Mapping of 79 loci for 83 plasma protein biomarkers in cardiovascular disease

Short title

Novel loci for the plasma proteome

Authors:

Lasse Folkersen[1,3], Eric Fauman[2], Maria Sabater-Lleal[3], Rona J Strawbridge[3], Mattias Frånberg[3], Bengt Sennblad[3], Damiano Baldassarre[4,5], Fabrizio Veglia[5], Steve E. Humphries[6], Rainer Rauramaa[7], Ulf de Faire[8], Andries J. Smit[9], Philippe Giral[10], Sudhir Kurl[11], Elmo Mannarino[12], Stefan Enroth[13], Åsa Johansson[13], Sofia Bosdotter Enroth[14], Stefan Gustafsson[15], Lars Lind[15], Cecilia Lindgren[16], Andrew P Morris[17], Vilmantas Giedraitis[16], Angela Silveira[3], Anders Franco-Cereceda[18], Elena Tremoli[4,5], the IMPROVE study group, Ulf Gyllensten[13], Erik Ingelsson[15,19], Søren Brunak[1], Per Eriksson[3], Daniel Ziemek[2], Anders Hamsten[3], Anders Mälarstig[3,20]

Affiliations:

1 Department of Systems Biology, Technical University of Denmark, Copenhagen, 2800, Denmark
2 Pfizer Worldwide Research & Development, Cambridge, MA, 02139, USA
3 Cardiovascular Medicine Unit, Department of Medicine Solna, Karolinska Institutet, Stockholm, 17176, Sweden
4 Dipartimento di Scienze Farmacologiche e Biomolecolari, Università di Milano, Milan, Italy.
5 Centro Cardiologico Monzino, IRCCS, Milan Italy.
6 British Heart Foundation Laboratories, University College of London, Department of Medicine, Rayne Building, London, WC1E 6BT, United Kingdom
7 Foundation for Research in Health Exercise and Nutrition, Kuopio Research Institute of Exercise Medicine, Kuopio, 70100, Finland.
8 Division of Cardiovascular Epidemiology, Institute of Environmental Medicine, Karolinska Institutet, and Department of Cardiology, Karolinska University Hospital, Solna, Karolinska Institutet, Stockholm, 17176, Sweden.
9 Department of Medicine, University Medical Center Groningen, Groningen, 30001, the Netherlands.
10 Assistance Publique - Hopitaux de Paris; Service Endocrinologie-Metabolisme, Groupe Hôpitalier Pitie-Salpetriere, Unités de Prévention Cardiovasculaire, Paris, 75013, France.
11 Institute of Public Health and Clinical Nutrition, University of Eastern Finland, Kuopio Campus, Kuopio, 80101, Finland
12 Internal Medicine, Angiology and Arteriosclerosis Diseases, Department of Clinical and Experimental Medicine, University of Perugia, Perugia, Italy.
13 Department of Immunology, Genetics and Pathology, Science for Life Laboratory Uppsala, Uppsala University, Uppsala, 75108, Sweden
14 Department of Internal Medicine, Uppsala University Hospital, 75185 Uppsala, Sweden
15 Department of Medical Sciences, Molecular Epidemiology and Science for Life Laboratory, Uppsala University, Uppsala, 75185, Sweden
16 Wellcome Trust Centre for Human Genetics, University of Oxford, Oxford, OX3 7BN, United Kingdom
17 Department of Biostatistics, University of Liverpool, Liverpool, L69 3BX, United Kingdom
18 Cardiothoracic Surgery Unit, Department of Molecular Medicine and Surgery, Karolinska Institutet,Stockholm, 17176, Sweden

19 Department of Medicine, Division of Cardiovascular Medicine, Stanford University School of Medicine, Stanford, CA 94305, USA.
20 Pfizer Worldwide Research and Development, Stockholm, 17176, Sweden

Name and complete address of corresponding author

Anders Mälarstig, Centre for Molecular Medicine, L8:02, 171 76 Stockholm, Sweden,

Fax number: + 46 8-31 31 47

Telephone: +46 8 55052514

Email: anders.malarstig@ki.se , anders.malarstig@pfizer.com ,

Abstract (max 300 words)

Recent advances in highly multiplexed immunoassays have allowed systematic large-scale measurement of hundreds of plasma proteins in large cohort studies. In combination with genotyping, such studies offer the prospect to 1) identify mechanisms involved with regulation of protein expression in plasma, and 2) determine whether the plasma proteins are likely to be causally implicated in disease. We report here the results of genome-wide association (GWA) studies of 83 proteins considered relevant to cardiovascular disease (CVD), measured in 3,394 individuals with multiple CVD risk factors. We identified 79 genome-wide significant ($p<5e-8$) association signals, 55 of which replicated at $P<0.0007$ in separate validation studies (n=2,639 individuals). Using automated text mining, manual curation, and network-based methods incorporating information on expression quantitative trait loci (eQTL), we propose plausible causal mechanisms for 25 trans-acting loci, including a potential post-translational regulation of stem cell factor by matrix metalloproteinase 9 and receptor-ligand pairs such as RANK-RANK ligand. Using public GWA study data, we further evaluate all 79 loci for their causal effect on coronary artery disease, and highlight several potentially causal associations. Overall, a majority of the plasma proteins studied showed evidence of regulation at the genetic level. Our results enable future studies of the causal architecture of human disease, which in turn should aid discovery of new drug targets.

Author Summary (150-200 words)

Several proteins that circulate in blood have been linked to cardiovascular disease through the use of classic epidemiology and correlation studies. If individuals with higher risk of disease have higher levels of a protein, the protein may be associated with disease. However, this does not necessarily mean that the protein causes disease; it may merely be an innocent bystander or a consequence of the disease process. To establish whether a protein causes disease, a genetic approach, insensitive to reverse causation, can be used. Instead of correlating the levels of the protein itself, gene variants that regulate the protein levels are used in the analysis. This approach requires prior knowledge of which genetic variants are linked to individual proteins. Therefore we completed a map of how common genetic variants affect the blood concentration levels of 83 proteins that have been implicated in cardiovascular disease. By using this map of cause-to-effect findings, we gained insights into the regulation of a majority of the proteins under study and how they relate to risk of coronary artery disease. This study provides a map of genetic regulation of important cardiovascular plasma proteins, insights into their upstream regulatory environment, as well as novel leads for cardiovascular drug development.

Introduction

Cardiovascular disease (CVD), especially coronary artery disease (CAD) is a leading cause of human morbidity and mortality. Data from the world health organization (WHO) showed that CVD caused approximately 17.5 million deaths in 2012, corresponding to 31% of all deaths globally. Of these 7.4 million were estimated to be due to coronary heart disease and 6.7 million to stroke [1].

Specific and mechanistically relevant biomarkers are important tools in risk prediction, disease diagnosis and successful development of new therapies [2]. Proteins in the circulation have been extensively explored as biomarkers across numerous disease conditions, not least because of the relative ease with which blood plasma and serum can be accessed, stored and analysed in observational studies and randomized controlled trials.

The usefulness of a plasma biomarker in disease prediction, or as surrogate endpoint in a clinical trial, depends on its specificity and sensitivity. These metrics reflect the relationship of the biomarker with a pre-specified disease endpoint, but are inherently influenced by biological factors such as the tissue expression, stability, regulation and variability of the biomarker. The genetic contribution to the variability of plasma biomarkers can be explored in genome-wide association (GWA) studies using single nucleotide polymorphisms (SNPs), and this approach has been applied to uncover numerous such relationships [3–5]. For distinct plasma biomarkers such as circulating proteins, the associations are also known as protein quantitative trait loci (pQTLs) [6–9].

Genetic loci for biomarkers and pQTLs have wide applicability in research. Firstly, pQTLs in trans can identify previously unknown regulatory pathways. Using trans-pQTLs to discover regulatory pathways is beneficial because it is based on in-vivo human observations that have well-established direction of causality, flowing from SNP to protein [7]. This approach has been extensively used in-vitro, for example in yeast studies [8], and the overall goal of such analysis is a deeper understanding of the regulatory check-points giving rise to a particularly biomarker concentration. For a biomarker that is causally involved in disease, e.g. low-density lipoprotein cholesterol (LDL-C), this is crucial knowledge as it allows targeting of upstream factors, e.g. HMG-CoA reductase.

Secondly, GWA study loci associated with circulating levels of plasma biomarkers that are predictive of disease risk enable evaluation of whether the biomarker association with disease is likely to be a causal relationship, using Mendelian randomization (MR). For example, although both c-reactive protein (CRP) and LDL-C predict risk of CVD and are lowered by treatment with statins, MR studies have concluded that plasma LDL-C is an aetiologically important factor, while plasma CRP is a biomarker that is not causally related to CVD [10,11]. Similarly, all efforts towards HDL-cholesterol lowering drugs have failed, consistent with MR results showing that SNPs affecting HDL-levels are unrelated to risk of CVD [12]. Based on these experiences of pharmacological treatment lowering the LDL-C concentration, one may suggest that a biomarker which is both predictive and causal provides a more attractive target for novel therapeutics. Numerous associations between biomarkers and disease have been described in the literature, but the potential causal involvement of these biomarkers has only been addressed for a limited number, partly due to a lack of robust genetic predictors for many plasma proteins.

In the present study, we analyzed 83 plasma proteins using the Olink ProSeek CVD array in 3,394 European subjects with at least 3 established CVD risk factors. The majority of these proteins are strong candidates for involvement in atherosclerosis, plaque rupture or thrombosis and many are upregulated in CVD patients compared to controls or predict future risk of CVD events, such as CAD.

The aims of the study were to i) identify genetic loci for circulating plasma proteins that have previously been connected with CVD, ii) explore the mechanisms underpinning novel loci by integrating genetics with other biological information and iii) apply the tools to test causality in CAD.

Results

Of 83 proteins selected for known involvement in vascular disease and inflammation [13] (supplementary table S3), we observed 79 SNP-trait associations, consisting of 78 SNPs and their associations with 56 proteins (figure 1 and table 1). Of the 79 associations, 41 were cis effects, where the index-SNP is within 500 kb of the gene encoding the measured plasma protein. The functional

effect at each of these 41 loci is likely to be a direct effect either on the sequence of the plasma protein or on regulatory variants proximal to the encoding gene. Additionally, we identified 38 trans effects, all acting over distances more than 100 MB or at different chromosomes from the gene encoding the associated protein. Both cis and trans findings represent new understanding of the direct regulation of candidate CVD proteins, with trans findings additionally providing an opportunity for new insight into regulatory pathways.

We could replicate all but 6 of the pQTLs at nominal significance ($P<0.05$) measured in three independent cohorts (n=2,639). All but 16 of the measureable pQTLs were found to be reproducible at $P<0.0007$ (Bonferroni corrected value). An additional 8 pQTLs were not measured in the replication cohorts. All 79 SNP-trait associations are reported in table 1 along with indication of replication status. Detailed replication statistics is available in supplementary table S1.

Protein quantitative trait loci acting in trans

For each of the reported trans associations, we evaluated the most likely cis-gene intermediary, and investigated pathways in the direction of the plasma protein (table 2). Analysis of coding proxies revealed that 10 trans loci had missense mutations in linkage disequilibrium (LD) with the index-SNPs, providing an obvious explanatory model for the mechanism of action.

The analysis of cis-eQTLs in 11 large eQTL data sets provided evidence for an additional 13 mediator cis-genes. Some of the findings were remarkably independent of tissue and cell-type, and showed concordant results in several of the 11 eQTL datasets under analysis, as indicated in table 2. At each locus with significant cis-eQTL association, we additionally investigated neighbouring eQTL and pQTL effects as LocusZoom plots (supplementary figure S2). In some cases, like rs4810479/KITLG, the index-SNP shows both the strongest association with KITLG and the strongest cis-gene association (PLTP in liver). However, cases also exist, like rs200373/CTSL1, where stronger eQTL effects for the candidate cis-gene intermediary exists from other SNPs, with low LD between the SNPs precluding straightforward interpretation. Further studies would be required to address this issue.

In pathway analysis using the String-database of protein interactions, an additional 6 trans-genes were highlighted as possible mediator genes through functional protein connections. The criterion in this analysis was that less than 5% of randomly re-wired networks had shorter distance, dictating simply that connections of length 1 from a cis-gene to the trait gene should be selected. Additionally, a more sophisticated weighted network analysis was performed where each path through the network was weighted by the strength of the (trans) eQTL of the index-SNP. The eQTL values were calculated using a large collection of eQTL databases with tissue and cells relevant to cardiovascular disease. Like in the unweighted network analysis permutation was used to determine significance threshold. Through this weighted network analysis approach we discovered 11 additional mediator candidates, examples being the rs61598054 -> *FOXO3* -> *AKT1* -> and the rs693918 -> *XDH* -> *TLR4* -> *IL18* that are illustrated in figure 2A and 2B.

Systematic literature mining suggested an additional 5 possible mediators. Co-occurrence in scientific abstracts can indicate real biological relationships that may be missing from the String network. Interestingly, across all trans-pQTL loci, the largest number of abstract co-occurrences was 626 for the receptor-ligand pair encoded by TNFSF11 and TNFRSF11B, a protein-protein interaction also reported in String-db.

The results of these five mediator-gene assigning approaches are summarised in table 2. While examples given above provide relatively clear indications of mechanism, more challenging cases do exist: the IL6-SNP rs10947260 for example gives evidence pointed at three candidate cis-mediator genes: *NOTCH4*, *AGER*, and *ATF6B*. *NOTCH4* and *ATF6B* were identified as containing FDR5% significant pathways to the destination IL6 (figure 2D). AGER was mentioned together with IL6 in 64 separate publications. *AGER* and *ATF6B* exhibit weak cis-eQTL effects with the index-SNP. All were found at some distance from the index-SNP (>180 kb). Further experimentation is required to establish the main mechanism in this case.

Pleiotropy of loci affecting protein levels

Inspection of potential pleiotropic effects of index SNPs on measured protein traits as described in Methods revealed 6 distinct candidate loci (supplemental figure S1). The *ABO* locus affecting *THBD*, *TEK*, *F3*, *PECAM1*, and *SELE* in our dataset and the *FUT2* locus affecting *MMP10*, *F3*, and *LGALS3* are well known for their pleiotropic effects [14]. Furthermore, all SNPs affecting BNP levels seem to impact NPPB levels as well indicating their effect on steps before cleavage of the precursor protein. NTproBNP is a prohormone with an inactive N-terminal part that is cleaved to produce the active BNP. However, because of its half-life NTproBNP is typically used as a prognostic biomarker. A locus within the *ZFPM2* gene seems to have a strong effect on *PDGFB*, *DDK1*, and, to a lesser extent, on *VEGFA*. Finally, the cluster of cis-acting variants in the *MMP1*, *MMP3*, and *MMP12* loci are not specific to only one of the proteins but seem to impact all three of the metalloproteinases in this genomic region.

Additionally, we investigated the known associations of the index-SNPs with a broad range of other phenotypes, as previously reported in literature (supplemental table S2)..

Associations between plasma proteins and cardiovascular risk

To assess a potential causal involvement of each protein in CAD, we calculated genetic risk scores from the publically available CARDIoGRAMplusC4D GWAS data with the aim to construct a more powerful genetic instrument for those markers for which there were multiple SNPs. First, a systematic look-up of all reported pQTL-SNPs was performed to test for association with CAD (table 3). Then, we further explored proteins with multiple independent loci by calculating pooled SNP scores per protein, thus creating more powerful instruments to analyze the causality for proteins with multiple SNPs. Results show that of the SNPs contributing to the concentrations of proteins (Table 1), eight were also significantly associated with risk of CAD (Table 3). These findings suggest a causal role for these proteins, and whilst the cis IL6R finding confirms previous observations [15], the other observations extend our knowledge of important factors in CVD. Results from pooled-scores include highlights such as the multi-SNP support of LGALS3 and the contradiction of CHI3L1 having a CAD-associated trans-effect but no CAD-association in the cis-loci.

Discussion

In this study, we identified 79 pQTLs by measuring 83 plasma proteins of cardiovascular interest in a cohort of 3,394 subjects with multiple risk CVD risk factors, which may increase the power to detect genetic variants associated with CAD-associated proteins. The study provided novel insights into 57 of the plasma proteins under investigation, including cis- and trans genetic regulation and effects of long-distance regulation networks and tentative evidence for causal involvement in CVD.To the best of our knowledge only a few of the findings were previously known; however reassuringly these replicated as expected: IL18/rs75649625 and rs4129267/IL6R [16], as well as AGER/sRAGE, CD40 and LGALS3 cis associations [14,17,18] and the rs8176741/TEK trans association [19], and the rs635634/SELE [20]

Insights into specific trans-effects

A proteomics GWA study provides an interesting opportunity for the study of trans-regulatory effects, because the trait is a well-defined biological entity. In some cases, the trans-pQTL investigating methods in table 2 converged on a very plausible candidate gene. For example, at the CCL4-rs62625034 locus the effector transcript is probably the *CCR5* gene, while at the TNFSF11-rs7813952 locus the effector transcript is likely the *TNFRSF11B* gene, two examples of known ligand-receptor pairs. Another example is the IL27-rs4905 variant, which sits within the *EBI3* gene. The *IL27* and *EBI3* genes encode the two subunits of the IL27 cytokine complex.

The effector transcript at the KITLG-rs4810479 locus may be *MMP9*, which encodes a metalloproteinase that cleaves the KITLG gene product, a membrane-bound stem cell factor [21]. Thus this trans pQTL may represent an example of genetic regulation via post-translational modification.

At a few loci, we found multiple lines of evidence suggesting different mediator genes at the same locus. This is not biologically impossible, nor is it uncommon in literature [22], but it does require more careful analysis. The challenge is illustrated by the IL6-SNP rs10947260, for which separate lines of evidence pointed to three candidate cis-mediator genes. As shown in figure 2D, a criticism

against concluding on the importance of a pathway to IL6 through the *CCND1* gene is that *NOTCH4* has many neighbours in the String-network, thereby increasing the risk of a spurious discovery.

While these examples seem specific, they illustrate challenges that have major consequences for the general interpretation of any genetic association results. Analyses such as these have driven the development of popular risk-gene assignment tools (e.g. [23]). Our findings illustrate the increased power of knowing a certain pathway destination through the use of pQTL.

Insights into potential causal involvement of the plasma proteins in CVD

The study provided an important opportunity to systematically test each of the plasma proteins for a potential causal role in CVD by investigating whether identified pQTLs also were associated with CAD risk. If an instrumental variable, e.g. a SNP or a set of SNPs, exclusively affects one factor, and also affects an overall phenotype, such as disease risk – then it may be deduced that the protein is causally involved in the development of this disease. According to this principle, eight proteins (PECAM1, SELE, F3, IL6R, CHI3L1, LGALS3, MMP12, and PDGFB) showed evidence of potentially causal involvement in CAD. The connection between IL6R and CAD has already been described [15], and several drug trials are underway to test whether an ILR6-inhibitor (tocilizumab) is effective in treatment of CAD (clinicaltrials.org). In light of this, the remaining proteins could be of interest as therapeutic targets.

There are some important limitations to the approach, as compared to a formal MR. A formal MR study requires that the genetic instrument is specific, is not in LD with other functional variants, and that there are no hidden population strata [24]. There is no reason to suspect that the second and third requirements were violated; the study was based on high-resolution imputation of cohorts that were ethnically homogeneous. Importantly, the specificity requirement was not always satisfied, weakening the findings for some proteins. This includes all the trans associations, as well as proteins for which pleiotropy was detected (supplemental figure S1 and table S2). In addition, association between

plasma protein concentrations per se and future CVD risk has not been carefully investigated for the majority of proteins included in the present study.

These limitations leave LGALS3, MMP12 and PDGFB as candidates for having a causal effect on CAD. Of the three SNPs affecting levels of LGALS3, rs1169306, rs7928577 and rs33988101 in trans, only the first two also contribute to CAD risk, resulting in a pooled CAD association P-value of P=1.46e-4. For MMP12 and PDGFB, the results are based on single SNPs showing associations with protein levels. Of the three, only MMP12 is a cis effect thereby strengthening the case for it being a specific MR instrument. These limitations notwithstanding, the map of pQTLs presented here, and in particular those acting in cis, should provide the means to systematically assess potential causal roles of these biomarkers in other common complex diseases. Additionally we highlight the online resource found at www.olink-improve.com where the data pQTL can be browsed in greater detail. This may in turn help to prioritise drug targets for development of disease-modifying therapies.

Conclusion

In conclusion, the main contributions of this paper are: i) identification of 79 pQTLs regulating important circulating cardiovascular plasma proteins, ii) novel evidence of the regulatory mechanisms underpinning at least half of these novel loci and iii) evidence of potential causal roles in CAD development for several plasma proteins. We believe that these three principal findings provide a strong contribution to the field of cardiovascular biomarkers and beyond.

Materials and Methods

The IMPROVE study

The IMPROVE study is a multicentre, observational study, which recruited 3,711 men and women aged between 55 to 79 years with at least three cardiovascular risk factors but without symptoms of CVD (previously described [25]). Serum and plasma from the study participants were collected at baseline, dispensed in polypropylene tubes and frozen at –80° C prior to shipment for centralized

biochemical analyses and biobanking at the Karolinska Institutet in Stockholm, Sweden. The study was approved by the ethics committees of the 7 centers of the discovery cohort [25] and by the Uppsala University ethics committee for the replication cohorts [26,27]. All participants gave written informed consent, and the study was conducted in accordance with the declaration of Helsinki.

Genotyping, quality control and imputation

DNA genotyping in the IMPROVE study was performed using the Illumina CardioMetabochip and Immunochip arrays. The combined SNP genotyping data from both platforms were merged and subjected to the following quality control (QC) using PLINK 1.7: SNPs were excluded for probe to genome mismatch, incorrect assignment of allelic variants in the array design, failed Hardy-Weinberg Equilibrium test at 1x10-6, call rate <95 % or failed Illumina genotype calling QC. Samples were excluded if they showed evidence of gender mismatch, abnormal inbreeding coefficient, failed cryptic relatedness test or had an overall sample call rate <95 %. After quality control, a total number of 3,394 subjects remained for analysis. Imputation was performed with IMPUTE 2.0 using the 1000 genomes version 5 as reference panel. The pre-imputation data set contained 244,814 SNPs and the post-imputation data set contained 5,270,624 SNPs

Plasma protein determinations and quality control

In total, there were 3,394 IMPROVE participants for whom quality controlled genotype information and plasma samples were available. Plasma concentrations were measured in baseline EDTA plasma samples using the ProSeek CVD array I (Olink Biosciences, Uppsala, Sweden), according to the standard protocol. The ProSeek method is based on the highly sensitive and specific proximity extension assay (PEA), which involves the binding of distinct polyclonal oligonucleotide-labelled antibodies to the target protein followed by quantification by real-time quantitative PCR [13]. In addition to the controls provided by Olink Biosciences, a pooled plasma control was included in all plates to enable further quality control (QC) such as calculation of variation coefficients. Prior to statistical analyses, we excluded individual assays with more than 20 % of samples below the lower detection limit and those with final inter-plate coefficients of variation above 25 %. After QC, a total number of 83 proteins out of the 92 remained for analysis. The native scale of Olink protein

measurements is log(2) but additional log(10) transformations were performed to ensure normally distributed variables. Validation of the OLINK method has been conducted [13], and the method has been used to validate previous findings obtained with established protein quantification methods [26,28].

Genome-wide quantitative trait locus discovery

Plasma protein readings were $\log_{10}$ transformed prior to analyses. Standardized residuals for each of the 83 plasma proteins were calculated using a linear model adjusting for age, sex, recruitment centre, protein analysis batch, smoking, diabetes and hypertension at baseline. To merge loci in table 1 and table S1, signals with R^2 higher than 0.1 and distance within 250 KB were omitted, retaining only the strongest signal in each block, referred to as the index SNP. The standardized residuals were used in a Wald-test in PLINK 1.9 to test association between genetic data and each plasma protein, using a significance threshold of $P < 5e\text{-}8$. All summary statistics can be downloaded at www.olink-improve.com [review pass-mail: *rev3@ki.se*].

