiden University Medical Center, Leiden , the Netherlands

[84]Department of General and Interventional Cardiology, University Heart Centre, University Medical Center Hamburg-Eppendorf, Hamburg, Germany

[85]German Center for Cardiovascular Research (DZHK), Partner Site Hamburg, Lübeck, Kiel, Hamburg, Germany

[86]Department of Internal Medicine II - Cardiology, University of Ulm Medical Centre, Ulm, Germany

[87]Department of Medicine, Epidemiology, and Health Services, University of Washington, Seattle WA, USA

[88]Group Health Research Institute, Group Health Cooperative, Seattle WA, USA

[89]Inflammatory Bowel & Immunobiology Research Institute, Cedars-Sinai Medical Center, Los Angeles, CA, USA

[90]Institute of Cardiovascular and Medical Sciences, Faculty of Medicine, University of Glasgow, Glasgow, UK

[91]Centre for Global Health, Department of Public Health, Faculty of Medicine, University of Split, Split, Croatia

[92]Centre for Population Health Sciences, University of Edinburgh, Teviot Place, Edinburgh, UK

[93]Laboratory of Haematology, La Timone Hospital, F-13385, Marseille, France

[94]INSERM, UMR_S 1062, Nutrition Obesity and Risk of Thrombosis,F-13385, Marseille, France

[95]Aix-Marseille University, UMR_S 1062, Nutrition Obesity and Risk of Thrombosis, F-13385, Marseille, France

[96]Department of General Practice and Primary health Care, University of Helsinki, Helsinki, Finland

[97]National Institute for Health and Welfare, Finland

[98]Helsinki University Central Hospital, Unit of General Practice, Helsinki, Finland

[99]Department of Psychology, University of Edinburgh, Edinburgh, UK

[100]Information based Medicine Program, Hunter Medical Research Institute, New Lambton Heights, Australia

[101]School of Biomedical Sciences and Pharmacy, University of Newcastle, New Lambton Heights, Australia

[102]Synlab Academy, Synlab Services LLC, Mannheim, Germany

[103]Clinical Institute of Medical and Chemical Laboratory Diagnostics, Medical University of Graz, Graz, Austria

[104]Institute for Immunology and Transfusion Medicine, University Medicine Greifswald, Greifswald, Germany

[105]Preventive Cardiology and Preventive Medicine, Department of Medicine 2, University Medical Center of the Johannes Gutenberg University Mainz, Mainz, Germany

[106]German Center for Cardiovascular Research (DZHK), Partner Site RheinMain, Mainz, Germany

[107]Durrer Center for Cardiogenetic Research, Amsterdam, the Netherlands

[108] Interuniversity Cardiology Institute of the Netherlands, Utrecht, the Netherlands

[109]Department of Pathology and Laboratory Medicine, Center for Clinical and Translational Sciences, University of Vermont College of Medicine, Colchester, VT, USA

[110]University of Washington, Seattle, WA, USA

[111]National Heart, Lung and Blood Institute Division of Intramural Research

[112]Cardiology Division, Massachusetts General Hospital, Boston, MA, USA

[113]Department of Epidemiology, University of Washington, Seattle WA, USA

[114]Seattle Epidemiologic Research and Information Center, Department of Veterans Affairs Office of Research and Development, Seattle, WA, USA

[115]Population Health Research Institute, St George's, University of London, London, UK

Corresponding author:

Abbas Dehghan MD PhD FESC

Department of Epidemiology

Erasmus Medical Center

Wytemaweg 80, 3015CN Rotterdam,

Phone: 0031 10 70 43997

Fax: 0031 10 70 44657

Email: a.dehghan@erasmusmc.nl

Abstract

Genome-wide association studies have previously identified 23 genetic loci associated with circulating fibrinogen concentration. These studies used HapMap imputation and did not examine the X chromosome. 1000 Genomes imputation provides better coverage of uncommon variants, and includes indels. We conducted a genome-wide association analysis of 34 studies imputed to the 1000 Genomes Project reference panel and including ~120,000 participants of European ancestry (95,806 participants with data on the X chromosome). Approximately 10.7 million SNPs and 1.2 million indels were examined. We identified 41 genome-wide significant fibrinogen loci of which 18 were newly identified. There were no genome-wide significant signals on the X chromosome. The lead variants of 5 significant loci were indels. We further identified 6 additional independent signals, including 3 rare variants, at two previously characterized loci: *FGB* and *IRF1*. Together the 41 loci explain 3% of the variance in plasma fibrinogen concentration.

Fibrinogen is a coagulation factor crucial to clot formation, and an active regulator of the inflammatory response (1). It is a strong and established predictor of cardiovascular disease, autoimmune disorders, and cancer (1-5). Circulating fibrinogen concentration has a moderate heritability of 34% to 46% (6-8). Previous genome-wide association studies (GWAS) have highlighted genetic loci involved in inflammatory pathways such as the acute-phase response and interleukin 1 and 6 signaling as main determinants of fibrinogen concentration (9-13).

The variance in fibrinogen concentration explained by genetic loci identified in these previous GWAS is less than one tenth of its estimated heritability (11). It is therefore likely that part of the heritability stems from genetic variants that are not well tagged by the single nucleotide polymorphisms (SNPs) found in HapMap, including further common, uncommon, and rare SNPs, and other types of variants such as insertions or deletions (indels). Additionally, part of the heritability could be explained by variants on the X chromosome, which has not previously been interrogated.

To better interrogate the full range of genetic variants, including those with low minor allele frequency that may have been poorly tagged by HapMap variants, we performed a meta-analysis of 34 GWAS imputed using 1000 Genomes Project reference panels (14), including the X chromosome. We performed a joint/conditional analysis to identify additional independent signals within known and new loci associated with plasma fibrinogen concentration.

Results

Autosomal meta-analysis

Participant characteristics in each study are shown in Supplementary Table 1, covariates adjusted for by each study are shown in Supplementary Table 2, and genomic inflation factors are shown in Supplementary Table 3. The meta-analysis of the autosomes included 9,492,263 SNPs and 841,128 indels, of which 4,354 SNPs and 420 indels at 41 loci were genome-wide significant. Of these, 18 loci are new signals (Table 1), while 23 have been associated with fibrinogen concentration by previous GWAS (Table 2). Among genome-wide significant variants, 14 of 4,354 were rare (MAF ≤ 0.01), and a further

477 were uncommon ($0.01 < \text{MAF} \leq 0.05$). The lead variants of known locus *SNX13*, and novel loci *ATXN2L*, *GYS2*, *GIMAP4*, and *IFT122* were indels. Separate QQ plots of all autosomal variants, common variants, uncommon variants, rare variants, SNPs, and indels are shown in Supplementary Figure 1. A Manhattan plot of all autosomal variants is shown in Supplementary Figure 2. Additionally, a Manhattan plot highlighting rare and uncommon variants is shown in Supplementary Figure 3. Heterogeneity I^2 and *P*-values are shown in Supplementary Table 4. Only rs7439150 at the fibrinogen gene cluster showed significant heterogeneity (I^2: 50.0, *P*-value: 0.0004). Regional plots are shown in Supplementary Figure 4, and forest plots are shown in Supplementary Figure 5. Associations with rare variants were found at the two most robust fibrinogen loci: the fibrinogen gene cluster and the *IRF1* locus (lead variant annotated to *C5orf56*). Associations with uncommon variants were also found at these loci, as well as at *SPPL2A* and *HNF4A*. At one known locus (*SNX13*) and four new loci (*IFT122*, *GIMAP4*, *GYS2*, and *ATXN2L*) the lead variant was an indel. At each of these loci there were also SNPs in linkage disequilibrium with the indel that reached genome-wide significance. *CD300LF* was the only previously identified locus that was not represented among our significant results. The previously reported lead variant in *CD300LF*, rs10512597 (*P*-value: 1.8×10^{-7}), had a smaller effect size (β: -0.006 ln(g/L)) than was previously reported (β: -0.008 ln(g/L)). There was no strong evidence of heterogeneity (I^2: 22.7, *P*-value: 0.11).

Conditional analysis

Two loci (fibrinogen gene cluster and *IRF1*) harbored multiple jointly significant variants (Table 3). Forest plots of the additional variants discovered through conditional analysis are shown in Supplementary Figure 6, and their heterogeneity I^2 and *P*-values are shown in Supplementary Table 5. At the fibrinogen gene cluster, five variants were jointly significant: the lead variant rs7439150, an additional common variant rs76289367, and three rare variants, rs150768229, rs6054, and rs148685782. rs148685782 showed significant heterogeneity ($I^2 = 65.0$, *P*-value = 0.0004). At the *IRF1* locus three variants were jointly significant: the lead variant, rs2057655, and two uncommon variants, rs12777 and 5:131786964. Of the secondary signals, rs12777 is in strong linkage disequilibrium with a previously

53 associated SNP, rs1242111 (R^2=0.8), while 5:131786964 is a new independent signal (R^2 = 0.0). The

54 uncommon variants near *SPPL2A* were not significant in the conditional analysis. The uncommon lead

55 variant rs141272690 was only marginally significant in the primary analysis (P-value = 1.89×10^{-8}), so

56 that even a small correlation with the lead common variant rs12913259 (R^2 = 0.02) raised the *P*-value

57 above the threshold in the conditional analysis.

58

59 *X-chromosome meta-analysis*

60 The meta-analysis of the X chromosome included 251,747 SNPs and 26,448 indels. There were no

61 genome-wide significant variants detected on the X chromosome. This was true in both sex-specific meta-

62 analyses, and in the combined meta-analyses, irrespective of whether the sex-specific results were

63 combined using inverse-variance weighted meta-analysis or sample size based meta-analyses. QQ plots

64 and Manhattan plots for the X chromosome are shown in Supplementary Figure 7 and 8.

