Title
Genetic Regulation of Adipose Gene Expression and Cardio-Metabolic Traits

Journal
American Journal of Human Genetics, 100(3)

ISSN
0002-9297

Authors
Civelek, Mete
Wu, Ying
Pan, Calvin
et al.

Publication Date
2017-03-01

DOI
10.1016/j.ajhg.2017.01.027

Peer reviewed

Genetic Regulation of Adipose Gene Expression and Cardio-Metabolic Traits

Mete Civelek,[1,2,3,13] Ying Wu,[4,13] Calvin Pan,[1] Chelsea K. Raulerson,[4] Arthur Ko,[5,6] Aiqing He,[7] Charles Tilford,[7] Niyas K. Saleem,[8] Alena Stančáková,[8] Laura J. Scott,[9] Christian Fuchsberger,[9] Heather M. Stringham,[9] Anne U. Jackson,[9] Narisu Narisu,[10] Peter S. Chines,[10] Kerrin S. Small,[11] Johanna Kuusisto,[8] Brian W. Parks,[1] Päivi Pajukanta,[5,6] Todd Kirchgessner,[7] Francis S. Collins,[10] Peter S. Gargalovic,[7] Michael Boehnke,[9] Markku Laakso,[8,14] Karen L. Mohlke,[4,14,*] and Aldons J. Lusis[1,5,6,12,14,*]

Subcutaneous adipose tissue stores excess lipids and maintains energy balance. We performed expression quantitative trait locus (eQTL) analyses by using abdominal subcutaneous adipose tissue of 770 extensively phenotyped participants of the METSIM study. We identified *cis*-eQTLs for 12,400 genes at a 1% false-discovery rate. Among an approximately 680 known genome-wide association study (GWAS) loci for cardio-metabolic traits, we identified 140 coincident *cis*-eQTLs at 109 GWAS loci, including 93 eQTLs not previously described. At 49 of these 140 eQTLs, gene expression was nominally associated ($p < 0.05$) with levels of the GWAS trait. The size of our dataset enabled identification of five loci associated ($p < 5 \times 10^{-8}$) with at least five genes located >5 Mb away. These *trans*-eQTL signals confirmed and extended the previously reported *KLF14*-mediated network to 55 target genes, validated the *CIITA* regulation of class II MHC genes, and identified *ZNF800* as a candidate master regulator. Finally, we observed similar expression-clinical trait correlations of genes associated with GWAS loci in both humans and a panel of genetically diverse mice. These results provide candidate genes for further investigation of their potential roles in adipose biology and in regulating cardio-metabolic traits.

Introduction

Genome-wide association studies (GWASs) have identified many loci for complex metabolic and cardiovascular traits, yet the underlying genes and mechanisms by which they affect disease remain poorly characterized.[1,2] The genetic analysis of gene expression by identification of expression quantitative trait loci (eQTLs) in relevant tissues has proven useful to predict candidate genes at GWAS loci and biological pathways that are perturbed in affected individuals.[3–6] Subcutaneous adipose tissue serves as a buffering system for lipid energy balance, particularly fatty acids,[7,8] and might play a protective role in metabolic and cardiovascular disease risk.[9]

Subcutaneous adipose eQTL studies have implicated genes involved in obesity and metabolic traits.[10–13] Recent GWASs for type 2 diabetes (T2D), cholesterol and triglyceride levels, body mass index, waist-hip ratio, and adiponectin have reported subcutaneous adipose eQTLs that are coincident with specific GWAS loci.[14–17] Similarly, a recent large GWAS for waist-hip ratio identified loci that were enriched for genes expressed in subcutaneous adipose tissue and for putative regulatory elements in adipocyte nuclei.[16] Many GWAS loci for these traits do not yet have clear candidate genes, in part because of the limited statistical power of existing eQTL studies.

trans-eQTL associations between variants and transcripts located far from each other or on different chromosomes can identify downstream disease genes, including those not implicated by GWASs. Identifying these distant relationships is difficult because of the multiple testing burden in humans. Studies of natural variation in mice have identified a number of "hotspot" loci associated with *trans* regulation of genes and clinical traits.[18,19] One of the first reported human *trans*-acting eQTLs involved the KLF14 transcription factor in adipose tissue.[20] The locus associated with T2D and HDL-cholesterol levels showed a *cis*-acting association with expression of *KLF14* and ten distal genes.[20] Studies of gene expression in circulating monocytes or whole-blood cells have also provided evidence of *trans* regulation of gene expression with linkage to traits relevant to lipid metabolism, type 1 diabetes, hypertension, celiac disease, and cancer.[21,22]

In light of the widespread use of mice to help validate and gain mechanistic understanding of genes in GWAS loci, commonalities and differences in the regulatory

[1]Department of Medicine, University of California, Los Angeles, Los Angeles, CA 90095, USA; [2]Center for Public Health Genomics, University of Virginia, Charlottesville, VA 22908, USA; [3]Department of Biomedical Engineering, University of Virginia, Charlottesville, VA 22908, USA; [4]Department of Genetics, University of North Carolina, Chapel Hill, NC 27599, USA; [5]Department of Human Genetics, University of California, Los Angeles, Los Angeles, CA 90095, USA; [6]Molecular Biology Institute, University of California, Los Angeles, Los Angeles, CA 90095, USA; [7]Bristol-Myers Squibb, Pennington, NJ 08534, USA; [8]Department of Medicine, University of Eastern Finland and Kuopio University Hospital, Kuopio 70210, Finland; [9]Department of Biostatistics and Center for Statistical Genetics, University of Michigan, Ann Arbor, MI 48109, USA; [10]National Human Genome Research Institute, National Institutes of Health, Bethesda, MD 20892, USA; [11]Department of Twin Research and Genetic Epidemiology, School of Medicine, King's College London, London SE1 7EH, UK; [12]Department of Microbiology, Immunology and Molecular Genetics, University of California, Los Angeles, Los Angeles, CA 90095, USA
[13]These authors contributed equally to this work
[14]These authors contributed equally to this work
*Correspondence: mohlke@med.unc.edu (K.L.M.), jlusis@mednet.ucla.edu (A.J.L.)
http://dx.doi.org/10.1016/j.ajhg.2017.01.027.

networks of mice and humans are of clear relevance for studies of common diseases.[6] A previous comparison between mouse and human adipose transcriptional networks detected a shared core-network module enriched for genes involved in the inflammatory and immune response causally associated with obesity-related traits.[10] Recent analysis of DNase I hypersensitive sites and occupancy profiles of transcription-factor binding in humans and mice suggests the preservation of similar regulatory mechanisms for adipose gene expression in both species, and this preservation could be leveraged to inform disease pathways.[23,24]

We describe here the analysis of gene expression in 770 subcutaneous adipose samples from Metabolic Syndrome in Men (METSIM), a study of 10,197 men, 45–73 years of age, living in the Kuopio area of Finland. Study data include dense genotypes and extensive metabolic and cardiovascular traits, such as plasma lipids, inflammatory markers, glycemic traits, and anthropometric traits.[25] We identified *cis*-eQTLs at GWAS loci for cardio-metabolic traits, as well as *trans*-eQTL hotspots with at least five distant target genes. We also identified associations of gene expression with human clinical traits and compared our results with those of analogous studies in a set of 120 inbred strains of mice.

Subjects and Methods

METSIM Study Participants and Sample Characteristics

We analyzed samples from 770 males who are part of the METSIM study.[25] The Ethics Committee of the Northern Savo Hospital District approved this study, and all participants gave written informed consent. The population-based METSIM study included 10,197 men, aged 45–73 years and randomly selected from the population register of Kuopio town in eastern Finland (population 95,000). Every participant had a 1 day outpatient visit to the Clinical Research Unit at the University of Kuopio, including an interview on the history of previous diseases and current drug treatment and an evaluation of glucose tolerance and cardiovascular risk factors. After 12 hr of fasting, a 2 hr oral 75 g glucose tolerance test was performed, and the blood samples were drawn at 0, 30, and 120 min. Plasma glucose was measured by enzymatic hexokinase photometric assay (Konelab Systems reagents; Thermo Fischer Scientific), and insulin and pro-insulin were determined by immunoassay (ADVIA Centaur Insulin IRI no. 02230141; Siemens Medical Solutions Diagnostics). Evaluation of insulin sensitivity (Matsuda index) and insulin secretion (calculated from area under the curve between 0 and 30 min of glucose tolerance test with the formula $InsAUC_{0–30}/GlucAUC_{0–30}$) have been previously described.[25,26] Plasma levels of lipids were determined via enzymatic colorimetric methods (Konelab System reagents, Thermo Fisher Scientific). Plasma adiponectin was measured with the Human Adiponectin Elisa Kit (Linco Research), C-reactive protein with high sensitive assay (Roche Diagnostics GmbH, Mannheim, Germany), and interleukin 1 receptor agonist with immunoassay (ELISA, Quantikine DRA00 Human IL-1RA, R&D Systems). Serum creatinine was measured by the Jaffe kinetic method (Konelab System reagents, Thermo Fisher Scientific) and was used for calculating the glomerular filtration rate. Height and weight were measured to the nearest 0.5 cm and 0.1 kg, respectively. Waist circumference (at the midpoint between the lateral iliac crest and lowest rib) and hip circumference (at the level of the trochanter major) were measured to the nearest 0.5 cm. Body composition was determined by bioelectrical impedance (RJL Systems) in participants in the supine position. The characteristics of the study participants are shown in Table S1. 770 participants were recruited for adipose-tissue needle biopsies. 61 participants were diagnosed with impaired glucose tolerance, and 27 participants had newly diagnosed type 2 diabetes at the time of the tissue collection.

Genotyping and Imputation

Genotyping of METSIM samples was performed with the Illumina HumanOmniExpress BeadChip array and the Illumina HumanCoreExome at the Center for Inherited Disease Research. Markers with poor mapping, no founder genotypes, call rate < 95%, deviation from Hardy-Weinberg equilibrium ($p < 10^{-6}$), or more than two alleles were removed from subsequent imputation. We carried out genotype imputation of the 681,789 directly genotyped variants that passed quality control by applying the Markov Chain Haplotyping algorithm (MaCH) and the reference panel from the Haplotype Reference Consortium (see Web Resources). After imputation, variants were filtered on the basis of imputation quality (MaCH $r^2 > 0.3$) and minor-allele frequency (MAF ≥ 0.01). 7,677,146 variants were retained for subsequent analysis.

Gene Expression Profiling

Total RNA from METSIM participants was isolated from adipose tissue via the QIAGEN miRNeasy kit, according to the manufacturer's instructions. RNA integrity numbers (RINs) were assessed with the Agilent Bioanalyzer 2100 instrument, and 770 samples with RIN > 7.0 were used for transcriptional profiling. Expression profiling with the Affymetrix U219 microarray was performed at the Department of Applied Genomics at Bristol-Myers Squibb according to the manufacturer's protocols. The probe sequences were re-annotated so that probes that mapped to multiple locations, contained variants with MAF > 0.01 in the 1000 Genomes Project European samples, or did not map to known transcripts on the basis of the RefSeq (version 59) and Ensembl (version 72) databases were removed; 6,199 probe sets were removed in this filtering step. For subsequent analyses, we used 43,145 probe sets that represent 18,155 unique genes. The microarray image data were processed with the Affymetrix GCOS algorithm via the robust multiarray average (RMA) method for determination of the specific hybridizing signal for each gene.