Narrow-sense heritability for all proteins was calculated using Genome-Wide Complex Trait Analysis [29]. A genetic relationship matrix was calculated using all imputed autosomal SNPs with imputation scores above R2=0.9, less than 1 % missingness and allele frequency above 5 %, in total 2,617,215 SNPs. The variants were hard called before estimating the phenotypic variance explained by all autosomal SNPs using the restricted maximum likelihood analysis (REML).

Replication of pQTL effects

Replication studies of all pQTLs were performed in three community-based cohorts in which Olink array protein data and genotypes were available. These cohorts were the NSPHS [27], the Prospective Investigation of the Vasculature in Uppsala Seniors (PIVUS) and the Uppsala Longitudinal Study of Adult Men (ULSAM) [26], consisting of samples from 976, 933 and 730 participants, respectively. Statistics were calculated according to additive association models, and findings were matched either directly on imputed SNP-id (96% of cases) or using a proxy with $R^2 > 0.8$ linkage disequilibrium. Replication P-values were calculated using the METAL meta-analysis software (version 2011-03-25).

Expression quantitative trait analysis

For each index-SNP, cis- and trans-eQTL data were calculated from the following sources: aorta intima-media, aorta adventitia, liver, mammary artery, and heart from the ASAP study [30], monocytes and B-cells from the Fairfax et al study [31], and monocytes stimulated with LPS-2h, LPS-24h and interferon-2h from another Fairfax et al study [32]. Each of these 11 data sets had information from gene expression microarrays and genotyping microarrays as described in the respective references. The mean sample size was 223 with a range of 89-367. Data from genotyping microarrays were imputed using the MACH 1.0 algorithm with 1000 genomes CEU data as reference (mean rsq quality score 0.89) [33]. The strength of eQTL association was calculated using a linear additive model between log2-transformed expression value and numerically encoded genotype data. For cis-eQTL associations, un-corrected p-values from cis-eQTL were reported if the association was stronger than $P < 0.0005$ (corresponding to a false discovery rate (FDR) <5%). For all significant cis-eQTL associations, locusZoom plots were generated showing regional effect differences between eQTL and pQTL studies [34].

Network analysis

The network analysis was performed based on the String database network (version 10) [35], using all edges with a confidence score above 400. For all genes within 0.5 MB of an effect-SNP ("cis-genes"), the shortest path length was calculated between the cis-gene and the gene encoding the measured protein biomarker ("trait-gene") using the igraph package in R (version 1.0.1). This was done both with an unweighted version of the Stringdb-network as well as with a weighted version, wherein each gene along the path was weighted by the trans-eQTL strength calculated from the effect-SNP (scored as 1, except if $P_{eQTL} < 0.05$ which gave score 0.8, and if $P_{eQTL} < 0.005$, which gave score 0.6).

For both weighted and unweighted networks, significance of a path was calculated as the fraction of 1000 randomly permuted networks that obtained a shorter path length than the one tested. Random networks were generated using permutation of the original scores and random rewiring of the network using the *igraph* rewire function, as detailed in code repository http://github.com/lassefolkersen/olink-improve. Given our data, only paths of length 1, i.e. direct links in String-db, were significant at a

0.05 level in the unweighted case. For the weighted case, only paths of length 2 with an intermediate trans-eQTL gene reached significance. Paths were subsequently checked for biological plausibility.

Literature analysis

To support the assignment of potential causal genes in pQTLs, we mined the literature for topical co-occurrences of each gene in a pQTL (defined by a 500kb window) with its associated protein. The Pfizer-internal LitMS tool can provide such matches based on all PubMed abstracts, a large synonym dictionary and manually curated rules that limit findings to more relevant articles, e.g. those in which gene and protein occur in the abstract's title [36]. The system outputs the number of co-occurrences and underlying article references for each gene-protein input pair. We then reviewed the literature findings to assign the most plausible causal genes where possible.

Pleiotropy

To understand the specificity of all reported index-SNPs we inspected all index SNPs that had at least 2 associations with distinct proteins at $P<0.05 / (83* 79) =7.7e-6$. This cutoff reflects a conservative approach to the multiple testing burden for all identified index SNPs (79) with all tested protein traits (83). The resulting association matrix was then clustered and visualized based on the negative log10 of the p-values of association. For the clustering, we used a complete-linkage hierarchical clustering approach based on correlation coefficients as a metric. In addition, index-SNPs were investigated for other associations in publically available GWAS databases.

Calculation of genetic risk scores

To assess the effect on disease, the publicly available CARDIoGRAMplusC4D 1000G imputed data was interrogated [37]. The goal was to perform *in silico* analysis for every SNP that showed significant associations with any of the measured traits. For traits that had multiple associated SNPs, pooled scores per affected protein were calculated using the R-package *gtx* version 0.0.8. Specifically for the pooled risk scores, the alleles of each protein were encoded so that the coded allele was increasing CAD risk regardless of its protein concentration effect. This ensured that pooled effect sizes reflected uniform directionality on CAD risk.

399

400

Acknowledgments

LF holds a grant from Innovation Fund Denmark (145-2014-5). APM is a Wellcome Trust Senior Fellow in Basic Biomedical Science (grant number WT098017)**.** SEH holds a Chair funded by the British Heart Foundation (PG08/008) and by the National Institute for Health Research University College London Hospitals Biomedical Research Centre.

Authorship Contributions

LF, EF, MSL, RJS, DZ and AM analysed the main data and wrote the manuscript. SE, ÅJ, SBE, SG, LL, CL, AM, VG, UG, and EI provided and analysed the replication data. All other authors contributed to the collection of samples and provided critical revision for intellectual content.

Competing interests

Ulf Gyllensten and Stefan Enroth are authors on a patent application entitled ‘‘Determination and analysis of Biomarkers in clinical samples’’; United Kingdom Patent Application Nos. 1414913.2 and 1410956.5 (2014, Pending).

References

Reference List

1. Shanthi Mendis, Pekka Puska, Bo Norrving (2011) Global Atlas on cardiovascular disease prevention and control. World Health Organization ISBN 978-92-4-156437-3.: 3-18.

2. Gashaw I, Ellinghaus P, Sommer A, Asadullah K (2012) What makes a good drug target? Drug Discov Today 17 Suppl: S24-S30. S1359-6446(11)00435-1 [pii];10.1016/j.drudis.2011.12.008 [doi].

3. Welter D, MacArthur J, Morales J, Burdett T, Hall P, Junkins H, Klemm A, Flicek P, Manolio T, Hindorff L, Parkinson H (2014) The NHGRI GWAS Catalog, a curated resource of SNP-trait associations. Nucleic Acids Res 42: D1001-D1006. gkt1229 [pii];10.1093/nar/gkt1229 [doi].

4. Enroth S, Johansson A, Enroth SB, Gyllensten U (2014) Strong effects of genetic and lifestyle factors on biomarker variation and use of personalized cutoffs. Nat Commun 5: 4684. ncomms5684 [pii];10.1038/ncomms5684 [doi].

5. Enroth S, Bosdotter ES, Johansson A, Gyllensten U (2015) Effect of genetic and environmental factors on protein biomarkers for common non-communicable disease and use of personally normalized plasma protein profiles (PNPPP). Biomarkers 20: 355-364. 10.3109/1354750X.2015.1093546 [doi].

6. Kettunen J, Tukiainen T, Sarin AP, Ortega-Alonso A, Tikkanen E, Lyytikainen LP, Kangas AJ, Soininen P, Wurtz P, Silander K, Dick DM, Rose RJ, Savolainen MJ, Viikari J, Kahonen M, Lehtimaki T, Pietilainen KH, Inouye M, McCarthy MI, Jula A, Eriksson J, Raitakari OT, Salomaa V, Kaprio J, Jarvelin MR, Peltonen L, Perola M, Freimer NB, Ala-Korpela M, Palotie A, Ripatti S (2012) Genome-wide association study identifies multiple loci influencing human serum metabolite levels. Nat Genet 44: 269-276. ng.1073 [pii];10.1038/ng.1073 [doi].

7. Hause RJ, Stark AL, Antao NN, Gorsic LK, Chung SH, Brown CD, Wong SS, Gill DF, Myers JL, To LA, White KP, Dolan ME, Jones RB (2014) Identification and validation of genetic variants that influence transcription factor and cell signaling protein levels. Am J Hum Genet 95: 194-208. S0002-9297(14)00314-0 [pii];10.1016/j.ajhg.2014.07.005 [doi].

8. Foss EJ, Radulovic D, Shaffer SA, Goodlett DR, Kruglyak L, Bedalov A (2011) Genetic variation shapes protein networks mainly through non-transcriptional mechanisms. PLoS Biol 9: e1001144. 10.1371/journal.pbio.1001144 [doi];PBIOLOGY-D-11-00024 [pii].

9. Sun W, Kechris K, Jacobson S, Drummond MB, Hawkins GA, Yang J, Chen TH, Quibrera PM, Anderson W, Barr RG, Basta PV, Bleecker ER, Beaty T, Casaburi R, Castaldi P, Cho MH, Comellas A, Crapo JD, Criner G, Demeo D, Christenson SA, Couper DJ, Curtis JL, Doerschuk CM, Freeman CM, Gouskova NA, Han MK, Hanania NA, Hansel NN, Hersh CP, Hoffman EA, Kaner RJ, Kanner RE, Kleerup EC, Lutz S, Martinez FJ, Meyers DA, Peters SP, Regan EA, Rennard SI, Scholand MB, Silverman EK, Woodruff PG, O'Neal WK, Bowler RP (2016) Common Genetic Polymorphisms Influence Blood Biomarker Measurements in COPD. PLoS Genet 12: e1006011. 10.1371/journal.pgen.1006011 [doi];PGENETICS-D-15-02045 [pii].

10. Casas JP, Shah T, Cooper J, Hawe E, McMahon AD, Gaffney D, Packard CJ, O'Reilly DS, Juhan-Vague I, Yudkin JS, Tremoli E, Margaglione M, Di MG, Hamsten A, Kooistra T, Stephens JW, Hurel SJ, Livingstone S, Colhoun HM, Miller GJ, Bautista LE, Meade T, Sattar N, Humphries SE, Hingorani AD (2006) Insight into the nature of the CRP-coronary event association using Mendelian randomization. Int J Epidemiol 35: 922-931. dyl041 [pii];10.1093/ije/dyl041 [doi].

11. Wensley F, Gao P, Burgess S, Kaptoge S, Di AE, Shah T, Engert JC, Clarke R, Davey-Smith G, Nordestgaard BG, Saleheen D, Samani NJ, Sandhu M, Anand S, Pepys MB, Smeeth L, Whittaker J, Casas JP, Thompson SG, Hingorani AD, Danesh J (2011) Association between C reactive protein and coronary heart disease: mendelian randomisation analysis based on individual participant data. BMJ 342: d548.

12. Voight BF, Peloso GM, Orho-Melander M, Frikke-Schmidt R, Barbalic M, Jensen MK, Hindy G, Holm H, Ding EL, Johnson T, Schunkert H, Samani NJ, Clarke R, Hopewell JC, Thompson JF, Li M, Thorleifsson G, Newton-Cheh C, Musunuru K, Pirruccello JP, Saleheen D, Chen L, Stewart A, Schillert A, Thorsteinsdottir U, Thorgeirsson G, Anand S, Engert JC, Morgan T, Spertus J, Stoll M, Berger K, Martinelli N, Girelli D, McKeown PP, Patterson CC, Epstein SE, Devaney J, Burnett MS, Mooser V, Ripatti S, Surakka I, Nieminen MS, Sinisalo J, Lokki ML, Perola M, Havulinna A, de FU, Gigante B, Ingelsson E, Zeller T, Wild P, de Bakker PI, Klungel OH, Maitland-van der Zee AH, Peters BJ, de BA, Grobbee DE, Kamphuisen PW, Deneer VH, Elbers CC, Onland-Moret NC, Hofker MH, Wijmenga C, Verschuren WM, Boer JM, van der Schouw YT, Rasheed A, Frossard P, Demissie S, Willer C, Do R, Ordovas JM, Abecasis GR, Boehnke M, Mohlke KL, Daly MJ, Guiducci C, Burtt NP, Surti A, Gonzalez E, Purcell S, Gabriel S, Marrugat J, Peden J, Erdmann J, Diemert P, Willenborg C, Konig IR, Fischer M, Hengstenberg C, Ziegler A, Buysschaert I, Lambrechts D, Van de Werf F, Fox KA, El Mokhtari NE, Rubin D, Schrezenmeir J, Schreiber S, Schafer A, Danesh J, Blankenberg S, Roberts R, McPherson R, Watkins H, Hall AS, Overvad K, Rimm E, Boerwinkle E, Tybjaerg-Hansen A, Cupples LA, Reilly MP, Melander O, Mannucci PM, Ardissino D, Siscovick D, Elosua R, Stefansson K, O'Donnell CJ, Salomaa V, Rader DJ, Peltonen L, Schwartz SM, Altshuler D, Kathiresan S (2012) Plasma HDL cholesterol and risk of myocardial infarction: a mendelian randomisation study. Lancet 380: 572-580. S0140-6736(12)60312-2 [pii];10.1016/S0140-6736(12)60312-2 [doi].

13. Assarsson E, Lundberg M, Holmquist G, Bjorkesten J, Thorsen SB, Ekman D, Eriksson A, Rennel DE, Ohlsson S, Edfeldt G, Andersson AC, Lindstedt P, Stenvang J, Gullberg M, Fredriksson S (2014) Homogenous 96-plex PEA immunoassay exhibiting high sensitivity, specificity, and excellent scalability. PLoS One 9: e95192. 10.1371/journal.pone.0095192 [doi];PONE-D-13-52390 [pii].

14. de Boer RA, Verweij N, van Veldhuisen DJ, Westra HJ, Bakker SJ, Gansevoort RT, Muller Kobold AC, van Gilst WH, Franke L, Mateo L, I, van der Harst P (2012) A genome-wide association study of circulating galectin-3. PLoS One 7: e47385. 10.1371/journal.pone.0047385 [doi];PONE-D-12-19830 [pii].

15. Interleukin-6 Receptor Mendelian Randomisation Analysis (IL6R MR) Consortium (2012) The interleukin-6 receptor as a target for prevention of coronary heart disease: a mendelian randomisation analysis. Lancet 379: 1214-1224. S0140-6736(12)60110-X [pii];10.1016/S0140-6736(12)60110-X [doi].

16. Melzer D, Perry JR, Hernandez D, Corsi AM, Stevens K, Rafferty I, Lauretani F, Murray A, Gibbs JR, Paolisso G, Rafiq S, Simon-Sanchez J, Lango H, Scholz S, Weedon MN, Arepalli S, Rice N, Washecka N, Hurst A, Britton A, Henley W, van de Leemput J, Li R, Newman AB, Tranah G, Harris T, Panicker V, Dayan C, Bennett A, McCarthy MI, Ruokonen A, Jarvelin MR, Guralnik J, Bandinelli S, Frayling TM, Singleton A, Ferrucci L (2008) A genome-wide association study identifies protein quantitative trait loci (pQTLs). PLoS Genet 4: e1000072. 10.1371/journal.pgen.1000072 [doi].

17. Jiang DK, Ma XP, Yu H, Cao G, Ding DL, Chen H, Huang HX, Gao YZ, Wu XP, Long XD, Zhang H, Zhang Y, Gao Y, Chen TY, Ren WH, Zhang P, Shi Z, Jiang W, Wan B, Saiyin H, Yin J, Zhou YF, Zhai Y, Lu PX, Zhang H, Gu X, Tan A, Wang JB, Zuo XB, Sun LD, Liu JO, Yi Q, Mo Z, Zhou G, Liu Y, Sun J, Shugart YY, Zheng SL, Zhang XJ, Xu J, Yu L (2015) Genetic variants in five novel loci including CFB and CD40 predispose to chronic hepatitis B. Hepatology 62: 118-128. 10.1002/hep.27794 [doi].

18. Cheng DT, Kim DK, Cockayne DA, Belousov A, Bitter H, Cho MH, Duvoix A, Edwards LD, Lomas DA, Miller BE, Reynaert N, Tal-Singer R, Wouters EF, Agusti A, Fabbri LM, Rames A, Visvanathan S, Rennard SI, Jones P, Parmar H, MacNee W, Wolff G, Silverman EK, Mayer RJ, Pillai SG (2013) Systemic soluble receptor for advanced glycation endproducts is a biomarker of emphysema and associated with AGER genetic variants in patients with chronic obstructive pulmonary disease. Am J Respir Crit Care Med 188: 948-957. 10.1164/rccm.201302-0247OC [doi].

19. Lieb W, Chen MH, Larson MG, Safa R, Teumer A, Baumeister SE, Lin H, Smith HM, Koch M, Lorbeer R, Volker U, Nauck M, Volzke H, Wallaschofski H, Sawyer DB, Vasan RS (2015) Genome-wide association study for endothelial growth factors. Circ Cardiovasc Genet 8: 389-397. CIRCGENETICS.114.000597 [pii];10.1161/CIRCGENETICS.114.000597 [doi].

20. Qi L, Cornelis MC, Kraft P, Jensen M, van Dam RM, Sun Q, Girman CJ, Laurie CC, Mirel DB, Hunter DJ, Rimm E, Hu FB (2010) Genetic variants in ABO blood group region, plasma soluble E-selectin levels and risk of type 2 diabetes. Hum Mol Genet 19: 1856-1862. ddq057 [pii];10.1093/hmg/ddq057 [doi].

21. Oriss TB, Krishnamoorthy N, Raundhal M, Morse C, Chakraborty K, Khare A, Huff R, Ray P, Ray A (2014) Cutting Edge: MMP-9 inhibits IL-23p19 expression in dendritic cells by targeting membrane stem cell factor affecting lung IL-17 response. J Immunol 192: 5471-5475. jimmunol.1303183 [pii];10.4049/jimmunol.1303183 [doi].

22. Musunuru K, Strong A, Frank-Kamenetsky M, Lee NE, Ahfeldt T, Sachs KV, Li X, Li H, Kuperwasser N, Ruda VM, Pirruccello JP, Muchmore B, Prokunina-Olsson L, Hall JL, Schadt EE, Morales CR, Lund-Katz S, Phillips MC, Wong J, Cantley W, Racie T, Ejebe KG, Orho-Melander M, Melander O, Koteliansky V, Fitzgerald K, Krauss RM, Cowan CA, Kathiresan S, Rader DJ (2010) From noncoding variant to phenotype via SORT1 at the 1p13 cholesterol locus. Nature 466: 714-719. nature09266 [pii];10.1038/nature09266 [doi].

23. Pers TH, Karjalainen JM, Chan Y, Westra HJ, Wood AR, Yang J, Lui JC, Vedantam S, Gustafsson S, Esko T, Frayling T, Speliotes EK, Boehnke M, Raychaudhuri S, Fehrmann RS, Hirschhorn JN, Franke L (2015) Biological interpretation of genome-wide association studies using predicted gene functions. Nat Commun 6: 5890. ncomms6890 [pii];10.1038/ncomms6890 [doi].

24. Burgess S, Malarstig A (2013) Using Mendelian randomization to assess and develop clinical interventions: limitations and benefits. J Comp Eff Res 2: 209-212. 10.2217/cer.13.14 [doi].

25. Strawbridge RJ, Deleskog A, McLeod O, Folkersen L, Kavousi M, Gertow K, Baldassarre D, Veglia F, Leander K, Gigante B, Kauhanen J, Rauramaa R, Smit AJ, Mannarino E, Giral P, Dehghan A, Hofman A, Franco OH, Humphries SE, Tremoli E, de FU, Gustafsson S, Ostensson CG, Eriksson P, Ohrvik J, Hamsten A (2014) A serum 25-hydroxyvitamin D concentration-associated genetic variant in DHCR7 interacts with type 2 diabetes status to influence subclinical atherosclerosis (measured by carotid intima-media thickness). Diabetologia 57: 1159-1172. 10.1007/s00125-014-3215-y [doi].

26. Nowak C, Sundstrom J, Gustafsson S, Giedraitis V, Lind L, Ingelsson E, Fall T (2016) Protein Biomarkers for Insulin Resistance and Type 2 Diabetes Risk in Two Large Community Cohorts. Diabetes 65: 276-284. db15-0881 [pii];10.2337/db15-0881 [doi].

27. Igl W, Johansson A, Gyllensten U (2010) The Northern Swedish Population Health Study (NSPHS)--a paradigmatic study in a rural population combining community health and basic research. Rural Remote Health 10: 1363. 1363 [pii].

28. Isgren A, Jakobsson J, Palsson E, Ekman CJ, Johansson AG, Sellgren C, Blennow K, Zetterberg H, Landen M (2015) Increased cerebrospinal fluid interleukin-8 in bipolar disorder patients associated with lithium and antipsychotic treatment. Brain Behav Immun 43: 198-204. S0889-1591(14)00474-7 [pii];10.1016/j.bbi.2014.10.001 [doi].

29. Yang J, Lee SH, Goddard ME, Visscher PM (2011) GCTA: a tool for genome-wide complex trait analysis. Am J Hum Genet 88: 76-82. S0002-9297(10)00598-7 [pii];10.1016/j.ajhg.2010.11.011 [doi].

30. Folkersen L, Wagsater D, Paloschi V, Jackson V, Petrini J, Kurtovic S, Maleki S, Eriksson MJ, Caidahl K, Hamsten A, Michel JB, Liska J, Gabrielsen A, Franco-Cereceda A, Eriksson P (2011) Unraveling divergent gene expression profiles in bicuspid and tricuspid aortic valve patients with thoracic aortic dilatation: the ASAP study. Mol Med 17: 1365-1373. molmed.2011.00286 [pii];10.2119/molmed.2011.00286 [doi].

31. Fairfax BP, Makino S, Radhakrishnan J, Plant K, Leslie S, Dilthey A, Ellis P, Langford C, Vannberg FO, Knight JC (2012) Genetics of gene expression in primary immune cells identifies cell type-specific master regulators and roles of HLA alleles. Nat Genet 44: 502-510. ng.2205 [pii];10.1038/ng.2205 [doi].

32. Fairfax BP, Humburg P, Makino S, Naranbhai V, Wong D, Lau E, Jostins L, Plant K, Andrews R, McGee C, Knight JC (2014) Innate immune activity conditions the effect of regulatory variants upon monocyte gene expression. Science 343: 1246949. 343/6175/1246949 [pii];10.1126/science.1246949 [doi].

33. Li Y, Willer CJ, Ding J, Scheet P, Abecasis GR (2010) MaCH: using sequence and genotype data to estimate haplotypes and unobserved genotypes. Genet Epidemiol 34: 816-834. 10.1002/gepi.20533 [doi].

34. Pruim RJ, Welch RP, Sanna S, Teslovich TM, Chines PS, Gliedt TP, Boehnke M, Abecasis GR, Willer CJ (2010) LocusZoom: regional visualization of genome-wide association scan results. Bioinformatics 26: 2336-2337. btq419 [pii];10.1093/bioinformatics/btq419 [doi].

35. Szklarczyk D, Franceschini A, Wyder S, Forslund K, Heller D, Huerta-Cepas J, Simonovic M, Roth A, Santos A, Tsafou KP, Kuhn M, Bork P, Jensen LJ, von MC (2015) STRING v10: protein-protein interaction networks, integrated over the tree of life. Nucleic Acids Res 43: D447-D452. gku1003 [pii];10.1093/nar/gku1003 [doi].

36. Roberts P, Bichko D, Crawford M, Klatte M, Xiang X (2016) LitMS v3.0 (Literature Mining System) (unpublished).

37. 2015) A comprehensive 1000 Genomes-based genome-wide association meta-analysis of coronary artery disease. Nat Genet 47: 1121-1130. ng.3396 [pii];10.1038/ng.3396 [doi].