65

66 *Functional annotation*

67 Genome-wide significant associations with other traits were found for 28 out of the 41 loci, of which 10

68 were associated with cholesterol levels, 7 were associated with C-reactive protein, and 5 were associated

69 with platelet count (Supplementary Table 5). Out of the 41 lead variants, 20 were associated with blood

70 expression levels of one or more neighboring genes (Supplementary Table 6). Notably, rs1035559 at

71 16q22.2 was exclusively associated with *HP* expression levels ($P = 9.8\times10^{-198}$), and rs7224737 at 17q21.2

72 was exclusively associated with *STAT3* expression levels ($P = 5.4\times10^{-12}$). Out of the 41 lead variants 36

73 were available in HaploReg V2. Detailed annotation of these variants as well as 457 correlated SNPs is

74 shown in Supplementary Table 7. Eight of these SNPs are predicted to influence the binding of miRNAs

75 to transcripts of their host gene. Further information about these SNPs and their effect on miRNA binding

76 is shown in Supplementary Table 8. Of these eight SNPs, two were lead variants. First, the fibrinogen

77 decreasing minor allele of lead variant rs715 in the 3'-UTR of *CPS1* is predicted to create a miRNA

78 binding site for miR-3154. Second, the fibrinogen increasing minor allele of lead variant rs6224634 in the

3'-UTR of *LHFPL4* is predicted to disrupt the binding site of miR-6761-3p. In both cases predicted successful miRNA-target gene binding is associated with lower fibrinogen concentration.

Variance explained

In the Women's Genome Health Study, the lead variant at the fibrinogen gene cluster explained 0.8% of the variance, and all five jointly significant variants together explained 1.6% of the variance. At 5q31.1 the lead variant explained 0.2% of the variance, while all three jointly significant variants together explained 0.3% of the variance. The 47 independently significant variants at 41 loci explained 3.0% of the variance in circulating fibrinogen concentration. The variance explained by the 23 previously identified loci was 2.6%.

Discussion

We identified 18 new autosomal loci associated with circulating fibrinogen concentration in individuals of European ancestry, increasing the variance explained from 2.6% to 3.0%. The small increase in the variance explained relative to the large number of new loci is suggestive of a highly polygenic genetic architecture. At two loci (fibrinogen gene cluster and *IRF1* locus) rare or uncommon variants were jointly significant alongside common lead variants. In five cases the lead variant at an associated locus was an indel. There were no significant associations on the X chromosome: this may be result of issues specific to the X chromosome rather than the absence of relevant signals. The most important issue is that the X chromosome is generally poorly covered by genotyping arrays (15).

Four of the 18 new loci implicate inflammatory pathways not previously linked to fibrinogen. First, the septin gene family is represented at two significant loci: *SEPT7* at 7p14.2 and *SEPT2* at 2q37.3. Proteins from the septin gene family form cage-like structures around bacteria to facilitate autophagy (16). The link between these processes and fibrinogen concentration is unclear. Second, our results also implicate genes from the GIMAP family, which are structurally similar to septins (17). The signal at 7q36.1 appears to be driven by one or more genes from a cluster of eight GIMAP genes, and the lead

105 variant is associated with blood expression levels of four of these. Through their involvement in
106 lymphocyte maturation, these genes influence lymphocyte counts and diversity, and thereby also the
107 inflammatory response (18). Finally, the lead variant at 16q22.2 is strongly associated with blood
108 expression levels of the neighboring *HP* (P-value $\leq 9.8\times10^{-198}$), the gene encoding haptoglobin. Like
109 fibrinogen, haptoglobin is an acute-phase reactant. The association of rs1035560 with fibrinogen
110 suggests that besides sharing upstream regulators, haptoglobin itself may be involved in the regulation of
111 circulating fibrinogen.

112 Six of the new loci appear to be closely related to STAT3, a transcription factor working
113 downstream of IL-6 that upregulates the expression of fibrinogen and other acute-phase proteins (19). At
114 17q21.2, lead variant rs7224737 (175 kb from *STAT3*) was associated with *STAT3* blood expression
115 levels ($P = 5.4\times10^{-12}$). At 9q22.2, the lead variant rs3138493 lies upstream of *GADD45G*. This gene is
116 expressed in the liver, where it has been shown to inhibit the Tyr705 phosphorylation of STAT3 (20). As
117 Tyr705 phosphorylation of STAT3 allows it to dimerize and move into the nucleus, it is essential for the
118 upregulation of STAT3 targets like the fibrinogen genes. At 10q26.13, the lead variant rs2420915 is an
119 intergenic SNP close to *FGFR2*. Over-expression of *FGFR2*, or the related *FGFR1* is required for the
120 Tyr705 phosphorylation of STAT3 (20). At 19q13.33, the lead variant rs73058052 is associated with
121 blood expression levels of *IRF3*. After activation in response to viral infection, *IRF3* enables the
122 expression of type I interferons *INFA* and *INFB*, leading to the upregulation of *STAT3* (21, 22).
123 Furthermore, our results point towards two SH2B adaptor proteins implicated in STAT3 signaling. At
124 12q24.12, the lead variant rs7310615 was associated with blood expression levels of *SH2B3*. Using
125 immortalized B lymphoblastoid cell lines, a loss of the SH2B3 protein was accompanied by increased
126 STAT3 phosphorylation (23). At 16p11.2, lead variant 16:28845027 lies close to *SH2B1*. The β variant
127 of SH2B1 appears to form a complex with STAT3, allowing STAT3 to cross through the membrane into
128 the nucleus as an alternative to STAT3 dimerization (24). Collectively, these findings suggest that a wide
129 range of disturbances to STAT3 may affect circulating fibrinogen concentration.

130 In addition to STAT3, our results highlight HNF4A, another transcription factor known to

131 regulate fibrinogen gene expression. The association between lead variant rs1800961 and circulating
132 fibrinogen has been previously been described by Wassel et al and Hufman et al (12, 25). rs1800961 is a
133 nonsynonymous coding variant that has been shown to decrease *HNF4A* expression in vitro (26).
134 The majority of rare and uncommon variants associated with fibrinogen concentration were found
135 at loci with common variant signals. Only the signal at *HNF4A* was led by an uncommon variant, and no
136 signals were led by rare variants. Conditional analysis suggests that there are two secondary signals at the
137 *IRF1* locus led by uncommon variants, and three secondary signals near the fibrinogen gene cluster led by
138 rare variants. The uncommon variants that were significant near *SPPL2A* were not significant in the
139 conditional analysis, but the linkage disequilibrium with the lead common variant was very low. Our
140 results suggest that common and rare variant signals are often independent of each other, and do not
141 support the hypothesis that associations with common variants are synthetic associations merely reflecting
142 linkage disequilibrium with rare variants (27, 28).
143 Absolute effect sizes of significant variants ranged from 0.005 to 0.033 ln(g/L) among common
144 variants, 0.013 to 0.087 ln(g/L) among uncommon variants, and 0.036 to 0.254 ln(g/L) among rare
145 variants. Despite their small effect size, common variants have helped discover biologically relevant
146 fibrinogen loci. Therefore, the complete lack of overlap between the effect sizes of significant common
147 and rare variants suggests that further rare variants with smaller effect sizes are likely to exist at important
148 and possibly unknown fibrinogen loci. While the rare variants with large effects we found were limited to
149 the two most important fibrinogen loci, rare variants with moderate effects may be more widespread.
150 When considering not only the primary signal at the fibrinogen gene cluster, but also the four
151 additional signals the variance explained by the locus doubles from 0.8% to 1.6%. Two of these
152 additional signals are driven by rare non-synonymous exonic variants (rs6054 and rs148685782) with
153 very large effect sizes (β=-0.12 and β =-0.21 ln(g/L) respectively). The association between rs6054 and
154 fibrinogen has been described earlier in a candidate gene study (12), and rs148685782 (also known as
155 γAla82Gly) has previously been reported as a causal variant for mild congenital hypofibrinogenaemia
156 (29-31). Furthermore, in a previous study we examined exome-wide genotypes using exome arrays and

157 identified independent associations of both rs6054 and rs148685782 with fibrinogen (25). In the present
158 study, however, two further variants, rs140473879 and rs149234484, are in strong linkage disequilibrium
159 with rs148685782 and tag this signal. These variants are intergenic, but each changes several regulatory
160 motifs. Thus, the identification of rs148685782 as a causal variant is not conclusive.

161 Strengths of this study include the use of a large ethnically homogenous sample, and coverage of
162 previously unexamined uncommon and rare variants, indels, and variants on the X chromosome. At the
163 same time, the lack of ethnic heterogeneity may also be a limitation, as including different ethnicities can
164 help narrow down the association signal to a smaller region (32). This study has other limitations that
165 should be acknowledged. To most effectively use the available data, we used all 34 studies in the
166 discovery sample (33). The results have thus not been replicated. Nevertheless, the consistent association
167 of these loci across the 34 studies and the strict Bonferroni correction enforcing a 5% false discovery rate
168 ensure that essentially all of the loci represent true associations. A second limitation is that an
169 approximation based on meta-analysis summary data was used to identify additional independently
170 associated variants at the identified loci rather than a stepwise conditional analysis using individual-level
171 data. Different methods were used to measure plasma fibrinogen across the studies: EDTA or citrate
172 plasma samples were used, and a variety of assays were used (34). While the association between
173 fibrinogen and cardiovascular disease has previously been shown to be independent of assay type, the
174 genetic etiology of fibrinogen may differ across assay types (35). However, to minimize the impact on our
175 results, studies that used multiple assays to measure fibrinogen performed their analyses stratified by the
176 assay.