PEER Factor Analysis

We applied the probabilistic estimation of expression residuals (PEER) method[27] to infer and account for complex non-genetic factors affecting gene expression levels. This method is designed to detect the maximum number of *cis*-eQTLs. To optimize the discovery of *trans*-eQTLs within the same analysis, we performed PEER analysis by examining 10–50 inferred factors (Nk) at increments of five factors. We then used Matrix eQTL[28] to assess the genetic association with inverse normal-transformed PEER-processed residuals from RMA-normalized expression data. The numbers of *cis*- and *trans*-eQTLs obtained at different Nk levels are shown in Table S2. We examined variants on chromosome 7, including rs4731702 at the known master regulator *KLF14*,[20] to determine the number of *trans*-eQTL target genes by using various numbers of PEER factors. We selected Nk = 35 as a single analysis to maximize the number of target genes at this locus; this

threshold captured the 94.8% of *cis*-eQTLs identified with 50 PEER factors. For downstream eQTL mapping, we used the inverse normal-transformed PEER-processed residuals after accounting for 35 factors.

eQTL Mapping

We performed eQTL mapping in 770 MESTIM individuals by using both FaST-LMM[29] and EPACTS and implementing a linear mixed model to account for the population structure among the samples.[30] Genotype dosages from all autosomal chromosomes and expression data that had passed the aforementioned quality-control measures were used. For FaST-LMM implementation, to improve power[31] when testing all the variants on chromosome N for association, we constructed the kinship matrix by using the variants from all other chromosomes besides N. This procedure allowed us to include the variant being tested for association in the regression equation only once. Results obtained with FaST-LMM and EPACTS were similar. Results obtained from EPACTS analysis were used in the identification of coincident eQTL and GWAS signals. Results from the FaST-LMM analysis are available on our website.

eQTLs were defined as *cis* (local) if the peak association was within 1 Mb on either side of the exon boundaries of the gene or as *trans* (distal) if the peak association was at least 5 Mb outside of the exon boundaries. We used all association p values in the *cis* region to estimate the false-discovery rate (FDR-qvalue) by using the qvalue package in R. Variants with association p values $< 2.46 \times 10^{-4}$ corresponding to 1% FDR were considered significant. Considering the large number of analyses we performed to calculate *trans*-eQTL associations, we used the conservative Bonferroni-corrected $p < 1.51 \times 10^{-13}$ (0.05/[7.67 million variants × 43,145 probe sets]) to identify *trans*-eQTLs. To detect possible *trans*-eQTL hotspots, we report variants that are associated with at least five genes at a more liberal threshold of $p < 5 \times 10^{-8}$. These hotspots were visualized with Circos-0.66.[32] LocusZoom was used for the regional visualization of eQTL results on the basis of linkage disequilibrium (LD) ascertained from the 770 METSIM samples.[33]

Heritability Calculations

Heritability was estimated via a linear mixed model with the GCTA software.[34,35] In this approach, gene expression phenotypes are assumed to be generated by genetic and environmental components. The assumption behind the linear-mixed-model approach is that the covariance of the genetic component of the phenotypic data is proportional to the kinship or genetic similarity matrix between the individuals. The analysis provides estimates of σ^2_u and σ^2_e, the variances corresponding to the genetic and environmental components, respectively. The heritability is then the fraction of the variance accounted for by the genetics

$$h^2 = \frac{\mathrm{var}(u)}{\mathrm{var}(u) + \mathrm{var}(e)},$$

where var(u) and var(e) are, respectively, the genetic and residual variance components estimated by the restricted-maximum-likelihood approach using related individuals, and is computed for each probe set. A standardized kinship matrix, which has a mean of 1 along the diagonal, is used for these estimates so that they are consistent with the classical definition of heritability.[36]

Evaluation of eQTLs in Another Adipose eQTL Cohort

eQTLs identified in METSIM and MuTHER cohorts[37] were compared in each study reciprocally. Of the ~5 million significant variant-probe pairs identified in the METSIM study for *cis* associations (Table S3), 453,484 had an exact match for the variant in a corresponding variant-probe pair in the MuTHER study. A majority of the genes are represented by multiple probes on the Affymetrix microarray used in the METSIM study and the Illumina microarray used in the MuTHER study. Therefore, we required at least 50% of the variant-probe pairs for a given gene to have the same direction of effect in both studies at various p value thresholds ($p < 5 \times 10^{-2}$ to 5×10^{-6}) in the MuTHER study. Of the 11,053 significant variant-probe pairs identified in the METSIM study for *trans* associations (Table S3), 2,312 had an exact match for the variant in a corresponding variant-probe pair in the MuTHER study. We report the METSIM-detected *trans*-eQTLs for which the exact variant-gene pair exists and shows the same direction of effect in MuTHER at several p value thresholds ($p < 5 \times 10^{-2}$ to 5×10^{-6}) for which MuTHER data were available. *cis*- and *trans*-eQTLs detected in the MuTHER study were compared to those in the METSIM study according to the same criteria.

For the three *trans*-eQTL hotspots for which the peak variant or its LD proxy was available in the MuTHER dataset, we were able to find corresponding variant-probe pairs for 83 of the 106 genes. We report the variant-gene pairs with the same direction of effect size at $p < 0.05$ in the MuTHER study.

cis-eQTLs at GWAS Loci

We focused on the GWAS index variants reported to be associated at genome-wide significance ($p < 5 \times 10^{-8}$) with any of 91 cardiometabolic diseases and related quantitative traits (Table S4). We downloaded these variants from the National Human Genome Research Institute–European Bioinformatics Institute GWAS Catalog (September 22, 2014) and included additional variants.[16,17] We estimated the number of loci for each trait group on the basis of the names of "reported gene(s)" and/or "mapped genes" shown in the catalog. Within the variants available in the METSIM eQTL data, we initially examined 115 variants at 68 loci for T2D, 114 variants at 68 loci for glycemic traits, 325 variants at 205 loci for obesity and related traits, 410 variants at 213 loci for lipid traits, 79 variants at 60 loci for metabolic syndrome, and 249 variants at 187 loci for other broadly defined cardiovascular risk traits. After removing duplicate variants present in at least one disease/trait group, we further examined the remaining 1,221 GWAS variants at approximately 680 loci for association with local gene expression levels.

For significant variant-transcript associations (FDR < 1%), if multiple probe sets mapped to the same transcript, we retained the probe sets for which the lead eSNP (the variant that exhibited the strongest evidence of association with the expression level of the corresponding transcript) had the largest pairwise LD r^2 estimate with the GWAS variant; LD was calculated from the 770 METSIM samples. If two or more highly correlated variants (LD $r^2 > 0.8$) were associated with the same transcript, we reported the variant with the largest LD r^2 with the lead eSNP. For two variants (pairwise LD $r^2 = 0.79$) associated with different *GALNT2* probe sets, we kept the variant exhibiting higher LD r^2 with the GWAS index variant.

We further examined whether each eQTL was coincident with the GWAS signal. We defined GWAS-coincident eQTLs as loci with pairwise LD $r^2 > 0.8$ between the GWAS index variant and

the lead eSNP. To evaluate association between the GWAS variant and the lead eQTL variant at each locus, we performed reciprocal conditional analyses; we tested association between the GWAS variant and transcript level when the lead eSNP was included in the model, and vice versa.

For 61 cardio-metabolic traits with publically available GWAS summary results[14,16,17,38–48] (Table S5), we applied the summary-data-based Mendelian randomization (SMR) method to propose relevant genes at the GWAS loci.[49] We performed a transcriptome-wide association for 61 cardio-metabolic traits and 24,383 probe sets with *cis*-eQTLs (<1 Mb) and focused on GWAS loci ($p < 5 \times 10^{-8}$). Probe sets with $p_{SMR} < 2.1 \times 10^{-6}$ (0.05/24,383 probe sets) were then tested for pleiotropy via heterogeneity in dependent instruments (HEIDI). Probe sets with $p_{HEIDI} \geq 0.05$ were identified as potential gene targets of GWAS loci.

We also quantified the uncertainty in the inference that the *cis*-eQTL genes are causally mediating the association between the GWAS loci with cardio-metabolic traits by using the causal inference test (CIT).[50] We used the GWAS index variants, expression level of the *cis*-eQTL genes, and the inverse normalized GWAS trait measured in METSIM to infer a causal relationship wherein for gene expression mediates the association between GWAS variant and trait when $p_{CIT} < 0.05$.[50]

Association between Gene Expression Level and Phenotypic Traits

We conducted regression analyses to evaluate the association between gene expression and 23 cardio-metabolic-related traits in up to 770 METSIM individuals. The RMA-normalized expression levels were inverse normal transformed after age and BMI were accounted for. We used non-PEER-corrected expression levels for this analysis because correction for PEER factors can remove broadly acting phenotypic effects on gene expression. Eight of the 35 PEER factors had significant correlation with BMI ($|r| = 0.10$ to 0.50, $p = 2.8 \times 10^{-3}$ to 3.5×10^{-50}). For genes coincident with GWAS loci, in addition to non-PEER-corrected associations, we calculated the association between PEER-corrected expression levels and traits. The 23 phenotypic traits examined included five obesity-related traits (BMI, waist-hip ratio, waist circumference, hip circumference, and fat-free mass), seven glycemic traits (glucose, insulin, proinsulin, HOMA-β, Hb1Ac, Matsuda index, and adiponectin), six lipid-related traits (total cholesterol, low-density lipoprotein cholesterol [LDL-C], high-density lipoprotein cholesterol [HDL-C], triglycerides, total fatty acids, and free fatty acids), two inflammatory traits (C-reactive protein and interleukin 1 receptor antagonist), two blood-pressure traits (systolic and diastolic blood pressure, respectively), and one kidney-function trait (glomerular filtration rate). The phenotypic traits were adjusted for age and BMI before inverse normal transformation. When examining the association between gene expression and BMI, we applied no adjustment for BMI for either the phenotypic trait (BMI) or gene expression level. We used FDR < 1% (equivalent $p < 6.5 \times 10^{-4}$) to define a significant association between gene expression level and phenotypic-trait level. Genes were clustered via hierarchical clustering based on Euclidian distance as implemented in the *heatmap2* function of the *gplots* package in R.

Mouse Expression and Phenotypic-Trait Data

The mouse studies using the Hybrid Mouse Diversity Panel (HMDP) have been described previously (GEO: GSE42890).[51,52] In brief, mice from 120 strains were obtained from The Jackson Laboratory and were bred at the University of California, Los Angeles. Mice were maintained on a chow diet (Ralston Purina Company) until 8 weeks of age, when they were either continued on a chow diet (n = 185 mice) or given a high-fat, high-sucrose diet (Research Diets-D12266B) for an additional 8 weeks (n = 227 mice). Body composition analysis was performed by nuclear magnetic resonance with a Bruker Minispec. Retro-orbital blood was collected under isoflurane anesthesia after mice fasted for 4 hr. Plasma lipids, insulin, glucose, and HOMA-IR were determined as previously described.[52,53] The animal protocol for the study was approved by the Institutional Care and Use Committee (IACUC) at the University of California, Los Angeles. Total RNA from flash-frozen epididymal adipose samples from 228 male mice was hybridized to Affymetrix HT_MG-430A arrays and scanned according to standard Affymetrix protocols. To reduce the chances of spurious association results, we performed RMA normalization after removing all individual probes with variants and all probe sets containing eight or more variant-containing probes, which resulted in 22,416 remaining probe sets. Correlations of non-PEER-corrected expression levels with phenotypic traits were calculated with the biweight midcorrelation, which is robust to outliers.[54]

Results

Genetic Variants Associated with Gene Expression in Subcutaneous Adipose Tissue

To identify genetic loci associated with transcript abundance in abdominal subcutaneous adipose tissue, we studied 770 extensively phenotyped men from the METSIM study. We analyzed ~7.67 million variants and abundance of 43,145 probe sets corresponding to 18,155 unique genes. The mean narrow-sense heritability, h^2, of the probe sets was 0.27, suggesting significant genetic effects on adipose gene expression (Figure S1).

eQTL mapping identified *cis*-eQTLs (<1 Mb) for 12,400 genes at a 1% FDR ($p < 2.46 \times 10^{-4}$) (Figure S2 and Table S3). The larger number of adipose eQTLs observed in comparison to that in previous adipose eQTL studies might be due to the larger sample size, denser imputation, and/or analysis of a larger number of transcripts; some differences might be due to different microarray platforms, statistical methodology, and p value thresholds (Table S6).[10,11,37]

30% of the *cis*-eQTLs discovered in METSIM at $p < 5 \times 10^{-6}$ showed consistent allelic direction of effect in the MuTHER study, and 79.1% of the *cis*-eQTLs reported in MuTHER at this threshold showed a consistent direction of effect in METSIM (Table S7).

Coincidence of *cis*-eQTLs and GWAS Loci

Hundreds of GWAS loci have been reported for cardio-metabolic diseases and related traits.[55] Given the value of identifying candidate genes at GWAS loci and the observation that many eQTLs are shared across tissues,[56] we investigated loci for a broad set of 91 cardio-metabolic diseases or traits (Table S4). Among the 1,221 GWAS signals ($p < 5 \times 10^{-8}$), we detected 944 initial unique

GWAS variant and eQTL gene pairs (FDR < 1%, equivalent $p < 2.4 \times 10^{-4}$) (Table S8).