Tables

Table 1

SNP id	Trait	-log(P)	SNP id	Trait	-log(P)
Cis-acting loci			Trans-acting loci		
rs1580006‡	ADM	14.69	**rs184243355***	CCL3	7.65
rs2070600‡	AGER (RAGE)	9.52	**rs73062378**	CCL4	12.35
rs549596*	BNP	13.76	**rs62625034**	CCL4	40.51
rs2188974	CCL3	17.31	**rs28601761**	CHI3L1	8.3
rs6607368	CCL4	30.2	**rs200373**	CTSL1	8.37
rs1569723	CD40	48.52	**rs6993770**	DKK1	8.79
rs2153101	CHI3L1	107.13	**rs495828**	F3	9.34
rs17610659	CSF1	9.19	**rs200433550***	F3	9.25
rs35285321	CSTB	42.93	**rs1260326**	FST (Follistatin)	8.69
rs111693235	CTSD	25.69	**rs4672375**	GAL	10.15
rs670211	CX3CL1	11.13	**rs76519098†**	GDF15	9.95
rs74544699	CXCL1	11.88	**rs693918**	IL18	10.62
rs35186877	CXCL16	8.76	**rs7599125‡**	IL18	7.95
rs72650832	CXCL6	41.21	**rs35166255**	IL1RL1	8.93
rs982764	FAS	11.7	**rs11599750**	IL27	9.85
rs3195944	GDF15	7.65	**rs10947260†**	IL6	9.74
rs6555820	HAVCR1	86.89	**rs4810479**	KITLG	10.35
rs13236526	HSPB1	16.96	**rs7928577**	LGALS3	8.67
rs139879640*	IL16	61.53	**rs1169306‡**	LGALS3	8.19
rs75649625	IL18	20.84	**rs33988101‡**	LGALS3	8.45
rs1420101	IL1RL1	131.69	**rs12570111†**	MMP1	7.33
rs4905	IL27	79.93	**rs492602**	MMP10	8.11
rs4129267	IL6R	264.67	**rs12469459**	MUC16	44.15
rs62115757	KLK11	61.91	**rs61598054***	NGF	7.42
rs11667946	KLK6	14.47	**rs75416436†**	NGF	7.38
rs9323280	LGALS3	61.25	**rs6557662***	NPPB	7.83
rs471994	MMP1	34.63	**rs140000161**	PAPPA	9.84
rs17368659	MMP12	96.26	**rs16873402‡**	PDGFB	7.62
rs7946057	MMP3	107.92	**rs635634**	PECAM1	44.72
rs56378716	MPO	8.73	**rs117538444†**	PGF	8.18
rs35207557*	NPPB	24.59	**rs635634**	SELE (E-selectin)	219.02
rs880949‡	PGF	7.8	**rs8176741**	TEK	49.06
rs116661163	REN (Renin)	7.99	**rs8176693**	THBD	9.95
rs1969539	SPON1	21.82	**rs241771‡**	TNFRSF11B	9.22
rs79250370	TEK (TIE2)	12.71	**rs142552223**	TNFSF11 (TRANCE)	16.47
rs3176123	THBD	23.64	**rs7813952**	TNFSF11 (TRANCE)	15.67

rs6469811	TNFRSF11B (Osteprotegerin)	10.54	**rs35538083**†	XPNPEP2	7.51
rs76769120‡	TNFRSF1B (TRAIL)	10.87	**rs11150189‡**	XPNPEP2	13.16
rs344560	TNFSF14	17.53			
rs2050011*	XPNPEP2	67.62			
rs2271025	AGRP	8.63			

573

574 **Overview of pQTL associations**

575 More commonly used non-systematic names indicated in parenthesis for some proteins. * pQTL that
576 was not measured in replication cohorts, † pQTL that was measured in replication cohorts, but did not
577 replicate at P<0.05, ‡ pQTL that did not replicate at Bonferroni corrected value of P<0.0007. A more
578 detailed version of this table is found as supplemental table S1.

579

580

581

582

583

584

585

586

587 **Table 2**

trait-gene	SNP	cis-gene	Distance (kb)	Dist-rank	Coding-proxy	Cis-eQTL	Un-weighted-pathway	eQTL-weighted-pathway	Literature-score
CCL4	rs62625034	CCR5	0	1	rs62625034 (R^2=1)				59
CTSL1	rs200373	IFI30	0	1		Monocytes+LPS (P=2.6e-05), Monocytes+IFN (P=1e-04)			
		MAST3	24	5	rs8108738 (R^2=0.64)				
F3	rs495828	SURF6	43	2		Monocytes (P=2.9e-05), B-cells (P=3.4e-05)			
		MED22	53	3				Via PPARD (P=0.00321)	
FST	rs1260326	GCKR	0	1	rs1260326 (R^2=1)				
		KRTCAP3	62	4		B-cells (P=3.4e-08)			
GDF15	rs76519098	MAPK8	283	4			Yes	Yes, short	
IL18	rs693918	XDH	-231	3				Via TLR4 (P=0.00085)	
IL18	rs7599125	LTBP1	-311	3				Via TGFB2 (P=0.00321)	
		NLRC4	-371	5			Yes	Yes, short	
IL1RL1	rs35166255	TIRAP	137	4			Yes	Yes, short	
		RPUSD4	-220	8		Monocytes+IFN (P=0.00034)			
IL27**	rs11599750	CWF19L1	187	6		4 eQTL-sets show cis-eQTL effect**			
IL6*	rs10947260	BTNL2	0	1	rs60263670 (R^2=1)				
		NOTCH4	-181	6				Via CCND1 (P=0.00427)	
		AGER	-221	9					64
		ATF6B	-277	18				Via ATF3 (P=0.00349)	
KITLG	rs4810479	PLTP	-4	1		Liver (P=4.2e-09), B-cells (P=4.3e-07)			
		PCIF1	-18	3		Monocytes+IFN (P=5.4e-05)			
		ACOT8	-59	9		Monocytes+IFN (P=0.00021)			
		MMP9	-92	12			Yes	Yes, short	
LGALS3	rs7928577	TIRAP	63	3				Via IL6 (P=0.000463)	
		CDON	-295	9				Via CTNNB1 (P=0.00494)	
LGALS3	rs1169306	HNF1A	0	1	rs2464196 (R^2=0.71)				
		C12orf43	3	2		5 eQTL-sets show cis-eQTL effect			
LGALS3	rs33988101	RASIP1	6	2	rs2287922 (R^2=0.88)				
		FUT2	9	3	rs602662 (R^2=0.68)				
		FGF21	-41	6				Via EGFR (P=0.000853)	
		BCAT2	80	10				Via GAPDH (P=0.000584)	

MMP10	rs492602	FUT2	0	1	rs601338 (R^2=0.99)				
		RASIP1	17	3	rs2287922 (R^2=0.68)				
		PPP1R15A	-169	18				Via GADD45A (P=0.0045)	
		BAX	-252	26				Via TNF (P=0.00461)	
MUC16	rs12469459	GAL3ST2	0	1	rs12469459 (R^2=1)				
		D2HGDH	8	2		Monocytes (P=9.6e-06)			
NGF	rs61598054	FOXO3	-70	2				Via AKT1 (P=0.00376)	
PAPPA	rs140000161	PRG2	0	1		Monocytes+IFN (P=5.4e-06)	Yes	Yes, short	
PECAM1	rs635634	SURF6	43	2		B-cells (P=1.7e-05), Monocytes (P=3.3e-05)			
SELE	rs635634	SURF6	43	2		B-cells (P=1.7e-05), Monocytes (P=3.3e-05)			
		MED22	53	3				Via PPARD (P=0.00277)	
TEK	rs8176741	ABO	0	1	rs8176747 (R^2=0.98)				
		MED22	76	5				Via ALB (P=0.00266)	
		RPL7A	-84	6				Via UBC (P=0.000421)	
		GBGT1	-92	9				Via ALB (P=0.00266)	
THBD	rs8176693	ABO	0	1	rs8176746 (R^2=1)				
TNFSF11	rs7813952	TNFRSF11B	-159	3			Yes	Yes, short	626

Systematic analysis of potential mechanisms behind trans-pQTL associations.

For each of 41 SNPs that had an effect in trans, cis-genes within 500 kb were analysed using 5 different methods for evaluation of mediator cis-gene: 1) presence of non-synonymous coding SNP in LD with index SNP at R^2>0.65, 2) presence of FDR5% cis-eQTL effect, 3) presence of significant pathway to trait-gene shorter than 95% of randomly permuted pathways, 4) presence of eQTL-weighted pathway to trait-gene shorter than 95% of randomly permuted pathways and/or 5) literature matching score above 50. A total of 1618 SNP-cis-gene pairs were considered, but only pairs that satisfied at least one of the tests are shown. * Shown in figure 1B. ** Shown in figure 1A.

Table 3

SNP	Trait-protein	Cis / trans	$P_{protein}$	β_{CAD}	P_{CAD}	$\beta_{CAD\text{-}pool}$	$P_{CAD\text{-}pool}$
rs635634	PECAM1	trans	1.9E-45	0.08	*4.47E-11*		
rs635634	SELE	trans	9.6E-220	0.08	*4.47E-11*		
rs495828	F3	trans	4.5E-10	0.07	*1.29E-10*		
rs4129267	IL6R	cis	2.1E-265	0.05	*2.21E-07*		
rs28601761	CHI3L1	trans	5.1E-09	0.05	*1.00E-06*	0.03	*2.3E-05*
rs1169306	LGALS3	trans	6.5E-09	0.03	*5.69E-04*	0.02	*5.9E-05*
rs7928577	LGALS3	trans	2.2E-09	0.06	*1.28E-03*	0.02	*5.9E-05*
rs17368659	MMP12	cis	5.5E-97	0.05	*1.39E-03*		
rs16873402	PDGFB	trans	2.4E-08	0.03	*1.47E-03*		
rs6993770	DKK1	trans	1.6E-09	0.03	6.90E-03		
rs880949	PGF	cis	1.6E-08	0.02	2.00E-02	0.02	2.0E-02
rs17610659	CSF1	cis	6.5E-10	0.02	2.25E-02		
rs112579976	CCL4	trans	2.5E-13	0.05	3.03E-02		
rs9323280	LGALS3	cis	5.6e-62	0.02	3.20E-01	0.02	*5.9E-05*
rs2153101	CHI3L1	cis	7.5E-108	0.01	4.68E-01	0.03	*2.3E-05*
rs33988101	LGALS3	trans	3.6E-09	0.01	5.16E-01	0.02	*5.9E-05*

rs117538444	PGF	trans	6.5E-09	0.01	7.64E-01	0.02	2.0E-02

Association between pQTLs and coronary artery disease (CAD) risk.

Each SNP from supplemental table S1 was investigated in the CARDIoGRAMplusC4D data, and the P-values for the pQTL and CAD risk were extracted. An additional pooled analysis was performed in cases where one plasma protein had multiple pQTLs,. The table shows all pQTLs for which either a single-SNP or pooled CAD association had a P<0.05. P-values highlighted in italics indicate that the association was also significant after correction for multiple testing.

Figure Legends

Figure 1

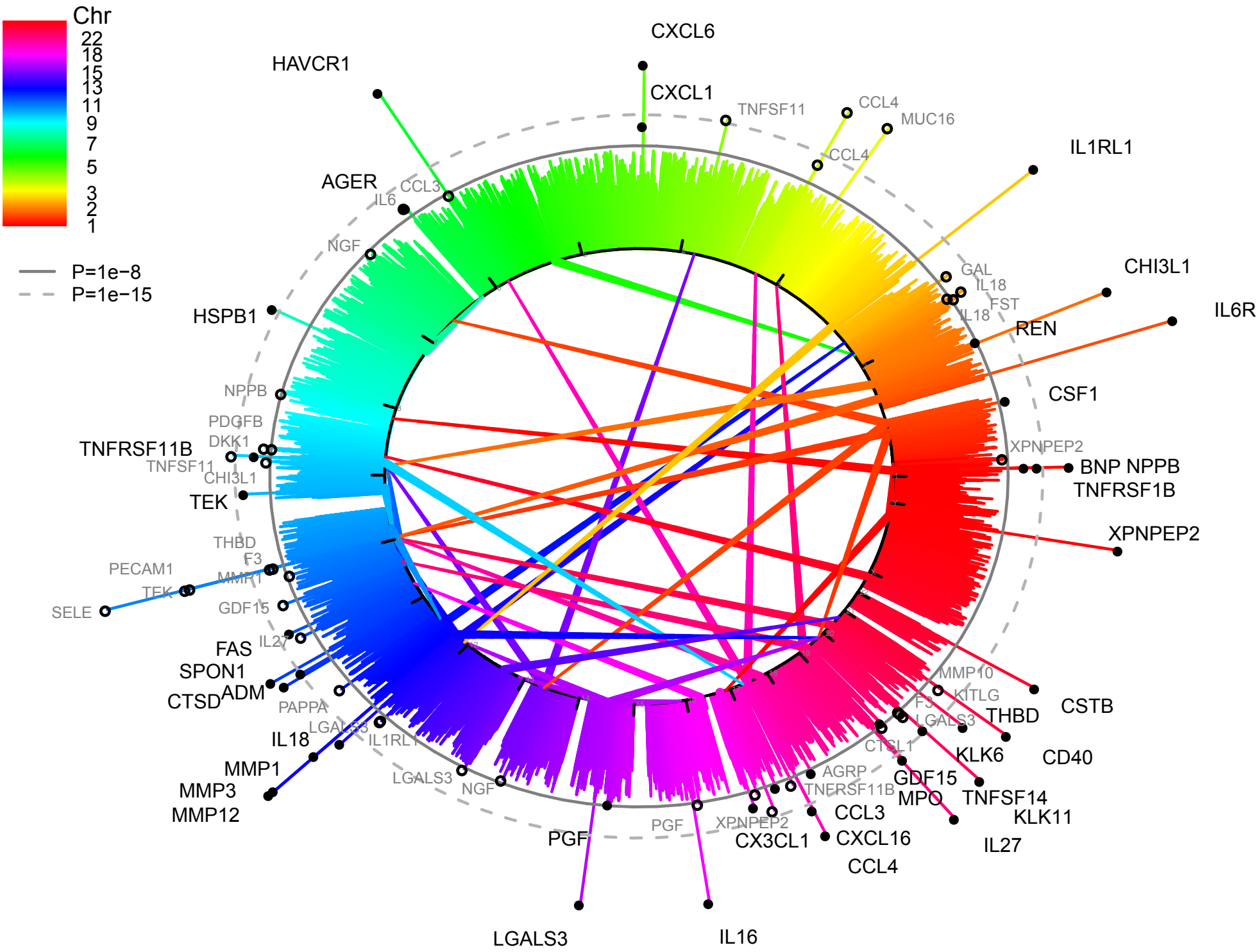
Genome-wide association strength of all measured plasma proteins.

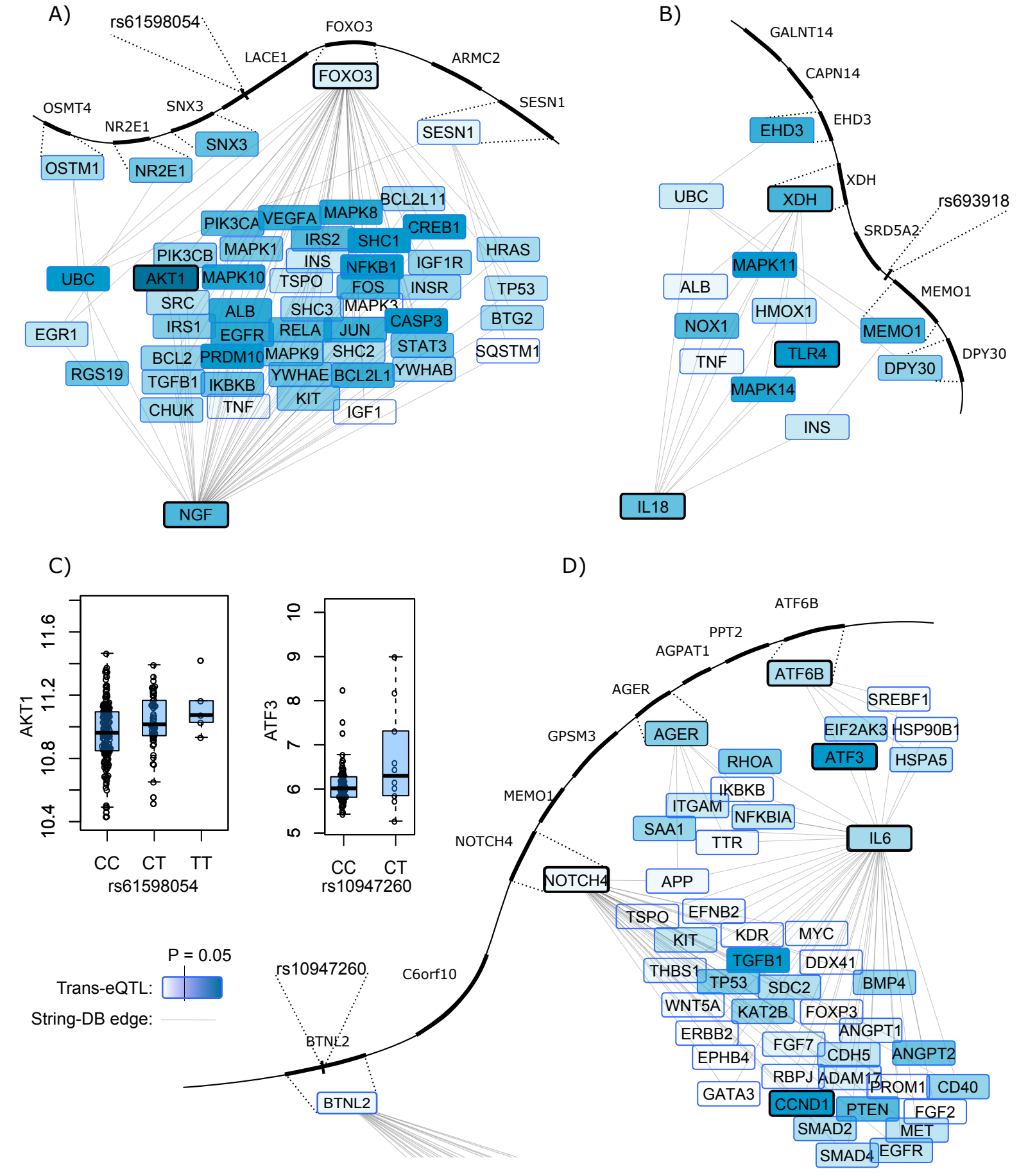
The extent of each stack indicates the negative log P of association between the plasma protein and SNPs. Stacks with black dots and black text labels indicate cis-associations. Stacks with hollow circles and grey text labels indicate trans-associations; their targets are indicated with central colour coded lines. Consequently, plasma proteins having both cis- and trans-effects can be identified as those with a black dot stack as well as connecting lines from hollow dots, e.g. XPNPEP2 or CCL4. Fully drawn circle shows P=5e-8. Dashed circle shows 1e-15. A detailed table of the genome-wide significant associations in this figure is available as supplemental table S1.

Figure 2

String-database network connections between proximal cis-gene and target plasma protein.

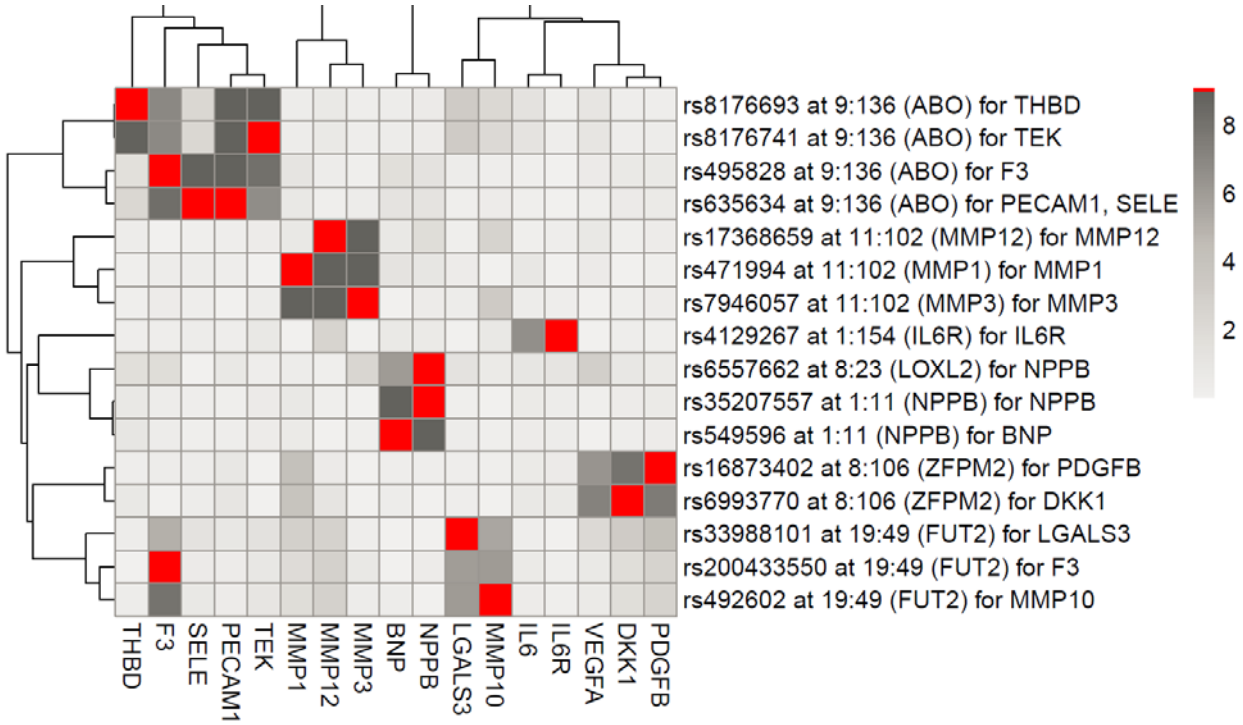
All short String paths that connect proximal cis-genes with the target plasma protein are shown. The colour intensity of each gene shows the eQTL association-strength with the index-SNP. The nodes highlighted with bold border show paths that satisfy P<0.05 in network permutation analysis. **A)** the rs61598054-SNP is harboured in an intron of the *LACE1* gene, but have no paths to the target gene *NGF* and a more likely mechanism is therefore *FOXO3 -> AKT1 -> NGF*, which involves a rs61598054-trans-eQTL effect on *AKT1*. In permutation analysis of re-wired networks this is stronger than 95% of random networks. **B)** Similarly for rs693918, while located between *SRD5A2* and *MEMO1*, the path *XDH -> TLR4 -> IL18* is a more likely mechanistic path, supported by eQTL effects on both *XDH* and *TLR4*. **C)** The rs61598054-*AKT1* trans-eQTL from panel A in 235 IFN-stimulated monocytes and the rs10947260-ATF3 trans-eQTL from panel D in 89 mammary artery samples. **D)** Although rs10947260 is found in an intron of *BTNL2*, no obvious path exists between *BTNL2* and *IL6*. However, both *ATF6B -> ATF3 -> IL6* and *NOTCH4 -> CCND1 -> IL6* are significant. DNA-strands are not to scale.