177 Finally, our ability to attribute these signals to causal genes remains limited. For each locus we
178 reported the gene closest to the lead variant, but proximity alone is not strong evidence that a gene is the
179 underlying causal gene. Thus, we also reported the genes whose expression levels in blood were most
180 strongly associated with the lead variant, and we reported genes with nonsynonymous exonic variants in
181 high linkage disequilibrium with the lead variant. Based on blood expression levels, some signals were
182 characterized by a single promising candidate causal gene, but other signals were associated with either

183 no candidate causal genes, or more than one. Furthermore, genetic variants can have effects on the
184 expression of multiple genes across different tissues, and these effects can be tissue specific.
185 We identified 41 loci that collectively explain 3% of the variance in plasma fibrinogen
186 concentration. Of these loci, 18 had not been identified previously through GWAS. The new loci
187 emphasize the importance of STAT3 to fibrinogen regulation, and highlight several new potential
188 pathways that should be experimentally confirmed. The use of 1000 Genomes Project imputation
189 increased our ability to assess the role of uncommon variants, resulting in an in depth characterization of
190 the two most important fibrinogen loci.
191
192 **Materials and Methods**
193 *Study sample*
194 This meta-analysis was conducted within the framework of the Cohorts for Heart and Aging Research in
195 Genetic Epidemiology (CHARGE) consortium (36). The study sample consists of 34 studies with
196 120,246 individuals of European ancestry. 12 studies with 25,453 participants were not included in the
197 previous fibrinogen GWAS (11). Fibrinogen concentration was measured in citrated or EDTA plasma
198 samples using a variety of methods including the Clauss method, immunonephelometric methods,
199 immunoturbidimetric methods, and prothrombin time derived methods as described in Supplementary
200 Table 1 and the Supplementary Methods, which further describe the studies. All studies were approved by
201 appropriate research ethics committees and all respondents signed informed consent prior to participation.
202
203 *Genotyping and imputation*
204 Genotyping, pre-imputation quality control, imputation, and analysis methods are presented in
205 Supplementary Table 2. All studies imputed variant dosages using reference panels from the 1000
206 Genomes Project using MACH or IMPUTE (14, 37-39). The phase I version 3 reference panel was used
207 by all studies except two, which used the phase I version 2 reference panel. Before meta-analysis, we
208 excluded variants with MACH imputation quality < 0.3 or IMPUTE imputation quality < 0.4, and

209 variants with effective minor allele count (minor allele count × imputation quality) < 10. These filters
210 were applied at the level of individual studies. Because we wanted to focus only on those variants that
211 passed these filters in a large proportion of the studies, we additionally excluded variants with a total
212 sample size of less than half of the maximum sample size at the meta-analysis level.
213
214 *Autosomal association analysis*
215 Plasma fibrinogen concentration was converted to g/L and natural log transformed. All studies adjusted
216 for age and sex. When necessary, analyses were also adjusted for study-specific covariates, such as center
217 or case/control status. In family studies, linear mixed models were used to account for family structure.
218 Analyses were adjusted for principal components to account for population structure and cryptic
219 relatedness. These adjustments are shown in Supplementary Table 2. To account for remaining
220 stratification, we applied a genomic control correction to the results of each of the studies before meta-
221 analysis. We used an inverse-variance model with fixed effects implemented in METAL to meta-analyze
222 association results (40). Heterogeneity was assessed using I^2 and corresponding *P*-values.
223 As proposed by Huang et al, variants with *P*-values lower than 2.5×10^{-8} were considered genome-
224 wide significant (based on a Bonferroni correction for 2,000,000 tests) (41). Significant variants were
225 assigned to loci in order of ascending *P*-value. A variant was assigned to a new locus when there were no
226 significant variants within 500 kb of it belonging to a previously defined locus. Variants were annotated
227 to genes using ANNOVAR version 2013Mar07 (42).
228
229 *X-chromosome association analysis*
230 Of the 120,246 participants, 95,806 had imputed data on the X chromosome. Dosages of variants on the X
231 chromosome were coded as [0,2] in men and [0,1,2] in women. This way one allele in men has the same
232 value as two alleles in women. Thus, we assume full inactivation of one of the two X chromosomes in
233 women. Variants in the pseudo-autosomal region were excluded. Analyses of the X chromosome were
234 stratified by sex in each study, and the studies then were meta-analyzed separately for men and women

235 using an inverse-variance model with fixed effects (40). We then combined the sex-specific meta-analysis
236 results for variants on the X chromosome using both an inverse variance weighted model with fixed
237 effects and a sample-size weighted model based on *P*-values and effect direction. The sample-size
238 weighted model does not take the effect size into account, and thus may work better when there are
239 different effects in men and women (43, 44), as can happen when there is incomplete inactivation in
240 women.
241
242 *Conditional analysis*
243 Some loci may harbor multiple independent variants that affect fibrinogen (11, 45). To putatively identify
244 these jointly significant variants, we used an approximate method for conditional and joint analysis using
245 meta-analysis summary statistics implemented in GCTA (46, 47). The method consists of a genome-wide
246 stepwise selection procedure selecting variants according to their conditional *P*-values and, after the
247 model has been optimized, the estimation of the joint effects of the selected variants. This method
248 depends on a reference panel to estimate linkage disequilibrium patterns between variants. We used best-
249 guess imputation for variants with imputation quality > 0.3 in 5,733 unrelated individuals from the
250 Rotterdam Study as the reference panel (48). A description of the Rotterdam Study is given in the
251 Supplementary Methods.
252
253 *Functional annotation*
254 For each locus, we searched the National Human Genome Research Institute GWAS catalog for genome-
255 wide significant associations with other traits within 100kb of the lead variant (49). We used the Blood
256 eQTL browser, a publicly available database, to examine whether any lead variants, or their most
257 correlated HapMap proxy (with $R^2 > 0.8$), were associated with expression levels of nearby genes in
258 blood. Results from the blood eQTL browser are based on non-transformed peripheral blood samples
259 from 5,311 individuals with replication in 2,775 individuals (50). For each lead SNP and its highly
260 correlated neighbors (with $R^2 > 0.9$), we used HaploReg V2 to determine the level of conservation,

261 association with gene expression in a range of tissues including the liver, and any overlap with ENCODE
262 transcription factor binding sites, and DNAse-hypersensitive, promoter , and enhancer regions in various
263 cell types (51, 52). Furthermore, we determined the overlap of these SNPs with microRNAs and
264 microRNA binding sites (see Supplementary Methods) (53-55).
265
266 *Variance explained*
267 In the Women's Genome Health Study, the largest contributor to the meta-analysis, we computed a
268 weighted genetic risk score based on the lead variants at each genome-wide significant locus, as well as
269 any jointly significant variants identified in the conditional analysis (56). A description of the Women's
270 Genome Health Study is given in the Supplementary Methods. Beta coefficients from the genome-wide
271 association meta-analysis including all studies were used as weights, except in loci with multiple jointly
272 significant variants. For variants at these loci, joint beta coefficients were obtained from the conditional
273 analysis. The genetic risk score was computed as the sum of the weighted variants dosages. The variance
274 in fibrinogen concentration explained was estimated using a linear regression model. Additionally, for
275 any loci with jointly significant variants we compared the variance explained by the lead variant to the
276 variance explained by the jointly significant variants. We were not able to directly compare our estimate
277 of the variance explained to previous estimates, as these had been computed in different populations and
278 were adjusted for age and sex. Thus, we re-calculated the variance explained without adjustment for age
279 and sex. For this we used HapMap-imputed dosages of the independently associated SNPs reported by
280 Sabater-Lleal et al (11). Since the variance explained is estimated on the basis of imperfectly imputed
281 dosages, we expect our estimates to be slightly lower than if they were based on measured genotypes.
282
283 **Acknowledgements**
284 The authors acknowledge the essential role of the Cohorts for Heart and Aging Research in Genome
285 Epidemiology (CHARGE) Consortium in development and support of this manuscript. Infrastructure for
286 the CHARGE Consortium is supported in part by the National Heart, Lung, and Blood Institute grant