On the basis of pairwise LD $r^2 > 0.8$ between the GWAS variant and the variant that exhibited the strongest association with the gene expression level (lead eSNP), 140 (15%) of the 944 *cis*-eQTLs appear to be coincident with the GWAS signals. The coincidence is supported by reciprocal conditional analysis: after conditioning was performed on each lead eSNP, no GWAS variant remained significant (FDR < 1%), and after conditioning on each GWAS variant, 124 eSNP signals were no longer significant, and 16 were strongly attenuated ($\Delta\log_{10}(p)$ of 8.5 to 112) (Table S8). Of the 140 *cis*-eQTLs at GWAS loci, 93 (66.4%) were not previously reported by large-scale GWASs that interrogated available *cis*-eQTLs[14–17,38,45,48,57–68] and 50 showed consistent direction of allelic effect at $p < 0.05$ in the MuTHER study. Table 1 shows 29 eQTLs for glycemic, obesity, and lipid traits at the LD threshold of $r^2 > 0.9$ between the GWAS variant and lead eSNP; three of these eQTLs were also identified with the SMR method. The full set of 944 adipose eQTLs at GWAS loci, evidence of their coincidence with the GWAS signals, and the corresponding p values in the MuTHER study are provided in Table S8.

To consider potential pleiotropic effects of the GWAS variants on gene expression and cardio-metabolic traits, we performed two additional tests. We applied the SMR Mendelian randomization method to loci for 61 traits and identified 46 genes ($p_{SMR} < 2.1 \times 10^{-6}$; $p_{HEIDI} \geq 0.05$) (Table S9), 19 of which were identified via LD ($r^2 > 0.8$) and conditional analyses. The SMR analysis also detected nine genes (*C18orf8*, *CABLES1*, *CADM1*, *CDK6*, *CENPW*, *LMOD1*, *MFAP2*, *MT1M*, and *STARD10*) that were not identified via the conditional analysis approach.

We applied a causal-inference test[50] and identified 15 variant-transcript pairs for seven traits that showed evidence of causal mediation ($p_{CIT} < 0.05$) (Table S10 and Figure S3). None of the target genes identified by the causal-inference test overlap with the genes identified by the SMR analysis.

The coincident GWAS and eQTL signals suggest candidate genes that might influence cardio-metabolic risk. For example, at the *ARL15* locus associated with adiponectin and HDL-C[14,38], the GWAS index variant rs6450176 was associated with expression of *FST* ($p = 3.3 \times 10^{-9}$), located ~500 kb away, and exhibited strong LD ($r^2 = 0.99$) with the lead eSNP rs59061738, associated with *FST* expression ($p = 1.8 \times 10^{-9}$, Figure 1A and Table 1). Reciprocal conditional analyses provided additional evidence that this eQTL was coincident with the GWAS signal ($p_{cond} = 0.43$ for rs6450176 and 0.17 for rs59061738). *FST* expression levels were negatively associated with HDL-C ($p = 1.2 \times 10^{-3}$) and adiponectin ($p = 6.0 \times 10^{-6}$, Figure 1B). Encoded by *FST*, the protein follistatin has been shown to promote adipocyte differentiation and reduce fat mass and insulin resistance.[70,71] The same GWAS variant showed no eQTL for the nearest gene, *ARL15* ($p = 0.43$), and the expression level of *ARL15* was not associated with HDL-C ($p = 0.85$) or adiponectin ($p = 0.084$). These data suggest that the associations with the levels of adiponectin and HDL-C at this GWAS locus might be mediated at least in part through the altered expression of *FST*.

Association between Gene Expression Level and Phenotypic Traits

We investigated the effects of gene expression levels on cardio-metabolic risk by evaluating the association between the expression levels of all 43,145 probe sets and 23 cardio-metabolic-related traits (Table S11). At FDR < 1% (equivalent $p < 6.5 \times 10^{-4}$), we observed 48,365 significant associations between probe set and trait, and these corresponded to 29,920 gene-trait associations and 7,643 genes that were associated with 1 to 16 cardio-metabolic traits (Table S12). 10,819 probe sets (6,064 genes) were associated with BMI, 6,640 probe sets (3,940 genes) with Matsuda index, and 4,933 probe sets (3,039 genes) with insulin levels.

We next examined associations between the identified eQTL genes and cardio-metabolic traits measured in extensively phenotyped participants of the METSIM study. Among the genes for the 140 GWAS-relevant eQTLs, the expression levels of 49 were associated with the corresponding GWAS traits at $p < 0.05$ (Figure 2; also Figure S4 and Table S8). For example, *TBX15* at the *TBX15-WARS2* locus was associated with waist-hip ratio adjusted for BMI ($p = 1.3 \times 10^{-8}$) and 12 other traits ($p = 4.3 \times 10^{-21}$ to 2.6×10^{-4}), and expression of *GPR146* at the lipid locus *GPR146* was associated with HDL-C ($p = 3.3 \times 10^{-13}$), triglycerides ($p = 3.9 \times 10^{-13}$), and ten other traits ($p = 2.3 \times 10^{-45}$ to 6.6×10^{-5}, Figure 2). These gene expression and trait associations at GWAS loci further suggest plausible roles of these genes in mediating variant effects on cardio-metabolic disorders.

Human GWAS eQTL Genes in Mice

We next sought additional evidence for the involvement of GWAS-relevant eQTL genes in regulating cardio-metabolic traits by comparing results with those from a diverse panel of 120 inbred mouse strains known as the Hybrid Mouse Diversity Panel (HMDP).[72] In the panel, we identified microarray probes for 107 of the 140 mouse orthologs and tested them for association between adipose expression level and metabolic traits, including plasma lipids, insulin, glucose, and body-fat composition. We observed a significant correlation between 70 genes and one of the metabolic traits ($p < 5.2 \times 10^{-5}$; 0.05/963 tests for nine traits × 107 genes) (Figure S5). Of these genes, 25 showed expression levels correlated with a similar metabolic trait in the same direction (Table S13). 13.1% of 10,771 orthologous genes had expression-trait correlations in the same direction in the adipose tissue of both species for waist-hip ratio, total cholesterol, insulin, and glucose. For example, GWAS variants associated with the homeostatic model of insulin resistance (HOMA-IR) are associated

Table 1. Selected Adipose eQTLs Coincident with GWAS Signals for Cardiometabolic Risk

				GWAS Variant					Lead eSNP						
GWAS Variant	**GWAS Trait**	**GWAS Locus**	**eQTL Gene**	**A1/A2**	$\beta_{initial}$	$p_{initial}$	β_{cond}	p_{cond}	**Lead eSNP**	**A1/A2**	$\beta_{initial}$	$p_{initial}$	β_{cond}	p_{cond}	**LD r^2**
rs2013208	HDL cholesterol	*RBM5*	*RBM6*	C/T	−0.987	4.4E−113	0.120	6.9E−01	rs11130233	G/T	−0.992	4.7E−116	−1.099	2.2E−04	0.98
rs12051272	adiponectin	*CDH13*	*CDH13*	T/G	1.370	4.3E−80	0.000	2.8E−01	rs12051272	T/G	1.370	4.3E−80	0.000	2.8E−01	1.00
rs6805251	HDL cholesterol	*GSK3B*	*GSK3B*	C/T	−0.696	1.5E−43	0.067	7.4E−01	rs334533	C/T	−0.718	9.2E−47	−0.786	1.3E−04	0.95
rs8077889	triglycerides	*MPP3*	*MPP3*	C/A	−0.696	9.2E−33	−0.174	7.4E−01	rs55768269	T/C	−0.695	5.6E−33	−0.569	2.8E−01	0.99
rs4148008	HDL cholesterol	*ABCA8*	*ABCA8*	G/C	−0.538	5.9E−26	0.415	2.1E−02	rs1156340	T/C	−0.598	1.8E−31	−1.019	2.5E−08	0.92
rs12489828	waist-hip ratio	*NT5DC2*	*NT5DC2*	G/T	0.503	4.0E−24	0.089	7.6E−01	rs6778735	T/C	0.510	1.8E−24	0.398	1.9E−01	0.97
rs138777	total cholesterol	*TOM1*	*HMGXB4*	A/G	−0.495	1.7E−22	0.141	8.4E−01	rs9306298	T/C	−0.498	1.1E−22	0.636	3.7E−01	0.99
rs2254287	LDL cholesterol	*B3GALT4*	*HSD17B8*	G/C	−0.417	3.5E−17	0.000	2.1E−01	rs2254287	G/C	−0.417	3.5E−17	0.000	2.1E−01	1.00
rs12679556	waist-hip ratio	*MSC*	*EYA1*	G/T	0.463	1.1E−16	−0.202	6.2E−01	rs4738141	G/A	0.469	3.4E−17	−0.677	9.7E−02	0.98
rs12748152	HDL cholesterol	*NR0B2-PIGV*	*PIGV*	T/C	−0.698	4.0E−14	−0.054	8.5E−01	rs6656815	A/G	−0.732	2.7E−15	−0.687	1.9E−02	0.90
rs10919388	waist-hip ratio	*GORAB*	*PRRX1*	C/A	−0.448	4.5E−14	−0.212	3.3E−01	rs6427242	G/C	−0.448	3.7E−14	−0.245	2.5E−01	0.93
rs8077889	triglycerides	*MPP3*	*DUSP3*	C/A	0.430	6.6E−13	0.000	5.6E−01	rs2342310	C/T	0.430	6.6E−13	0.000	5.6E−01	1.00
rs11136341	LDL cholesterol	*PLEC1*	*PLEC*	G/A	−0.358	9.5E−13	0.084	6.5E−01	rs10107388	C/T	−0.379	4.9E−14	−0.455	1.5E−02	0.92
rs2590838	adiponectin	*GNL3*	*NEK4*	G/A	0.346	8.4E−12	−0.018	9.9E−01	rs35212380	C/G	0.346	7.8E−12	−0.415	6.7E−01	1.00
rs439401	triglycerides	*APOE-TOMM40*	*APOE*	T/C	0.384	2.1E−11	0.000	6.7E−01	rs439401	T/C	0.384	2.1E−11	0.000	6.7E−01	1.00
rs181362	HDL cholesterol	*UBE2L3*	*YDJC*	T/C	0.327	7.6E−11	−0.254	1.3E−01	rs11089620	G/C	0.368	3.1E−13	−0.621	2.3E−04	0.91
rs7134375	HDL cholesterol	*PDE3A*	*PDE3A*	C/A	0.342	1.6E−10	0.000	3.5E−01	rs7134375	C/A	0.342	1.6E−10	0.000	3.5E−01	1.00
rs11869286	HDL cholesterol	*STARD3*	*STARD3*	G/C	0.333	1.1E−09	0.000	9.8E−01	rs11869286	G/C	0.333	1.1E−09	0.000	9.8E−01	1.00
rs9400239	body mass index	*FOXO3*	*FOXO3*	C/T	0.323	3.0E−09	0.008	9.7E−01	rs3800228	G/T	0.337	5.9E−10	0.350	5.5E−02	0.92
rs12748152	HDL cholesterol	*NR0B2-PIGV*	*ARID1A*	T/C	−0.521	2.2E−08	−0.091	7.8E−01	rs34217609	C/T	−0.515	1.1E−08	−0.388	2.2E−01	0.92
rs3812316	triglycerides	*MLXIPL*	*BCL7B*	C/G	−0.411	6.5E−08	−0.059	8.2E−01	rs799166	C/G	−0.431	2.1E−08	−0.395	1.4E−01	0.92
rs4846914	HDL cholesterol	*GALNT2*	*GALNT2*	G/A	−0.267	4.1E−07	0.221	4.4E−01	rs4631704	C/T	−0.279	1.3E−07	0.503	8.4E−02	0.97
rs12145743	HDL cholesterol	*PMVK-HDGF*	*RRNAD1*	T/G	−0.262	7.9E−07	−0.115	8.1E−01	rs3806415	C/T	−0.263	7.8E−07	−0.139	7.7E−01	0.99
rs10501320	proinsulin	*MADD*	*ACP2*	G/C	0.309	1.3E−06	0.024	9.7E−01	rs11039149	A/G	0.310	1.2E−06	0.259	6.5E−01	0.99
rs6784615	waist-hip ratio	*NISCH-STAB1*	*NISCH*	T/C	0.649	1.6E−06	0.000	6.6E−01	rs728408	A/G	0.649	1.5E−06	0.000	7.4E−01	1.00
rs780094	glucose, triglycerides	*GCKR*	*EMILIN1*	T/C	0.240	3.3E−06	0.000	7.2E−01	rs780094	T/C	0.240	3.3E−06	0.000	7.2E−01	1.00
rs10761731	triglycerides	*JMJD1C*	*NRBF2*	A/T	−0.231	5.8E−06	0.000	5.8E−02	rs10761739	G/C	−0.233	5.2E−06	0.000	4.3E−02	1.00