Chr
22
18
15
13
11
9
7
5
3
2
1
P=1e−8
P=1e−15
CXCL6
HAVCR1
CXCL1
TNFSF11
CCL4
MUC16
CCL4
IL1RL1
AGER
CCL3
IL6
NGF
GAL
IL18
FST
IL18
CHI3L1
IL6R
REN
HSPB1
NPPB
CSF1
PDGFB
DKK1
TNFRSF11B
TNFSF11
CHI3L1
TEK
XPNPEP2
BNP NPPB
TNFRSF1B
XPNPEP2
THBD
F3
PECAM1
MMP1
TEK
SELE
GDF15
IL27
FAS
SPON1
CTSD
ADM
PAPPA
LGALS3
IL18
IL1RL1
MMP1
MMP3
MMP12
LGALS3
NGF
PGF
PGF
XPNPEP2
CX3CL1
AGRP
TNFRSF11B
CCL3
CXCL16
CCL4
MMP10
KITLG
F3
LGALS3
THBD
CSTB
CTSL1
KLK6
CD40
GDF15
MPO
TNFSF14
KLK11
IL27
LGALS3
IL16


A)
rs61598054
FOXO3
LACE1
ARMC2
OSMT4
SNX3
NR2E1
SESN1
OSTM1
BCL2L11
PIK3CA
VEGFA
MAPK8
CREB1
IRS2
SHC1
PIK3CB
MAPK1
HRAS
INS
NFKB1
IGF1R
UBC
AKT1
MAPK10
TSPO
FOS
INSR
TP53
SRC
ALB
SHC3
MAPK3
CASP3
BTG2
EGR1
IRS1
EGFR
RELA
JUN
STAT3
SQSTM1
BCL2
PRDM10
MAPK9
SHC2
RGS19
YWHAE
BCL2L1
YWHAB
TGFB1
IKBKB
KIT
CHUK
TNF
IGF1
NGF
B)
GALNT14
CAPN14
EHD3
XDH
UBC
rs693918
SRD5A2
MAPK11
ALB
MEMO1
HMOX1
NOX1
TLR4
TNF
DPY30
MAPK14
INS
IL18
C)
AKT1
11.6
11.2
10.8
10.4
CC
CT
TT
rs61598054
ATF3
10
9
8
7
6
5
CC
CT
rs10947260
P = 0.05
Trans-eQTL:
String-DB edge:
rs10947260
C6orf10
BTNL2
BTNL2
D)
ATF6B
PPT2
AGPAT1
ATF6B
AGER
SREBF1
AGER
EIF2AK3
HSP90B1
GPSM3
ATF3
HSPA5
RHOA
IKBKB
MEMO1
ITGAM
NFKBIA
SAA1
IL6
NOTCH4
TTR
NOTCH4
APP
EFNB2
TSPO
KIT
KDR
MYC
TGFB1
DDX41
THBS1
TP53
SDC2
BMP4
WNT5A
KAT2B
FOXP3
ANGPT1
ERBB2
FGF7
CDH5
ANGPT2
EPHB4
RBPJ
ADAM17
PROM1
CD40
GATA3
CCND1
PTEN
FGF2
SMAD2
MET
SMAD4
EGFR

SUPPLEMENTAL MATERIAL

Figure S1



Potential pleiotropy between genome-wide significant SNPs and measured trait proteins. This figure shows all lead SNPs that have at least 2 associations with distinct proteins at P<0.05 / (83* 79) =7.7e-6. This cutoff reflects a conservative approach to the multiple testing burden for all identified lead SNPs (79) with all tested protein traits (83). Protein traits are not displayed if they have no associations with the selected SNPs at the defined threshold. Red colour indicates the main effect as reported in table S1. Grey-scale colours indicate the effect strength on a –log10(P) scale as indicated.

Figure S2

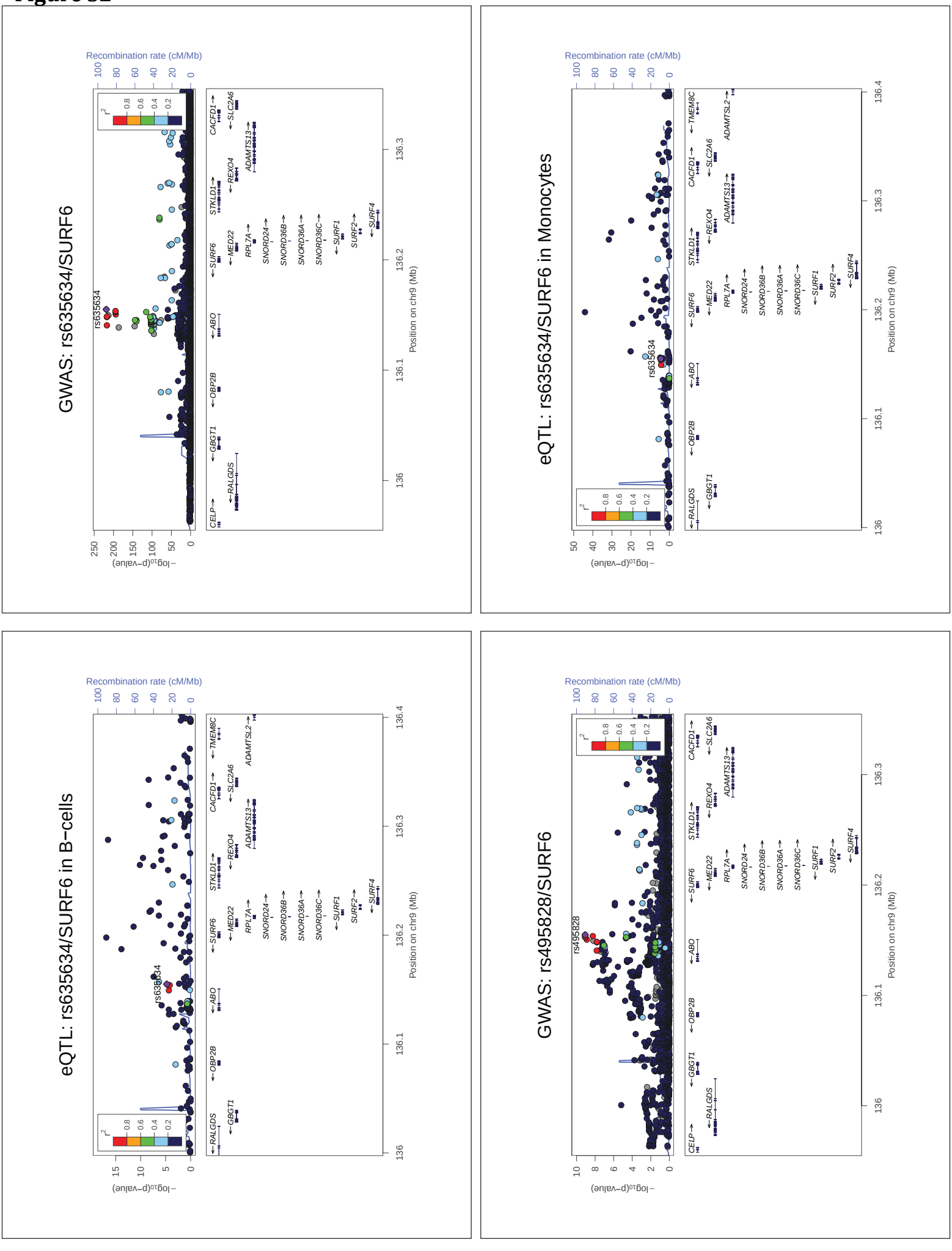

GWAS: rs635634/SURF6
eQTL: rs635634/SURF6 in Monocytes
eQTL: rs635634/SURF6 in B−cells
GWAS: rs495828/SURF6
Recombination rate (cM/Mb)
rs635634
rs495828
r²
−log10(p−value)
Position on chr9 (Mb)
CELP
RALGDS
GBGT1
OBP2B
ABO
SURF6
MED22
RPL7A
SNORD24
SNORD36B
SNORD36A
SNORD36C
SURF1
SURF2
SURF4
STKLD1
REXO4
ADAMTS13
CACFD1
SLC2A6
ADAMTSL2
TMEM8C

Figure S2

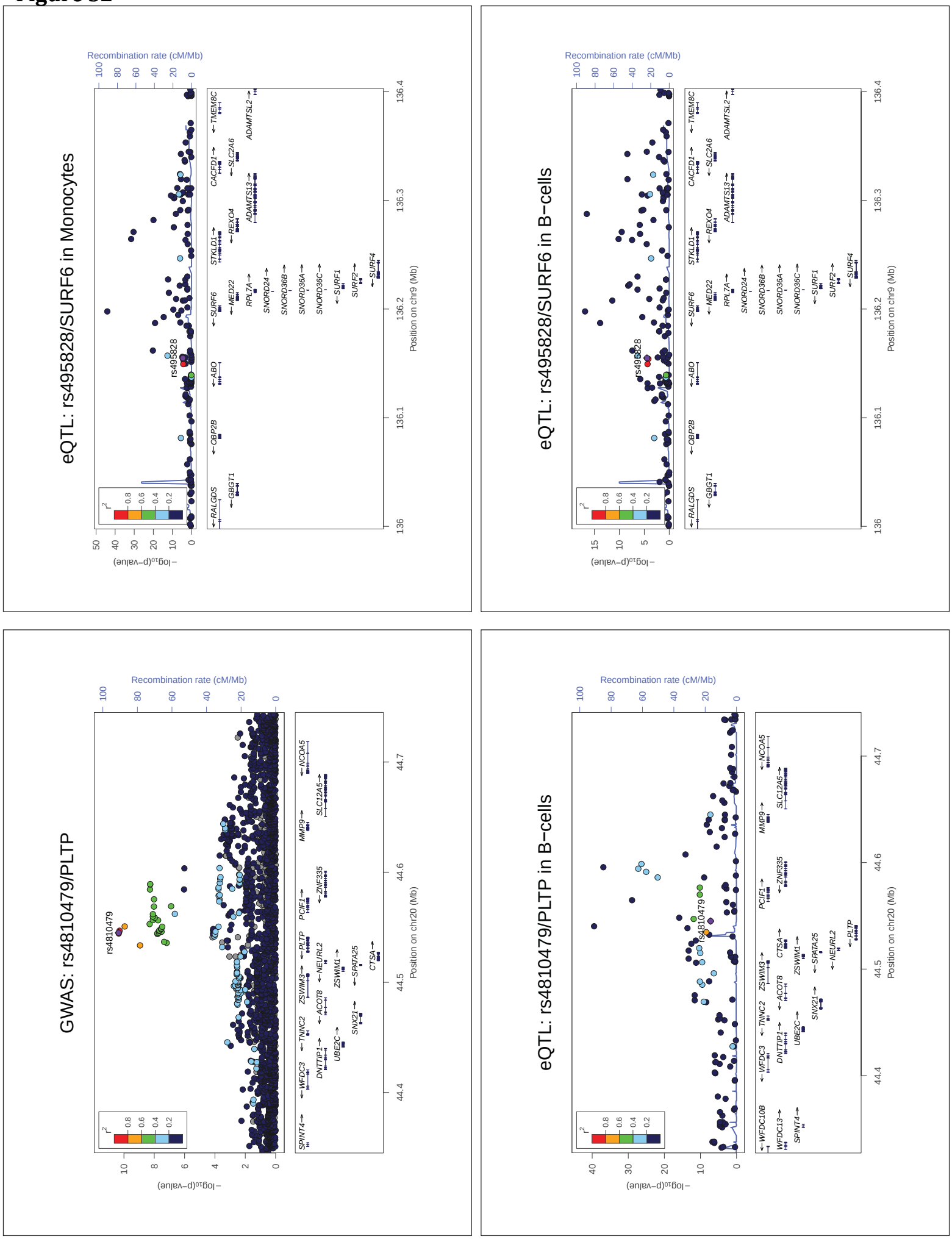

eQTL: rs495828/SURF6 in Monocytes
eQTL: rs495828/SURF6 in B−cells
GWAS: rs4810479/PLTP
eQTL: rs4810479/PLTP in B−cells
Recombination rate (cM/Mb)
rs495828
rs4810479
Position on chr9 (Mb)
Position on chr20 (Mb)

Figure S2

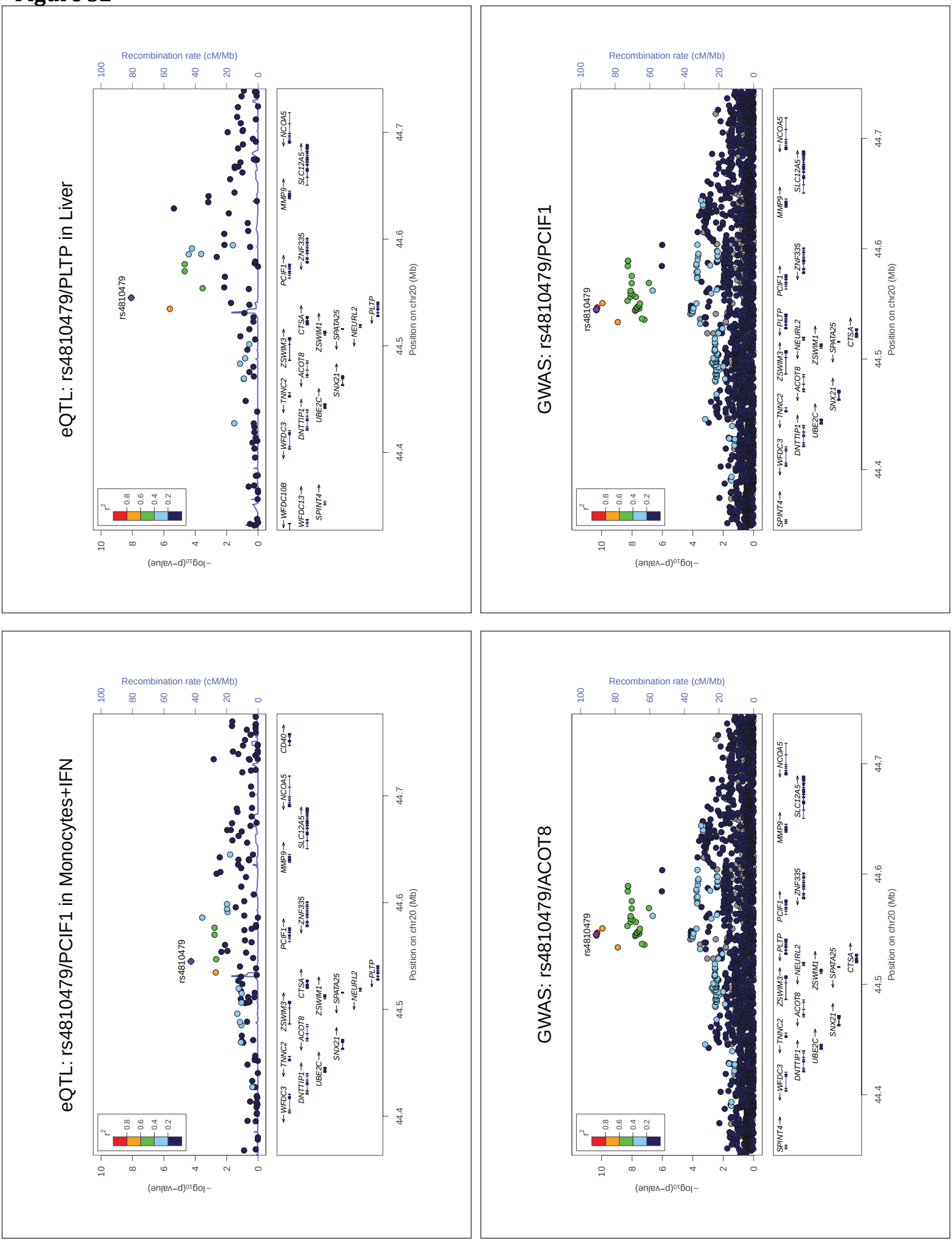
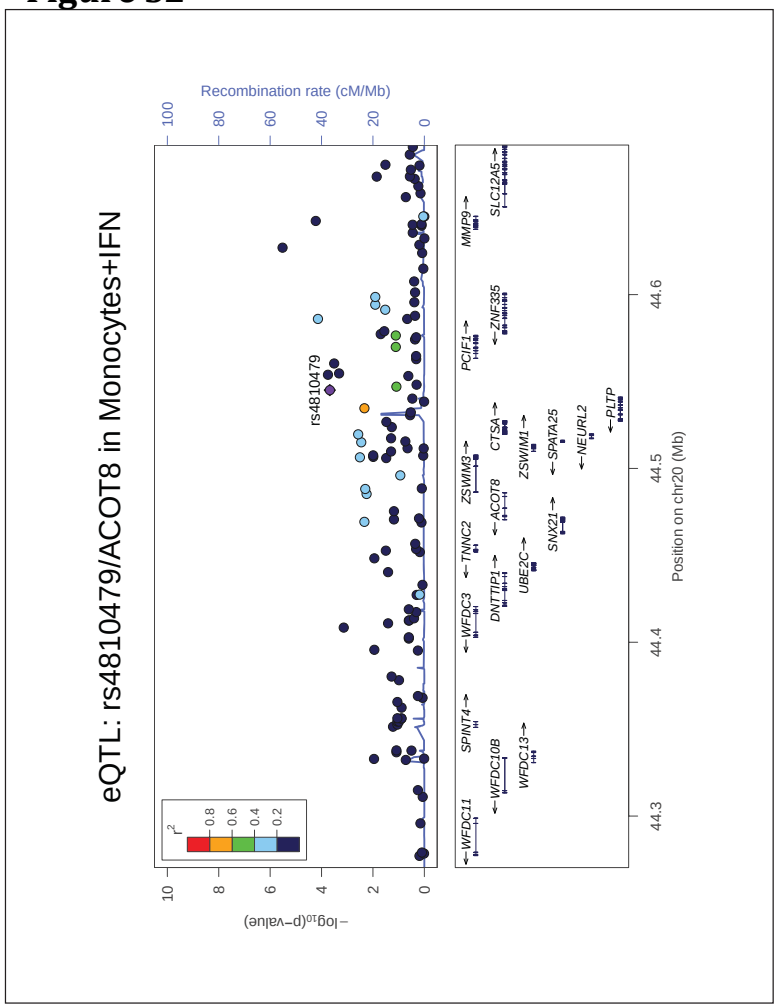
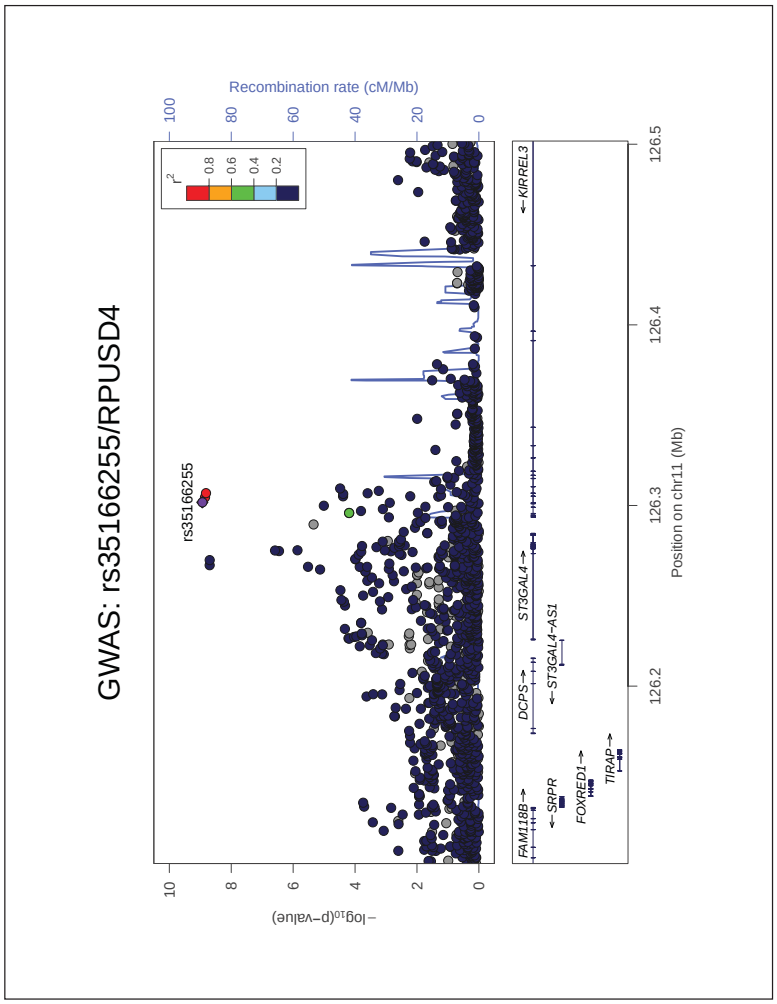
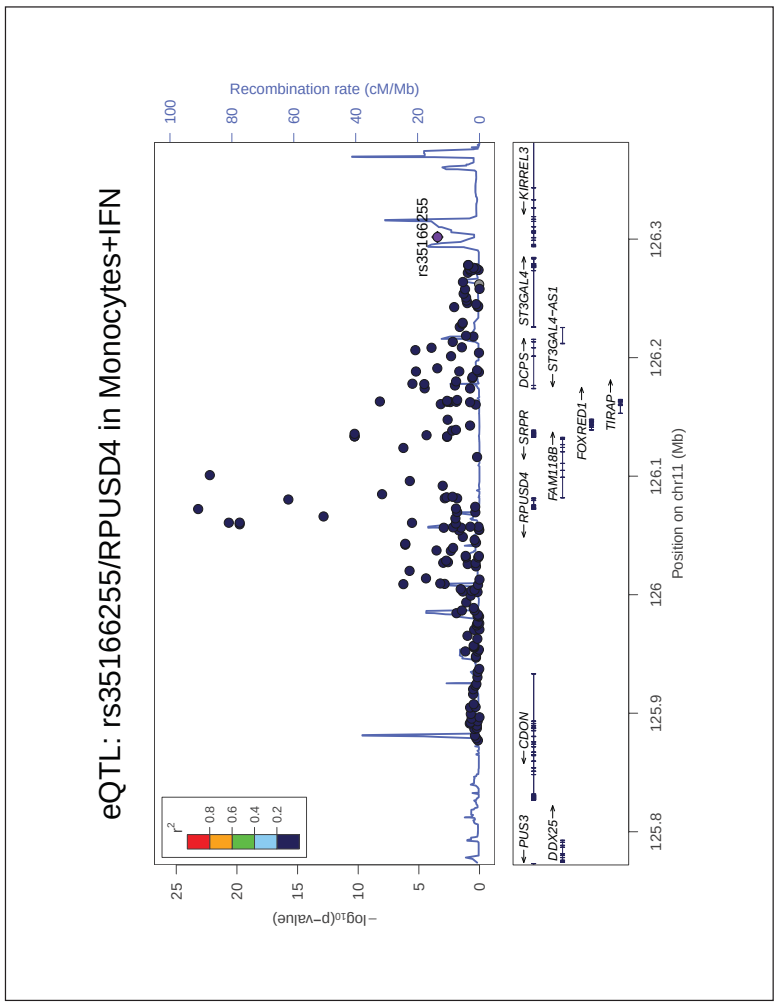

eQTL: rs4810479/PLTP in Liver
GWAS: rs4810479/PCIF1
eQTL: rs4810479/PCIF1 in Monocytes+IFN
GWAS: rs4810479/ACOT8
Recombination rate (cM/Mb)
rs4810479
−log10(p−value)
Position on chr20 (Mb)
r2
WFDC10B
WFDC13
SPINT4
WFDC3
TNNC2
ZSWIM3
DNTTIP1
ACOT8
CTSA
UBE2C
ZSWIM1
SNX21
SPATA25
NEURL2
PLTP
PCIF1
ZNF335
MMP9
SLC12A5
NCOA5
CD40

Figure S2


eQTL: rs4810479/ACOT8 in Monocytes+IFN
Recombination rate (cM/Mb)
rs4810479
r²
0.8
0.6
0.4
0.2
−log10(p−value)
WFDC11
SPINT4
WFDC3
TNNC2
ZSWIM3
PCIF1
MMP9
WFDC10B
DNTTIP1
ACOT8
CTSA
ZNF335
SLC12A5
WFDC13
UBE2C
ZSWIM1
SNX21
SPATA25
NEURL2
PLTP
44.3
44.4
44.5
44.6
Position on chr20 (Mb)