287 R01HL105756. **ARIC** is carried out as a collaborative study supported by National Heart, Lung, and
288 Blood Institute (NHLBI) contracts HHSN268201100005C, HHSN268201100006C,
289 HHSN268201100007C, HHSN268201100008C, HHSN268201100009C, HHSN268201100010C,
290 HHSN268201100011C, and HHSN268201100012C, R01HL087641, R01HL59367 and R01HL086694;
291 National Human Genome Research Institute contract U01HG004402; and National Institutes of Health
292 contract HHSN268200625226C. The authors thank the staff and participants of the ARIC study for their
293 important contributions. Infrastructure was partly supported by Grant Number UL1RR025005, a
294 component of the National Institutes of Health and NIH Roadmap for Medical Research. We would like
295 to thank the University of Minnesota Supercomputing Institute for use of the calhoun supercomputers.
296 **B58C** acknowledges use of phenotype and genotype data from the British 1958 Birth Cohort
297 DNA collection, funded by the Medical Research Council grant G0000934 and the Wellcome
298 Trust grant 068545/Z/02. Genotyping for the B58C-WTCCC subset was funded by the
299 Wellcome Trust grant 076113/B/04/Z. The B58C-T1DGC genotyping utilized resources
300 provided by the Type 1 Diabetes Genetics Consortium, a collaborative clinical study sponsored
301 by the National Institute of Diabetes and Digestive and Kidney Diseases (NIDDK), National
302 Institute of Allergy and Infectious Diseases (NIAID), National Human Genome Research
303 Institute (NHGRI), National Institute of Child Health and Human Development (NICHD), and
304 Juvenile Diabetes Research Foundation International (JDRF) and supported by U01 DK062418.
305 B58C-T1DGC GWAS data were deposited by the Diabetes and Inflammation Laboratory,
306 Cambridge Institute for Medical Research (CIMR), University of Cambridge, which is funded by
307 Juvenile Diabetes Research Foundation International, the Wellcome Trust and the National
308 Institute for Health Research Cambridge Biomedical Research Centre; the CIMR is in receipt of
309 a Wellcome Trust Strategic Award (079895). The B58C-GABRIEL genotyping was supported
310 by a contract from the European Commission Framework Programme 6 (018996) and grants

from the French Ministry of Research. **BMES** has been supported by the Australian National Health & Medical Research Council, Canberra Australia (Grant Numbers 974159, 211069, 457349, 512423, 475604, 529912, and the funding for Centre for Clinical Research Excellence in Translational Clinical Research in Eye Diseases, CCRE in TCR-Eye, grant ID 529923); In addition, funding by the Wellcome Trust, UK (to A Viswanathan, P McGuffin, P Mitchell, F Topouzis, P Foster) has supported the genotyping costs of the entire BMES population. This **CHS** research was supported by NHLBI contracts HHSN268201200036C, HHSN268200800007C, N01HC55222, N01HC85079, N01HC85080, N01HC85081, N01HC85082, N01HC85083, N01HC85086; and NHLBI grants U01HL080295, R01HL087652, R01HL105756, R01HL103612, and R01HL120393 with additional contribution from the National Institute of Neurological Disorders and Stroke (NINDS). Additional support was provided through R01AG023629 from the National Institute on Aging (NIA). A full list of principal CHS investigators and institutions can be found at CHS-NHLBI.org. The provision of genotyping data was supported in part by the National Center for Advancing Translational Sciences, CTSI grant UL1TR000124, and the National Institute of Diabetes and Digestive and Kidney Disease Diabetes Research Center (DRC) grant DK063491 to the Southern California Diabetes Endocrinology Research Center. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health. The **CROATIA-Split** study was funded by grants from the Medical Research Council (UK), European Commission Framework 6 project EUROSPAN (Contract No. LSHG-CT-2006-018947) and Republic of Croatia Ministry of Science, Education and Sports research grants to I.R. (108-1080315-0302). We would like to acknowledge the staff of several institutions in Croatia that supported the field work, including but not limited to The University of Split and

334 Zagreb Medical Schools and the Croatian Institute for Public Health. The SNP genotyping for
335 the CROATIA-Split cohort was performed by AROS Applied Biotechnology, Aarhus, Denmark.
336 We would like to thank Jared O'Connell for performing the pre-phasing for CROATIA-Split
337 prior to 1000G imputation. The **CROATIA-Korcula** study was funded by grants from the
338 Medical Research Council (UK), European Commission Framework 6 project EUROSPAN
339 (Contract No. LSHG-CT-2006-018947) and Republic of Croatia Ministry of Science, Education
340 and Sports research grants to I.R. (108-1080315-0302). We would like to acknowledge the
341 invaluable contributions of the recruitment team in Korcula, the administrative teams in Croatia
342 and Edinburgh and the people of Korcula. The SNP genotyping for the CROATIA-Korcula
343 cohort was performed in Helmholtz Zentrum München, Neuherberg, Germany. We would like to
344 thank Jared O'Connell for performing the pre-phasing for CROATIA-Korcula prior to 1000G
345 imputation. The **CROATIA-Vis** study was funded by grants from the Medical Research Council
346 (UK) and Republic of Croatia Ministry of Science, Education and Sports research grants to I.R.
347 (108-1080315-0302). We would like to acknowledge the staff of several institutions in Croatia
348 that supported the field work, including but not limited to The University of Split and Zagreb
349 Medical Schools, the Institute for Anthropological Research in Zagreb and Croatian Institute for
350 Public Health. The SNP genotyping for the CROATIA-Vis cohort was performed in the core
351 genotyping laboratory of the Wellcome Trust Clinical Research Facility at the Western General
352 Hospital, Edinburgh, Scotland. We would like to thank Jared O'Connell for performing the pre-
353 phasing for CROATIA-Vis prior to 1000G imputation. **EPIC-Norfolk**. The EPIC Norfolk Study
354 is funded by program grants from the Medical Research Council UK and Cancer Research UK,
355 and by additional support from the European Union, Stroke Association, British Heart
356 Foundation, Department of Health, Food Standards Agency, and the Wellcome Trust. **FHS** was

357 partially supported by the National Heart, Lung, and Blood Institute's (NHLBI's) Framingham
358 Heart Study (Contract No. N01-HC-25195) and its contract with Affymetrix, Inc. for genotyping
359 services (Contract No. N02-HL-6-4278). A portion of this research utilized the Linux Cluster for
360 Genetic Analysis (LinGA-II) funded by the Robert Dawson Evans Endowment of the
361 Department of Medicine at Boston University School of Medicine and Boston Medical Center.
362 The analyses reflect intellectual input and resource development from the Framingham Heart
363 Study investigators participating in the SNP Health Association Resource (SHARe) project.
364 Partial investigator support was provided by the National Institute of Diabetes and Digestive and
365 Kidney Diseases K24 DK080140 (JB Meigs), the National Institute on Aging and National
366 Institute for Neurological Disorders and Stroke R01 AG033193, NS017950 (S Seshadri).
367 **GeneSTAR** was supported by grants from the National Institutes of Health/National Heart,
368 Lung, and Blood Institute (U01 HL72518, HL097698, HL59684, HL071025-01A1), by a grant
369 from the National Institutes of Health/National Institute of Nursing Research (NR0224103), and
370 by a grant from the National Institutes of Health/National Center for Research Resources (M01-
371 RR000052) to the Johns Hopkins General Clinical Research Center. The authors would like to
372 thank the men and women participating in the **HCS** as well as The University of Newcastle,
373 Vincent Fairfax Family Foundation and The Hunter
374 Medical Research Institute. The **Genetic Epidemiology Network of Arteriopathy (GENOA)**
375 study is supported by the National Institutes of Health, grant numbers HL087660 and HL100245
376 from National Heart, Lung, Blood Institute. We thank Eric Boerwinkle, PhD from the Human
377 Genetics Center and Institute of Molecular Medicine and Division of Epidemiology, University
378 of Texas Health Science Center, Houston, Texas, USA and Julie Cunningham, PhD from the

379 Department of Health Sciences Research, Mayo Clinic College of Medicine, Rochester, MN,
380 USA for their help with genotyping.
381 **The GOYA Male** study was conducted as part of the activities of the Gene-diet Interactions in
382 Obesity project (GENDINOB, www.gendinob.dk) and the MRC centre for Causal Analyses in
383 Translational Epidemiology (MRC CAiTE). We thank the staff of the Copenhagen City Heart
384 Study for their skillful examination of the study subjects in collection of baseline and follow-up
385 data. Tarunveer Singh Ahluwalia received his Postdoctoral Research funding from GENDINOB
386 project and acknowledges the same. The **Gutenberg Health Study** is funded through the
387 government of Rhineland-Palatinate („Stiftung Rheinland-Pfalz für Innovation", contract AZ
388 961-386261/733), the research programs "Wissen schafft Zukunft" and "Center for Translational
389 Vascular Biology (CTVB)" of the Johannes Gutenberg-University of Mainz, and its contract
390 with Boehringer Ingelheim and PHILIPS Medical Systems, including an unrestricted grant for
391 the Gutenberg Health Study. VG, PSW are funded by the Federal Ministry of Education and
392 Research (BMBF 01EO1003). We thank all study participants as well as everybody involved in
393 the **Helsinki Birth Cohort Study**. Helsinki Birth Cohort Study has been supported by grants
394 from the Academy of Finland, the Finnish Diabetes Research Society, Folkhälsan Research
395 Foundation, Novo Nordisk Foundation, Finska Läkaresällskapet, Signe and Ane Gyllenberg
396 Foundation, University of Helsinki, Ministry of Education, Ahokas Foundation, Emil Aaltonen
397 Foundation. **The InCHIANTI study** baseline (1998-2000) was supported as a "targeted project"
398 (ICS110.1/RF97.71) by the Italian Ministry of Health and in part by the U.S. National Institute
399 on Aging (Contracts: 263 MD 9164 and 263 MD 821336); This research was supported in part
400 by the Intramural Research Program of the NIH, National Institute on Aging. The **KORA**
401 **studies** were financed by the Helmholtz Zentrum München, German Research Center for