(*Continued on next page*)

Table 1. *Continued*

				GWAS Variant					Lead eSNP					
GWAS Variant	**GWAS Trait**	**GWAS Locus**	**eQTL Gene**	**A1/A2**	**$\beta_{initial}$**	**$p_{initial}$**	**β_{cond}**	**p_{cond}**	**Lead eSNP**	**A1/A2**	**$\beta_{initial}$**	**$p_{initial}$**	**β_{cond}**	**p_{cond}**	**LD r^2**
rs1532624	HDL cholesterol	*CETP*	*CETP*	C/A	0.237	5.8E−06	−0.549	2.1E−01	rs4784741	C/T	0.246	2.4E−06	−0.777	7.4E−02	0.99
rs4311394	adiponectin	*ARL15*	*FST*	A/G	0.287	6.2E−06	−0.707	3.1E−01	rs59061738	A/G	0.292	3.7E−06	−1.011	1.5E−01	0.99

This table contains the 29 GWAS-relevant eQTLs for which the pairwise LD $r^2 > 0.9$ between GWAS index and lead eSNP, for which $p_{initial} < 10^{-5}$ for the association between GWAS SNP and gene expression level, and which had not been previously described in large-scale genome-wide association studies that had interrogated available databases/resources for *cis*-eQTLs searches. The GWAS traits were limited to the most common cardio-metabolic traits, including body mass index, waist-hip ratio adjusted for body mass index, triglycerides, HDL cholesterol, LDL cholesterol, total cholesterol, glucose, proinsulin, adiponectin, and risk of type 2 diabetes. A full list of the 944 adipose eQTLs identified at GWAS loci for cardio-metabolic diseases and traits, including the 140 GWAS coincident eQTLs with pairwise LD $r^2 > 0.8$, are shown in Table S8. A1/A2: the effect/non-effect alleles for association tests in METSIM eQTL analyses. A1 for the GWAS variant was the allele previously reported to increase the cardio-metabolic risk. LD r^2 between the GWAS variant and lead eSNP was calculated from 770 METSIM individuals.

with decreased expression level of *IRS1* ($p = 2.0 \times 10^{-50}$). *IRS1* expression is negatively correlated with HOMA-IR ($r = -0.40$, $p = 3.0 \times 10^{-30}$) in METSIM and the HMDP ($r = -0.50$, $p = 1 \times 10^{-13}$) (Figure S6).

The HMDP data support a biological contribution of 25 eQTL genes at GWAS loci; 18 of these genes did not have significant associations in humans ($p > 0.05$) but had significant associations in mice, possibly because of the heterogeneous environmental effects in humans and controlled experimental environment in mice (Table S13). For example, the chromosome 2q35 variants reported by the GIANT consortium to be associated with BMI[17] span ~300 kb and nine genes. In METSIM, the BMI risk allele rs492400 was significantly associated with increased expression of *RQCD1* ($p = 8.9 \times 10^{-43}$, Figure 1C), although *RQCD1* expression was only nominally correlated with BMI ($r = 0.09$, $p = 0.01$, Figure 1D). In the HMDP, *Rqcd1* expression was more strongly positively correlated with subcutaneous fat weight ($r = 0.45$, $p = 1.2 \times 10^{-12}$, Figure 1E), providing further support for the role of *RQCD1* in BMI. Consistent correlation between gene expression and clinical traits in mice for *cis*-regulated genes at human GWAS loci supports the fruitful use of mouse models to further understand the roles of candidate genes for cardio-metabolic traits.

trans-eQTLs

Associations of variants with expression of distant genes in *trans* allow the identification of regulatory networks in adipose tissue. We observed *trans*-eQTLs at a Bonferroni-corrected significance threshold ($p < 1.51 \times 10^{-13}$) for 90 genes (Table S3). To better identify regulatory networks in human samples, we searched for variants that were associated with at least five distal genes at a more liberal threshold of $p < 5 \times 10^{-8}$. These variants were located on chromosomes 3, 7, 11, and 16 in five independent loci (pairwise LD $r^2 = 0$) that we termed *trans*-eQTL hotspots (Table S14, Figure 3). Four of the hotspot loci were associated with local (*cis*) gene expression at $p < 1 \times 10^{-3}$ (Figure S7).

40 of the 55 *trans*-mediated genes at the *KLF14* locus were significantly associated with metabolic phenotypes at 1% FDR, representing a significant enrichment with respect to all the genes assayed on the microarray (Chi-square p value = 8.4×10^{-6}) (Figure S8). The *trans*-regulated genes were co-expressed with pairwise correlations of −0.48 to 0.67 ($p = 8.9 \times 10^{-4}$ to 1.6×10^{-100}); however, they were not enriched for known biological pathways according to the analysis involving the DAVID database.[73] Some of the distal genes have been previously studied in relation to adipose biology, but their association with the *KLF14* locus was not known. Among the *trans*-regulated genes, *PLIN5* was associated with 13 of the metabolic phenotypes (Figure S8). *PLIN5* encodes Perilipin 5, a lipid-droplet-associated protein that helps maintain a balance between lipolysis and lipogenesis and has been shown to play a role in fatty-acid oxidation in adipose tissue.[74]

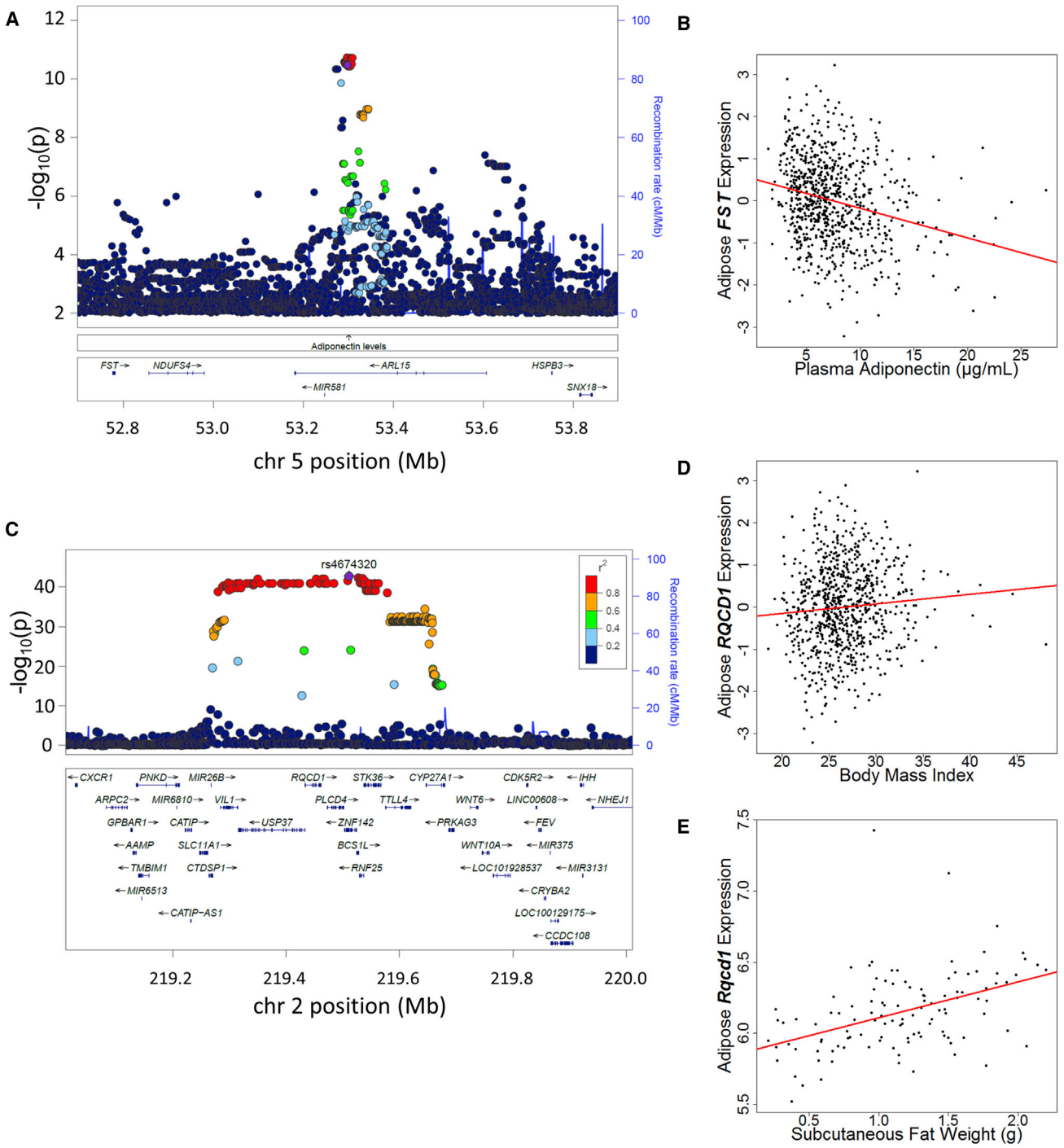

Figure 1. Example Subcutaneous Adipose eQTL Genes at GWAS Loci

(A) The adipose eQTL of *FST* is coincident with the adiponectin GWAS locus *ARL15*.[69] Regional association of variants with expression level of *FST* is shown with the GWAS variant rs6450176 plotted as the index (purple diamond). LD is colored on the basis of the METSIM population.

(B) Association between adiponectin level and *FST* expression level in METSIM.

(C) The adipose eQTL of *RQCD1* is coincident with the *USP37* GWAS locus for BMI.[17] Regional association of variants with expression level of *RQCD1* is shown with the eSNP rs4674320 ($r^2 = 1.0$ with BMI index SNP rs492400) plotted as the index.

(D and E) Association between (D) BMI and *RQCD1* expression level in humans from the METSIM study and (E) body fat and *Rqcd1* expression level in mice from the HMDP study.

The expanded *KLF14*-mediated *trans* network of genes suggests that these targets might also influence cardiometabolic traits.

Variants located near *CIITA* (class II, major histocompatibility transactivator), which is known to activate the MHC class II genes at the HLA locus[75] (Figure 3B), were

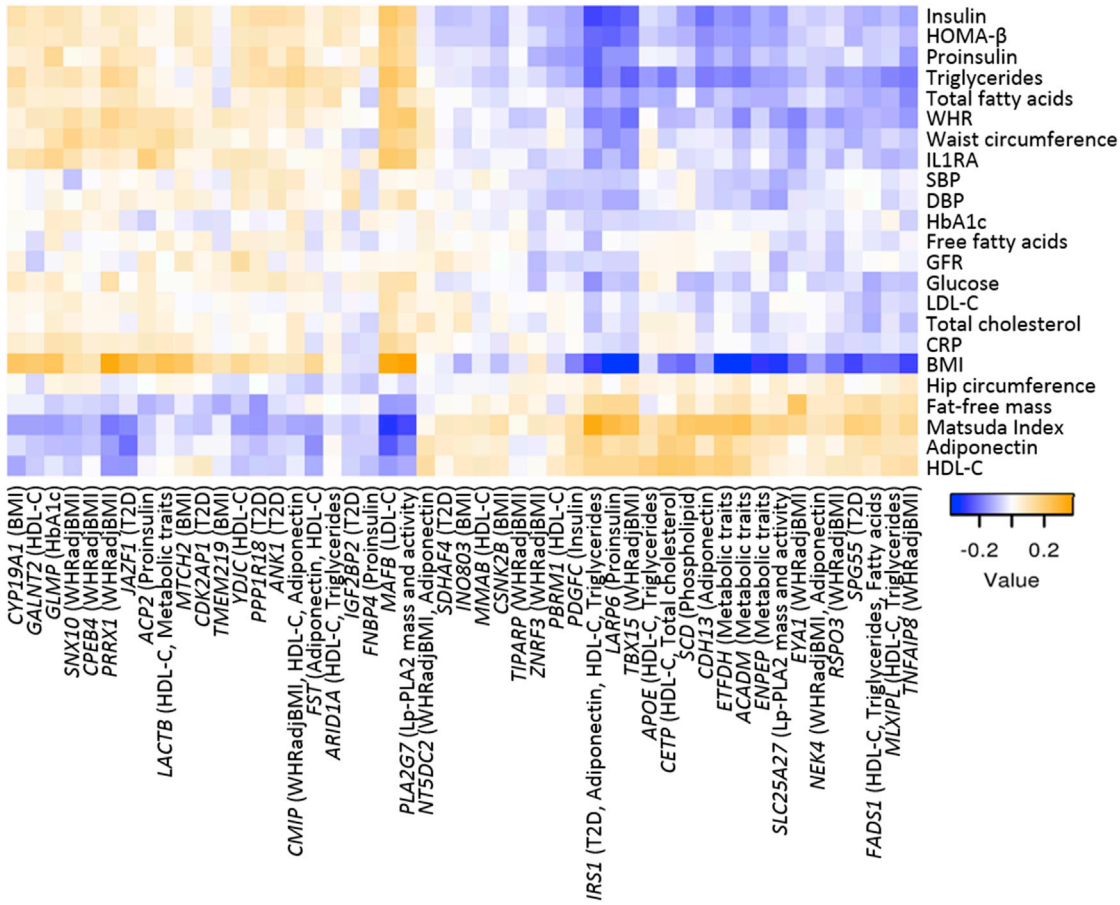

Figure 2. Heatmap of Effect Sizes for Significant Associations between Gene Expression Level and Cardio-Metabolic-Trait Levels at GWAS Loci with Coincident eQTLs