GWAS: rs35166255/RPUSD4
Recombination rate (cM/Mb)
rs35166255
r²
0.8
0.6
0.4
0.2
−log10(p−value)
FAM118B
DCPS
ST3GAL4
KIRREL3
SRPR
ST3GAL4-AS1
FOXRED1
TIRAP
126.2
126.3
126.4
126.5
Position on chr11 (Mb)


eQTL: rs35166255/RPUSD4 in Monocytes+IFN
Recombination rate (cM/Mb)
rs35166255
r²
0.8
0.6
0.4
0.2
−log10(p−value)
PUS3
CDON
RPUSD4
SRPR
DCPS
ST3GAL4
KIRREL3
DDX25
FAM118B
ST3GAL4-AS1
FOXRED1
TIRAP
125.8
125.9
126
126.1
126.2
126.3
Position on chr11 (Mb)

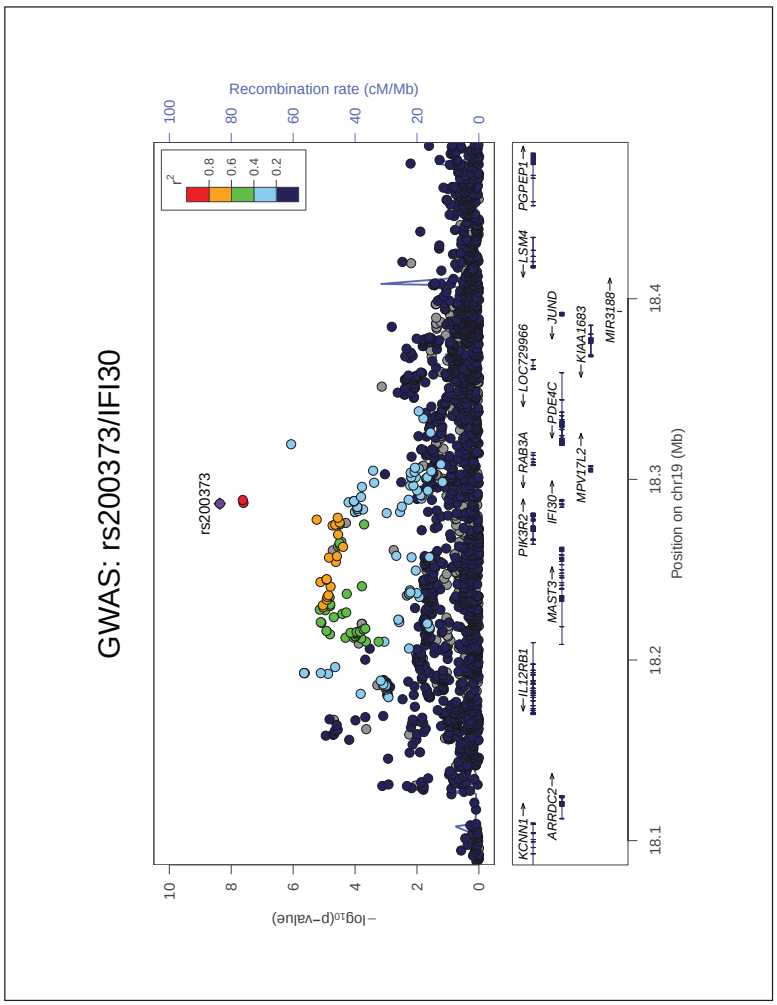

GWAS: rs200373/IFI30
Recombination rate (cM/Mb)
rs200373
r²
0.8
0.6
0.4
0.2
−log10(p−value)
KCNN1
IL12RB1
PIK3R2
RAB3A
LOC729966
LSM4
PGPEP1
ARRDC2
MAST3
IFI30
PDE4C
JUND
MPV17L2
KIAA1683
MIR3188
18.1
18.2
18.3
18.4
Position on chr19 (Mb)

Figure S2

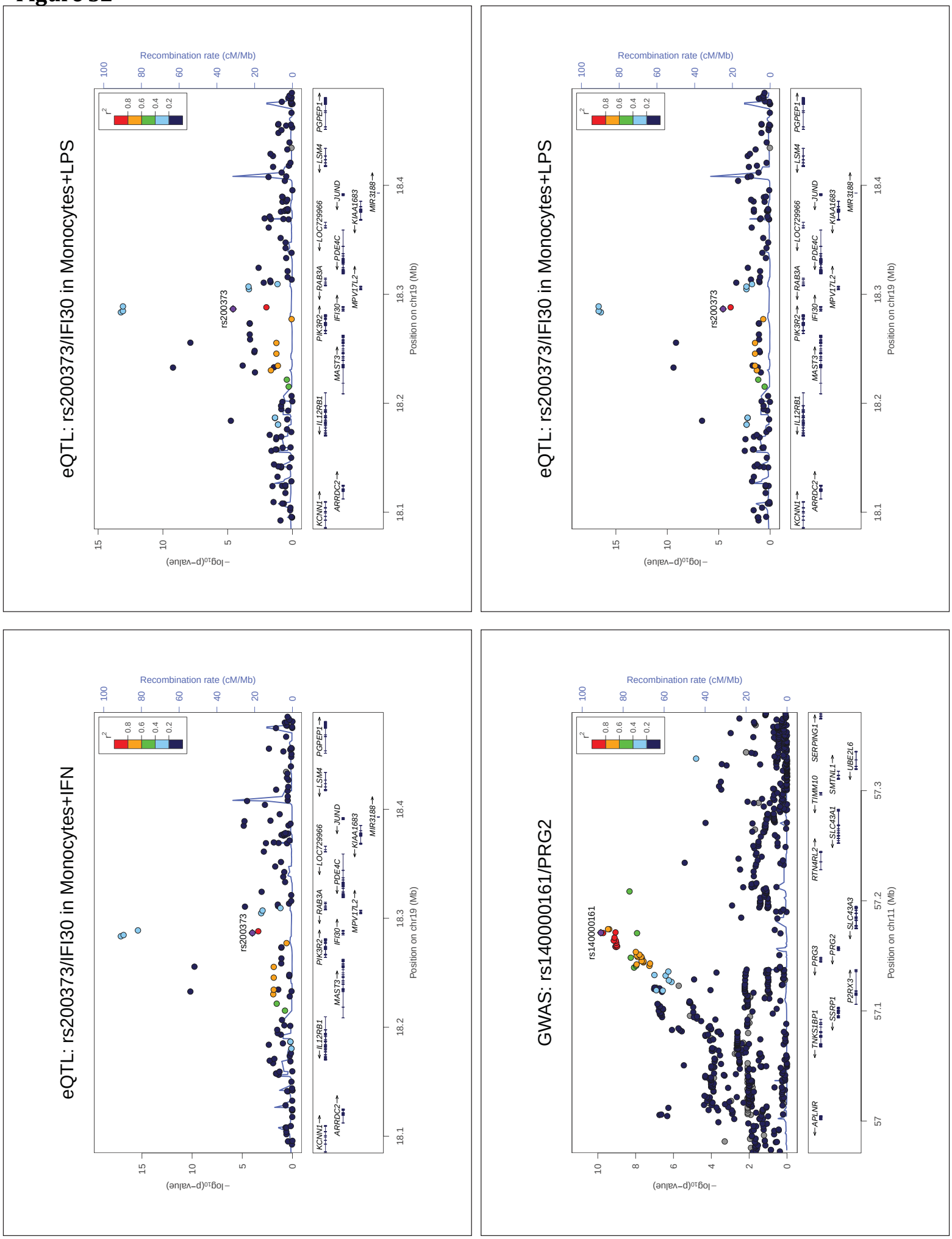

eQTL: rs200373/IFI30 in Monocytes+LPS
eQTL: rs200373/IFI30 in Monocytes+IFN
GWAS: rs140000161/PRG2
Recombination rate (cM/Mb)
rs200373
rs140000161
−log10(p−value)
Position on chr19 (Mb)
Position on chr11 (Mb)

Figure S2

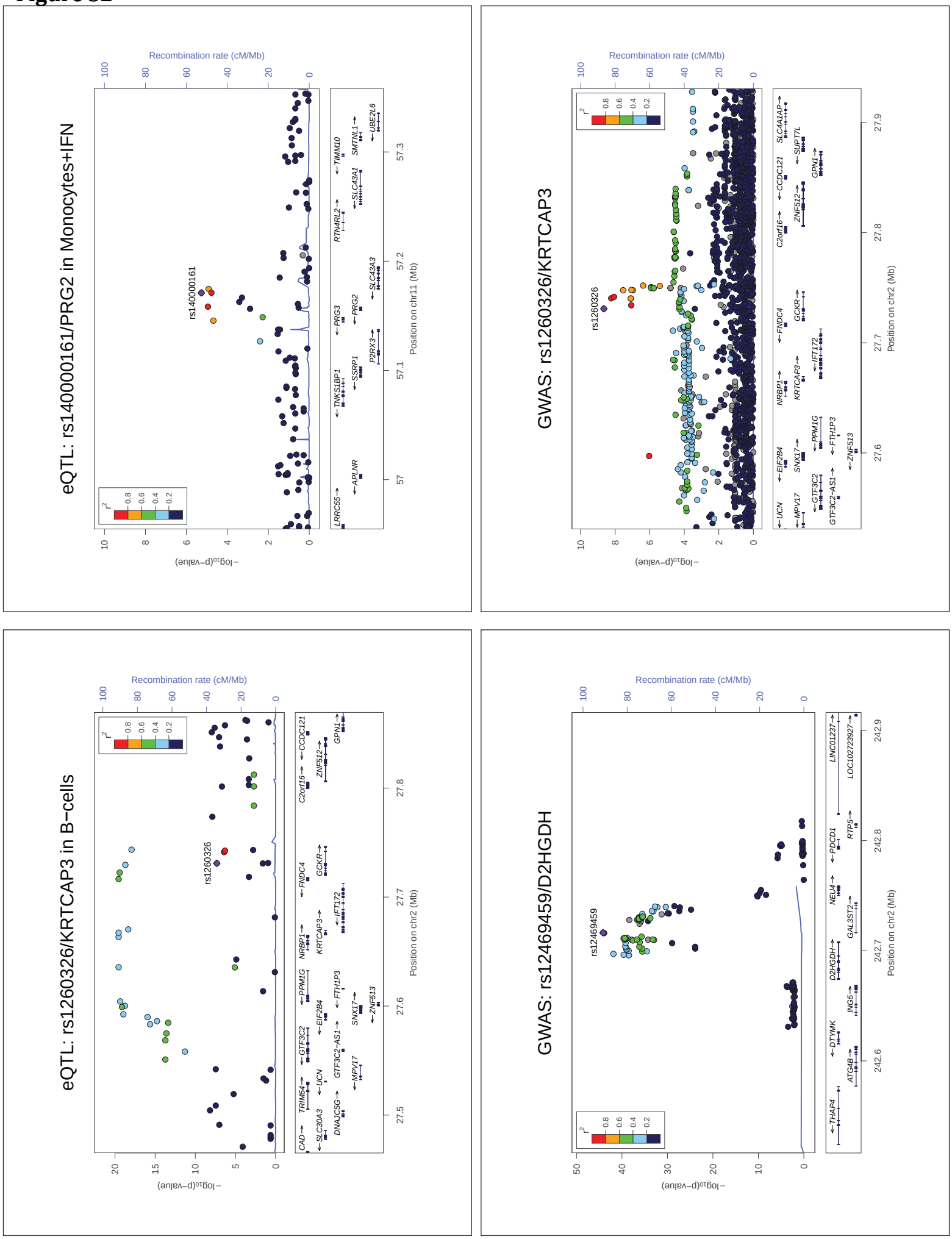

eQTL: rs140000161/PRG2 in Monocytes+IFN
GWAS: rs1260326/KRTCAP3
eQTL: rs1260326/KRTCAP3 in B−cells
GWAS: rs12469459/D2HGDH
Recombination rate (cM/Mb)
−log10(p−value)
Position on chr11 (Mb)
Position on chr2 (Mb)
rs140000161
rs1260326
rs12469459
r2

Figure S2

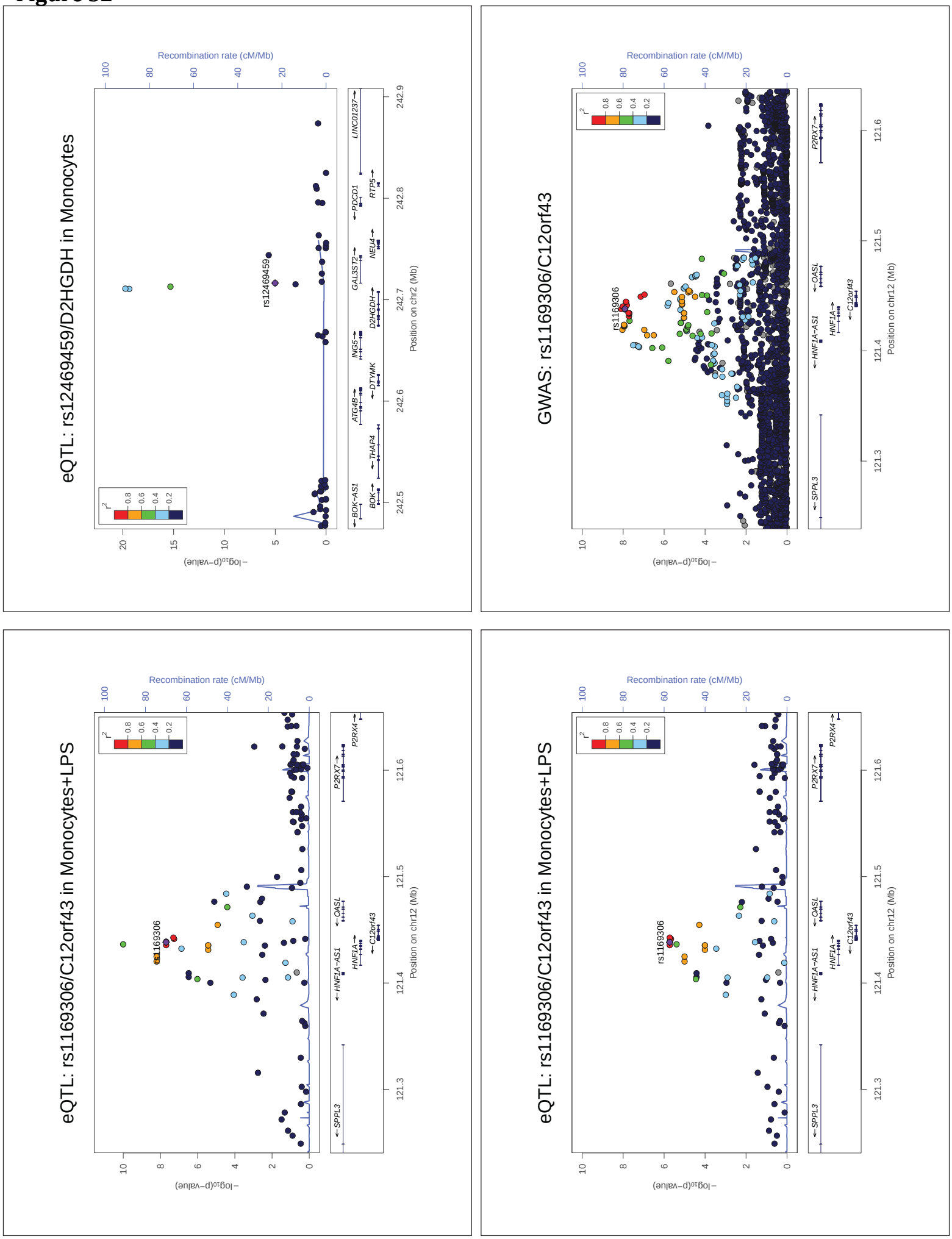

eQTL: rs12469459/D2HGDH in Monocytes
GWAS: rs1169306/C12orf43
eQTL: rs1169306/C12orf43 in Monocytes+LPS
eQTL: rs1169306/C12orf43 in Monocytes+LPS
Recombination rate (cM/Mb)
−log10(p−value)
Position on chr2 (Mb)
Position on chr12 (Mb)
rs12469459
rs1169306
r2
BOK-AS1
BOK
THAP4
ATG4B
DTYMK
ING5
D2HGDH
GAL3ST2
NEU4
PDCD1
RTP5
LINC01237
SPPL3
HNF1A-AS1
HNF1A
C12orf43
OASL
P2RX7
P2RX4

Figure S2

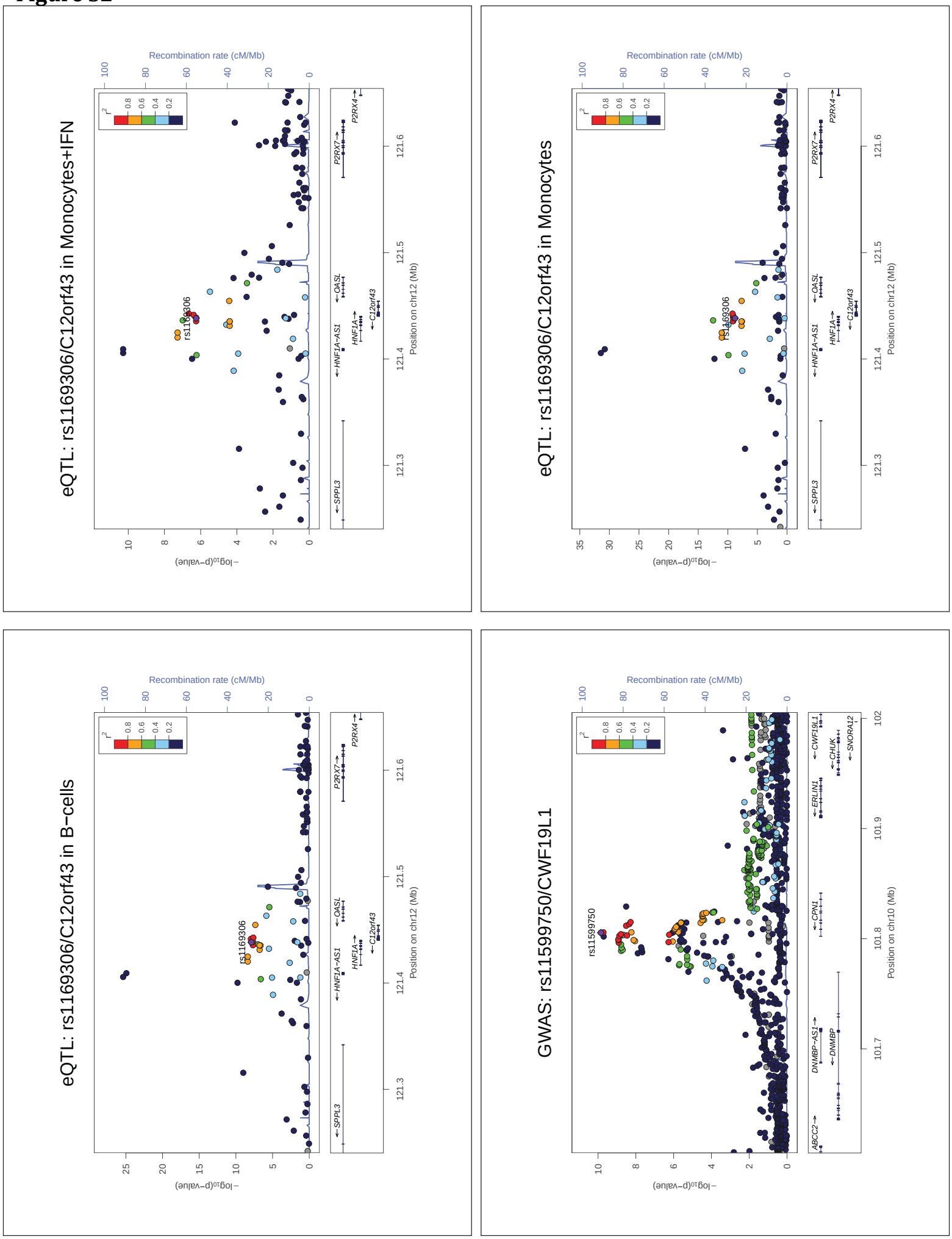

eQTL: rs1169306/C12orf43 in Monocytes+IFN
eQTL: rs1169306/C12orf43 in Monocytes
eQTL: rs1169306/C12orf43 in B−cells
GWAS: rs11599750/CWF19L1
Recombination rate (cM/Mb)
rs1169306
rs11599750
Position on chr12 (Mb)
Position on chr10 (Mb)
SPPL3
HNF1A−AS1
OASL
HNF1A
C12orf43
P2RX7
P2RX4
ABCC2
DNMBP−AS1
DNMBP
CPN1
ERLIN1
CHUK
CWF19L1
SNORA12

Figure S2

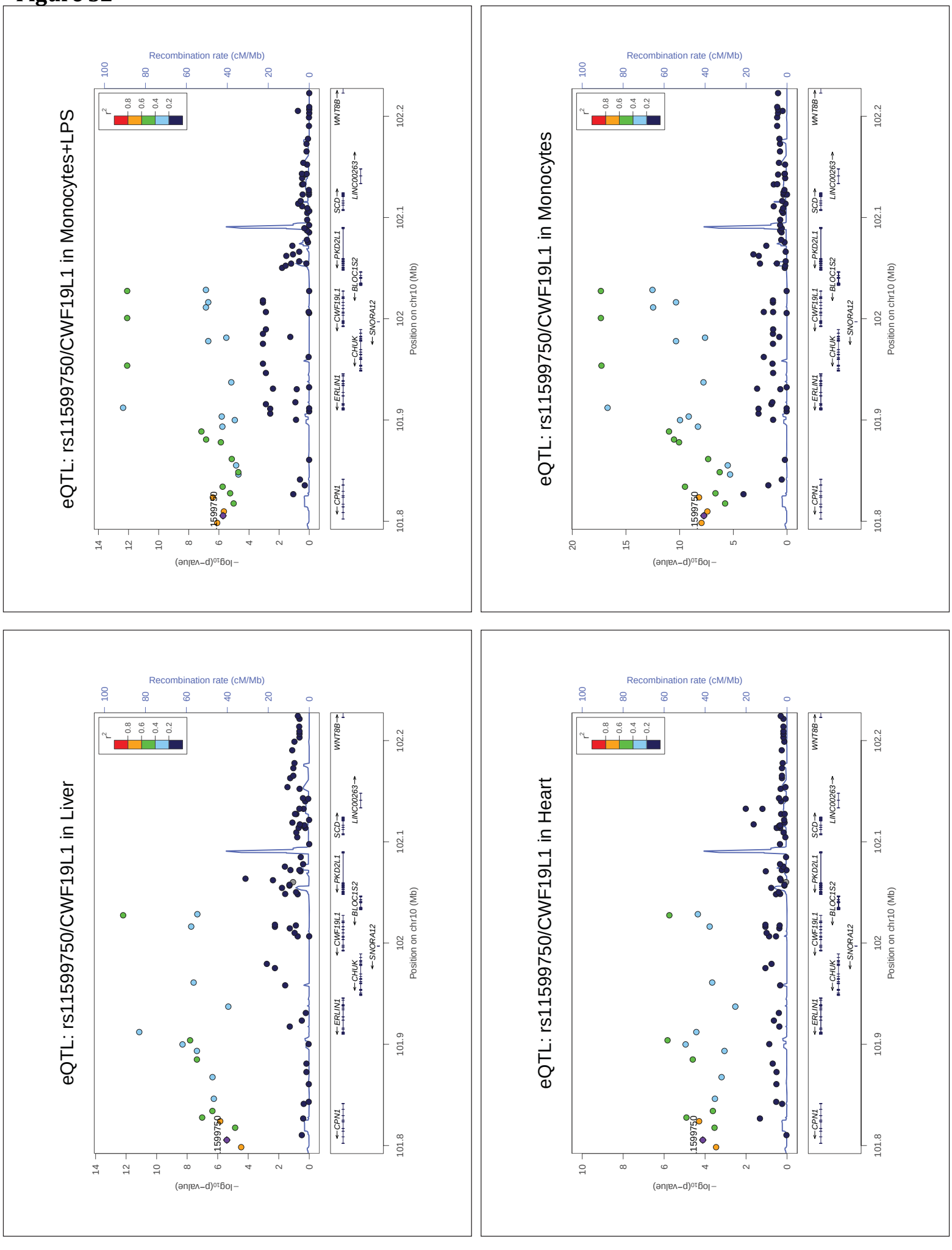

eQTL: rs11599750/CWF19L1 in Monocytes+LPS
eQTL: rs11599750/CWF19L1 in Monocytes
eQTL: rs11599750/CWF19L1 in Liver
eQTL: rs11599750/CWF19L1 in Heart
Recombination rate (cM/Mb)
r²
0.8
0.6
0.4
0.2
11599750
−log10(p−value)
Position on chr10 (Mb)
CPN1
ERLIN1
CWF19L1
CHUK
BLOC1S2
SNORA12
PKD2L1
SCD
LINC00263
WNT8B