402 Environmental Health, Neuherberg, Germany, and supported by grants from the German Federal
403 Ministry of Education and Research (BMBF) and by the state of Bavaria. Part of this work was
404 financed by the German National Genome Research Network (NGFNPlus, project number
405 01GS0834) and through additional funds from the University of Ulm. Furthermore, the research
406 was supported within the Munich Center of Health Sciences (MC Health) as part of LMU
407 innovative. We thank the **LBC1936 and LBC1921** participants and research team members. We
408 thank the nurses and staff at the Wellcome Trust Clinical Research Facility, where subjects were
409 tested and the genotyping was performed. The whole genome association study was funded by
410 the Biotechnology and Biological Sciences Research Council (BBSRC; Ref. BB/F019394/1).
411 The LBC1936 research was supported by Age UK. The LBC1921 data collection was funded by
412 the BBSRC. The work was undertaken by The University of Edinburgh Centre for Cognitive
413 Ageing and Cognitive Epidemiology, part of the cross council Lifelong Health and Wellbeing
414 Initiative (G0700704/84698). Funding from the Biotechnology and Biological Sciences Research
415 Council (BBSRC), Engineering and Physical Sciences Research Council (EPSRC), Economic
416 and Social Research Council (ESRC) and Medical Research Council (MRC) is gratefully
417 acknowledged. **LURIC** has received funding from the 6th Framework Program (integrated
418 project Bloodomics, grant LSHM-CT-2004-503485) and from the 7th Framework Program
419 (Atheroremo, grant agreement number 201668 and RiskyCAD, grant agreement number 305739)
420 of the European Union as well as from the INTERREG IV Oberrhein Program (Project A28,
421 Genetic mechanisms of cardiovascular diseases) with support from the European Regional
422 Development Fund (ERDF) and the Wissenschaftsoffensive TMO. We thank the LURIC study
423 team who were either temporarily or permanently involved in patient recruitment as well as
424 sample and data handling, in addition to the laboratory staff at the Ludwigshafen General

Hospital and the Universities of Freiburg and Ulm, Germany. The **MARTHA** project was supported by grants from the Program Hospitalier de Recherche Clinique. Statistical analyses of the MARTHA dataset were performed using the C2BIG computing cluster, funded by the Région Ile de France, Pierre and Marie Curie University, and the ICAN Institute for Cardiometabolism and Nutrition (ANR-10-IAHU-05). **MESA** and the MESA SHARe project are conducted and supported by the National Heart, Lung, and Blood Institute (NHLBI) in collaboration with MESA investigators. Support is provided by grants and contracts N01 HC-95159, N01-HC-95160, N01-HC-95161, N01-HC-95162, N01-HC-95163, N01-HC-95164, N01-HC-95165, N01-HC-95166, N01-HC-95167, N01-HC-95168, N01-HC-95169 and RR-024156. Funding for CARe genotyping was provided by NHLBI Contract N01-HC-65226. The provision of genotyping data was supported in part by the National Center for Advancing Translational Sciences, CTSI grant UL1TR000124, and the National Institute of Diabetes and Digestive and Kidney Disease Diabetes Research Center (DRC) grant DK063491 to the Southern California Diabetes Endocrinology Research Center. The authors thank the participants of the MESA study, the Coordinating Center, MESA investigators, and study staff for their valuable contributions. A full list of participating MESA investigators and institutions can be found at http://www.mesa-nhlbi.org. **NTR:** Funding was obtained from the Netherlands Organization for Scientific Research (NWO) and MagW/ZonMW grants 904-61-090, 985-10-002, 904-61-193,480-04-004, 400-05-717, Addiction-31160008, Middelgroot-911-09-032, Spinozapremie 56-464-14192, Center for Medical Systems Biology (CSMB, NWO Genomics), NBIC/BioAssist/RK(2008.024), Biobanking and Biomolecular Resources Research Infrastructure (BBMRI –NL, 184.021.007). VU University's Institute for Health and Care Research (EMGO+) and Neuroscience Campus Amsterdam (NCA); the European Science Foundation (ESF, EU/QLRT-2001-01254), the

448 European Community's Seventh Framework Program (FP7/2007-2013), ENGAGE (HEALTH-
449 F4-2007-201413); the European Science Council (ERC Advanced, 230374), Rutgers University
450 Cell and DNA Repository (NIMH U24 MH068457-06), the Avera Institute, Sioux Falls, South
451 Dakota (USA) and the National Institutes of Health (NIH, R01D0042157-01A, MH081802,
452 Grand Opportunity grants 1RC2 MH089951). Part of the genotyping and analyses were funded
453 by the Genetic Association Information Network (GAIN) of the Foundation for the National
454 Institutes of Health. Computing was supported by BiG Grid, the Dutch e-Science Grid, which is
455 financially supported by NWO. **ORCADES** was supported by the Chief Scientist Office of the
456 Scottish Government, the Royal Society, the MRC Human Genetics Unit, Arthritis Research UK
457 and the European Union framework program 6 EUROSPAN project (contract no. LSHG-CT-
458 2006-018947). DNA extractions were performed at the Wellcome Trust Clinical Research
459 Facility in Edinburgh. We would like to acknowledge the invaluable contributions of Lorraine
460 Anderson and the research nurses in Orkney, the administrative team in Edinburgh and the
461 people of Orkney. **PROCARDIS** was supported by the European Community Sixth Framework
462 Program (LSHM-CT- 2007-037273), AstraZeneca, the British Heart Foundation, the Wellcome
463 Trust (Contract No. 075491/Z/04), the Swedish Research Council, the Knut and Alice
464 Wallenberg Foundation, the Swedish Heart-Lung Foundation, the Torsten and Ragnar Söderberg
465 Foundation, the Strategic Cardiovascular and Diabetes Programs of Karolinska Institutet and
466 Stockholm County Council, the Foundation for Strategic Research and the Stockholm County
467 Council. Jemma C Hopewell and Robert Clarke acknowledge support from the BHF Centre of
468 Research Excellence, Oxford. M.Sabater-Lleal is supported by the Swedish Heart-Lung
469 Foundation (20130399), and acknowledges funding from Åke Wiberg and Tore Nilssons
470 foundations. B.Sennblad acknowledges funding from the Magnus Bergvall Foundation and the

471 Foundation for Old Servants. **PROSPER-PHASE** received funding from the European Union's
472 Seventh Framework Programme (FP7/2007-2013) under grant agreement n° HEALTH-F2-2009-
473 223004. For a part of the genotyping we received funding from the Netherlands Consortium of
474 Healthy Aging (NGI: 05060810). Measurement of serum fibrinogen was supported by a grant
475 from the Scottish Executive Chief Scientist Office, Health Services Research Committee grant
476 number CZG/4/306. This work was performed as part of an ongoing collaboration of the
477 PROSPER study group in the universities of Leiden, Glasgow and Cork. Prof. Dr. J.W. Jukema
478 is an Established Clinical Investigator of the Netherlands Heart Foundation (2001 D 032). The
479 generation and management of GWAS genotype data for the **Rotterdam Study** is supported by
480 the Netherlands Organisation of Scientific Research NWO Investments (nr. 175.010.2005.011,
481 911-03-012). This study is funded by the Research Institute for Diseases in the Elderly (014-93-
482 015; RIDE2), the Netherlands Genomics Initiative (NGI)/Netherlands Organisation for Scientific
483 Research (NWO) project nr. 050-060-810. We thank Pascal Arp, Mila Jhamai, Marijn Verkerk,
484 Lizbeth Herrera, Marjolein Peters and Carolina Medina-Gomez for their help in creating the
485 GWAS database, and Karol Estrada and Carolina Medina-Gomez for the creation and analysis of
486 imputed data. The Rotterdam Study is funded by Erasmus Medical Center and Erasmus
487 University, Rotterdam, Netherlands Organization for the Health Research and Development
488 (ZonMw), the Research Institute for Diseases in the Elderly (RIDE), the Ministry of Education,
489 Culture and Science, the Ministry for Health, Welfare and Sports, the European Commission
490 (DG XII), and the Municipality of Rotterdam. Abbas Dehghan is supported by NWO grant (veni,
491 916.12.154) and the EUR Fellowship. The authors are grateful to the study participants, the staff
492 from the Rotterdam Study and the participating general practitioners and pharmacists. The
493 **SardiNIA** (''ProgeNIA'') team was supported by Contract NO1-AG-1–2109 from the NIA. We

thank the many individuals who generously participated in this study, the Mayors and citizens of the Sardinian towns involved, the head of the Public Health Unit ASL4, and the province of Ogliastra for their volunteerism and cooperation. In addition, we are grateful to the Mayor and the administration in Lanusei for providing and furnishing the clinic site. We are grateful to the physicians Angelo Scuteri, Marco Orrù, Maria Grazia Pilia, Liana Ferreli, Francesco Loi, nurses Paola Loi, Monica Lai and Anna Cau who carried out participant physical exams; the recruitment personnel Susanna Murino; Mariano Dei, Sandra Lai, Antonella Mulas, Andrea Maschio, Fabio Busonero for genotyping; Maria Grazia Piras and Monia Lobina for fibrinogen phenotyping. This research was supported by the Intramural Research Program of the NIH, National Institute on Aging, by Sardinian Autonomous Region (L.R. no. 7/2009) grant cRP3-154, and by grant FaReBio2011 “Farmaci e Reti Biotecnologiche di Qualità”. **SHIP** is part of the Community Medicine Research net of the University of Greifswald, Germany, which is funded by the Federal Ministry of Education and Research (grants no. 01ZZ9603, 01ZZ0103, and 01ZZ0403), the Ministry of Cultural Affairs as well as the Social Ministry of the Federal State of Mecklenburg - West Pomerania. Genome- wide data have been supported by the Federal Ministry of Education and Research (grant no. 03ZIK012) and a joint grant from Siemens Healthcare, Erlangen, Germany and the Federal State of Mecklenburg West Pomerania. Computing resources have been made available by the Leibniz Supercomputing Centre of the Bavarian Academy of Sciences and Humanities (HLRB project h1231). The University of Greifswald is a member of the 'Center of Knowledge Interchange' program of the Siemens AG and the Caché Campus program of the InterSystems GmbH. This work is also part of the research project Greifswald Approach to Individualized Medicine (GANI_MED). The GANI_MED consortium is funded by the Federal Ministry of Education and Research and the