Rows show 23 selected cardio-metabolic traits, and columns show the eQTL genes (and reported GWAS trait at the coincident locus). Negative values (blue) indicate that increased gene expression level was associated ($p < 0.05$) with decreased trait level, whereas positive values (orange) indicate that increased gene expression level was associated with increased trait level.

associated with six MHC class II genes (*HLA-DPA1*, *HLA-DMA*, *HLA-DPB1*, *HLA-DOA*, *HLA-DRA*, and *HLA-DMB*) located on chromosome 6 and with *CD74* on chromosome 5. *CD74* encodes for a protein that associates with the class II MHC complex and regulates antigen presentation.[76] The variants were also suggestively associated with *CIITA* expression ($p = 1.0 \times 10^{-3}$, Figure S7), suggesting that variation in its expression might be responsible for the *trans*-eQTL signals. Our eQTL results captured the known biology of CIITA regulation of MHC class II genes. Further, expression of *CIITA* and the seven *trans*-mediated genes was significantly associated with 14 of the metabolic traits ($|\beta| = 0.12$ to 0.37, $p = 6.4 \times 10^{-4}$ to 7.7×10^{-26}), suggesting a role for this network of genes in adipose tissue (Figure S8).

The second largest *trans*-eQTL hotspot was located on chromosome 3 and was associated with the expression of 44 genes (Figure 3B). This locus has been reported in other tissues, including liver and omental fat,[77] but the target genes identified in our study do not overlap with the previous ones, suggesting a unique *trans*-regulatory network in subcutaneous adipose tissue.[11] Variants in this locus showed the strongest local association with the expression of *SLC25A38*, located ~1 Mb away ($p = 1.7 \times 10^{-7}$, Figure S7). Missense mutations in *SLC25A38*, which encodes a mitochondrial solute carrier, lead to congenital sideroblastic anemia characterized by defective erythropoiesis and mitochondrial iron overload.[78] The function of *SLC25A38* in adipose tissue remains unclear. The signal observed at this locus might represent cell types other than adipocytes in light of the cellular heterogeneity of adipose tissue. The adipose expression level of *SLC25A38* was nominally associated with 24 of the *trans*-mediated genes ($|\beta| = 0.07$ to 0.29, $p = 6.8 \times 10^{-16}$ to 4.7×10^{-2}) and with 14 cardio-metabolic-related traits ($|\beta| = 0.08$ to 0.33, $p = 6.7 \times 10^{-21}$ to 3.1×10^{-2}) (Figure S7). For example, *SLC25A38* is negatively associated with BMI in METSIM participants ($\beta = -0.33$, $p = 6.7 \times 10^{-21}$). On the basis of structural similarity with other SLC25 proteins, SLC25A38 is predicted to transport glycine into the mitochondrial matrix for condensation with succinyl-coenzyme A to form 5-aminolevulinic acid, which is exported to the cytoplasm for haem synthesis.[79,80] Only a few of the *trans*-mediated genes have been shown to play

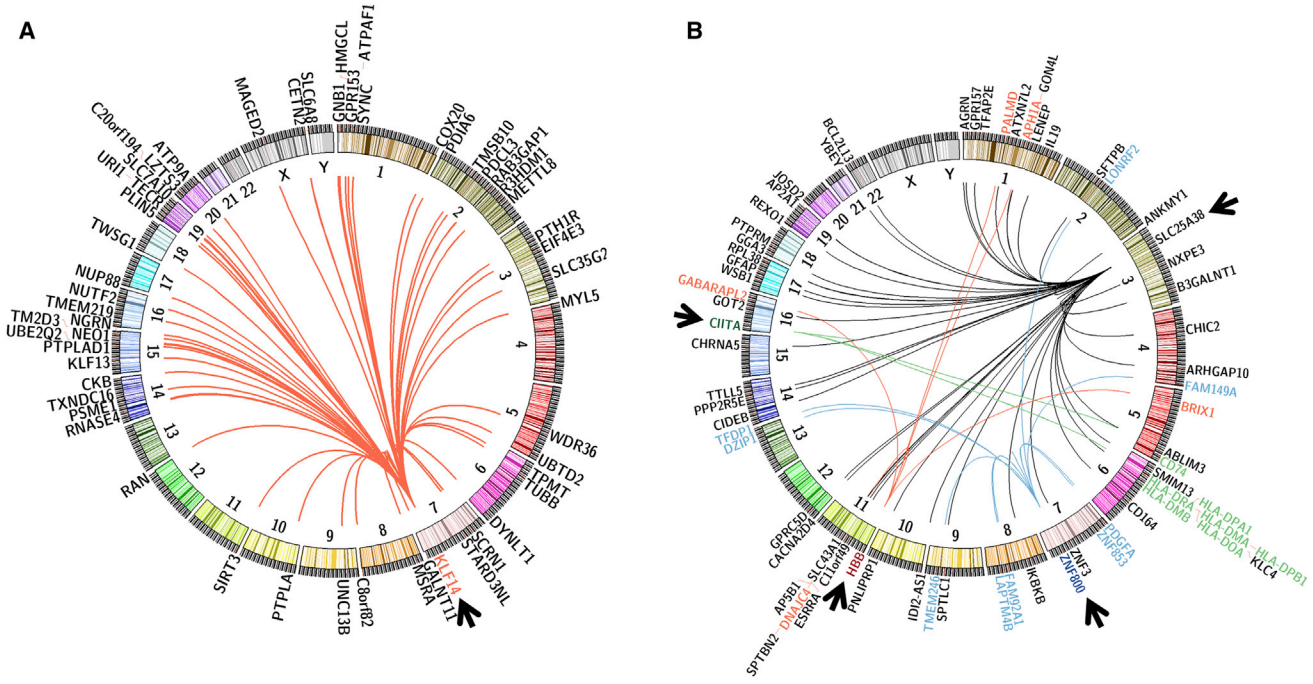

Figure 3. *trans*-eQTL Hotspots in Subcutaneous Adipose Tissue
(A) *KLF14* *trans*-eQTL hotspot and target genes in subcutaneous adipose tissue. Representation of the location of 55 distal genes for which expression level is associated with rs12154627 at the *KLF14* locus.
(B) Representation of the location of 65 distal genes for which expression level is associated with one of four *trans*-eQTL hotspots: *SLC25A38* locus (black), *ZNF800* locus (blue), *CIITA* locus (green), and *HBB* locus (red).
Arrowheads point to the *trans*-eQTL loci, and curves indicate the associations with the four sets of target genes.

important roles in adipocytes. For example, *ESRRA* encodes estrogen-related receptor alpha, which modulates the expression of adipogenesis genes during adipocyte differentiation.[81] *SPTLC1* encodes one of the subunits of serine palmitoyltransferase, the key enzyme in sphingolipid biosynthesis.[82] These bioactive lipids are altered with obesity and can regulate inflammatory gene expression in adipocytes.[83] Another target gene, *GOT2*, encodes glutamic-oxaloacetic transaminase, which functions in the inner mitochondrial membrane and has been shown to play a key role in amino acid metabolism.[84] Although the function in adipose of the genes associated with *SLC25A38* variants remains unclear, our results suggest a metabolic role.

We also observed a *trans*-eQTL hotspot, located on chromosome 7 but ~3.3 Mb away from *KLF14*, associated with the expression of nine genes. The variants at this locus were independent from those at *KLF14* and had distinct target genes (Figure 3B). The lead variant, rs62621812 in *ZNF800*, encodes a missense change (p.Pro103Ser) and is also associated with that gene's expression level ($p = 2.8 \times 10^{-16}$), providing evidence suggesting that this gene regulates the target genes. *ZNF800* encodes a C2H2 zinc finger protein and is a putative transcription factor.[85] *ZNF800* expression level is associated with the Matsuda Index ($\beta = 0.12$, $p = 1.3 \times 10^{-3}$), suggesting a role in insulin sensitivity.

A fifth *trans*-eQTL, located on chromosome 11, was associated with expression of five *trans* genes. The peak variant, rs10742583, is located ~2 kb upstream of *HBB*, which encodes beta hemoglobin, and is in complete LD ($r^2 = 1.0$) with a HBB synonymous variant, rs713040, although rs10742583 was not associated with the expression level of any local gene ($p < 2.4 \times 10^{-4}$). A functional role for *HBB* in adipose biology is not clear. Although *HBB* is distinctly and highly expressed in whole blood, the *trans*-mediated target genes are expressed in a range of tissues.[56] This *trans*-eQTL signal might reflect cellular heterogeneity, especially blood cells present in adipose tissue.

To validate these *trans*-eQTL hotspots, we asked whether eQTLs were also observed in the MuTHER study. Variant information was available for three of the loci, near *CIITA*, *SLC25A38*, and *KLF14*. All seven genes that mapped to the *CIITA* locus, 35 of the 44 genes that mapped to the *SLC25A38* locus, and 41 of the 55 genes that mapped to the *KLF14* locus had corresponding probes available in the MuTHER study. Although only two of the 35 target genes for the *SLC25A38* locus had evidence of association (FDR < 1%) in the MuTHER study, many *KLF14* and *CIITA* target genes were replicated (Table S14). For the *KLF14* *trans*-eQTL hotspot, originally described in MuTHER,[20] 25 of the 41 target genes showed consistent direction of effect and nominal associations ($p < 0.05$) in the MuTHER study, whereas none of the genes showed an opposite effect. For the *CIITA* locus, six of the seven target genes were validated in MuTHER with a consistent direction of effect ($p < 0.05$), indicating strong support for these two *trans*-eQTL hotspots.

Discussion

We report an analysis of genetic variants, adipose gene expression, and cardio-metabolic traits in a deeply phenotyped cross-sectional sample of 770 participants in the METSIM study and compare the results to those of studies performed in 120 diverse inbred strains of mice. These results help to prioritize potential target genes at GWAS loci, and they identify several *trans*-regulatory networks associated with metabolic syndrome and its component traits. In addition, the results reveal concordance between mice and humans in terms of correlations between gene expression and traits.

A main challenge with regard to gaining biological insights from genetic associations is to identify which genes and pathways explain the associations. Systems genetics aims to address this challenge by integrating genetic variants with molecular phenotypes to more comprehensively define the relationship between genotype and phenotype. Because a large fraction of the variation underlying common diseases appears to be regulatory,[4,86,87] eQTLs that coincide with GWAS loci can link trait-associated loci with molecular perturbations. Overall, using reciprocal conditional analyses, we provided evidence for 140 genes at 105 loci that may be involved in metabolic traits (Table 1 and Table S8). This analysis of gene expression in a single tissue might not be relevant for the mechanisms of some of the cardio-metabolic traits. Although our analysis provides indirect evidence of an association, because the colocalization of an eQTL with a disease locus could be coincidental, the cardio-metabolic phenotypes available in the METSIM cohort provided additional evidence for 49 genes for which expression level was associated with the relevant trait. In addition, a subset of the 140 genes was also linked to GWAS traits via the summary-data-based Mendelian-randomization and causal-inference tests. Of note, we limited our analysis to lead eQTL variants and lead GWAS variants in high pairwise LD ($r^2 \geq 0.8$). This threshold might be conservative if the GWAS lead variant imperfectly tags a causal variant and might underestimate the number of GWAS loci that are coincident with eQTLs. Consistent with this possibility, we found evidence of association for 579 genes at 415 loci for which conditioning on the GWAS variant partially attenuated the eQTL effect (Table S8). Furthermore, secondary independent associations for both eQTLs and GWAS signals have been shown to play an important role in altering the expression of a gene in a locus;[22] however, additional secondary and tertiary eQTL signals were not separately analyzed in this study. These genes provide a larger set of possible targets for GWAS loci.