Table S1

SNP	Position (chr:MB)	Trait	Dist (kb)	Likely mediator gene	Disco-very P	Disco-very Beta	A1/A2	A1 freq	INFO	Protein Name	Replication P	Combined P	Directions
rs2153101	chr1:203.2	CHI3L1	-12	CHI3L1	7E-108	-0.62	A/T	0.21	1.05	Chitinase-3-like protein 1	7.01e-136	1.8e-236	---
rs635634	chr9:136.2	SELE	trans	SURF6	1E-219	-0.86	T/C	0.22	0.99	E-selectin	8.1e-133	<1e-256	---
rs1420101	chr2:103	IL1RL1 (ST2)	0	IL1RL1	2E-132	0.58	C/T	0.63	1.03	Interleukin-1 receptor-like 1	5.06e-125	5.86e-254	+++
rs72650832	chr4:74.7	CXCL6	-29	CXCL6	6.2E-42	-0.53	T/C	0.65	0.41	C-X-C motif chemokine 6	8.79e-104	3.12e-130	---
rs635634	chr9:136.2	PECAM1	trans		1.9E-45	-0.41	T/C	0.22	0.99	Platelet endothelial cell adhesion molecule	3.27e-90	6.91e-125	---
rs1569723	chr20:44.7	CD40	-5	CD40	3E-49	-0.39	C/A	0.28	1.00	Tumor necrosis factor receptor superfamily member 5	1.156e-79	1.81e-121	---
rs17368659	chr11:102.7	MMP12	0	MMP12	5.5E-97	0.82	G/T	0.85	0.70	Macrophage metalloelastase	2.15e-76	2.84e-171	+++
rs6555820	chr5:156.5	HAVCR1 (TIM)	0	HAVCR1	1.3E-87	0.46	C/A	0.52	1.03	Hepatitis A virus cellular receptor1	8.443e-60	1.76e-145	+++
rs4129267	chr1:154.4	IL6R (IL6RA)	0	IL6R	2E-265	0.81	T/C	0.36	0.99	Interleukin-6 receptor subunit alpha	1.617e-56	1.53e-297	+-+
rs7946057	chr11:102.7	MMP3	7	MMP3	1E-108	0.53	A/T	0.53	0.97	Stromelysin-1	1.41e-53	5.21e-160	+?+
rs471994	chr11:102.7	MMP1	-29	MMP1	2.3E-35	0.32	G/A	0.65	0.97	Interstitial collagenase	9.129e-52	2.182e-82	+++
rs62625034	chr3:46.4	CCL4	trans	CCR5	3.1E-41	-0.55	G/T	0.9	0.99	C-C motif chemokine 4	2.178e-44	7.618e-83	---
rs4905	chr19:4.2	IL27 (IL27A)	0	EBI3	1.2E-80	-0.49	G/A	0.3	1.02	Interleukin-27 subunit alpha	1.59e-40	4.52e-119	-?-
rs9323280	chr14:56	LGALS3 (GAL3)	189	LGALS3	5.6E-62	0.77	A/C	0.87	0.60	Galectin-3	4.052e-34	6.6e-94	+++
rs12469459	chr2:242.7	MUC16 (CA125)	trans	GAL3ST2	7.1E-45	0.65	A/T	0.68	0.31	Mucin-16	1.05e-33	3.202e-75	+?+
rs8176741	chr9:136.1	TEK (TIE2)	trans	RALGDS	8.7E-50	-0.61	G/A	0.91	1.02	Angiopoietin-1 receptor	1.769e-30	2.264e-78	---
rs74544699	chr4:74.7	CXCL1	0	CXCL1	1.3E-12	-0.45	A/G	0.95	0.78	Growth-regulated alpha protein	5.48e-25	1.042e-30	-?-

rs3176123	chr20:23	THBD (TM)	0	THBD	2.3E-24	-0.3	T/G	0.81	1.04	Thrombomodulin	1.441e-23	5.828e-46	---
rs982764	chr10:90.8	FAS	0	FAS	2E-12	-0.18	C/T	0.32	1.04	Tumor necrosis factor receptor superfamily member 6	1.06e-21	1.792e-28	-?-
rs79250370	chr9:27.2	TEK (TIE2)	0	TEK	1.9E-13	-0.58	G/A	0.94	0.44	Angiopoietin-1 receptor	5.353e-20	1.141e-30	---
rs2188974	chr17:34.4	CCL3	1	CCL3	4.9E-18	-0.31	A/G	0.81	0.75	C-C motif chemokine 3	9.466e-19	7.919e-35	---
rs6607368	chr17:34.8	CCL4	386	CCL4	6.3E-31	0.51	A/C	0.8	0.46	C-C motif chemokine 4	1.78e-18	2.081e-47	+?+
rs111693235	chr11:1.8	CTSD	4	CTSD	2E-26	0.35	C/G	0.71	0.65	Cathepsin D	2.1e-17	1.071e-41	+?+
rs4810479	chr20:44.5	KITLG (SCF)	trans	MMP9	4.5E-11	-0.18	T/C	0.74	1.01	Kit ligand	8.921e-17	2.802e-25	---
rs62115757	chr19:51.5	KLK11 (HK11)	4	KLK11	1.2E-62	0.77	T/G	0.79	0.39	Kallikrein-11	6.13e-16	6.693e-76	+?+
rs670211	chr16:57.4	CX3CL1	0	CX3CL1	7.4E-12	0.18	G/A	0.41	0.84	Fractalkine	3.654e-15	7.084e-25	+++
rs344560	chr19:6.7	TNFSF14	0	TNFSF14	3E-18	0.75	C/T	0.96	0.46	Tumor necrosis factor ligand superfamily member 14	5.478e-15	1.313e-31	+-+
rs35285321	chr21:45.2	CSTB	-6	CSTB	1.2E-43	-0.43	A/G	0.38	0.64	Cystatin-B	7.56e-15	7.873e-57	--?
rs75649625	chr11:112.1	IL18	-17	IL18	1.4E-21	0.29	G/A	0.76	0.86	Interleukin-18	3.323e-13	4.094e-33	+++
rs1969539	chr11:14	SPON1	0	SPON1	1.5E-22	-0.25	G/A	0.5	0.89	Spondin-1	5.531e-13	8.121e-34	---
rs693918	chr2:31.9	IL18	trans		2.4E-11	0.19	G/A	0.55	0.72	Interleukin-18	2.905e-12	7.815e-22	+++
rs1260326	chr2:27.7	FST (FS)	trans	GCKR	2E-09	0.14	T/C	0.44	1.07	Follistatin	8.217e-11	1.863e-18	+++
rs4672375	chr2:60.5	GAL	trans		7.1E-11	0.17	G/A	0.58	0.90	Galanin peptides	3.79e-10	8.019e-19	+?+
rs116661163	chr1:204.6	REN	-475	REN	1E-08	-0.72	C/G	0.98	0.38	Renin	4.168e-10	4.455e-17	---
rs142552223	chr3:172.2	TNFSF11 (TRANCE)	trans		3.4E-17	0.81	G/A	0.97	0.58	Tumor necrosis factor ligand superfamily member 11	6.042e-10	1.636e-25	+++
rs13236526	chr7:75.9	HSPB1 (HSP27)	0	HSPB1	1.1E-17	0.36	A/G	0.7	0.44	Heat shock protein beta-1	1.59e-09	2.35e-25	++?
rs3195944	chr19:18.5	GDF15	-20	GDF15	2.2E-08	-0.33	A/G	0.88	0.39	Growth differentiation factor 15	8.05e-09	5.618e-15	-?-
rs56378716	chr17:56.4	MPO	0	MPO	1.9E-09	-0.5	G/A	0.02	0.99	Myeloperoxidase	4.798e-08	5.002e-16	---
rs73062378	chr3:45.8	CCL4	trans		4.5E-13	-0.29	T/C	0.81	0.61	C-C motif chemokine 4	7.16e-08	2.257e-19	-?-
rs140000161	chr11:57.2	PAPPA	trans		1.4E-10	-0.37	A/G	0.91	0.55	Pappalysin-1	9.1e-08	1.27e-16	-?-
rs28601761	chr8:126.5	CHI3L1	trans		5E-09	0.14	C/G	0.61	1.03	Chitinase-3-like protein 1	3.746e-07	9.569e-15	+++

rs8176693	chr9:136.1	THBD (TM)	trans	MED22	1.1E-10	-0.27	C/T	0.91	1.03	Thrombomodulin	4.021e-07	2.427e-16	---
rs7813952	chr8:120.1	TNFSF11 (TRANCE)	trans	TNFRSF11B	2.1E-16	0.22	C/T	0.58	0.81	Tumor necrosis factor ligand superfamily member 11	4.092e-07	1.284e-21	+++
rs200373	chr19:18.3	CTSL1	trans		4.3E-09	-0.14	T/A	0.48	1.03	Cathepsin L1	5.46e-07	2.116e-14	-?-
rs6993770	chr8:106.6	DKK1	trans		1.6E-09	-0.17	T/A	0.27	0.99	Dickkopf-related protein 1	8.545e-06	7.278e-14	---
rs2271025	chr16:67	AGRP	530†		2.3E-09	-0.36	G/A	0.92	0.55	Agouti-related protein	1.866e-05	2.346e-13	---
rs35186877	chr17:4.6	CXCL16	11	CXCL16	1.7E-09	0.21	G/A	0.81	0.76	C-X-C motif chemokine 16	3.07e-05	2.54e-13	+?+
rs35166255	chr11:126.3	IL1RL1 (ST2)	trans		1.2E-09	-0.46	G/A	0.96	0.68	Interleukin-1 receptor-like 1	3.69e-05	2.02e-13	-?-
rs492602	chr19:49.2	MMP10	trans	FUT2, RASIP1	7.8E-09	0.14	G/A	0.45	1.03	Stromelysin-2	5.71e-05	2.364e-12	+++
rs11667946	chr19:51.5	KLK6	-1	KLK6	3.4E-15	0.3	C/T	0.53	0.41	Kallikrein-6	0.000116	2.909e-18	+?+
rs7928577	chr11:126.2	LGALS3 (GAL3)	trans	TIRAP	2.1E-09	0.26	T/G	0.08	0.98	Galectin-3	0.0001312	1.834e-12	+++
rs495828	chr9:136.2	F3 (TF)	trans	SURF6	4.6E-10	-0.17	T/G	0.25	0.99	Tissue factor	0.0001693	6.249e-13	---
rs17610659	chr1:110.5	CSF1	30	CSF1	6.5E-10	0.15	T/C	0.48	0.97	Macrophage colony-stimulating factor 1	0.000228	1.202e-12	+++
rs11599750	chr10:101.8	IL27 (IL27A)	trans	CWF19L1	1.4E-10	0.16	C/T	0.62	1.00	Interleukin-27 subunit alpha	0.000429	2.816e-13	+?+
rs6469811	chr8:120.1	TNFRSF11B (OPG)	-137	TNFRSF11B	2.9E-11	-0.18	G/A	0.55	0.81	Osteprotegerin	0.0005594	2.52e-13	---
rs76769120	chr1:12.2	TNFRSF1B (TRAIL/TNFR2)	0	TNFRSF1B	1.3E-11	0.42	G/T	0.91	0.46	Tumor necrosis factor receptor superfamily member 1B	0.000869	6.649e-14	+?+
rs11150189	chr16:79.7	XPNPEP2 (MAMP)	trans		6.9E-14	0.2	A/G	0.69	0.97	Xaa-Pro aminopeptidase 2	0.00104	3.739e-16	++?
rs2070600	chr6:32.2	AGER (RAGE)	0	AGER	3E-10	-0.37	T/C	0.04		Advanced glycosylation end product-specific receptor	0.001181	4.749e-12	-+-
rs241771	chr17:26.6	TNFRSF11B (OPG)	trans		6E-10	-0.15	T/C	0.45	0.98	Osteprotegerin	0.003457	3.469e-11	---
rs1169306	chr12:121.4	LGALS3 (GAL3)	trans	HNF1A, C12orf43	6.5E-09	0.14	T/C	0.38	1.02	Galectin-3	0.003868	2.847e-10	+-+
rs16873402	chr8:106.6	PDGFB	trans		2.4E-08	0.15	C/T	0.71	0.94	Platelet-derived growth factor subunit B	0.005719	1.407e-09	+++
rs7599125	chr2:32.9	IL18	trans	NLRC4	1.1E-08	-0.23	G/A	0.44	0.35	Interleukin-18	0.006416	8.746e-10	---
rs33988101	chr19:49.2	LGALS3 (GAL3)	trans	FUT2, RASIP1	3.5E-09	0.14	T/G	0.48	1.04	Galectin-3	0.009141	5.557e-10	+++

rs880949	chr14:75.4	PGF (PLGF)	6	PGF	1.6E-08	-0.17	G/A	0.58	0.68	Placenta growth factor	0.01379	3.281e-09	---
rs1580006	chr11:10.4	ADM (AM)	42	ADM	2E-15	-0.19	A/T	0.54	1.00	Pro-adrenomedullin	0.02987	7.124e-14	---
rs10947260	chr6:32.4	IL6	trans	NOTCH4, AGER, ATF6B	1.8E-10	0.25	C/T	0.1		Interleukin-6	0.1524	6.21e-09	+++
rs35538083	chr1:27.1	XPNPEP2 (MAMP)	trans	PIGV	3.1E-08	-0.27	T/C	0.93	0.95	Xaa-Pro aminopeptidase 2	0.181	3.717e-08	--?
rs75416436	chr13:42.6	NGF (BETANGF)	trans		4.2E-08	-0.5	G/A	0.96	0.44	Beta-nerve growth factor	0.286	9.912e-08	--?
rs12570111	chr10:12.3	MMP1	trans		4.7E-08	0.14	T/C	0.56	0.94	Interstitial collagenase	0.616	1.4e-05	+?-
rs76519098	chr10:49.9	GDF15	trans	MAPK8	1.1E-10	-0.83	C/T	0.98	0.39	Growth differentiation factor 15	0.6989	2.571e-06	--+
rs117538444	chr15:89.9	PGF (PLGF)	trans		6.6E-09	0.78	C/T	0.98	0.37	Placenta growth factor	0.9649	9.685e-06	+-+
rs549596	chr1:11.9	BNP	1	NPPB	1.7E-14	-0.19	T/C	0.59	0.97	Binatriuretic peptides		1.738e-14	-??
rs35207557	chr1:11.9	NPPB (NTPROBNP)	0	NPPB	2.6E-25	-0.26	T/TA	0.6	0.98	Natriuretic peptides B		2.57e-25	-??
rs184243355	chr5:153.2	CCL3	trans		2.2E-08	-0.41	T/C	0.94	0.47	C-C motif chemokine 3		2.239e-08	-??
rs61598054	chr6:108.8	NGF (BETANGF)	trans	FOXO3	3.8E-08	-0.31	C/T	0.87	0.42	Beta-nerve growth factor		3.802e-08	-??
rs6557662	chr8:23.2	NPPB (NTPROBNP)	trans		1.5E-08	0.23	A/G	0.73	0.44	Natriuretic peptides B		1.479e-08	+??
rs139879640	chr15:81.6	IL16	0	IL16	3E-62	0.86	TCTCA/	0.94	0.96	Pro-interleukin-16 [Cleaved into: Interleukin-16		2.951e-62	+??
rs200433550	chr19:49.2	F3 (TF)	trans		5.6E-10	0.16	TA/T	0.58	0.91	Tissue factor		5.623e-10	+??
rs2050011	chrX:128.9	XPNPEP2 (MAMP)	-1	XPNPEP2	2.4E-68	-0.36	T/G	0.33	1.05	Xaa-Pro aminopeptidase 2		2.399e-68	-??

Overview of all associations between plasma protein and SNPs significant at genome-wide level. Trait – the plasma protein target; Dist (kb) – if cis, the distance between SNP and protein encoding gene; Likely mediator gene – most likely cis-mediator gene, protein-encoding gene in cis-cases, based on table 2 in trans cases; Discovery P – the pQTL association P-value from the Olink-Improve discovery cohort (n=3,394); Discovery Beta – the Olink-Improve effect size; A1/A2 – encoded allele and alternative allele; A1 freq – frequency of encoded allele; INFO – imputation quality score (MACH 1.0). Protein name; Replication P - the pQTL association P-value from the replication cohorts (n=976, n=933,n=730); Combined P – the meta-analysis P-value of both discovery and replication; Directions - for replication meta-analysis are indicated as IMPROVE (discovery), NSPHS (replication), ULSAM-PIVUS (merged replication). †while 530.7 kb is formally outside of the pre-defined cis-limit of 500 kb, the AGRP association was classified as cis-acting; all other pQTL associations were either acting across chromosomes or at distances more than 100 MB. SE: standard error.

Table S2

Other trait	Other SNP	r^2 (EUR 1000G)	Other P-value	Pubmed ID	Olink SNP	Olink Trait Protein	Olink P-value
DBP	rs6668659	0.67	5.1E-10	24560520	rs549596	BNP	1.75E-14
Mean Arterial Pressure	rs6668659	0.67	1.2E-09	24560520	rs549596	BNP	1.75E-14
SBP	rs6668659	0.67	4.5E-09	24560520	rs549596	BNP	1.75E-14
HDL cholesterol	rs12748152	0.92	1.0E-15	24097068	rs35538083	XPNPEP2	3.11E-08
LDL cholesterol	rs12748152	0.92	3.0E-12	24097068	rs35538083	XPNPEP2	3.11E-08
Triglycerides	rs12748152	0.92	1.0E-09	24097068	rs35538083	XPNPEP2	3.11E-08
Interleukin-6_receptor|CSF|(gene=IL6R)	rs4129267	1.00	2.7E-62	25340798	rs4129267	IL6R	2.14E-265
IL6RA|plasma|(gene=IL6R)	rs4129267	1.00	4.4E-58	25147954	rs4129267	IL6R	2.14E-265
Protein quantitative trait loci (sIL-6R)	rs4129267	1.00	2.0E-57	18464913	rs4129267	IL6R	2.14E-265
C-reactive protein	rs4129267	1.00	2.0E-48	21300955	rs4129267	IL6R	2.14E-265
Fibrinogen|plasma	rs61812598	1.00	2.7E-36	26561523	rs4129267	IL6R	2.14E-265
Fibrinogen (EA)	rs4129267	1.00	6.0E-27	23969696	rs4129267	IL6R	2.14E-265
Ankylosing spondylitis	rs4129267	1.00	3.4E-13	23749187	rs4129267	IL6R	2.14E-265
Asthma	rs4129267	1.00	2.0E-08	21907864	rs4129267	IL6R	2.14E-265
Fibrinogen	rs2228145	0.99	2.0E-11	20031577	rs4129267	IL6R	2.14E-265
Rheumatoid arthritis	rs2228145	0.99	4.0E-09	24390342	rs4129267	IL6R	2.14E-265
YKL-40 levels	rs4950928	0.90	1.0E-13	18403759	rs2153101	CHI3L1	7.50E-108
Triglycerides	rs1260326	1.00	2.0E-239	24097068	rs1260326	FST	2.02E-09
Blood metabolite ratios (glucose/mannose)	rs1260326	1.00	3.0E-148	24816252	rs1260326	FST	2.02E-09
Blood metabolite levels (mannose)	rs1260326	1.00	1.0E-77	24816252	rs1260326	FST	2.02E-09
Urate levels	rs1260326	1.00	1.0E-44	23263486	rs1260326	FST	2.02E-09
Tryglycerides transethnic	rs1260326	1.00	1.6E-42	23555291	rs1260326	FST	2.02E-09
Cholesterol, total	rs1260326	1.00	3.0E-42	24097068	rs1260326	FST	2.02E-09
Fasting glucose- BMI-adjusted	rs1260326	1.00	1.0E-40	22885924	rs1260326	FST	2.02E-09
Fasting glucose	rs1260326	1.00	1.0E-40	22885924	rs1260326	FST	2.02E-09
C-reactive protein	rs1260326	1.00	5.0E-40	21300955	rs1260326	FST	2.02E-09
Lipid metabolism phenotypes (TG.assay, whole)	rs1260326	1.00	1.0E-37	19936222	rs1260326	FST	2.02E-09

Lipid metabolism phenotypes (TG.by.NMR, whole)	rs1260326	1.00	3.0E-35	19936222	rs1260326	FST	2.02E-09
Lipid metabolism phenotypes (TG.assay, fasting)	rs1260326	1.00	4.0E-32	19936222	rs1260326	FST	2.02E-09
Lipid metabolism phenotypes (TG.by.NMR, fasting)	rs1260326	1.00	3.0E-29	19936222	rs1260326	FST	2.02E-09
Lipid metabolism phenotypes (VLDL.large, whole)	rs1260326	1.00	3.0E-28	19936222	rs1260326	FST	2.02E-09
Lipid metabolism phenotypes (VLDL.large, fasting)	rs1260326	1.00	4.0E-24	19936222	rs1260326	FST	2.02E-09
Fasting Insulin- BMI adjusted	rs1260326	1.00	2.7E-22	22885924	rs1260326	FST	2.02E-09
Serum albumin level	rs1260326	1.00	4.0E-19	23022100	rs1260326	FST	2.02E-09
Metabolite levels (Ala, Gln)	rs1260326	1.00	3.0E-18	22286219	rs1260326	FST	2.02E-09
2 hr glucose -FG and BMI adjusted	rs1260326	1.00	9.0E-15	22885924	rs1260326	FST	2.02E-09
Chronic kidney disease (eGFRcrea)	rs1260326	1.00	3.0E-14	20383146	rs1260326	FST	2.02E-09
Serum albumin level (EA)	rs1260326	1.00	3.0E-14	23022100	rs1260326	FST	2.02E-09
Fasting Insulin	rs1260326	1.00	3.8E-14	22885924	rs1260326	FST	2.02E-09
Blood metabolite levels (alanine)	rs1260326	1.00	6.0E-14	24816252	rs1260326	FST	2.02E-09
Hypertriglyceridemia	rs1260326	1.00	2.0E-13	23505323	rs1260326	FST	2.02E-09
Liver enzyme levels (gamma-glutamyl transferase)	rs1260326	1.00	4.0E-13	22001757	rs1260326	FST	2.02E-09
Glycemic traits (pregnancy) (FPG)	rs1260326	1.00	6.0E-13	23903356	rs1260326	FST	2.02E-09
Metabolite levels	rs1260326	1.00	1.0E-12	22916037	rs1260326	FST	2.02E-09
Alanine	rs1260326	1.00	7.6E-12	23823483	rs1260326	FST	2.02E-09
PC aa C40:5	rs1260326	1.00	1.3E-11	26068415	rs1260326	FST	2.02E-09
Lactate	rs1260326	1.00	3.3E-11	23823483	rs1260326	FST	2.02E-09
Glycemic traits (pregnancy) (FCP)	rs1260326	1.00	6.0E-11	23903356	rs1260326	FST	2.02E-09
Two-hour glucose challenge	rs1260326	1.00	3.0E-10	20081857	rs1260326	FST	2.02E-09
Metabolic traits (TG)	rs1260326	1.00	4.0E-10	19060910	rs1260326	FST	2.02E-09
Platelet counts	rs1260326	1.00	9.0E-10	22139419	rs1260326	FST	2.02E-09
α-hydroxybutyrate	rs1260326	1.00	1.3E-09	23823483	rs1260326	FST	2.02E-09
Non-albumin protein levels (ALB)	rs1260326	1.00	3.0E-09	22558069	rs1260326	FST	2.02E-09
TAG 50:4	rs1260326	1.00	3.4E-09	23823483	rs1260326	FST	2.02E-09