Ministry of Cultural Affairs of the Federal State of Mecklenburg – West Pomerania (03IS2061A). **TwinsUK**. The study was funded by the Wellcome Trust; European Community's Seventh Framework Programme (FP7/2007-2013). The study also receives support from *the National Institute for Health Research (NIHR) Clinical Research Facility at Guy's & St Thomas' NHS Foundation Trust and NIHR Biomedical Research Centre based at Guy's and St Thomas' NHS Foundation Trust and King's College London.* Tim Spector is an NIHR senior Investigator and is holder of an ERC Advanced Principal Investigator award. SNP Genotyping was performed by The Wellcome Trust Sanger Institute and National Eye Institute via NIH/CIDR. The **WGHS** is supported by HL043851 and HL080467 from the National Heart, Lung, and Blood Institute and CA047988 from the National Cancer Institute, the Donald W. Reynolds Foundation and the Fondation Leducq, with collaborative scientific support and funding for genotyping provided by Amgen. The **WHI** program is funded by the National Heart, Lung, and Blood Institute, National Institutes of Health, U.S. Department of Health and Human Services through contracts HHSN268201100046C, HHSN268201100001C, HHSN268201100002C, HHSN268201100003C, HHSN268201100004C, and HHSN271201100004C. The authors thank the WHI investigators and staff for their dedication, and the study participants for making the program possible. A full listing of WHI investigators can be found at: http://www.whi.org/researchers/Documents%20%20Write%20a%20Paper/WHI%20Investigator%20Short%20List.pdf. Genotyping and analysis were supported through the Women's Health Initiative Sequencing Project (NHLBI RC2 HL-102924), the Genetics and Epidemiology of Colorectal Cancer Consortium (NCI CA137088), the Genomics and Randomized Trials Network (NHGRI U01-HG005152), and an NCI training grant (R25CA094880).

Conflicts of interest

Dr. BM Psaty serves on the DSMB for a clinical trial of a device funded by the manufacturer (Zoll LifeCor) and on the Steering Committee of the Yale Open Data Access Project funded by Johnson & Johnson. OH Franco works in ErasmusAGE, a center for aging research across the life course funded by Nestlé Nutrition (Nestec Ltd.), Metagenics Inc., and AXA. Nestlé Nutrition (Nestec Ltd.), Metagenics Inc., and AXA had no role in the design and conduct of the study, the collection, management, analysis, and interpretation of the data, nor in the preparation, review or approval of the manuscript.

References

1. Davalos, D. and Akassoglou, K. (2012) Fibrinogen as a key regulator of inflammation in disease. *Semin Immunopathol,* **34,** 43-62.
2. Seebacher, V., Polterauer, S., Grimm, C., Husslein, H., Leipold, H., Hefler-Frischmuth, K., Tempfer, C., Reinthaller, A. and Hefler, L. (2010) The prognostic value of plasma fibrinogen levels in patients with endometrial cancer: a multi-centre trial. *Br. J. Cancer,* **102,** 952-956.
3. Yapijakis, C., Bramos, A., Nixon, A.M., Ragos, V. and Vairaktaris, E. (2012) The interplay between hemostasis and malignancy: the oral cancer paradigm. *Anticancer Res.,* **32,** 1791-1800.
4. Emerging Risk Factors Collaboration, Kaptoge, S., Di Angelantonio, E., Pennells, L., Wood, A.M., White, I.R., Gao, P., Walker, M., Thompson, A., Sarwar, N. *et al.* (2012) C-reactive protein, fibrinogen, and cardiovascular disease prediction. *N. Engl. J. Med.,* **367,** 1310-1320.
5. Fibrinogen Studies Collaboration and Danesh, J. and Lewington, S. and Thompson, S.G. and Lowe, G.D. and Collins, R. and Kostis, J.B. and Wilson, A.C. and Folsom, A.R. and Wu, K. *et al.* (2005) Plasma fibrinogen level and the risk of major cardiovascular diseases and nonvascular mortality: an individual participant meta-analysis. *JAMA,* **294,** 1799-1809.
6. de Lange, M., Snieder, H., Ariens, R.A., Spector, T.D. and Grant, P.J. (2001) The genetics of haemostasis: a twin study. *Lancet,* **357,** 101-105.
7. Souto, J.C., Almasy, L., Borrell, M., Gari, M., Martinez, E., Mateo, J., Stone, W.H., Blangero, J. and Fontcuberta, J. (2000) Genetic determinants of hemostasis phenotypes in Spanish families. *Circulation,* **101,** 1546-1551.
8. Neijts, M., van Dongen, J., Kluft, C., Boomsma, D.I., Willemsen, G. and de Geus, E.J. (2013) Genetic architecture of the pro-inflammatory state in an extended twin-family design. *Twin Res Hum Genet,* **16,** 931-940.
9. Danik, J.S., Pare, G., Chasman, D.I., Zee, R.Y., Kwiatkowski, D.J., Parker, A., Miletich, J.P. and Ridker, P.M. (2009) Novel loci, including those related to Crohn disease, psoriasis, and inflammation, identified in a genome-wide association study of fibrinogen in 17 686 women: the Women's Genome Health Study. *Circ Cardiovasc Genet,* **2,** 134-141.
10. Dehghan, A., Yang, Q., Peters, A., Basu, S., Bis, J.C., Rudnicka, A.R., Kavousi, M., Chen, M.H., Baumert, J., Lowe, G.D. *et al.* (2009) Association of novel genetic Loci with circulating fibrinogen levels: a genome-wide association study in 6 population-based cohorts. *Circ Cardiovasc Genet,* **2,** 125-133.
11. Sabater-Lleal, M. and Huang, J. and Chasman, D. and Naitza, S. and Dehghan, A. and Johnson, A.D. and Teumer, A. and Reiner, A.P. and Folkersen, L. and Basu, S. *et al.* (2013) Multiethnic meta-analysis of genome-wide association studies in >100 000 subjects identifies 23 fibrinogen-associated Loci but no strong evidence of a causal association between circulating fibrinogen and cardiovascular disease. *Circulation,* **128,** 1310-1324.
12. Wassel, C.L., Lange, L.A., Keating, B.J., Taylor, K.C., Johnson, A.D., Palmer, C., Ho, L.A., Smith, N.L., Lange, E.M., Li, Y. *et al.* (2011) Association of genomic loci from a cardiovascular gene SNP array with fibrinogen levels in European Americans and African-Americans from six cohort studies: the Candidate Gene Association Resource (CARe). *Blood,* **117,** 268-275.
13. Baumert, J. and Huang, J. and McKnight, B. and Sabater-Lleal, M. and Steri, M. and Chu, A.Y. and Trompet, S. and Lopez, L.M. and Fornage, M. and Teumer, A. *et al.* (2014) No evidence for genome-wide interactions on plasma fibrinogen by smoking,

alcohol consumption and body mass index: results from meta-analyses of 80,607 subjects. *PLoS One,* **9,** e111156.

14. 1000 Genomes Project Consortium, Abecasis, G.R., Auton, A., Brooks, L.D., DePristo, M.A., Durbin, R.M., Handsaker, R.E., Kang, H.M., Marth, G.T. and McVean, G.A. (2012) An integrated map of genetic variation from 1,092 human genomes. *Nature,* **491,** 56-65.
15. Wise, A.L., Gyi, L. and Manolio, T.A. (2013) eXclusion: toward integrating the X chromosome in genome-wide association analyses. *Am J Hum Genet,* **92,** 643-647.
16. Mostowy, S., Bonazzi, M., Hamon, M.A., Tham, T.N., Mallet, A., Lelek, M., Gouin, E., Demangel, C., Brosch, R., Zimmer, C. *et al.* (2010) Entrapment of intracytosolic bacteria by septin cage-like structures. *Cell Host Microbe,* **8,** 433-444.
17. Schwefel, D., Frohlich, C., Eichhorst, J., Wiesner, B., Behlke, J., Aravind, L. and Daumke, O. (2010) Structural basis of oligomerization in septin-like GTPase of immunity-associated protein 2 (GIMAP2). *Proc. Natl. Acad. Sci. U. S. A.,* **107,** 20299-20304.
18. Ciucci, T. and Bosselut, R. (2014) Gimap and T cells: a matter of life or death. *Eur. J. Immunol.,* **44,** 348-351.
19. Zhang, L., Yang, Z., Ma, A., Qu, Y., Xia, S., Xu, D., Ge, C., Qiu, B., Xia, Q., Li, J. *et al.* (2014) Growth arrest and DNA damage 45G down-regulation contributes to Janus kinase/signal transducer and activator of transcription 3 activation and cellular senescence evasion in hepatocellular carcinoma. *Hepatology,* **59,** 178-189.
20. Dudka, A.A., Sweet, S.M. and Heath, J.K. (2010) Signal transducers and activators of transcription-3 binding to the fibroblast growth factor receptor is activated by receptor amplification. *Cancer Res.,* **70,** 3391-3401.
21. Hiscott, J., Pitha, P., Genin, P., Nguyen, H., Heylbroeck, C., Mamane, Y., Algarte, M. and Lin, R. (1999) Triggering the interferon response: the role of IRF-3 transcription factor. *J. Interferon Cytokine Res.,* **19,** 1-13.
22. Schindler, C., Levy, D.E. and Decker, T. (2007) JAK-STAT signaling: from interferons to cytokines. *J. Biol. Chem.,* **282,** 20059-20063.
23. Perez-Garcia, A., Ambesi-Impiombato, A., Hadler, M., Rigo, I., LeDuc, C.A., Kelly, K., Jalas, C., Paietta, E., Racevskis, J., Rowe, J.M. *et al.* (2013) Genetic loss of SH2B3 in acute lymphoblastic leukemia. *Blood,* **122,** 2425-2432.
24. Chang, Y.J., Chen, K.W., Chen, C.J., Lin, M.H., Sun, Y.J., Lee, J.L., Chiu, I.M. and Chen, L. (2014) SH2B1beta interacts with STAT3 and enhances fibroblast growth factor 1-induced gene expression during neuronal differentiation. *Mol. Cell. Biol.,* **34,** 1003-1019.
25. Huffman, J.E., de Vries, P.S., Morrison, A.C., Sabater-Lleal, M., Kacprowski, T., Auer, P.L., Brody, J.A., Chasman, D.I., Chen, M.H., Guo, X. *et al.* (2015) Rare and low-frequency variants and their association with plasma levels of fibrinogen, FVII, FVIII, and vWF. *Blood,* **126,** e19-29.
26. Ek, J., Rose, C.S., Jensen, D.P., Glumer, C., Borch-Johnsen, K., Jorgensen, T., Pedersen, O. and Hansen, T. (2005) The functional Thr130Ile and Val255Met polymorphisms of the hepatocyte nuclear factor-4alpha (HNF4A): gene associations with type 2 diabetes or altered beta-cell function among Danes. *J. Clin. Endocrinol. Metab.,* **90,** 3054-3059.
27. Dickson, S.P., Wang, K., Krantz, I., Hakonarson, H. and Goldstein, D.B. (2010) Rare variants create synthetic genome-wide associations. *PLoS Biol,* **8,** e1000294.
28. Wray, N.R., Purcell, S.M. and Visscher, P.M. (2011) Synthetic associations created by rare variants do not explain most GWAS results. *PLoS Biol,* **9,** e1000579.
29. Brennan, S.O., Fellowes, A.P., Faed, J.M. and George, P.M. (2000) Hypofibrinogenemia in an individual with 2 coding (gamma82 A-->G and Bbeta235 P-->L) and 2 noncoding mutations. *Blood,* **95,** 1709-1713.