Employing genotype imputation with a dense reference panel and analyzing 43,145 transcripts on a microarray to measure expression, we were able to identify more eQTLs than were identified in other subcutaneous adipose studies with similar sample sizes.[10,11,37] Using these eQTLs, we implicated 50 genes that were distinct from the initial locus annotations used as signposts for the GWAS loci (Table S8). For example, rs11231693 is associated with waist-hip ratio adjusted for BMI[16] and is located in the intron of *MACROD1*. Our results showed a strong association with the expression level of *VEGFB* ($p = 2.2 \times 10^{-67}$), located 140 kb from the variant. For ten loci, we identified multiple genes associated with the risk variants. For example, BMI-associated variants in the *INO80E* locus[17] are also the strongest variants associated with the expression of five nearby genes: *YPEL3*, *INO80E*, *TMEM219*, *TBX6*, and *HIRIP3* ($p = 1.4 \times 10^{-8}$ to 1.8×10^{-4}), suggesting that variation in one or more of these genes might influence BMI. Finally, at the *GCKR, LCAT*, and *LDLR* genes that harbor common and rare coding variants that can affect plasma lipid levels, we observed eQTL effects on nearby genes, including *EMILIN1* at the *GCKR* locus ($p = 3.3 \times 10^{-6}$), *GFOD2* and *NUTF2* at the *LCAT* locus ($p = 1.9 \times 10^{-6}$ and 1.6×10^{-4}, respectively), and *YIPF2* at the *LDLR* locus ($p = 1.5 \times 10^{-6}$). Some coincident associations between GWAS variants and eQTLs might not contribute strongly to trait variation.

Recent studies demonstrated the evolutionary conservation of regulatory networks that increase susceptibility to atherosclerosis and other cardio-metabolic traits in disease-relevant tissues in humans and mice.[88] Genetically diverse mouse populations allow for the control of confounding factors, such as environmental exposure that are difficult to assess in humans. The HMDP has also helped researchers to understand contributions of genetic factors to cardio-metabolic traits.[72] The results of our cross-species analysis showed consistent association between traits and the expression of 25 genes in humans and mice, providing support for further study of these genes in mouse models (Figure S5 and Table S13). Of course, the possibility exists that mechanisms might differ between species or that the wrong tissues are being compared. In addition to prioritizing candidate genes, studies of natural variation in gene expression in mice should be useful in elucidating mechanisms relating to the gene-by-gene, gene-by-environment, and gene-by-sex interactions.[51,52,89] For example, although it is clear that genetic background contributes to weight gain and related traits in response to dietary challenge,[90] the inability to accurately ascertain environmental factors over a lifetime in humans has made molecular dissection difficult. In contrast, studies of natural variation in mice have revealed that gene-by-environment interactions can contribute substantially to the overall variance in traits such as obesity and that the underlying genes and pathways can be identified.[51,52]

Whereas numerous studies have successfully mapped *cis* associations, few studies have reported *trans*-eQTLs because of small effect sizes, multiple testing thresholds, and computational burden.[22,37,91] These issues are consistent with the low reciprocal replication rates of *trans*-eQTLs identified in MuTHER and METSIM cohorts. However, in-depth characterization of the architecture of *trans* regulation of gene expression is useful in attempting to understand complex biological mechanisms, as shown

by the expanded network of *trans*-regulated genes at the *KLF14* locus, which has been associated with T2D and several metabolic traits. The 55 genes that map to this locus demonstrate the extensive downstream effects of an individual association signal and provide additional genes that might influence disease manifestation. The identification of *trans*-regulatory networks also offers opportunities for understanding fundamental biology. For example, ZNF800 is a putative transcription factor, and although little is known about the function of the encoded protein, our results implicate a novel role for ZNF800 in adipose biology. In addition, some *trans*-eQTL associations might be influenced by the cellular composition of the adipose tissue; for example, adipose tissue of obese individuals tends to have higher macrophage content.[90,91] Although we failed to observe any association between the *trans*-acting loci and expression of macrophage markers such as *ABCG1* and *CD68* (data not shown), we cannot rule out a role for cell-type heterogeneity. We also observed weaker *cis* associations compared to *trans* associations in the *trans*-eQTL hotspots. This may be an artifact of data pre-processing or inaccuracy in measuring low levels of expression. Future studies involving single-cell analyses might clarify the role of adipose tissue by identifying cell-type-specific eQTLs. Our study highlights the power of a systems-genetics approach in dissecting complex traits and identifying causal genes and pathways.

Accession Numbers

Newly deposited human expression data can be obtained from the Gene Expression Omnibus with the accession number GEO: GSE70353.

Supplemental Data

Supplemental Data include 8 figures and 14 tables and can be found with this article online at http://dx.doi.org/10.1016/j.ajhg.2017.01.027.

Conflicts of Interest

A.H., C.T., T.K., and P.S.G. are employees and shareholders of Bristol-Myers Squibb.

Acknowledgments

This study was supported by NIH grants R00HL121172 (M.C.), P01HL28481 (A.J.L. and P.P.), R01DK093757 (K.L.M.), U01DK105561 (K.L.M.), R01DK072193 (K.L.M.), U01DK062370 (M.B.), 1-ZIA-HG000024 (F.S.C.), R01HL095056 (P.P.), F31HL127921 (A.K.), and R00HL123021 (B.W.P.); Academy of Finland grants 77299 and 124243 (M.L.); the Finnish Heart Foundation (M.L.); the Finnish Diabetes Foundation (M.L.); Finnish Funding Agency for Technology and Innovation (TEKES) contract 1510/31/06 (M.L.); and the Commission of the European Community HEALTH-F2-2007-201681 (M.L.). C.K.R. was supported by NIH T32GM067553. Bristol-Myers Squibb supported the generation of METSIM and HMDP microarray data.

Received: July 24, 2016
Accepted: January 12, 2017
Published: March 2, 2017

Web Resources

Efficient and Parallelizable Association Container Toolbox (EPACTS), http://genome.sph.umich.edu/wiki/EPACTS
GWAS Catalog, https://www.ebi.ac.uk/gwas/
The Haplotype Reference Consortium, http://www.haplotype-reference-consortium.org
METSIM and HMDP, https://systems.genetics.ucla.edu/

References

1. Edwards, S.L., Beesley, J., French, J.D., and Dunning, A.M. (2013). Beyond GWASs: Illuminating the dark road from association to function. Am. J. Hum. Genet. *93*, 779–797.
2. Visscher, P.M., Brown, M.A., McCarthy, M.I., and Yang, J. (2012). Five years of GWAS discovery. Am. J. Hum. Genet. *90*, 7–24.
3. Brænne, I., Civelek, M., Vilne, B., Di Narzo, A., Johnson, A.D., Zhao, Y., Reiz, B., Codoni, V., Webb, T.R., Foroughi Asl, H., et al.; Leducq Consortium CAD Genomics‡ (2015). Prediction of causal candidate genes in coronary artery disease loci. Arterioscler. Thromb. Vasc. Biol. *35*, 2207–2217.
4. Albert, F.W., and Kruglyak, L. (2015). The role of regulatory variation in complex traits and disease. Nat. Rev. Genet. *16*, 197–212.
5. Conde, L., Bracci, P.M., Richardson, R., Montgomery, S.B., and Skibola, C.F. (2013). Integrating GWAS and expression data for functional characterization of disease-associated SNPs: an application to follicular lymphoma. Am. J. Hum. Genet. *92*, 126–130.
6. Civelek, M., and Lusis, A.J. (2014). Systems genetics approaches to understand complex traits. Nat. Rev. Genet. *15*, 34–48.
7. Coelho, M., Oliveira, T., and Fernandes, R. (2013). Biochemistry of adipose tissue: An endocrine organ. Arch. Med. Sci. *9*, 191–200.
8. Gustafson, B., Hedjazifar, S., Gogg, S., Hammarstedt, A., and Smith, U. (2015). Insulin resistance and impaired adipogenesis. Trends Endocrinol. Metab. *26*, 193–200.
9. Porter, S.A., Massaro, J.M., Hoffmann, U., Vasan, R.S., O'Donnel, C.J., and Fox, C.S. (2009). Abdominal subcutaneous adipose tissue: a protective fat depot? Diabetes Care *32*, 1068–1075.
10. Emilsson, V., Thorleifsson, G., Zhang, B., Leonardson, A.S., Zink, F., Zhu, J., Carlson, S., Helgason, A., Walters, G.B., Gunnarsdottir, S., et al. (2008). Genetics of gene expression and its effect on disease. Nature *452*, 423–428.
11. Greenawalt, D.M., Dobrin, R., Chudin, E., Hatoum, I.J., Suver, C., Beaulaurier, J., Zhang, B., Castro, V., Zhu, J., Sieberts, S.K., et al. (2011). A survey of the genetics of stomach, liver, and adipose gene expression from a morbidly obese cohort. Genome Res. *21*, 1008–1016.
12. Nica, A.C., Parts, L., Glass, D., Nisbet, J., Barrett, A., Sekowska, M., Travers, M., Potter, S., Grundberg, E., Small, K., et al.; MuTHER Consortium (2011). The architecture of gene regulatory variation across multiple human tissues: the MuTHER study. PLoS Genet. *7*, e1002003.