Hematological and biochemical traits (ALB)	rs1260326	1.00	4.0E-09	20139978	rs1260326	FST	2.02E-09
Cardiovascular disease risk factors (TRIG)	rs1260326	1.00	2.0E-08	21943158	rs1260326	FST	2.02E-09
Serum albumin level (Japanese)	rs1260326	1.00	2.0E-08	23022100	rs1260326	FST	2.02E-09
TAG 48:2	rs1260326	1.00	2.5E-08	23823483	rs1260326	FST	2.02E-09
TAG 50:3	rs1260326	1.00	2.6E-08	23823483	rs1260326	FST	2.02E-09
PC 34:3	rs1260326	1.00	2.9E-08	23823483	rs1260326	FST	2.02E-09
PC 32:2	rs1260326	1.00	3.9E-08	23823483	rs1260326	FST	2.02E-09
Waist circumference and related phenotypes (triglycerides)	rs1260326	1.00	4.0E-08	18454146	rs1260326	FST	2.02E-09
TAG 48:3	rs1260326	1.00	4.9E-08	23823483	rs1260326	FST	2.02E-09
Urate levels (Urate)	rs780093	0.91	4.0E-17	20884846	rs1260326	FST	2.02E-09
Sex hormone-binding globulin levels (Men + Women)	rs780093	0.91	2.0E-16	22829776	rs1260326	FST	2.02E-09
Waist Circumference - Triglycerides (WC-TG)	rs780093	0.91	2.0E-12	21386085	rs1260326	FST	2.02E-09
Crohn's disease	rs780093	0.91	5.0E-11	21102463	rs1260326	FST	2.02E-09
Sex hormone-binding globulin levels (Women)	rs780093	0.91	9.0E-11	22829776	rs1260326	FST	2.02E-09
Triglycerides-Blood Pressure (TG-BP)	rs780093	0.91	3.0E-10	21386085	rs1260326	FST	2.02E-09
Palmitoleic acid (16:1n-7) plasma levels	rs780093	0.91	1.0E-09	23362303	rs1260326	FST	2.02E-09
Metabolic traits (glucose/mannose + 54 other traits)	rs780094	0.90	6.0E-53	21886157	rs1260326	FST	2.02E-09
Fasting glucose-related traits (FPG)	rs780094	0.90	6.0E-38	20081858	rs1260326	FST	2.02E-09
Fasting glucose-related traits (HOMA-IR)	rs780094	0.90	3.0E-24	20081858	rs1260326	FST	2.02E-09
Fasting insulin-related traits (HOMA-IR)	rs780094	0.90	3.0E-24	20081858	rs1260326	FST	2.02E-09
Fasting glucose-related traits (interaction with BMI)	rs780094	0.90	4.0E-24	22581228	rs1260326	FST	2.02E-09
Fasting insulin-related traits (FI)	rs780094	0.90	4.0E-20	20081858	rs1260326	FST	2.02E-09
Fasting glucose-related traits (FI)	rs780094	0.90	4.0E-20	20081858	rs1260326	FST	2.02E-09
Metabolic syndrome (TG)	rs780094	0.90	6.0E-20	22399527	rs1260326	FST	2.02E-09
height	rs780094	0.90	7.5E-12	25282103	rs1260326	FST	2.02E-09
Calcium levels	rs780094	0.90	1.0E-10	24068962	rs1260326	FST	2.02E-09
Fasting insulin-related traits (interaction with BMI)	rs780094	0.90	3.0E-10	22581228	rs1260326	FST	2.02E-09

Uric acid levels	rs780094	0.90	1.0E-09	19503597	rs1260326	FST	2.02E-09
Fasting plasma glucose (East Asian)	rs780094	0.90	3.6E-09	25187374	rs1260326	FST	2.02E-09
Triglycerides (Hispanic)	rs780094	0.90	7.0E-09	23726366	rs1260326	FST	2.02E-09
Phospholipid levels (plasma) (DPA)	rs780094	0.90	9.0E-09	21829377	rs1260326	FST	2.02E-09
Triglycerides (AA)	rs4665972	0.87	1.0E-08	23726366	rs1260326	FST	2.02E-09
Inflammatory biomarkers (IL18)	rs7577696	0.77	3.0E-19	24182552	rs693918	IL18	2.38E-11
Eosinophil counts (EA)	rs1420101	1.00	5.0E-14	19198610	rs1420101	IL1RL1	2.03E-132
Serum protein levels (sST2)	rs950880	0.97	7.0E-94	23999434	rs1420101	IL1RL1	2.03E-132
Pulmonary function (interaction) (FEV1/FVC, Pack-years)	rs2070600	1.00	1.0E-21	23284291	rs2070600	AGER	3.01E-10
Pulmonary function (FEV1/FVC)	rs2070600	1.00	3.0E-14	20010835	rs2070600	AGER	3.01E-10
Crohn's disease	rs10947261	0.97	3.0E-12	23850713	rs10947260	IL6	1.83E-10
Knee osteoarthritis	rs10947262	0.97	5.0E-09	20305777	rs10947260	IL6	1.83E-10
Vascular endothelial growth factor levels	rs6993770	1.00	5.0E-23	21757650	rs6993770	DKK1	1.61E-09
Platelet counts	rs6993770	1.00	4.0E-17	22139419	rs6993770	DKK1	1.61E-09
Vascular endothelial growth factor levels	rs6993770	0.80	5.0E-23	21757650	rs16873402	PDGFB	2.42E-08
Platelet counts	rs6993770	0.80	4.0E-17	22139419	rs16873402	PDGFB	2.42E-08
Triglycerides	rs2954029	0.74	1.0E-107	24097068	rs28601761	CHI3L1	5.06E-09
Cholesterol, total	rs2954029	0.74	2.0E-65	24097068	rs28601761	CHI3L1	5.06E-09
LDL cholesterol	rs2954029	0.74	2.0E-50	24097068	rs28601761	CHI3L1	5.06E-09
HDL cholesterol	rs2954029	0.74	3.0E-29	24097068	rs28601761	CHI3L1	5.06E-09
CAD	rs2954029	0.74	4.8E-09	23202125	rs28601761	CHI3L1	5.06E-09
Lipid metabolism phenotypes (APOB.assay, whole)	rs6982636	0.71	7.0E-12	19936222	rs28601761	CHI3L1	5.06E-09
Lipid metabolism phenotypes (TG.assay, whole)	rs6982636	0.71	1.0E-09	19936222	rs28601761	CHI3L1	5.06E-09
Liver enzyme levels (alkaline phosphatase)	rs2954021	0.64	2.0E-13	22001757	rs28601761	CHI3L1	5.06E-09
Liver enzyme levels (alanine transaminase)	rs2954021	0.64	5.0E-09	22001757	rs28601761	CHI3L1	5.06E-09
Elevated serum carcinoembryonic antigen levels	rs8176741	1.00	2.0E-24	24941225	rs8176741	TEK	8.63E-50
Tumor biomarkers (CEA)	rs8176749	0.98	7.0E-105	23300138	rs8176741	TEK	8.63E-50
sTie-2|plasma|(gene=TEK)	rs8176693	0.98	1.8E-33	25552591	rs8176741	TEK	8.63E-50
vWF	rs8176693	0.98	1.6E-17	23381943	rs8176741	TEK	8.63E-50

End-stage coagulation (vWF)	rs8176743	0.98	2.0E-17	23381943	rs8176741	TEK	8.63E-50
FVIII Ag	rs8176693	0.98	1.2E-14	23381943	rs8176741	TEK	8.63E-50
Urinary metabolites (H-NMR features) (5.2625, Unknown)	rs8176749	0.98	4.0E-12	24586186	rs8176741	TEK	8.63E-50
Ang-2|plasma|(gene=MCPH1)	rs8176746	0.98	2.1E-08	25552591	rs8176741	TEK	8.63E-50
Mean corpuscular hemoglobin concentration	rs8176746	0.98	4.0E-08	20139978	rs8176741	TEK	8.63E-50
Malaria	rs8176722	0.70	9.0E-10	23717212	rs8176741	TEK	8.63E-50
Tumor biomarkers (CEA)	rs8176749	1.00	7.0E-105	23300138	rs8176693	THBD	1.12E-10
sTie-2|plasma|(gene=TEK)	rs8176693	1.00	1.8E-33	25552591	rs8176693	THBD	1.12E-10
vWF	rs8176693	1.00	1.6E-17	23381943	rs8176693	THBD	1.12E-10
End-stage coagulation (vWF)	rs8176743	1.00	2.0E-17	23381943	rs8176693	THBD	1.12E-10
FVIII Ag	rs8176693	1.00	1.2E-14	23381943	rs8176693	THBD	1.12E-10
Urinary metabolites (H-NMR features) (5.2625, Unknown)	rs8176749	1.00	4.0E-12	24586186	rs8176693	THBD	1.12E-10
Ang-2|plasma|(gene=MCPH1)	rs8176746	1.00	2.1E-08	25552591	rs8176693	THBD	1.12E-10
Mean corpuscular hemoglobin concentration	rs8176746	1.00	4.0E-08	20139978	rs8176693	THBD	1.12E-10
Elevated serum carcinoembryonic antigen levels	rs8176741	0.98	2.0E-24	24941225	rs8176693	THBD	1.12E-10
Malaria	rs8176722	0.71	9.0E-10	23717212	rs8176693	THBD	1.12E-10
Liver enzyme levels (alkaline phosphatase)	rs579459	1.00	3.0E-123	22001757	rs495828	F3	4.54E-10
E-selectin|plasma|(gene=SELE)	rs651007	1.00	1.9E-103	23300549	rs495828	F3	4.54E-10
E-selectin levels	rs651007	1.00	2.0E-82	20147318	rs495828	F3	4.54E-10
Hematological and biochemical traits (ALP)	rs495828	1.00	4.0E-59	20139978	rs495828	F3	4.54E-10
Serum alkaline phosphatase levels	rs651007	1.00	1.0E-56	24094242	rs495828	F3	4.54E-10
Soluble levels of adhesion molecules (P-Selectin)	rs579459	1.00	2.0E-41	20167578	rs495828	F3	4.54E-10
Blood metabolite ratios (ADSGEGDFXAEGGGVR/ADpSGEGDFXAEGGGVR)	rs649129	1.00	9.0E-37	24816252	rs495828	F3	4.54E-10
Blood metabolite ratios (DSGEGDFXAEGGGVR/ADpSGEGDFXAEGGGVR)	rs495828	1.00	6.0E-34	24816252	rs495828	F3	4.54E-10
Urinary metabolites (H-NMR features) (5.1825, Unknown)	rs579459	1.00	2.0E-32	24586186	rs495828	F3	4.54E-10

Soluble E-selectin levels	rs579459	1.00	1.0E-29	19729612	rs495828	F3	4.54E-10
Blood metabolite ratios (ADpSGEGDFXAEGGGVR/X-14304--leucylalanine)	rs579459	1.00	1.0E-28	24816252	rs495828	F3	4.54E-10
Urinary metabolites (H-NMR features) (2.0525, Unknown)	rs579459	1.00	1.0E-28	24586186	rs495828	F3	4.54E-10
End-stage coagulation (FVIII)	rs651007	1.00	2.0E-25	23381943	rs495828	F3	4.54E-10
Blood metabolite levels (ADpSGEGDFXAEGGGVR)	rs651007	1.00	6.0E-20	24816252	rs495828	F3	4.54E-10
Red blood cell traits (EA, RBCC)	rs579459	1.00	9.0E-18	23222517	rs495828	F3	4.54E-10
Venous thromboembolism	rs495828	1.00	3.0E-16	22672568	rs495828	F3	4.54E-10
Soluble levels of adhesion molecules (ICAM)	rs649129	1.00	1.0E-15	20167578	rs495828	F3	4.54E-10
Coronary heart disease	rs579459	1.00	4.0E-14	21378990	rs495828	F3	4.54E-10
vWF	rs495828	1.00	9.4E-13	23381943	rs495828	F3	4.54E-10
Red blood cell count	rs495828	1.00	3.0E-12	20139978	rs495828	F3	4.54E-10
Hematological and biochemical traits (Hb)	rs495828	1.00	1.0E-11	20139978	rs495828	F3	4.54E-10
Hematological and biochemical traits (Ht)	rs495828	1.00	6.0E-10	20139978	rs495828	F3	4.54E-10
Coronary artery disease or ischemic stroke	rs579459	1.00	2.0E-09	24262325	rs495828	F3	4.54E-10
Metabolite levels (LDL)	rs651007	1.00	6.0E-09	21909109	rs495828	F3	4.54E-10
Ferritin (log)|serum|(gene=FTL|FTH1)	rs651007	1.00	1.3E-08	25352340	rs495828	F3	4.54E-10
CAD	rs579459	1.00	2.7E-08	23202125	rs495828	F3	4.54E-10
Coronary artery disease or large artery stroke	rs579459	1.00	3.0E-08	24262325	rs495828	F3	4.54E-10
Angiotensin-converting enzyme activity	rs495828	1.00	3.0E-08	20066004	rs495828	F3	4.54E-10
FVIII Ag	rs495828	1.00	3.7E-08	23381943	rs495828	F3	4.54E-10
LDL cholesterol	rs635634	0.83	2.0E-41	24097068	rs495828	F3	4.54E-10
Cholesterol, total	rs635634	0.83	3.0E-35	24097068	rs495828	F3	4.54E-10
invasive epithelial ovarian cancer	rs635634	0.83	4.4E-09	25581431	rs495828	F3	4.54E-10
Soluble ICAM-1	rs507666	0.83	3.0E-91	21533024	rs495828	F3	4.54E-10
Lipid traits (LDL)	rs507666	0.83	2.0E-11	24386095	rs495828	F3	4.54E-10
Lipid traits (TC)	rs507666	0.83	4.0E-11	24386095	rs495828	F3	4.54E-10
LDL-C transethnic	rs2519093	0.82	2.2E-13	23555291	rs495828	F3	4.54E-10

LDL cholesterol	rs635634	1.00	2.0E-41	24097068	rs635634	PECAM1	1.89E-45
Cholesterol, total	rs635634	1.00	3.0E-35	24097068	rs635634	PECAM1	1.89E-45
E-selectin|plasma|(gene=SELE)	rs635634	1.00	5.0E-16	25147954	rs635634	PECAM1	1.89E-45
invasive epithelial ovarian cancer	rs635634	1.00	4.4E-09	25581431	rs635634	PECAM1	1.89E-45
Soluble ICAM-1	rs507666	0.99	3.0E-91	21533024	rs635634	PECAM1	1.89E-45
Lipid traits (LDL)	rs507666	0.99	2.0E-11	24386095	rs635634	PECAM1	1.89E-45
Lipid traits (TC)	rs507666	0.99	4.0E-11	24386095	rs635634	PECAM1	1.89E-45
Venous thromboembolism	rs2519093	0.98	8.0E-16	22672568	rs635634	PECAM1	1.89E-45
LDL-C transethnic	rs2519093	0.98	2.2E-13	23555291	rs635634	PECAM1	1.89E-45
CAD	rs2519093	0.98	1.2E-11	26343387	rs635634	PECAM1	1.89E-45
Liver enzyme levels (alkaline phosphatase)	rs579459	0.83	3.0E-123	22001757	rs635634	PECAM1	1.89E-45
E-selectin levels	rs651007	0.83	2.0E-82	20147318	rs635634	PECAM1	1.89E-45
Hematological and biochemical traits (ALP)	rs495828	0.83	4.0E-59	20139978	rs635634	PECAM1	1.89E-45
Serum alkaline phosphatase levels	rs651007	0.83	1.0E-56	24094242	rs635634	PECAM1	1.89E-45
Soluble levels of adhesion molecules (P-Selectin)	rs579459	0.83	2.0E-41	20167578	rs635634	PECAM1	1.89E-45
Blood metabolite ratios (ADSGEGDFXAEGGGVR/ADpSGEGDFXAEGGGVR)	rs649129	0.83	9.0E-37	24816252	rs635634	PECAM1	1.89E-45
Blood metabolite ratios (DSGEGDFXAEGGGVR/ADpSGEGDFXAEGGGVR)	rs495828	0.83	6.0E-34	24816252	rs635634	PECAM1	1.89E-45
Urinary metabolites (H-NMR features) (5.1825, Unknown)	rs579459	0.83	2.0E-32	24586186	rs635634	PECAM1	1.89E-45
Soluble E-selectin levels	rs579459	0.83	1.0E-29	19729612	rs635634	PECAM1	1.89E-45
Blood metabolite ratios (ADpSGEGDFXAEGGGVR/X-14304--leucylalanine)	rs579459	0.83	1.0E-28	24816252	rs635634	PECAM1	1.89E-45
Urinary metabolites (H-NMR features) (2.0525, Unknown)	rs579459	0.83	1.0E-28	24586186	rs635634	PECAM1	1.89E-45
End-stage coagulation (FVIII)	rs651007	0.83	2.0E-25	23381943	rs635634	PECAM1	1.89E-45
Blood metabolite levels (ADpSGEGDFXAEGGGVR)	rs651007	0.83	6.0E-20	24816252	rs635634	PECAM1	1.89E-45

Red blood cell traits (EA, RBCC)	rs579459	0.83	9.0E-18	23222517	rs635634	PECAM1	1.89E-45
Soluble levels of adhesion molecules (ICAM)	rs649129	0.83	1.0E-15	20167578	rs635634	PECAM1	1.89E-45
Coronary heart disease	rs579459	0.83	4.0E-14	21378990	rs635634	PECAM1	1.89E-45
vWF	rs579459	0.83	9.9E-13	23381943	rs635634	PECAM1	1.89E-45
Red blood cell count	rs495828	0.83	3.0E-12	20139978	rs635634	PECAM1	1.89E-45
Hematological and biochemical traits (Hb)	rs495828	0.83	1.0E-11	20139978	rs635634	PECAM1	1.89E-45
Hematological and biochemical traits (Ht)	rs495828	0.83	6.0E-10	20139978	rs635634	PECAM1	1.89E-45
Coronary artery disease or ischemic stroke	rs579459	0.83	2.0E-09	24262325	rs635634	PECAM1	1.89E-45
Metabolite levels (LDL)	rs651007	0.83	6.0E-09	21909109	rs635634	PECAM1	1.89E-45
Ferritin (log)|serum|(gene=FTL|FTH1)	rs651007	0.83	1.3E-08	25352340	rs635634	PECAM1	1.89E-45
FVIII Ag	rs579459	0.83	2.3E-08	23381943	rs635634	PECAM1	1.89E-45
Coronary artery disease or large artery stroke	rs579459	0.83	3.0E-08	24262325	rs635634	PECAM1	1.89E-45
Angiotensin-converting enzyme activity	rs495828	0.83	3.0E-08	20066004	rs635634	PECAM1	1.89E-45
LDL cholesterol	rs635634	1.00	2.0E-41	24097068	rs635634	SELE	9.64E-220
Cholesterol, total	rs635634	1.00	3.0E-35	24097068	rs635634	SELE	9.64E-220
E-selectin|plasma|(gene=SELE)	rs635634	1.00	5.0E-16	25147954	rs635634	SELE	9.64E-220
invasive epithelial ovarian cancer	rs635634	1.00	4.4E-09	25581431	rs635634	SELE	9.64E-220
Soluble ICAM-1	rs507666	0.99	3.0E-91	21533024	rs635634	SELE	9.64E-220
Lipid traits (LDL)	rs507666	0.99	2.0E-11	24386095	rs635634	SELE	9.64E-220
Lipid traits (TC)	rs507666	0.99	4.0E-11	24386095	rs635634	SELE	9.64E-220
Venous thromboembolism	rs2519093	0.98	8.0E-16	22672568	rs635634	SELE	9.64E-220
LDL-C transethnic	rs2519093	0.98	2.2E-13	23555291	rs635634	SELE	9.64E-220
CAD	rs2519093	0.98	1.2E-11	26343387	rs635634	SELE	9.64E-220
Liver enzyme levels (alkaline phosphatase)	rs579459	0.83	3.0E-123	22001757	rs635634	SELE	9.64E-220
E-selectin levels	rs651007	0.83	2.0E-82	20147318	rs635634	SELE	9.64E-220
Hematological and biochemical traits (ALP)	rs495828	0.83	4.0E-59	20139978	rs635634	SELE	9.64E-220
Serum alkaline phosphatase levels	rs651007	0.83	1.0E-56	24094242	rs635634	SELE	9.64E-220
Soluble levels of adhesion molecules (P-Selectin)	rs579459	0.83	2.0E-41	20167578	rs635634	SELE	9.64E-220
Blood metabolite ratios (ADSGEGDFXAEGGGVR/ADpSGEGDFXAEGGGVR	rs649129	0.83	9.0E-37	24816252	rs635634	SELE	9.64E-220

)							
Blood metabolite ratios (DSGEGDFXAEGGGVR/ADpSGEGDFXAEGGGVR)	rs495828	0.83	6.0E-34	24816252	rs635634	SELE	9.64E-220
Urinary metabolites (H-NMR features) (5.1825, Unknown)	rs579459	0.83	2.0E-32	24586186	rs635634	SELE	9.64E-220
Soluble E-selectin levels	rs579459	0.83	1.0E-29	19729612	rs635634	SELE	9.64E-220
Blood metabolite ratios (ADpSGEGDFXAEGGGVR/X-14304--leucylalanine)	rs579459	0.83	1.0E-28	24816252	rs635634	SELE	9.64E-220
Urinary metabolites (H-NMR features) (2.0525, Unknown)	rs579459	0.83	1.0E-28	24586186	rs635634	SELE	9.64E-220
End-stage coagulation (FVIII)	rs651007	0.83	2.0E-25	23381943	rs635634	SELE	9.64E-220
Blood metabolite levels (ADpSGEGDFXAEGGGVR)	rs651007	0.83	6.0E-20	24816252	rs635634	SELE	9.64E-220
Red blood cell traits (EA, RBCC)	rs579459	0.83	9.0E-18	23222517	rs635634	SELE	9.64E-220
Soluble levels of adhesion molecules (ICAM)	rs649129	0.83	1.0E-15	20167578	rs635634	SELE	9.64E-220
Coronary heart disease	rs579459	0.83	4.0E-14	21378990	rs635634	SELE	9.64E-220
vWF	rs579459	0.83	9.9E-13	23381943	rs635634	SELE	9.64E-220
Red blood cell count	rs495828	0.83	3.0E-12	20139978	rs635634	SELE	9.64E-220
Hematological and biochemical traits (Hb)	rs495828	0.83	1.0E-11	20139978	rs635634	SELE	9.64E-220
Hematological and biochemical traits (Ht)	rs495828	0.83	6.0E-10	20139978	rs635634	SELE	9.64E-220
Coronary artery disease or ischemic stroke	rs579459	0.83	2.0E-09	24262325	rs635634	SELE	9.64E-220
Metabolite levels (LDL)	rs651007	0.83	6.0E-09	21909109	rs635634	SELE	9.64E-220
Ferritin (log)|serum|(gene=FTL|FTH1)	rs651007	0.83	1.3E-08	25352340	rs635634	SELE	9.64E-220
FVIII Ag	rs579459	0.83	2.3E-08	23381943	rs635634	SELE	9.64E-220
Coronary artery disease or large artery stroke	rs579459	0.83	3.0E-08	24262325	rs635634	SELE	9.64E-220
Angiotensin-converting enzyme activity	rs495828	0.83	3.0E-08	20066004	rs635634	SELE	9.64E-220
Height	rs11599750	1.00	2.0E-13	20881960	rs11599750	IL27	1.40E-10
height	rs11599750	1.00	4.7E-13	25282103	rs11599750	IL27	1.40E-10
Bone mineral density (FNBMD)	rs7084921	0.88	9.0E-10	22504420	rs11599750	IL27	1.40E-10
Age-related hearing impairment (interaction)	rs4462272	0.62	8.0E-09	24939585	rs11599750	IL27	1.40E-10
Circulating vasoactive peptide levels (ADM)	rs2957692	0.63	1.0E-12	23381795	rs1580006	ADM	2.04E-15