30. Ivaskevicius, V., Jusciute, E., Steffens, M., Geisen, C., Hanfland, P., Wienker, T.F., Seifried, E. and Oldenburg, J. (2005) gammaAla82Gly represents a common fibrinogen gamma-chain variant in Caucasians. *Blood Coagul. Fibrinolysis,* **16,** 205-208.
31. Wyatt, J., Brennan, S.O., May, S. and George, P.M. (2000) Hypofibrinogenaemia with compound heterozygosity for two gamma chain mutations - gamma 82 Ala-->Gly and an intron two GT-->AT splice site mutation. *Thromb. Haemost.,* **84,** 449-452.
32. DIAbetes Genetics Replication and Meta-analysis (DIAGRAM) Consortium and Asian Genetic Epidemiology Network Type 2 Diabetes (AGEN-T2D) Consortium and South Asian Type 2 Diabetes (SAT2D) Consortium and Mexican American Type 2 Diabetes (MAT2D) Consortium and Type 2 Diabetes Genetic Exploration by Nex-generation sequencing in muylti-Ethnic Samples (T2D-GENES) Consortium and Mahajan, A. and Go, M.J. and Zhang, W. and Below, J.E. and Gaulton, K.J. *et al.* (2014) Genome-wide trans-ancestry meta-analysis provides insight into the genetic architecture of type 2 diabetes susceptibility. *Nat. Genet.,* **46,** 234-244.
33. Skol, A.D., Scott, L.J., Abecasis, G.R. and Boehnke, M. (2006) Joint analysis is more efficient than replication-based analysis for two-stage genome-wide association studies. *Nat. Genet.,* **38,** 209-213.
34. Skeppholm, M., Wallen, N.H., Blomback, M. and Kallner, A. (2008) Can both EDTA and citrate plasma samples be used in measurements of fibrinogen and C-reactive protein concentrations? *Clin. Chem. Lab. Med.,* **46,** 1175-1179.
35. Peters, S.A., Woodward, M., Rumley, A., Koenig, W., Tunstall-Pedoe, H. and Lowe, G.D. (2013) Direct comparisons of three alternative plasma fibrinogen assays with the von Clauss assay in prediction of cardiovascular disease and all-causes mortality: the Scottish Heart Health Extended Cohort. *Br. J. Haematol.,* **162,** 392-399.
36. Psaty, B.M., O'Donnell, C.J., Gudnason, V., Lunetta, K.L., Folsom, A.R., Rotter, J.I., Uitterlinden, A.G., Harris, T.B., Witteman, J.C., Boerwinkle, E. *et al.* (2009) Cohorts for Heart and Aging Research in Genomic Epidemiology (CHARGE) Consortium: Design of prospective meta-analyses of genome-wide association studies from 5 cohorts. *Circ Cardiovasc Genet,* **2,** 73-80.
37. Howie, B.N., Donnelly, P. and Marchini, J. (2009) A flexible and accurate genotype imputation method for the next generation of genome-wide association studies. *PLoS Genet,* **5,** e1000529.
38. Li, Y., Willer, C., Sanna, S. and Abecasis, G. (2009) Genotype imputation. *Annu Rev Genomics Hum Genet,* **10,** 387-406.
39. Li, Y., Willer, C.J., Ding, J., Scheet, P. and Abecasis, G.R. (2010) MaCH: using sequence and genotype data to estimate haplotypes and unobserved genotypes. *Genet. Epidemiol.,* **34,** 816-834.
40. Willer, C.J., Li, Y. and Abecasis, G.R. (2010) METAL: fast and efficient meta-analysis of genomewide association scans. *Bioinformatics,* **26,** 2190-2191.
41. Huang, J., Ellinghaus, D., Franke, A., Howie, B. and Li, Y. (2012) 1000 Genomes-based imputation identifies novel and refined associations for the Wellcome Trust Case Control Consortium phase 1 Data. *Eur. J. Hum. Genet.,* **20,** 801-805.
42. Wang, K., Li, M. and Hakonarson, H. (2010) ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data. *Nucleic Acids Res,* **38,** e164.
43. Magi, R., Lindgren, C.M. and Morris, A.P. (2010) Meta-analysis of sex-specific genome-wide association studies. *Genet. Epidemiol.,* **34,** 846-853.
44. Magi, R. and Morris, A.P. (2010) GWAMA: software for genome-wide association meta-analysis. *BMC Bioinformatics,* **11,** 288.
45. Gusev, A., Bhatia, G., Zaitlen, N., Vilhjalmsson, B.J., Diogo, D., Stahl, E.A., Gregersen, P.K., Worthington, J., Klareskog, L., Raychaudhuri, S. *et al.* (2013) Quantifying missing heritability at known GWAS loci. *PLoS Genet,* **9,** e1003993.

46. Yang, J., Ferreira, T., Morris, A.P., Medland, S.E., Genetic Investigation of, A.T.C., Replication, D.I.G., Meta-analysis, C., Madden, P.A., Heath, A.C., Martin, N.G. *et al.* (2012) Conditional and joint multiple-SNP analysis of GWAS summary statistics identifies additional variants influencing complex traits. *Nat. Genet.,* **44,** 369-375, S361-363.
47. Yang, J., Lee, S.H., Goddard, M.E. and Visscher, P.M. (2011) GCTA: a tool for genome-wide complex trait analysis. *Am J Hum Genet,* **88,** 76-82.
48. Hofman, A., Brusselle, G.G., Darwish Murad, S., van Duijn, C.M., Franco, O.H., Goedegebure, A., Ikram, M.A., Klaver, C.C., Nijsten, T.E., Peeters, R.P. *et al.* (2015) The Rotterdam Study: 2016 objectives and design update. *Eur. J. Epidemiol.,* **30,** 661-708.
49. Welter, D., MacArthur, J., Morales, J., Burdett, T., Hall, P., Junkins, H., Klemm, A., Flicek, P., Manolio, T., Hindorff, L. *et al.* (2014) The NHGRI GWAS Catalog, a curated resource of SNP-trait associations. *Nucleic Acids Res,* **42,** D1001-1006.
50. Westra, H.J., Peters, M.J., Esko, T., Yaghootkar, H., Schurmann, C., Kettunen, J., Christiansen, M.W., Fairfax, B.P., Schramm, K., Powell, J.E. *et al.* (2013) Systematic identification of trans eQTLs as putative drivers of known disease associations. *Nat. Genet.,* **45,** 1238-1243.
51. Ward, L.D. and Kellis, M. (2012) HaploReg: a resource for exploring chromatin states, conservation, and regulatory motif alterations within sets of genetically linked variants. *Nucleic Acids Res,* **40,** D930-934.
52. Encode Project Consortium (2012) An integrated encyclopedia of DNA elements in the human genome. *Nature,* **489,** 57-74.
53. Ghanbari, M., de Vries, P.S., de Looper, H., Peters, M.J., Schurmann, C., Yaghootkar, H., Dorr, M., Frayling, T.M., Uitterlinden, A.G., Hofman, A. *et al.* (2014) A genetic variant in the seed region of miR-4513 shows pleiotropic effects on lipid and glucose homeostasis, blood pressure, and coronary artery disease. *Hum. Mutat.,* **35,** 1524-1531.
54. Ghanbari, M., Franco, O.H., de Looper, H., Hofman, A., Erkeland, S. and Dehghan, A. (2015) Genetic Variations in miRNA Binding Sites Affect miRNA-Mediated Regulation of Several Genes Associated with Cardiometabolic Phenotypes. *Circ Cardiovasc Genet.*
55. Ghanbari, M., Sedaghat, S., de Looper, H.W., Hofman, A., Erkeland, S.J., Franco, O.H. and Dehghan, A. (2015) The association of common polymorphisms in miR-196a2 with waist to hip ratio and miR-1908 with serum lipid and glucose. *Obesity (Silver Spring),* **23,** 495-503.
56. Ridker, P.M., Chasman, D.I., Zee, R.Y., Parker, A., Rose, L., Cook, N.R., Buring, J.E. and Women's Genome Health Study Working, G. (2008) Rationale, design, and methodology of the Women's Genome Health Study: a genome-wide association study of more than 25,000 initially healthy american women. *Clin. Chem.,* **54,** 249-255.