13. Zhong, H., Beaulaurier, J., Lum, P.Y., Molony, C., Yang, X., Macneil, D.J., Weingarth, D.T., Zhang, B., Greenawalt, D., Dobrin, R., et al. (2010). Liver and adipose expression associated SNPs are enriched for association to type 2 diabetes. PLoS Genet. *6*, e1000932.
14. Willer, C.J., Schmidt, E.M., Sengupta, S., Peloso, G.M., Gustafsson, S., Kanoni, S., Ganna, A., Chen, J., Buchkovich, M.L., Mora, S., et al.; Global Lipids Genetics Consortium (2013). Discovery and refinement of loci associated with lipid levels. Nat. Genet. *45*, 1274–1283.
15. Mahajan, A., Go, M.J., Zhang, W., Below, J.E., Gaulton, K.J., Ferreira, T., Horikoshi, M., Johnson, A.D., Ng, M.C., Prokopenko, I., et al.; DIAbetes Genetics Replication And Meta-analysis (DIAGRAM) Consortium; Asian Genetic Epidemiology Network Type 2 Diabetes (AGEN-T2D) Consortium; South Asian Type 2 Diabetes (SAT2D) Consortium; Mexican American Type 2 Diabetes (MAT2D) Consortium; and Type 2 Diabetes Genetic Exploration by Nex-generation sequencing in muylti-Ethnic Samples (T2D-GENES) Consortium (2014). Genome-wide trans-ancestry meta-analysis provides insight into the genetic architecture of type 2 diabetes susceptibility. Nat. Genet. *46*, 234–244.
16. Shungin, D., Winkler, T.W., Croteau-Chonka, D.C., Ferreira, T., Locke, A.E., Mägi, R., Strawbridge, R.J., Pers, T.H., Fischer, K., Justice, A.E., et al.; ADIPOGen Consortium; CARDIOGRAMplusC4D Consortium; CKDGen Consortium; GEFOS Consortium; GENIE Consortium; GLGC; ICBP; International Endogene Consortium; LifeLines Cohort Study; MAGIC Investigators; MuTHER Consortium; PAGE Consortium; and ReproGen Consortium (2015). New genetic loci link adipose and insulin biology to body fat distribution. Nature *518*, 187–196.
17. Locke, A.E., Kahali, B., Berndt, S.I., Justice, A.E., Pers, T.H., Day, F.R., Powell, C., Vedantam, S., Buchkovich, M.L., Yang, J., et al.; LifeLines Cohort Study; ADIPOGen Consortium; AGEN-BMI Working Group; CARDIOGRAMplusC4D Consortium; CKDGen Consortium; GLGC; ICBP; MAGIC Investigators; MuTHER Consortium; MIGen Consortium; PAGE Consortium; ReproGen Consortium; GENIE Consortium; and International Endogene Consortium (2015). Genetic studies of body mass index yield new insights for obesity biology. Nature *518*, 197–206.
18. Orozco, L.D., Bennett, B.J., Farber, C.R., Ghazalpour, A., Pan, C., Che, N., Wen, P., Qi, H.X., Mutukulu, A., Siemers, N., et al. (2012). Unraveling inflammatory responses using systems genetics and gene-environment interactions in macrophages. Cell *151*, 658–670.
19. van Nas, A., Ingram-Drake, L., Sinsheimer, J.S., Wang, S.S., Schadt, E.E., Drake, T., and Lusis, A.J. (2010). Expression quantitative trait loci: Replication, tissue- and sex-specificity in mice. Genetics *185*, 1059–1068.
20. Small, K.S., Hedman, A.K., Grundberg, E., Nica, A.C., Thorleifsson, G., Kong, A., Thorsteindottir, U., Shin, S.Y., Richards, H.B., Soranzo, N., et al.; GIANT Consortium; MAGIC Investigators; DIAGRAM Consortium; and MuTHER Consortium (2011). Identification of an imprinted master trans regulator at the KLF14 locus related to multiple metabolic phenotypes. Nat. Genet. *43*, 561–564.
21. Rotival, M., Zeller, T., Wild, P.S., Maouche, S., Szymczak, S., Schillert, A., Castagné, R., Deiseroth, A., Proust, C., Brocheton, J., et al.; Cardiogenics Consortium (2011). Integrating genome-wide genetic variations and monocyte expression data reveals trans-regulated gene modules in humans. PLoS Genet. *7*, e1002367.
22. Westra, H.J., Peters, M.J., Esko, T., Yaghootkar, H., Schurmann, C., Kettunen, J., Christiansen, M.W., Fairfax, B.P., Schramm, K., Powell, J.E., et al. (2013). Systematic identification of trans eQTLs as putative drivers of known disease associations. Nat. Genet. *45*, 1238–1243.
23. Cheng, Y., Ma, Z., Kim, B.H., Wu, W., Cayting, P., Boyle, A.P., Sundaram, V., Xing, X., Dogan, N., Li, J., et al.; Mouse ENCODE Consortium (2014). Principles of regulatory information conservation between mouse and human. Nature *515*, 371–375.
24. Vierstra, J., Rynes, E., Sandstrom, R., Zhang, M., Canfield, T., Hansen, R.S., Stehling-Sun, S., Sabo, P.J., Byron, R., Humbert, R., et al. (2014). Mouse regulatory DNA landscapes reveal global principles of cis-regulatory evolution. Science *346*, 1007–1012.
25. Stancáková, A., Javorský, M., Kuulasmaa, T., Haffner, S.M., Kuusisto, J., and Laakso, M. (2009). Changes in insulin sensitivity and insulin release in relation to glycemia and glucose tolerance in 6,414 Finnish men. Diabetes *58*, 1212–1221.
26. Matsuda, M., and DeFronzo, R.A. (1999). Insulin sensitivity indices obtained from oral glucose tolerance testing: comparison with the euglycemic insulin clamp. Diabetes Care *22*, 1462–1470.
27. Stegle, O., Parts, L., Piipari, M., Winn, J., and Durbin, R. (2012). Using probabilistic estimation of expression residuals (PEER) to obtain increased power and interpretability of gene expression analyses. Nat. Protoc. *7*, 500–507.
28. Shabalin, A.A. (2012). Matrix eQTL: Ultra fast eQTL analysis via large matrix operations. Bioinformatics *28*, 1353–1358.
29. Lippert, C., Listgarten, J., Liu, Y., Kadie, C.M., Davidson, R.I., and Heckerman, D. (2011). FaST linear mixed models for genome-wide association studies. Nat. Methods *8*, 833–835.
30. Kang, H.M., Sul, J.H., Service, S.K., Zaitlen, N.A., Kong, S.Y., Freimer, N.B., Sabatti, C., and Eskin, E. (2010). Variance component model to account for sample structure in genome-wide association studies. Nat. Genet. *42*, 348–354.
31. Yang, J., Zaitlen, N.A., Goddard, M.E., Visscher, P.M., and Price, A.L. (2014). Advantages and pitfalls in the application of mixed-model association methods. Nat. Genet. *46*, 100–106.
32. Krzywinski, M., Schein, J., Birol, I., Connors, J., Gascoyne, R., Horsman, D., Jones, S.J., and Marra, M.A. (2009). Circos: an information aesthetic for comparative genomics. Genome Res. *19*, 1639–1645.
33. Pruim, R.J., Welch, R.P., Sanna, S., Teslovich, T.M., Chines, P.S., Gliedt, T.P., Boehnke, M., Abecasis, G.R., and Willer, C.J. (2010). LocusZoom: regional visualization of genome-wide association scan results. Bioinformatics *26*, 2336–2337.
34. Yang, J., Benyamin, B., McEvoy, B.P., Gordon, S., Henders, A.K., Nyholt, D.R., Madden, P.A., Heath, A.C., Martin, N.G., Montgomery, G.W., et al. (2010). Common SNPs explain a large proportion of the heritability for human height. Nat. Genet. *42*, 565–569.
35. Yang, J., Lee, S.H., Goddard, M.E., and Visscher, P.M. (2011). GCTA: A tool for genome-wide complex trait analysis. Am. J. Hum. Genet. *88*, 76–82.
36. Kostem, E., and Eskin, E. (2013). Improving the accuracy and efficiency of partitioning heritability into the contributions of genomic regions. Am. J. Hum. Genet. *92*, 558–564.

37. Grundberg, E., Small, K.S., Hedman, A.K., Nica, A.C., Buil, A., Keildson, S., Bell, J.T., Yang, T.P., Meduri, E., Barrett, A., et al.; Multiple Tissue Human Expression Resource (MuTHER) Consortium (2012). Mapping cis- and trans-regulatory effects across multiple tissues in twins. Nat. Genet. *44*, 1084–1089.
38. Dastani, Z., Hivert, M.F., Timpson, N., Perry, J.R., Yuan, X., Scott, R.A., Henneman, P., Heid, I.M., Kizer, J.R., Lyytikäinen, L.P., et al.; DIAGRAM+ Consortium; MAGIC Consortium; GLGC Investigators; MuTHER Consortium; DIAGRAM Consortium; GIANT Consortium; Global B Pgen Consortium; Procardis Consortium; MAGIC investigators; and GLGC Consortium (2012). Novel loci for adiponectin levels and their influence on type 2 diabetes and metabolic traits: a multi-ethnic meta-analysis of 45,891 individuals. PLoS Genet. *8*, e1002607.
39. Winkler, T.W., Justice, A.E., Graff, M., Barata, L., Feitosa, M.F., Chu, S., Czajkowski, J., Esko, T., Fall, T., Kilpeläinen, T.O., et al.; CHARGE Consortium; DIAGRAM Consortium; GLGC Consortium; Global-BPGen Consortium; ICBP Consortium; and MAGIC Consortium (2015). The influence of age and sex on genetic associations with adult body size and shape: A large-scale genome-wide interaction study. PLoS Genet. *11*, e1005378.
40. Berndt, S.I., Gustafsson, S., Mägi, R., Ganna, A., Wheeler, E., Feitosa, M.F., Justice, A.E., Monda, K.L., Croteau-Chonka, D.C., Day, F.R., et al. (2013). Genome-wide meta-analysis identifies 11 new loci for anthropometric traits and provides insights into genetic architecture. Nat. Genet. *45*, 501–512.
41. Nikpay, M., Goel, A., Won, H.H., Hall, L.M., Willenborg, C., Kanoni, S., Saleheen, D., Kyriakou, T., Nelson, C.P., Hopewell, J.C., et al.; CARDIoGRAMplusC4D Consortium (2015). A comprehensive 1,000 Genomes-based genome-wide association meta-analysis of coronary artery disease. Nat. Genet. *47*, 1121–1130.
42. Walford, G.A., Gustafsson, S., Rybin, D., Stančáková, A., Chen, H., Liu, C.T., Hong, J., Jensen, R.A., Rice, K., Morris, A.P., et al. (2016). Genome-wide association study of the modified Stumvoll insulin sensitivity index identifies BCL2 and FAM19A2 as novel insulin sensitivity loci. Diabetes *65*, 3200–3211.
43. Prokopenko, I., Poon, W., Mägi, R., Prasad B, R., Salehi, S.A., Almgren, P., Osmark, P., Bouatia-Naji, N., Wierup, N., Fall, T., et al. (2014). A central role for GRB10 in regulation of islet function in man. PLoS Genet. *10*, e1004235.
44. Manning, A.K., Hivert, M.F., Scott, R.A., Grimsby, J.L., Bouatia-Naji, N., Chen, H., Rybin, D., Liu, C.T., Bielak, L.F., Prokopenko, I., et al.; DIAbetes Genetics Replication And Meta-analysis (DIAGRAM) Consortium; and Multiple Tissue Human Expression Resource (MUTHER) Consortium (2012). A genome-wide approach accounting for body mass index identifies genetic variants influencing fasting glycemic traits and insulin resistance. Nat. Genet. *44*, 659–669.
45. Strawbridge, R.J., Dupuis, J., Prokopenko, I., Barker, A., Ahlqvist, E., Rybin, D., Petrie, J.R., Travers, M.E., Bouatia-Naji, N., Dimas, A.S., et al.; DIAGRAM Consortium; GIANT Consortium; MuTHER Consortium; CARDIoGRAM Consortium; and C4D Consortium (2011). Genome-wide association identifies nine common variants associated with fasting proinsulin levels and provides new insights into the pathophysiology of type 2 diabetes. Diabetes *60*, 2624–2634.
46. Soranzo, N., Sanna, S., Wheeler, E., Gieger, C., Radke, D., Dupuis, J., Bouatia-Naji, N., Langenberg, C., Prokopenko, I., Stolerman, E., et al.; WTCCC (2010). Common variants at 10 genomic loci influence hemoglobin A_1(C) levels via glycemic and nonglycemic pathways. Diabetes *59*, 3229–3239.
47. Saxena, R., Hivert, M.F., Langenberg, C., Tanaka, T., Pankow, J.S., Vollenweider, P., Lyssenko, V., Bouatia-Naji, N., Dupuis, J., Jackson, A.U., et al.; GIANT consortium; and MAGIC investigators (2010). Genetic variation in GIPR influences the glucose and insulin responses to an oral glucose challenge. Nat. Genet. *42*, 142–148.
48. Dupuis, J., Langenberg, C., Prokopenko, I., Saxena, R., Soranzo, N., Jackson, A.U., Wheeler, E., Glazer, N.L., Bouatia-Naji, N., Gloyn, A.L., et al.; DIAGRAM Consortium; GIANT Consortium; Global BPgen Consortium; Anders Hamsten on behalf of Procardis Consortium; and MAGIC investigators (2010). New genetic loci implicated in fasting glucose homeostasis and their impact on type 2 diabetes risk. Nat. Genet. *42*, 105–116.
49. Zhu, Z., Zhang, F., Hu, H., Bakshi, A., Robinson, M.R., Powell, J.E., Montgomery, G.W., Goddard, M.E., Wray, N.R., Visscher, P.M., and Yang, J. (2016). Integration of summary data from GWAS and eQTL studies predicts complex trait gene targets. Nat. Genet. *48*, 481–487.
50. Millstein, J., Zhang, B., Zhu, J., and Schadt, E.E. (2009). Disentangling molecular relationships with a causal inference test. BMC Genet. *10*, 23.
51. Parks, B.W., Nam, E., Org, E., Kostem, E., Norheim, F., Hui, S.T., Pan, C., Civelek, M., Rau, C.D., Bennett, B.J., et al. (2013). Genetic control of obesity and gut microbiota composition in response to high-fat, high-sucrose diet in mice. Cell Metab. *17*, 141–152.
52. Parks, B.W., Sallam, T., Mehrabian, M., Psychogios, N., Hui, S.T., Norheim, F., Castellani, L.W., Rau, C.D., Pan, C., Phun, J., et al. (2015). Genetic architecture of insulin resistance in the mouse. Cell Metab. *21*, 334–346.
53. Castellani, L.W., Nguyen, C.N., Charugundla, S., Weinstein, M.M., Doan, C.X., Blaner, W.S., Wongsiriroj, N., and Lusis, A.J. (2008). Apolipoprotein AII is a regulator of very low density lipoprotein metabolism and insulin resistance. J. Biol. Chem. *283*, 11633–11644.
54. Langfelder, P., and Horvath, S. (2012). Fast R functions for robust correlations and hierarchical clustering. J. Stat. Softw. *46*.
55. Welter, D., MacArthur, J., Morales, J., Burdett, T., Hall, P., Junkins, H., Klemm, A., Flicek, P., Manolio, T., Hindorff, L., and Parkinson, H. (2014). The NHGRI GWAS Catalog, a curated resource of SNP-trait associations. Nucleic Acids Res. *42*, D1001–D1006.
56. GTEx Consortium (2015). Human genomics. The Genotype-Tissue Expression (GTEx) pilot analysis: Multitissue gene regulation in humans. Science *348*, 648–660.
57. Thorleifsson, G., Walters, G.B., Gudbjartsson, D.F., Steinthorsdottir, V., Sulem, P., Helgadottir, A., Styrkarsdottir, U., Gretarsdottir, S., Thorlacius, S., Jonsdottir, I., et al. (2009). Genome-wide association yields new sequence variants at seven loci that associate with measures of obesity. Nat. Genet. *41*, 18–24.
58. Kathiresan, S., Willer, C.J., Peloso, G.M., Demissie, S., Musunuru, K., Schadt, E.E., Kaplan, L., Bennett, D., Li, Y., Tanaka, T., et al. (2009). Common variants at 30 loci contribute to polygenic dyslipidemia. Nat. Genet. *41*, 56–65.
59. Levy, D., Ehret, G.B., Rice, K., Verwoert, G.C., Launer, L.J., Dehghan, A., Glazer, N.L., Morrison, A.C., Johnson, A.D., Aspelund, T., et al. (2009). Genome-wide association study