Matrix_Metalloproteinase-3|CSF|(gene=MMP3)	rs471994	1.00	2.3E-21	25340798	rs471994	MMP1	2.34E-35
Matrix metalloproteinase levels	rs495366	0.72	6.0E-34	20031604	rs471994	MMP1	2.34E-35
Matrix_Metalloproteinase-3|CSF|(gene=MMP3)	rs7946057	1.00	2.4E-42	25340798	rs7946057	MMP3	1.22E-108
Matrix metalloproteinase levels	rs11225434	0.94	9.0E-29	20031604	rs7946057	MMP3	1.22E-108
Chronic obstructive pulmonary disease (severe)	rs626750	0.64	3.0E-09	24621683	rs17368659	MMP12	5.52E-97
Interleukin-18 levels	rs1834481	0.79	1.0E-08	20150558	rs75649625	IL18	1.46E-21
C-reactive protein	rs1169310	1.00	2.0E-08	18439552	rs1169306	LGALS3	6.47E-09
N-glycan levels (DG11)	rs735396	1.00	4.0E-08	21203500	rs1169306	LGALS3	6.47E-09
C-reactive protein (HA women)	rs2259816	0.99	3.0E-10	22939635	rs1169306	LGALS3	6.47E-09
Liver enzyme levels (GGT)	rs1169313	0.98	2.0E-10	18940312	rs1169306	LGALS3	6.47E-09
Gamma glutamyl transpeptidase	rs2393791	0.73	7.0E-30	21909109	rs1169306	LGALS3	6.47E-09
C-reactive protein levels	rs2393791	0.73	3.0E-10	24763700	rs1169306	LGALS3	6.47E-09
C-reactive protein and white blood cell count (CRP)	rs2393791	0.73	3.0E-09	22788528	rs1169306	LGALS3	6.47E-09
Liver enzyme levels (gamma-glutamyl transferase)	rs7310409	0.73	7.0E-45	22001757	rs1169306	LGALS3	6.47E-09
Urate levels	rs7188445	0.98	2.0E-09	23263486	rs11150189	XPNPEP2	7.00E-14
Thyroid hormone levels (TSH)	rs3813582	0.92	8.0E-18	23408906	rs11150189	XPNPEP2	7.00E-14
Thyroid hormone levels (TSH - Males)	rs3813582	0.92	6.0E-17	23408906	rs11150189	XPNPEP2	7.00E-14
Thyroid volume (thyroid volume)	rs17767419	0.92	9.0E-15	21565293	rs11150189	XPNPEP2	7.00E-14
Thyroid function	rs3813582	0.92	6.0E-10	22494929	rs11150189	XPNPEP2	7.00E-14
Vitamin B12 levels	rs492602	1.00	5.0E-17	18776911	rs492602	MMP10	7.69E-09
Cholesterol, total	rs492602	1.00	1.0E-16	24097068	rs492602	MMP10	7.69E-09
Blood metabolite ratios (ADSGEGDFXAEGGGVR/ADpSGEGDFXAEGGGVR)	rs601338	0.99	2.0E-20	24816252	rs492602	MMP10	7.69E-09
Crohn's disease	rs516246	0.99	1.0E-15	23128233	rs492602	MMP10	7.69E-09
Blood metabolite levels (ADpSGEGDFXAEGGGVR)	rs601338	0.99	3.0E-11	24816252	rs492602	MMP10	7.69E-09
Liver enzyme levels (gamma-glutamyl	rs516246	0.99	8.0E-10	22001757	rs492602	MMP10	7.69E-09

transferase)							
Folate pathway vitamin levels (vitamin B12)	rs602662	0.89	3.0E-20	19303062	rs492602	MMP10	7.69E-09
Folate pathway vitamin levels (Plasma Vitamin B12)	rs602662	0.89	2.0E-15	19744961	rs492602	MMP10	7.69E-09
Metabolic traits (SM-11 + 2 other traits)	rs503279	0.89	4.0E-20	21886157	rs492602	MMP10	7.69E-09
Blood metabolite ratios (DSGEGDFXAEGGGVR/ADpSGEGDFXAEGGGVR)	rs503279	0.89	1.0E-13	24816252	rs492602	MMP10	7.69E-09
Pubertal anthropometrics (Single Height-males)	rs281379	0.83	5.0E-08	23449627	rs492602	MMP10	7.69E-09
Liver enzyme levels (alkaline phosphatase)	rs281377	0.70	1.0E-15	22001757	rs492602	MMP10	7.69E-09
Urinary metabolites (H-NMR features) (5.2125, Fucose)	rs2287921	0.68	3.0E-36	24586186	rs492602	MMP10	7.69E-09
Retinal vascular caliber (Retinal venular caliber)	rs2287921	0.68	2.0E-25	21060863	rs492602	MMP10	7.69E-09
Urinary metabolites (H-NMR features) (5.2825, Fucose)	rs2287921	0.68	7.0E-19	24586186	rs492602	MMP10	7.69E-09
Urinary metabolites (H-NMR features) (5.2275, Fucose)	rs2287921	0.68	1.0E-12	24586186	rs492602	MMP10	7.69E-09
Dietary macronutrient intake (Fat)	rs838145	0.64	4.0E-10	23636237	rs492602	MMP10	7.69E-09
Urinary metabolites (H-NMR features) (5.2125, Fucose)	rs2287921	0.89	3.0E-36	24586186	rs33988101	LGALS3	3.57E-09
Retinal vascular caliber (Retinal venular caliber)	rs2287921	0.89	2.0E-25	21060863	rs33988101	LGALS3	3.57E-09
Urinary metabolites (H-NMR features) (5.2825, Fucose)	rs2287921	0.89	7.0E-19	24586186	rs33988101	LGALS3	3.57E-09
Urinary metabolites (H-NMR features) (5.2275, Fucose)	rs2287921	0.89	1.0E-12	24586186	rs33988101	LGALS3	3.57E-09
Crohn's disease	rs281379	0.73	7.0E-12	21102463	rs33988101	LGALS3	3.57E-09
Pubertal anthropometrics (Single Height-males)	rs281379	0.73	5.0E-08	23449627	rs33988101	LGALS3	3.57E-09
Dietary macronutrient intake (Fat)	rs838145	0.71	4.0E-10	23636237	rs33988101	LGALS3	3.57E-09
Folate pathway vitamin levels (vitamin B12)	rs602662	0.68	3.0E-20	19303062	rs33988101	LGALS3	3.57E-09
Folate pathway vitamin levels (Plasma Vitamin B12)	rs602662	0.68	2.0E-15	19744961	rs33988101	LGALS3	3.57E-09
Metabolic traits (SM-11 + 2 other traits)	rs503279	0.68	4.0E-20	21886157	rs33988101	LGALS3	3.57E-09
Blood metabolite ratios (DSGEGDFXAEGGGVR/ADpSGEGDFXAEGGGVR)	rs503279	0.68	1.0E-13	24816252	rs33988101	LGALS3	3.57E-09

Blood metabolite ratios (ADSGEGDFXAEGGGVR/ADpSGEGDFXAEGGGVR)	rs601338	0.67	2.0E-20	24816252	rs33988101	LGALS3	3.57E-09
Blood metabolite levels (ADpSGEGDFXAEGGGVR)	rs601338	0.67	3.0E-11	24816252	rs33988101	LGALS3	3.57E-09
Liver enzyme levels (gamma-glutamyl transferase)	rs516246	0.67	8.0E-10	22001757	rs33988101	LGALS3	3.57E-09
Vitamin B12 levels	rs492602	0.66	5.0E-17	18776911	rs33988101	LGALS3	3.57E-09
Cholesterol, total	rs492602	0.66	1.0E-16	24097068	rs33988101	LGALS3	3.57E-09
Metabolite levels	rs4810479	1.00	2.0E-42	22916037	rs4810479	KITLG	4.44E-11
Lipid metabolism phenotypes (HDL.large, fasting)	rs6065904	0.72	4.0E-40	19936222	rs4810479	KITLG	4.44E-11
Lipid metabolism phenotypes (L-HDL-L/M-HDL-L)	rs6065904	0.72	2.0E-31	22286219	rs4810479	KITLG	4.44E-11
Inflammatory bowel disease	rs1569723	1.00	1.0E-13	23128233	rs1569723	CD40	3.01E-49
Kawasaki disease	rs1569723	1.00	6.0E-09	22446961	rs1569723	CD40	3.01E-49
Rheumatoid arthritis	rs4810485	0.99	3.0E-09	20453842	rs1569723	CD40	3.01E-49
Rheumatoid arthritis (EA)	rs4239702	0.85	1.0E-16	24390342	rs1569723	CD40	3.01E-49
Multiple sclerosis	rs2425752	0.64	5.0E-10	21833088	rs1569723	CD40	3.01E-49

Pleiotropy of reported trait protein SNPs with findings from previously published GWAS studies. Publically available studies were investigated and associations were reported for proxy SNPs with r^2 LD above 0.6 and association P-value stronger than 5e-8. Other trait – the trait investigated in the published GWAS; Other SNP – the index SNP in the published GWAS; r2 (EUR 1000G) – linkage disequilibrium between Olink-improve study index SNP and the other SNP; Other P-value – P-value as reported in published GWAS; Pubmed ID – the pubmed ID of the published GWAS; Olink SNP – the index SNP of the Olink-improve study; Olink Trait Protein – the trait protein associated in the Olink-improve study; Olink P-value – the P-value as also reported in table 1.

Table S3

Gene	Protein Name	#samples below LOD	CV%	Incl-uded	Mean (SD)	Median (IQR)	V(G)/Vp	R2
ADM	Pro-adrenomedullin	1	0.2	yes	0.87 (0.029)	0.87 (0.85-0.88)	-4.3% -4.3%	
AGER	Advanced glycosylation end product-specific receptor	0	0.15	yes	0.66 (0.043)	0.67 (0.64-0.69)	-3.9% -3.9%	
AGRP	Agouti-related protein	0	0.17	yes	0.64 (0.049)	0.64 (0.61-0.67)	-2.5% -0.9%	
BNP	Natriuretic peptides B	1434	0.2	yes	0.15 (0.16)	0.14 (0.001-0.22)	0.8% 6%	
CASP8	Caspase-8	518		no				
CCL2	Monocyte chemoattractant protein 1	0	0.2	yes	0.6 (0.069)	0.61 (0.56-0.65)	3.3% 10.9%	
CCL20	C-C motif chemokine 20	0	0.15	yes	0.79 (0.07)	0.78 (0.74-0.83)	34.4% 74.1%	
CCL3	C-C motif chemokine 3	0	0.17	yes	0.37 (0.095)	0.38 (0.31-0.44)	6.3% 17.3%	
CCL4	C-C motif chemokine 4	0	0.13	yes	0.88 (0.036)	0.88 (0.86-0.9)	-2.9% -2.1%	
CD40	Tumor necrosis factor receptor superfamily member 5	0	0.15	yes	0.97 (0.03)	0.96 (0.95-0.98)	-4.1% -4.1%	
CD40LG	CD40 ligand	0	0.2	yes	0.77 (0.1)	0.76 (0.69-0.84)	-3.5% -3.1%	
CHI3L1	Chitinase-3-like protein 1	0	0.16	yes	0.8 (0.066)	0.8 (0.76-0.84)	1.1% 6.6%	
CSF1	Macrophage colony-stimulating factor 1	0	0.15	yes	0.95 (0.015)	0.95 (0.94-0.96)	7.6% 20.1%	
CSTB	Cystatin-B	0	0.16	yes	0.74 (0.04)	0.74 (0.71-0.76)	-3.7% -3.7%	
CTSD	Cathepsin D	1	0.15	yes	0.87 (0.03)	0.87 (0.85-0.89)	3.4% 11.3%	

CTSL1	Cathepsin L1	0	0.16	yes	0.78 (0.034)	0.78 (0.75-0.8)	-4.1% -4.1%
CX3CL1	Fractalkine	0	0.26	yes	0.73 (0.045)	0.73 (0.7-0.76)	-0.6% 2.9%
CXCL1	Growth-regulated alpha protein	0	0.15	yes	0.79 (0.093)	0.8 (0.73-0.86)	1.4% 7.0%
CXCL16	C-X-C motif chemokine 16	0	0.17	yes	0.58 (0.041)	0.58 (0.56-0.61)	10.9% 26.6%
CXCL6	C-X-C motif chemokine 6	0	0.15	yes	0.78 (0.066)	0.78 (0.74-0.83)	-1.8% 0.5%
DKK1	Dickkopf-related protein 1	0	0.13	yes	0.72 (0.063)	0.72 (0.68-0.77)	-0.8% 2.7%
EGF	Pro-epidermal growth factor	0	0.15	yes	0.65 (0.15)	0.66 (0.55-0.76)	-4.1% -4.1%
ESM1	Endothelial cell-specific molecule 1	0	0.17	yes	0.58 (0.053)	0.58 (0.54-0.61)	-0.1% 4.1%
F2R	Proteinase-activated receptor 1	0	0.18	yes	0.84 (0.034)	0.84 (0.82-0.86)	3.4% 11.4%
F3	Tissue factor	0	0.18	yes	0.81 (0.03)	0.81 (0.79-0.83)	-3.8% -3.8%
FABP4	Fatty acid-binding protein, adipocyte	4	0.2	yes	0.48 (0.096)	0.49 (0.43-0.55)	2.5% 9.4%
FAS	Tumor necrosis factor receptor superfamily member 6	0	0.13	yes	0.89 (0.023)	0.9 (0.88-0.91)	8% 21%
FGF23	Fibroblast growth factor 23	0	0.17	yes	0.46 (0.08)	0.47 (0.41-0.51)	-3.9% -3.9%
FIGF	Vascular endothelial growth factor D	0	0.19	yes	0.84 (0.038)	0.84 (0.82-0.86)	6.1% 16.6%
FST	Follistatin	0	0.18	yes	0.74 (0.041)	0.74 (0.71-0.77)	0.7% 5.7%
GAL	Galanin peptides	0	0.15	yes	0.72 (0.06)	0.73 (0.69-0.76)	3.6% 11.6%
GDF15	Growth/differentiation factor	0	0.13	yes	0.99	0.99	-0.7%

	15				(0.027)	(0.97-1.0)	2.8%
GH1	Somatotropin	0	0.15	yes	0.92 (0.11)	0.93 (0.84-1.0)	-3.7% -3.2%
HAVCR1	Hepatitis A virus cellular receptor 1	0	0.16	yes	0.78 (0.06)	0.78 (0.74-0.82)	-2.7% -0.8%
HBEGF	Proheparin-binding EGF-like growth factor	0	0.18	yes	0.67 (0.042)	0.67 (0.65-0.7)	11% 26.6%
HGF	Hepatocyte growth factor	0	0.13	yes	0.83 (0.03)	0.83 (0.82-0.85)	2.8% 10%
HSPB1	Heat shock protein beta-1	303	0.16	yes	0.34 (0.24)	0.32 (0.14-0.53)	-0.7% 2.8%
IKBKG	NF-kappa-B essential modulator	163	0.22	yes	0.43 (0.22)	0.43 (0.29-0.58)	1.3% 7%
IL16	Pro-interleukin-16 [Cleaved into: Interleukin-16	0	0.2	yes	0.63 (0.053)	0.63 (0.6-0.67)	6.4% 17.3%
IL18	Interleukin-18	0	0.12	yes	1.0 (0.026)	1.0 (0.99-1.01)	15.6% 36.1%
IL1RL1	Interleukin-1 receptor-like 1	0	0.13	yes	0.54 (0.072)	0.55 (0.5-0.59)	6.4% 17.3%
IL1RN	Interleukin-1 receptor antagonist protein	0		no			
IL27	Interleukin-27 subunit alpha	0	0.2	yes	0.42 (0.071)	0.43 (0.38-0.47)	-3.3% -2.3%
IL4	Interleukin-4	2559		no			
IL6	Interleukin-6	0	0.13	yes	0.7 (0.078)	0.69 (0.65-0.74)	-1.3% 1.7%
IL6R	Interleukin-6 receptor subunit alpha	0	0.14	yes	0.85 (0.029)	0.85 (0.83-0.87)	-3.7% -3.7%
IL8	Interleukin-8	1	0.15	yes	0.73 (0.055)	0.73 (0.69-0.76)	3.6% 11.9%
ITGB1BP2	Integrin beta-1-binding protein 2	1788		no			
KITLG	Kit ligand	0	0.12	yes	0.87 (0.031)	0.87 (0.85-0.89)	2.9% 10.3%

KLK11	Kallikrein-11	0	0.19	yes	0.72 (0.039)	0.72 (0.7-0.75)	12.7% 30.3%
KLK6	Kallikrein-6	0	0.15	yes	0.83 (0.028)	0.83 (0.81-0.85)	-1.2% 1.7%
LEP	Leptin	82	0.18	yes	0.47 (0.16)	0.5 (0.39-0.59)	-4.2% -4.2%
LGALS3	Galectin-3	0	0.16	yes	0.78 (0.03)	0.78 (0.76-0.8)	-4% -4%
MB	Myoglobin	3	0.17	yes	0.74 (0.048)	0.74 (0.71-0.77)	7.6% 19.6%
MMP1	Interstitial collagenase	7	0.19	yes	0.49 (0.19)	0.51 (0.4-0.61)	54.8% 114.2%
MMP10	Stromelysin-2	0	0.16	yes	0.82 (0.043)	0.82 (0.79-0.85)	1.6% 7.6%
MMP12	Macrophage metalloelastase	0	0.14	yes	0.85 (0.043)	0.85 (0.82-0.88)	-3.8% -3.8%
MMP3	Stromelysin-1	20	0.13	yes	-0.027 (0.3)	0.031 (-0.15-0.17)	6.3% 17.3%
MMP7	Matrilysin	1	0.2	yes	0.83 (0.11)	0.86 (0.76-0.92)	5.9% 16.5%
MPO	Myeloperoxidase	0	0.15	yes	0.57 (0.04)	0.57 (0.54-0.59)	-4.1% -4.1%
MUC16	Mucin-16	1	0.22	yes	0.71 (0.076)	0.72 (0.67-0.76)	-4.2% -4.2%
NGF	Beta-nerve growth factor	510	0.18	yes	-0.059 (0.13)	-0.064 (-0.16-0.029)	0.5% 5.5%
NPPB	Pro-Natriuretic peptides B	135	0.2	yes	0.56 (0.12)	0.57 (0.48-0.65)	7.2% 19.1%
OLR1	Oxidized low-density lipoprotein receptor 1	0	0.17	yes	0.67 (0.059)	0.66 (0.63-0.7)	7.1% 18.6%
PAPPA	Pappalysin-1	60	0.17	yes	0.32 (0.14)	0.34 (0.25-0.41)	4.6% 13.8%
PDGFB	Platelet-derived growth factor subunit B	0	0.13	yes	0.82 (0.091)	0.83 (0.77-0.88)	-3.9% -3.9%

PECAM1	Platelet endothelial cell adhesion molecule	0	0.2	yes	0.78 (0.032)	0.77 (0.75-0.8)	-3.1% -2.6%
PGF	Placenta growth factor	1	0.15	yes	0.88 (0.024)	0.88 (0.87-0.9)	6.5% 17.7%
PLAT	Tissue-type plasminogen activator	0	0.24	yes	0.89 (0.037)	0.89 (0.87-0.91)	-4% -4%
PLAUR	Urokinase plasminogen activator surface receptor	0	0.14	yes	1 (0.015)	1.0 (0.99-1.01)	4% 12.5%
PRL	Prolactin	3	0.15	yes	0.62 (0.081)	0.62 (0.57-0.67)	8.8% 22.1%
PTX3	Pentraxin-related protein PTX3	914		no			
REN	Renin	0	0.17	yes	0.87 (0.056)	0.87 (0.83-0.91)	0.9% 6%
RETN	Resistin	0	0.16	yes	0.84 (0.036)	0.84 (0.82-0.87)	21% 47%
RNASE3	Eosinophil cationic protein	2	0.16	yes	0.77 (0.053)	0.77 (0.74-0.8)	-2.8% -1.8%
S100A12	Protein S100-A12	669		no			
SELE	E-selectin	0	0.16	yes	0.77 (0.052)	0.77 (0.73-0.8)	5.5% 15.6%
SELPLG	P-selectin glycoprotein ligand 1	1081		no			
SIRT2	NAD-dependent protein deacetylase sirtuin-2	585		no			
SPON1	Spondin-1	0	0.17	yes	0.69 (0.033)	0.69 (0.67-0.71)	-3.6% -3.6%
SRC	Proto-oncogene tyrosine-protein kinase Src	0	0.21	yes	0.74 (0.12)	0.77 (0.65-0.85)	-2.8% -1.5%
TEK	Angiopoietin-1 receptor	0	0.14	yes	0.82 (0.023)	0.82 (0.8-0.83)	-4.2% -4.2%
THBD	Thrombomodulin	0	0.16	yes	1.0 (0.018)	1.0 (0.99-1)	2% 8.4%
TNFRSF10B	Tumor necrosis factor	0	0.24	yes	0.19	0.26	0.9%

	receptor superfamily member 10B				(0.21)	(0.086-0.33)	6.4%
TNFRSF11B	Osteoprotegerin	1	0.13	yes	1 (0.023)	1 (0.99-1)	10.7% 26.2%
TNFRSF1A	Tumor necrosis factor receptor superfamily member 1A	0		no			
TNFRSF1B	Tumor necrosis factor receptor superfamily member 1B	6	0.2	yes	0.63 (0.071)	0.64 (0.59-0.68)	6.6% 17.9%
TNFSF10	Tumor necrosis factor ligand superfamily member 10	0	0.16	yes	0.35 (0.11)	0.36 (0.28-0.43)	-3.9% -3.9%
TNFSF11	RANK ligand	0	0.15	yes	1.1 (0.014)	1.1 (1.1-1.1)	13.6% 32.1%
TNFSF14	Tumor necrosis factor ligand superfamily member 14	0	0.15	yes	0.73 (0.038)	0.73 (0.7-0.75)	3.5% 11.4%
VEGFA	Vascular endothelial growth factor A	0	0.13	yes	1 (0.017)	1 (1-1)	0.5% 5.4%
XPNPEP2	Xaa-Pro aminopeptidase 2	462	0.19	yes	0.47 (0.14)	0.48 (0.33-0.58)	-0.1% 4.1%

Overview of all 92 measured proteins, with quality control parameters, descriptive statistics and heritability estimates. All descriptive statistics are reported on the log10-transformed data that was used for analysis; *#samples below LOD* – the number of samples below limit of detection; *CV%* - coefficient of variation; *Included* – final choice on inclusion in analysis; *Mean (SD)* – mean and standard-deviation; *Median (IQR)* – median and inter-quartile range; *V(G)/Vp*– The GCTA calculated narrow-sense heritability. Missing values correspond to GCTA invertable matrix errors - meaning that too few samples were available for the V(G)/Vp estimation ; R2 – description still missing.