Tables

Table 1: Association of the lead variants at 18 newly identified loci with natural log transformed plasma fibrinogen concentration (g/L).

Locus	Variant	Position	Closest Gene	eQTL	NSYN variants	A1/A2	Frequency	β	*P*-value
2p25.3	rs7588285	3648186	*COLEC11*			C/G	0.20	0.0074	1.2×10^{-08}
3p25.3	rs62246343	9543642	*LHFPL4*			T/C	0.17	0.0071	2.2×10^{-08}
3q21.1	rs1976714	122864771	*PDIA5*			T/G	0.35	-0.0055	2.3×10^{-08}
3q21.3	3:129228166	129228166	*IFT122*	*RPL32P3*		D/R	0.10	0.009	1.0×10^{-08}
7p14.2	rs2710804	36084529	*EEPD1*			C/T	0.37	0.0055	2.9×10^{-09}
7q36.1	7:150289652	150289652	*GIMAP4*	*GIMAP4*		D/R	0.21	-0.0073	9.3×10^{-11}
8p23.1	rs7012814	9173358	*LOC157273*			A/G	0.47	0.0060	2.1×10^{-10}
9q22.2	rs3138493	92219260	*GADD45G*	*SEMA4D*		T/C	0.48	-0.0054	2.5×10^{-09}
10q23.31	rs2250644	91008879	*LIPA*			T/C	0.33	0.0054	2.2×10^{-08}
10q26.13	rs2420915	122840277	*MIR5694*	*WDR11*		A/G	0.09	-0.0094	5.2×10^{-09}
11p12	rs7934094	43505707	*TTC17*			G/T	0.22	-0.0083	2.5×10^{-13}

12p12.1	12:21703935	21703935	*GYS2*			R/D	0.37	0.0062	8.4×10^{-09}
12q24.12	rs7310615	111865049	*SH2B3*	*SH2B3*	*SH2B3*	C/G	0.50	-0.0069	1.5×10^{-13}
15q15.1	rs56702977	42671308	*CAPN3*	*ZFP106*		A/G	0.13	0.0080	2.1×10^{-09}
16p11.2	16:28845027	28845027	*ATXN2L*	*TUFM*		D/R	0.39	0.0061	7.7×10^{-10}
16q22.2	rs1035560	72032730	*PKD1L3*	*HP*		C/T	0.40	0.0064	2.6×10^{-12}
17q21.2	rs7224737	40289364	*RAB5C*	*STAT3*	*HSPB9*	A/G	0.24	0.0061	6.1×10^{-09}
19q13.33	rs73058052	50099422	*PRR12*	*IRF3*	*PRRG2*	T/C	0.16	0.0074	2.0×10^{-08}

Abbreviations: eQTL indicates the gene with the strongest significant association between its expression levels in blood and the lead variant or its proxy. NSYN variants indicates genes containing nonsynonymous variant correlated to the lead variant ($R^2 > 0.9$). A1 indicates the coded allele. A2 indicates the other allele. Frequency is the frequency of the coded allele. β indicates the β coefficient adjusted for age, sex, population structure, and study-specific covariates, such as center or case/control status. The β coefficient can be interpreted as the ln(g/L) change in fibrinogen per 1 unit change in the dosage of the coded allele.

Table 2: Association of the lead variants at 23 known loci with natural log transformed plasma fibrinogen concentration (g/L).

Locus	Variant	Position	Closest Gene	eQTL	NSYN variants	A1/A2	Frequency	β	*P*-value
1p31.3	rs1892534	66105944	*LEPR*			T/C	0.38	-0.0073	4.3×10^{-15}
1q21.3	rs61812598	154420087	*IL6R*		*IL6R*	A/G	0.39	-0.0115	2.7×10^{-36}
1q44	rs10157379	247605599	*NLRP3*	*NLRP3*		C/T	0.38	-0.0103	6.3×10^{-29}
2q12	rs1558643	102731691	*IL1R1*			T/C	0.40	0.0058	3.1×10^{-10}
2q13	rs6734238	113841030	*IL1F10*	*IL1RN*		G/A	0.41	0.0106	6.7×10^{-30}
2q34	rs715	211543055	*CPS1*		*CPS1*	C/T	0.32	-0.0082	4.3×10^{-16}
2q37.3	rs59104589	242237902	*HDLBP*	*STK25*		T/C	0.34	-0.0083	8.2×10^{-19}
3q22.2	rs9840812	135843162	*PPP2R3A*	*PCCB*		C/T	0.23	0.0117	1.7×10^{-27}
4p16.3	rs59950280	3452345	*HGFAC*			A/G	0.34	0.0075	1.7×10^{-12}
4q31.3	rs7439150	155481541	*FGB*		*FBG*	A/G	0.20	0.0313	9.5×10^{-181}
5q31.1	rs2057655	131807624	*C5orf56*	*SLC22A4*		A/G	0.21	-0.0203	1.8×10^{-73}
7p21.1	7:17904452	17904452	*SNX13*			R/D	0.48	0.0067	1.3×10^{-13}

7p15.3	rs71520386	22853521	*TOMM7*			T/C	0.20	0.0066	5.1×10^{-09}
8q24.3	rs11780978	145034852	*PLEC*	*GRINA*		A/G	0.40	0.0059	5.5×10^{-10}
10q21.3	rs7916868	64988931	*JMJD1C*			A/T	0.49	0.0089	1.6×10^{-22}
11q12.2	rs11230201	59996994	*MS4A6A*	*MS4A6A*		G/C	0.41	-0.0057	4.5×10^{-10}
12q13.12	rs2731439	51060350	*DIP2B*	*DIP2B*		T/C	0.36	-0.0064	8.7×10^{-12}
14q24.1	rs367677	69273090	*ZFP36L1*			G/A	0.22	0.0077	1.8×10^{-12}
15q21.2	rs12913259	51014716	*SPPL2A*			T/C	0.30	-0.0068	2.3×10^{-12}
16q12.2	rs11859517	53181247	*CHD9*			T/C	0.29	-0.0074	8.9×10^{-14}
20q13.12	rs1800961	43042364	*HNF4A*		*HNF4A*	T/C	0.03	-0.0170	1.2×10^{-10}
21q22.2	rs9808651	40466468	*PSMG1*			A/G	0.27	-0.0095	2.5×10^{-20}
22q13.33	rs75347843	51112361	*SHANK3*	*ARSA*		A/G	0.19	0.0084	1.8×10^{-10}

Abbreviations: eQTL indicates the gene with the strongest significant association between its expression levels in blood and the lead variant or its proxy. NSYN variants indicates genes containing nonsynonymous variant correlated to the lead variant ($R^2 > 0.9$). A1 indicates the coded allele. A2 indicates the other allele. Frequency is the frequency of the coded allele. β indicates the β coefficient adjusted for age, sex, population structure, and study-specific covariates, such as center or case/control status. The β coefficient can be interpreted as the ln(g/L) change in

fibrinogen per 1 unit change in the dosage of the coded allele.

Table 3: Joint/conditional association of 8 variants at 2 loci with natural log transformed plasma fibrinogen concentration (g/L).

Locus	Variant	Position	Closest Gene	Annotation	A1/A2	Frequency	β	*P*-value	Joint β	Joint *P*-value
4q31.3	rs7439150	155481541	*FGB*	intergenic	A/G	0.205	0.0313	9.5×10^{-181}	0.0259	1.9×10^{-92}
4q31.3	rs150768229	155488301	*FGB*	intronic	C/A	0.009	-0.0458	6.4×10^{-12}	-0.0385	9.3×10^{-09}
4q31.3	rs6054	155489608	*FGB*	NSYN	T/C	0.005	-0.1228	2.4×10^{-53}	-0.1222	4.9×10^{-52}
4q31.3	rs148685782	155533035	*FGG*	NSYN	C/G	0.005	-0.2239	1.2×10^{-87}	-0.2179	4.0×10^{-82}
4q31.3	rs76289367	155546159	*FGG*	intergenic	G/T	0.148	0.0263	2.0×10^{-76}	0.0109	1.6×10^{-11}
5q31.1	rs12777	131671662	*SLC22A4*	SYN	G/C	0.044	0.0240	9.3×10^{-27}	0.0207	6.9×10^{-21}
5q31.1	5:131786964	131786964	*C5orf56*	ncRNA	I/R	0.015	-0.0543	2.5×10^{-14}	-0.0428	2.0×10^{-09}
5q31.1	rs2057655	131807624	*C5orf56*	ncRNA	A/G	0.207	-0.0203	1.8×10^{-73}	-0.0188	1.9×10^{-64}

Abbreviations: A1 indicates the coded allele. A2 indicates the other allele. Frequency is the frequency of the coded allele. NSYN indicates a nonsynonymous exonic variant. SYN indicates a synonymous exonic variant. β indicates the β coefficient adjusted for age, sex, population structure, and study-specific covariates, such as center or case/control status. Joint β indicates the β coefficient of the jointly significant variants, adjusted for the above and for each other. All β coefficients can be interpreted as the ln(g/L) change in fibrinogen per 1 unit change in the dosage of the coded allele.