of blood pressure and hypertension. Nat. Genet. *41*, 677–687.
60. Voight, B.F., Scott, L.J., Steinthorsdottir, V., Morris, A.P., Dina, C., Welch, R.P., Zeggini, E., Huth, C., Aulchenko, Y.S., Thorleifsson, G., et al.; MAGIC investigators; and GIANT Consortium (2010). Twelve type 2 diabetes susceptibility loci identified through large-scale association analysis. Nat. Genet. *42*, 579–589.
61. Speliotes, E.K., Willer, C.J., Berndt, S.I., Monda, K.L., Thorleifsson, G., Jackson, A.U., Lango Allen, H., Lindgren, C.M., Luan, J., Mägi, R., et al.; MAGIC; and Procardis Consortium (2010). Association analyses of 249,796 individuals reveal 18 new loci associated with body mass index. Nat. Genet. *42*, 937–948.
62. Heid, I.M., Jackson, A.U., Randall, J.C., Winkler, T.W., Qi, L., Steinthorsdottir, V., Thorleifsson, G., Zillikens, M.C., Speliotes, E.K., Mägi, R., et al.; MAGIC (2010). Meta-analysis identifies 13 new loci associated with waist-hip ratio and reveals sexual dimorphism in the genetic basis of fat distribution. Nat. Genet. *42*, 949–960.
63. Teslovich, T.M., Musunuru, K., Smith, A.V., Edmondson, A.C., Stylianou, I.M., Koseki, M., Pirruccello, J.P., Ripatti, S., Chasman, D.I., Willer, C.J., et al. (2010). Biological, clinical and population relevance of 95 loci for blood lipids. Nature *466*, 707–713.
64. Kooner, J.S., Saleheen, D., Sim, X., Sehmi, J., Zhang, W., Frossard, P., Been, L.F., Chia, K.S., Dimas, A.S., Hassanali, N., et al.; DIAGRAM; and MuTHER (2011). Genome-wide association study in individuals of South Asian ancestry identifies six new type 2 diabetes susceptibility loci. Nat. Genet. *43*, 984–989.
65. C4D Consortium (2011). A genome-wide association study in Europeans and South Asians identifies five new loci for coronary artery disease. Nat. Genet. *43*, 339–344.
66. Morris, A.P., Voight, B.F., Teslovich, T.M., Ferreira, T., Segrè, A.V., Steinthorsdottir, V., Strawbridge, R.J., Khan, H., Grallert, H., Mahajan, A., et al.; Wellcome Trust Case Control Consortium; Meta-Analyses of Glucose and Insulin-related traits Consortium (MAGIC) Investigators; Genetic Investigation of ANthropometric Traits (GIANT) Consortium; Asian Genetic Epidemiology Network–Type 2 Diabetes (AGEN-T2D) Consortium; South Asian Type 2 Diabetes (SAT2D) Consortium; and DIAbetes Genetics Replication And Meta-analysis (DIAGRAM) Consortium (2012). Large-scale association analysis provides insights into the genetic architecture and pathophysiology of type 2 diabetes. Nat. Genet. *44*, 981–990.
67. Kettunen, J., Tukiainen, T., Sarin, A.P., Ortega-Alonso, A., Tikkanen, E., Lyytikäinen, L.P., Kangas, A.J., Soininen, P., Würtz, P., Silander, K., et al. (2012). Genome-wide association study identifies multiple loci influencing human serum metabolite levels. Nat. Genet. *44*, 269–276.
68. Sabater-Lleal, M., Huang, J., Chasman, D., Naitza, S., Dehghan, A., Johnson, A.D., Teumer, A., Reiner, A.P., Folkersen, L., Basu, S., et al.; VTE Consortium; STROKE Consortium; Wellcome Trust Case Control Consortium 2 (WTCCC2); C4D Consortium; and CARDIoGRAM Consortium (2013). Multiethnic meta-analysis of genome-wide association studies in >100 000 subjects identifies 23 fibrinogen-associated Loci but no strong evidence of a causal association between circulating fibrinogen and cardiovascular disease. Circulation *128*, 1310–1324.
69. Richards, J.B., Waterworth, D., O'Rahilly, S., Hivert, M.F., Loos, R.J., Perry, J.R., Tanaka, T., Timpson, N.J., Semple, R.K., Soranzo, N., et al.; GIANT Consortium (2009). A genome-wide association study reveals variants in ARL15 that influence adiponectin levels. PLoS Genet. *5*, e1000768.
70. Brown, M.L., Bonomi, L., Ungerleider, N., Zina, J., Kimura, F., Mukherjee, A., Sidis, Y., and Schneyer, A. (2011). Follistatin and follistatin like-3 differentially regulate adiposity and glucose homeostasis. Obesity (Silver Spring) *19*, 1940–1949.
71. Braga, M., Reddy, S.T., Vergnes, L., Pervin, S., Grijalva, V., Stout, D., David, J., Li, X., Tomasian, V., Reid, C.B., et al. (2014). Follistatin promotes adipocyte differentiation, browning, and energy metabolism. J. Lipid Res. *55*, 375–384.
72. Lusis, A.J., Seldin, M.M., Allayee, H., Bennett, B.J., Civelek, M., Davis, R.C., Eskin, E., Farber, C.R., Hui, S., Mehrabian, M., et al. (2016). The Hybrid Mouse Diversity Panel: a resource for systems genetics analyses of metabolic and cardiovascular traits. J. Lipid Res. *57*, 925–942.
73. Huang, W., Sherman, B.T., and Lempicki, R.A. (2009). Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. Nat. Protoc. *4*, 44–57.
74. Mason, R.R., and Watt, M.J. (2015). Unraveling the roles of PLIN5: linking cell biology to physiology. Trends Endocrinol. Metab. *26*, 144–152.
75. LeibundGut-Landmann, S., Waldburger, J.M., Krawczyk, M., Otten, L.A., Suter, T., Fontana, A., Acha-Orbea, H., and Reith, W. (2004). Mini-review: Specificity and expression of CIITA, the master regulator of MHC class II genes. Eur. J. Immunol. *34*, 1513–1525.
76. Wong, D., Lee, W., Humburg, P., Makino, S., Lau, E., Naranbhai, V., Fairfax, B.P., Chan, K., Plant, K., and Knight, J.C. (2014). Genomic mapping of the MHC transactivator CIITA using an integrated ChIP-seq and genetical genomics approach. Genome Biol. *15*, 494.
77. Zhang, X., Gierman, H.J., Levy, D., Plump, A., Dobrin, R., Goring, H.H., Curran, J.E., Johnson, M.P., Blangero, J., Kim, S.K., et al. (2014). Synthesis of 53 tissue and cell line expression QTL datasets reveals master eQTLs. BMC Genomics *15*, 532.
78. Guernsey, D.L., Jiang, H., Campagna, D.R., Evans, S.C., Ferguson, M., Kellogg, M.D., Lachance, M., Matsuoka, M., Nightingale, M., Rideout, A., et al. (2009). Mutations in mitochondrial carrier family gene SLC25A38 cause nonsyndromic autosomal recessive congenital sideroblastic anemia. Nat. Genet. *41*, 651–653.
79. Kannengiesser, C., Sanchez, M., Sweeney, M., Hetet, G., Kerr, B., Moran, E., Fuster Soler, J.L., Maloum, K., Matthes, T., Oudot, C., et al. (2011). Missense SLC25A38 variations play an important role in autosomal recessive inherited sideroblastic anemia. Haematologica *96*, 808–813.
80. Wong, W.S., Wong, H.F., Cheng, C.K., Chang, K.O., Chan, N.P., Ng, M.H., and Wong, K.F. (2015). Congenital sideroblastic anaemia with a novel frameshift mutation in SLC25A38. J. Clin. Pathol. *68*, 249–251.
81. Ijichi, N., Ikeda, K., Horie-Inoue, K., Yagi, K., Okazaki, Y., and Inoue, S. (2007). Estrogen-related receptor alpha modulates the expression of adipogenesis-related genes during adipocyte differentiation. Biochem. Biophys. Res. Commun. *358*, 813–818.
82. Weiss, B., and Stoffel, W. (1997). Human and murine serine-palmitoyl-CoA transferase—Cloning, expression and characterization of the key enzyme in sphingolipid synthesis. Eur. J. Biochem. *249*, 239–247.
83. Samad, F., Hester, K.D., Yang, G., Hannun, Y.A., and Bielawski, J. (2006). Altered adipose and plasma sphingolipid metabolism

in obesity: A potential mechanism for cardiovascular and metabolic risk. Diabetes *55*, 2579–2587.

84. Arrio-Dupont, M., Coulet, P.R., and Gautheron, D.C. (1985). Coupled reaction of immobilized aspartate aminotransferase and malate dehydrogenase. A plausible model for the cellular behaviour of these enzymes. Biochim. Biophys. Acta *829*, 58–68.
85. Najafabadi, H.S., Mnaimneh, S., Schmitges, F.W., Garton, M., Lam, K.N., Yang, A., Albu, M., Weirauch, M.T., Radovani, E., Kim, P.M., et al. (2015). C2H2 zinc finger proteins greatly expand the human regulatory lexicon. Nat. Biotechnol. *33*, 555–562.
86. Maurano, M.T., Humbert, R., Rynes, E., Thurman, R.E., Haugen, E., Wang, H., Reynolds, A.P., Sandstrom, R., Qu, H., Brody, J., et al. (2012). Systematic localization of common disease-associated variation in regulatory DNA. Science *337*, 1190–1195.
87. Nicolae, D.L., Gamazon, E., Zhang, W., Duan, S., Dolan, M.E., and Cox, N.J. (2010). Trait-associated SNPs are more likely to be eQTLs: Annotation to enhance discovery from GWAS. PLoS Genet. *6*, e1000888.
88. Talukdar, H.A., Foroughi Asl, H., Jain, R.K., Ermel, R., Ruusalepp, A., Franzén, O., Kidd, B.A., Readhead, B., Giannarelli, C., Kovacic, J.C., et al. (2016). Cross-tissue regulatory gene networks in coronary artery disease. Cell Syst. *2*, 196–208.
89. Tyler, A.L., Donahue, L.R., Churchill, G.A., and Carter, G.W. (2016). Weak epistasis generally stabilizes phenotypes in a mouse intercross. PLoS Genet. *12*, e1005805.
90. Bouchard, C., Tremblay, A., Després, J.P., Nadeau, A., Lupien, P.J., Thériault, G., Dussault, J., Moorjani, S., Pinault, S., and Fournier, G. (1990). The response to long-term overfeeding in identical twins. N. Engl. J. Med. *322*, 1477–1482.
91. Zhang, X., Johnson, A.D., Hendricks, A.E., Hwang, S.J., Tanriverdi, K., Ganesh, S.K., Smith, N.L., Peyser, P.A., Freedman, J.E., and O'Donnell, C.J. (2014). Genetic associations with expression for genes implicated in GWAS studies for atherosclerotic cardiovascular disease and blood phenotypes. Hum. Mol. Genet. *23*, 782–795.